IN PRESS: JEP-HPP, 07/04/2024

Measuring learning and attention to irrelevant distractors in contextual cueing

Miguel A. Vadillo[1], Patricia Aniento[1], David Hernández-Gutiérrez[1,2],
Luca Saini[1,3], & M. Pilar Aivar[1]

[1]Department of Basic Psychology, Universidad Autónoma de Madrid, Spain
[2]Basque Center on Cognition, Brain and Language, Spain
[3]Department of Psychology, Albert-Ludwigs-Universität Freiburg, Germany

Mailing address:
Miguel A. Vadillo
Departamento de Psicología Básica
Facultad de Psicología
Universidad Autónoma de Madrid
28049 Madrid, Spain
e-mail: miguel.vadillo@uam.es

Autor note: MAV was supported by grant 2020-5A/SOC-19723 from Comunidad de Madrid, and grants PID2020-118583GB-I00 and CNS2022-135346 from Agencia Estatal de Investigación and European Regional Development Fund. MPA was supported by grant PID2021-125162NB-I00 from Agencia Estatal de Investigación and European Regional Development Fund. We would like to thank David Shanks for his valuable comments to an earlier version of this article. Correspondence concerning this article should be addressed to Miguel A. Vadillo, Departamento de Psicología Básica, Facultad de Psicología, Universidad Autónoma de Madrid, 28049-Madrid, Spain. E-mail: miguel.vadillo@uam.es

Abstract

Visual search usually improves with repeated exposure to a search display. Previous research suggests that such "contextual cueing" effect may be supported even by aspects of the search display that participants have been explicitly asked to ignore. Based on this evidence, it has been suggested that the development of contextual cueing over trials does not depend on selective attention. In the present series of experiments, we show that the most common strategy used to prevent participants from paying attention to the task-irrelevant distractors often results in suboptimal selection. Specifically, we show that visual search is slower when search displays include many irrelevant distractors. Eye-tracking data shows that this happens, at least in part, because participants fixate on them. These results cast doubts on previous demonstrations that contextual cueing is independent from selective attention.

Keywords: Contextual cueing; Implicit learning; Selective attention; Visual search

Public Significance Statement

Can we learn about stimuli we are trying to ignore? Previous research suggests that, occasionally, people learn statistical relations between visual stimuli they have been instructed to ignore, even if paying attention to them is counterproductive. This is consistent with the common assumption that some types of learning are relatively automatic and require no attention. However, can we be sure that people actually ignore stimuli just because they are instructed to do so? In the present study we show that participants do pay some attention to task-irrelevant stimuli, putting a question mark on previous studies concluding that learning does not require attention.

Not everything that happens around us is relevant or interesting. Statistical regularities in the environment often provide an excellent guide to decide how to make the best use of our limited attentional resources. Over the last couple of decades, an extensive body of research has shown that visual search is extremely sensitive to such regularities (e.g., Chun & Turk-Browne, 2007; Pearson et al., 2022; Theeuwes et al., 2022). For instance, the detection and identification of visual targets is facilitated when they appear in familiar contexts (Biederman, 1972; Chun & Jiang, 1998) or when they are presented in predictable locations (Geng & Behrmann, 2005, Jiang et al., 2013). In the same vein, salient distractors are suppressed more effectively when their likely locations can be predicted (Wang & Theeuwes, 2018). The scientific consensus is that our ability to detect regularities in visual input and use them to improve visual search relies on automatic processes that demand few cognitive resources and give rise to hard-wired attentional habits that, once established, are difficult to override by controlled processes (Awh et al. 2012; Goujon et al., 2015; Jiang, 2018; Jiang & Sisk, 2019; Pollmann, 2019; Theeuwes, 2018).

The present study focuses on a particular experimental task that has become a popular paradigm to explore visual statistical learning, namely, *contextual cueing* of visual search (Chun & Jiang, 1998). In a typical experiment, participants are instructed to find a T-shaped target among a series of L-shaped distractors over hundreds of trials. Unknown to them, some search displays are presented several times during the experiment, while others are completely random arrangements of distractors and targets. Eventually, participants seem to detect this regularity, as revealed by the fact that search times are significantly faster for repeated search displays. It is unclear whether participants become aware of this information though. When debriefed at the end of the experiment, many participants claim that they did not notice the repetitions and they show a remarkably poor performance if asked to tell repeated displays from random lures (Chun & Jiang, 1998, 2003; Colagiuri & Livesey, 2016). Research suggests that contextual cueing shows other features of automaticity. In particular, contextual cueing is still observed if participants are distracted by a secondary, but cognitively demanding, task while they complete the visual search task (Vicente-Conesa et al., 2022; Vickery et al., 2010).

Most shockingly, some experimental results suggest that contextual cueing can take place even for stimuli that participants are actively trying to ignore. The question of whether contextual cueing depends on selective attention was initially explored in two, now classic, studies by Jiang and Chun (2001) and Jiang and Leung (2005). The first of these studies presented participants with visual search displays including stimuli in two colors. Participants were instructed that the target would always appear in one of these colors (e.g., red) and that consequently they did not need to pay attention to the stimuli presented in the other color (e.g., green), which would always be distractors. This procedure allowed Jiang and Chun (2001) to independently manipulate whether only attended distractors, only ignored distractors[1] or both predicted the location of the target. One of their experiments suggested that, counterintuitively, participants could show a contextual cueing effect even for the *ignored* set of distractors, although the effect was much smaller in magnitude that the contextual cueing effect elicited by the attended set of distractors.

Jiang and Chun (2001) hypothesized that the reason for the small contextual cueing effect observed for ignored distractors might be that the task was too easy for

[1] Throughout the manuscript, we refer to the set of distractors that participants are asked to ignore as "ignored distractors" or "irrelevant distractors". We use these labels because they are easy to understand and because they are consistent with the terminology used in previous studies approaching this question. But note that, strictly speaking, these stimuli may not be completely "ignored", in the sense that participants might pay attention to them even if they are instructed not to do so. And they are not completely "irrelevant" either, in the sense that in some experimental conditions they are actually predictive of the target location.

participants, allowing them to pay some residual attention even to the irrelevant set of distractors (Lavie, 1995). Consistent with this hypothesis, in a subsequent experiment using a slightly more demanding version of the task, they found no evidence whatsoever of contextual cueing for the ignored distractors. Taken alone, the results of this study suggest that irrelevant distractors can sometimes elicit a small contextual cueing effect, possibly resulting from residual attention spilling over to irrelevant stimuli. In other words, people seem to learn about irrelevant stimuli because, contrary to the experimental instructions, they might not succeed at filtering them out. It follows from this that contextual cueing does require some amount of selective attention to distractors.

A second study by Jiang and Leung (2005) reached a quite different conclusion. Following up on previous research with other implicit learning paradigms (Frensch et al., 1998), Jiang and Leung (2005) observed that the fact that contextual cueing was small or non-existent for irrelevant distractors did not necessarily imply that participants were not learning anything about those stimuli. It is possible that attention is necessary for the *expression* of the contextual cueing effect but not for the (latent) *learning* process underlying it. Following this logic, even if participants have been asked to ignore some stimuli, it should still be possible to uncover evidence of learning as soon as participants are invited to pay attention to these distractors. To test this prediction, Jiang and Leung (2005) run essentially a replication of the studies reported by Jiang and Chun (2001) with the only exception that after the training stage, the colors of distractors were swapped in some trials, so that the distractors that had previously been presented in the unattended color were now attended and vice versa. Consistent with their hypothesis, they found that the previously ignored distractors elicited a significant contextual cueing effect as soon as they were presented in the attended color. Selective attention seemed to be important only for the expression of contextual cueing, which could nevertheless be learned latently even without participants' paying attention to distractors. This conclusion dovetailed with the results of experiments conducted in other implicit learning paradigms, also suggesting that lack of attention might prevent the expression but not the acquisition of implicit knowledge (Frensch et al., 1998, 1998). It is also consistent with the prevalent view of contextual cueing as a highly automatic type of learning (Goujon et al., 2015; Jiang, 2018).

This leaves us with two different (but not necessarily exclusive) explanations for why contextual cueing is sometimes observed for task-irrelevant distractors: (1) participants might accidentally pay some residual attention to them; and (2) the learning process underlying contextual cueing might not require selective attention to stimuli. In a previous high-powered preregistered study conducted in our laboratory (Vadillo et al., 2020), we have explored the replicability of the empirical findings supporting these claims. Our results revealed some evidence of contextual cueing for irrelevant distractors, although the size of the effect was relatively small and did not reach statistical significance in all the experiments. In contrast, we did not find any evidence of latent learning; that is, swapping the colors of distractors, so that previously ignored stimuli became fully attended, did not uncover any evidence of learning. While our failure to find evidence of latent learning casts doubts on explanation (2), our results do not provide positive evidence in favor of explanation (1) either. It is still possible that the small amount of contextual cueing elicited by irrelevant distractors is driven by automatic processes that operate in the absence of selective attention. But it could also be the case that distractors receive attention. Therefore (1) remains a plausible explanation. Previous research shows, for instance, that distractor inhibition is often a suboptimal and costly process (Cunningham & Egeth, 2016; Moher & Egeth, 2012).

The goal of the present series of experiments was to directly test explanation (1) by taking different measures of attention to irrelevant stimuli and confirming whether or not participants succeed at filtering them out. At the same time, the present experiments provide yet another opportunity to confirm whether task-irrelevant distractors can also support contextual cueing, either directly or “latently”.

Experiments 1-3 were essentially conceptual replications of Jiang and Leung (2005) and Vadillo et al. (2020), with the only exception that we manipulated the set size of the irrelevant distractors. We reasoned that if participants were able to completely filter out irrelevant distractors, then the number of items presented in the irrelevant color should make no difference in their response times (Kaptein et al., 1995). In addition, in Experiment 4 we also tracked participants’ eye movements while they were completing a simplified version of the general task employed in Experiments 1-3. Our intention, in this case, was to obtain a direct measure of overt attention given to irrelevant distractors, on the assumption that fixating an item implies a certain level of attention to it. To foreshadow, our results revealed some evidence of contextual cueing for task-irrelevant stimuli, although we failed to find any clear evidence of latent learning. Most importantly, we found that the number of stimuli presented in the irrelevant color made a significant difference in search times, suggesting that participants were not able to completely ignore these stimuli. Eye-tracking data confirmed that participants needed to make more fixations to find the target when the search display included many distractors in the irrelevant color and that some of these fixations were on irrelevant items. Overall, our results suggest that task-irrelevant visual distractors are not completely ignored and, therefore, contextual cueing for irrelevant distractors need not imply that learning can take place in the absence of selective attention.

Experiment 1

As in Jiang and Chun (2001), Jiang and Leung (2005) and Vadillo et al. (2020), Experiment 1 manipulated the amount of attention paid to distractors by presenting search displays with stimuli in two different colors and instructing participants that the target would always be presented in one of them, so that distractors in the other color could be safely ignored. Specifically, participants were presented with four types of search displays that were randomly intermixed. In condition *Both-Old*, all the elements of the search display appeared repeatedly in the same position over blocks of trials. In condition *Attended-Old*, only the distractors in the relevant color kept their position over repetitions of the search display, while the distractors in the ignored color appeared in completely random locations on each block. Conversely, in condition *Ignored-Old*, the distractors in the irrelevant color appeared in the same location over repetitions of the search display, while distractors in the relevant color appeared in random locations. Finally, in condition *Both-New*, all distractors appeared in random locations. **Figure 1** shows an example of search displays in each of these experimental conditions, taken from Experiment 1. If participants cannot learn about task-irrelevant distractors (i.e., because they are attentionally inhibited), then the repetition of the ignored context should make no difference in search times; only the repetition of the attended context (i.e., in *Attended-Old* and *Both-Old* conditions) should facilitate search times. In contrast, if participants do learn about task-irrelevant distractors, then search times should be faster in condition *Ignored-Old* than in condition *Both-New*, and faster in condition *Both-Old* than in condition *Attended-Old*. In other words, a main effect of ignored context on search times would indicate that instructing participants to ignore some distractors does not completely prevent learning about them.

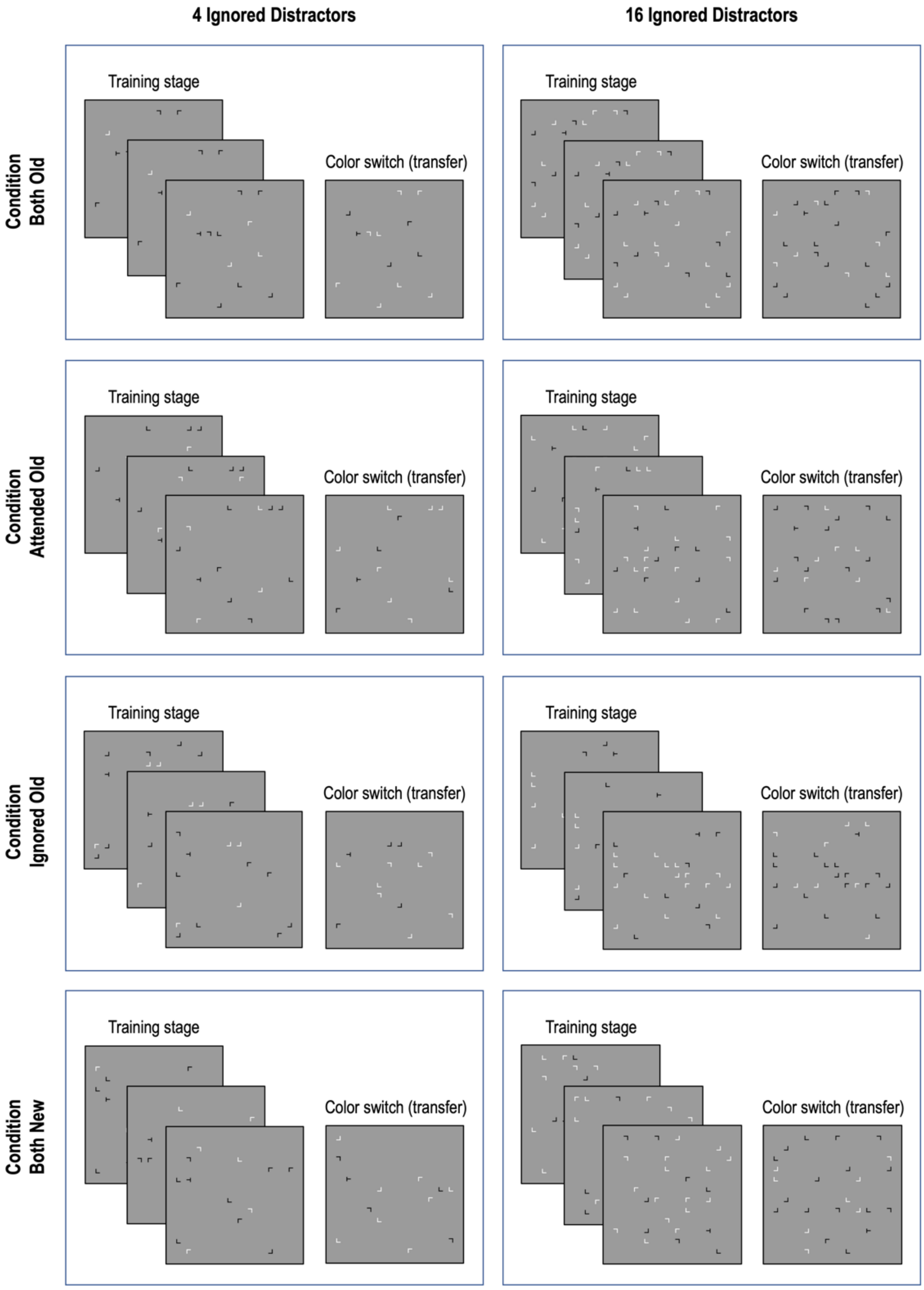

Figure 1. Example of search displays used in Experiment 1. In this particular example, the attended color is black, but for half of the participants the attended color was white.

Orthogonally, in one half of the trials the search displays included only four distractors in the irrelevant color, while in the other half of the trials there were 16 distractors in the irrelevant color. The number of relevant distractors was kept constant over trials. If participants are able to completely ignore the irrelevant distractors, as instructed by the task, then the number of items presented in the irrelevant color should make no difference in search times, since set size for relevant items does not change. Alternatively, faster search times for search displays with fewer irrelevant distractors would be indicative of participants paying some attention to irrelevant distractors.

Finally, as in Jiang and Leung (2005), Experiment 1 included a *transfer stage* where half of the trials were identical to those in the training stage (i.e., condition Color Stay), while in the other half of the trials, the colors of distractors were swapped (condition Color Switch), so that previously irrelevant distractors now became relevant and vice versa. This allowed us to test the hypothesis that, even if participants fail to show any evidence of learning for ignored distractors during the training stage, they may still have learned "latently" about them. Such knowledge might then be expressed if participants are forced to pay attention to these stimuli, by presenting them in the task-relevant color. Following Jiang and Leung (2005), a successful demonstration of these latent-learning effect would imply finding evidence of contextual cueing in the *Ignored-Old* condition relative to the control *Both-New* condition in the Color Switch trials.

Method

Transparency and Openness

All the experiments reported in the present article comply with the TOP guidelines. All materials, data, and scripts are publicly available at the Open Science Framework (Vadillo et al., 2023). The methods and analysis plan of Experiments 1-3 were preregistered before any data collection took place. The registered protocols are publicly available at https://osf.io/964na, https://osf.io/5htym, and https://osf.io/7js3d, respectively. The method sections in the present article merely paraphrase the information provided in the registered protocols. Unless noted otherwise, all the analyses and data preprocessing followed the protocol. The data were collected between October 2020 and June 2022.

Participants

To the best of our knowledge, no previous experiment has manipulated set size of the ignored distractors in a contextual cueing experiment. Therefore, we could not plan our sample size based on this effect. However, it was possible to conduct a power analysis based on the effect of repeating the ignored context in previous experiments. In the training stage of Experiment 3 by Vadillo et al. (2020), we found a small (and indeed non-significant) effect of Ignored Context, $F(1, 46) = 3.93$, $p = .053$, with an effect size of $d_z = 0.29$. Using G*Power, we estimated that 96 participants would be needed to replicate this effect with .80 power in a two-tailed t-test with $\alpha = .05$. Considering that some participants might fail to meet the selection criteria described below, we decided to collect data from 100 participants in Experiment 1. Participants were psychology students from UAM, who were rewarded with course credit for their contribution. On average, participants were 19.36 years old ($SD = 1.43$) and 85% of them were female. All participants conducted the experiment in small groups in a laboratory equipped with 12 individual cubicles. The study was approved by the UAM ethics committee (ref. CEI-80-1473) and all participants provided informed consent.

Stimuli and Apparatus

The method of Experiment 1 was identical to Experiment 3 of Vadillo et al. (2020), except for the manipulation of the set size of the ignored context. On each trial, participants were shown a search display with a total of 12 or 24 L-shaped distractors (depending on set size condition) and one T-shaped target presented against a gray background. All search displays contained eight distractors in the attended color (two per quadrant) and either four or sixteen distractors in the ignored color (i.e., one or four per quadrant). In all cases, distractors were L-shaped stimuli, which could be rotated 0º, 90º, 180º and 270º. The T-shaped target was always rotated 90º or 270º and was presented in the same color across trials (black or white, randomly chosen for each participant). Distractors and targets were positioned in a 12 × 12 grid, invisible to participants. At the beginning of the experiment, 32 locations (eight per quadrant) of the grid, roughly equidistant from the center of the screen, were preselected to contain the targets. Distractors never appeared in these locations. Stimuli were presented in 18.5-inche computer monitors driven at 1366 × 768-pixel resolution. Each cell of the 12 × 12 grid spanned 60 pixels (180 mm) square. Distractors and targets occupied 24 pixels (72 mm) square and were always positioned in the center of their cell. This means that the empty space between two adjacent stimuli (distractors or targets) was at least 108 mm. Stimulus presentation and response collection were controlled with Matlab (The MathWorks, Natick, MA) with the Psychophysics Toolbox extension (Brainard, 1997; Kleiner, Brainard & Pelli, 2007; Pelli, 1997).

Procedure and Design

Participants were instructed to search for the target as fast as possible and press key <z> if the stem of the T pointed to the left and <m> if the stem of the T pointed to the right. Instructions encouraged them to be as fast as possible, but without making errors. Before starting the experiment, participants were told that the target would always be presented in one color (black or white) and that, to improve their performance, they should ignore all the stimuli presented in the other color.

The experiment began with a training stage consisting of 24 blocks of trials, each of them comprising 32 trials. Each block contained eight search displays for each of the four experimental conditions: *Both-Old*, *Attended-Old*, *Ignored-Old*, and *Both-New*. The eight search displays in the *Both-Old* condition were presented repeatedly over the experiment, once per block. In the *Attended-Old* condition, only the distractors presented in the same color as the target were presented in the same location and orientation across blocks. Distractors presented in the ignored color were presented in random locations across blocks. In contrast, in the *Ignored-Old* condition, only the distractors presented in the ignored color were presented in the same location and orientation across blocks, while distractors in the attended color were presented in random locations across blocks. Finally, in the *Both-New* condition all the distractors, regardless of color, were presented in random locations. Orthogonally, half of search displays in each condition contained 4 distractors in the ignored context and the other half contained 16 distractors in the ignored context. The left/right orientation of the target was determined randomly in each trial, so that participants could not learn a direct association between patterns and responses.

Immediately after the training stage and without any interruption, participants completed two transfer blocks, each comprising 64 trials. On each block, participants were presented with the same 32 search displays used during the training stage (i.e., Color Stay trials), in addition to 32 new search displays created by reversing the colors of the distractors (i.e., Color Switch trials).

Each trial began with a 1-sec fixation cross presented at the center of the screen, followed by the search display, which remained visible until participants responded pressing either <z> or <m> in the computer keyboard. After an incorrect response, the message "Wrong!" appeared on the screen for two seconds. Trials were separated by a 1-sec blank screen. Participants were given the opportunity to make a 20-second pause after every 100 trials.

Results and Discussion

Data Pre-processing

Following the preregistered protocol, two participants in Experiment 1 with overall accuracies below 95% in the visual search task were removed from the analyses. Trials immediately following a rest break, trials with a response time (RT) above 10 seconds and trials with incorrect responses were removed from the analyses. Then, for each participant we estimated the mean and standard deviation (SD) of the valid RTs and removed any RT departing 3 or more SDs from each participant's mean. To further reduce noise in the data, we collapsed data from adjacent blocks into two-block epochs.

Visual Search Performance during the Training Stage

The top panel of **Figure 2** shows mean RTs during the training stage of Experiment 1. As can be seen, RTs declined over the course of the experiment in all conditions. That is, participants became generally faster at finding the target as they gained experience with the visual search task. The decrease was steeper for experimental conditions where the attended context predicted the location of the target (denoted by green lines in **Figure 2**) than for experimental conditions where the attended context did not predict the location of the target (denoted by red lines). In other words, a contextual cueing effect was found for distractors presented in the attended context. In contrast, RTs were similar in experimental conditions where the ignored context predicted the location of the target (solid lines) and the experimental conditions where the ignored context was not predictive (dashed lines), suggesting that distractors in the ignored color did not elicit a strong contextual cueing effect. Perhaps the most relevant pattern is that RTs were generally slower when there were many distractors in the ignored color (top right panel) than when there were only a few (top left panel), suggesting that participants were not completely able to filter out distractors in the task-irrelevant color.

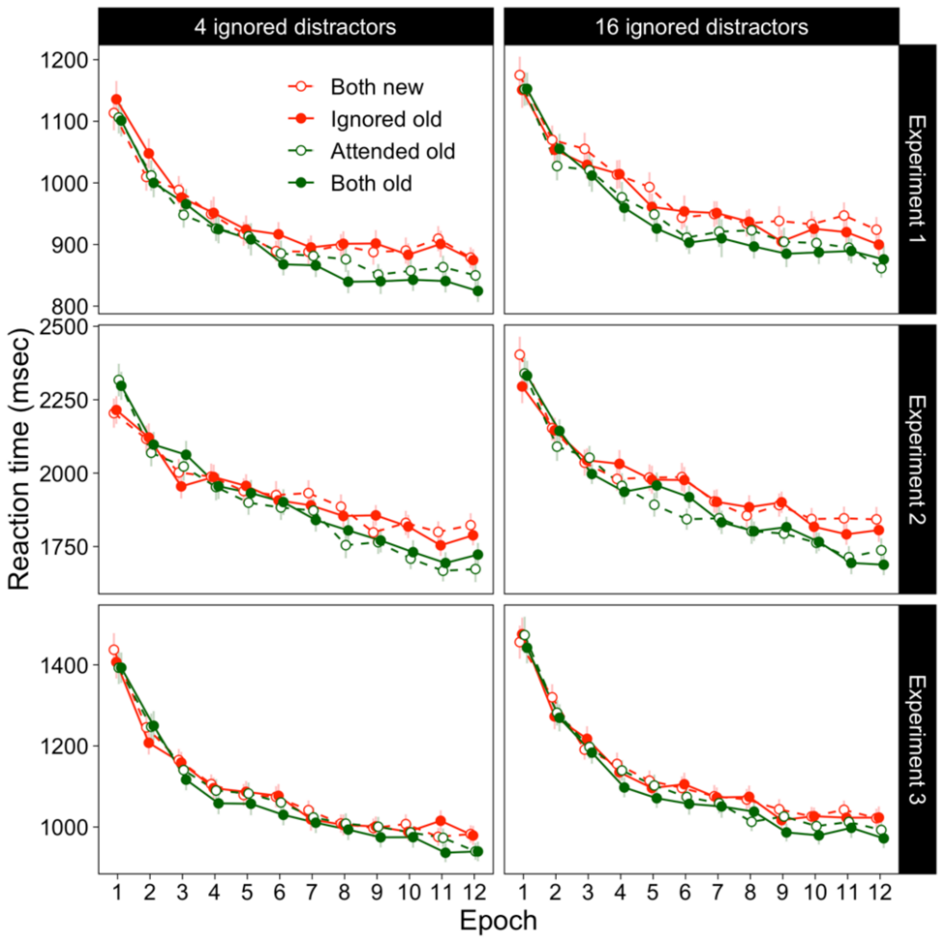

Figure 2. Reaction times during the training stage in Experiments 1-3. Error bars denote the standard error of the mean.

Following the preregistered protocol, RTs from the training stage were analyzed with a 2 (Attended Context: repeated vs. new) × 2 (Ignored Context: repeated vs. new)

× 2 (Ignored Set Size: 4 vs. 16) × 12 (Epoch) repeated-measures analysis of variance (ANOVA). For the sake of brevity, given the large number of effects and interactions, the results of the ANOVA are reported in detail in **Table S1** of the Supplementary Material. The main effects of Epoch, $F(3.75, 363.52) = 108.24$, $p < .001$, $\eta_p^2 = .527$, and Attended Context, $F(1, 97) = 73.22$, $p < .001$, $\eta_p^2 = .430$, were significant, showing that RTs declined over the course of the experiment and that, on average, participants were faster when the stimuli presented in the attended color repeated over trials.

The crucial main effect of Ignored Context did not reach significance, $F(1, 97) = 3.04$, $p = .084$, $\eta_p^2 = .030$. Given that this result is particularly relevant for the present series of experiments, to facilitate the visualization of the Ignored Context effect **Figure 3** plots search times in the two conditions where the irrelevant-distractors were predictive (i.e., *Both-Old* and *Ignored-Old*) against search times in the two conditions where the irrelevant distractors were nonpredictive (i.e., *Attended-Old* and *Both-New*). The results of Experiment 1 are plotted on the leftmost section of the figure. On average, the data points are slightly below the diagonal, which means that participants were a bit faster at finding the target when the irrelevant distractors were presented in the same locations repeatedly. The effect was very small though, barely noticeable to the unaided eye and, in fact, non-significant in this experiment. We will return to this non-significant effect in later sections of this article.

Perhaps most importantly, the main effect of Ignored Set Size was significant and relatively large, $F(1, 97) = 110.70$, $p < .001$, $\eta_p^2 = .533$. Participants were generally slower when the search display included 16 distractors in the irrelevant color (right top panel in **Figure 2**) than when there were only four distractors in the irrelevant color (left top panel in **Figure 2**), showing that they could not completely ignore these stimuli.

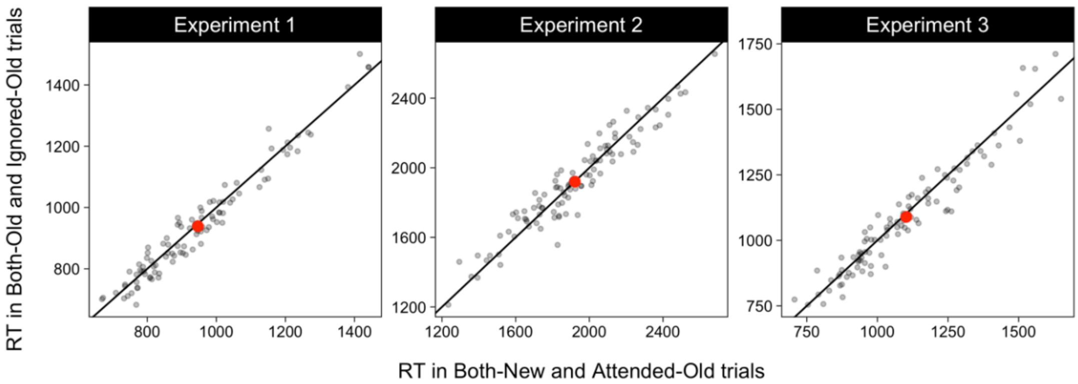

Figure 3. Reaction times during the training stage in Experiments 1-3 grouped by condition. Experimental conditions where the irrelevant distractors were nonpredictive (i.e., *Both-New* and *Attended-Old*) are plotted against experimental conditions where the irrelevant distractors were predictive (i.e., *Both-Old* and *Ignored-Old*). The grey points denote average RTs from each individual participant, while the red points denote the mean across all participants in each experiment.

Visual Search Performance during the Transfer Stage

The top panel of **Figure 4** summarizes participants' performance in the transfer stage of Experiment 1. As explained in the Method section, in Color Stay trials the task was identical to the training stage and, not surprisingly, search times in Color Stay trials (left panel) largely show the same trends observed in the previous analyses. As can be seen, participants were faster at finding the target when the attended context was predictive of the target location (green lines) compared to when the attended context was non-predictive (red lines). This suggests that, once again, a strong contextual cueing effect was observed for distractors presented in the attended color. Participants also seemed to be slightly faster at finding the target when the ignored context predicted the location of the target (solid lines) than when the ignored context was non-predictive (dashed lines). But this contextual cueing effect for ignored distractors was noticeably weaker, if present at all. Finally, and perhaps more importantly, search times were slower when the search display included 16 distractors in the irrelevant color than when it included only four, suggesting that participants did not completely ignore these distractors.

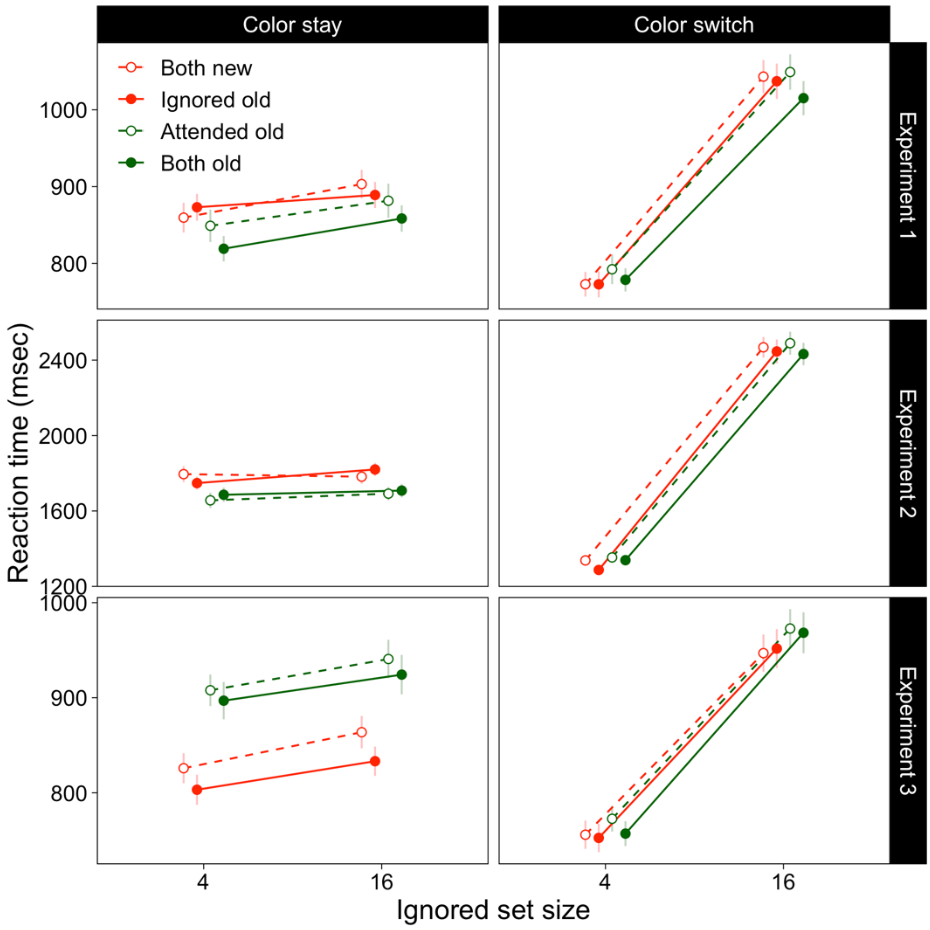

Figure 4. Reaction times during the transfer stage in Experiments 1-3. Error bars denote the standard error of the mean. Note that the difference between green lines and red lines is reversed in Color Stay trials of Experiment 3 (compared to Experiments 1 and 2) due to a programming error (described in the main text).

The top right panel of **Figure 4** shows search times in Color Switch trials in Experiment 1. As explained in the Method, these trials were identical to Color Stay trials, except that the distractors that had previously been presented in black where now presented in white and vice versa. That is, distractors that had been ignored in the training stage (and Color Stay trials) were now attended and distractors that have been attended in the training stage (and Color Stay trials) were now ignored. Of note, the labels in **Figure 4** denote whether distractors had been attended or not in the training stage. For example, in Color Switch trials, the *Ignored-Old* condition denotes trials in which previously ignored but predictive distractors were now presented in the attended color. Perhaps the most noticeable result in Color Switch trials is that the manipulation of set size made a stronger difference in search times, compared to Color Stay trials. This is a logical consequence of the fact that the colors of distractors were reversed. In the training stage and in Color Stay trials, the manipulation of set size affected distractors presented in the irrelevant color, but when colors are reversed the manipulation affects distractors in the relevant color. This has a strong impact on search times. Beyond the effect of set size, the right panel of **Figure 4** also suggests some evidence of contextual cueing for distractor previously presented in the attended color (green vs. red lines) and for distractors previously presented in the ignored color (solid vs. dashed lines), although the trend is less clear than in Color Stay trials, suggesting that, in general, reversing the color of distractors has a negative impact on contextual cueing.

Following the preregistered protocol, RTs from the transfer stage were first analyzed with a 2 (Attended Context: repeated vs. new) × 2 (Ignored Context: repeated vs. new) × 2 (Ignored set size: 4 vs. 16) × 2 (Switch: color stay vs. color switch) ANOVA. The results of this ANOVA are reported in detail in **Table S2** of the Supplementary Material. The interpretation of the results is somewhat obscured by the large number of factors, but for our present purposes the most important result is that we found large main effects of Ignored Set Size, $F(1, 97) = 448.49$, $p < .001$, $\eta_p^2 = .822$, and Switch, $F(1, 97) = 67.26$, $p < .001$, $\eta_p^2 = .409$, qualified by a significant Ignored Set Size × Switch interaction, $F(1, 97) = 331.18$, $p < .001$, $\eta_p^2 = .774$. As can be seen in the top panel of **Figure 4**, the large interaction between Switch and Set size is due to the fact that the effect of Set size was substantially larger in Color Switch than in Color Stay trials. Switch also interacted with Attended Context, $F(1, 97) = 11.56$, $p < .001$, $\eta_p^2 = .106$. Because we expected that Switch would interact with other factors, the preregistered protocol included separate follow-up ANOVAs for Color Stay and Color Switch trials.

As explained above, Color Stay trials were identical in the learning and transfer stages and, therefore, we expected the same pattern of results. An Ignored set size × Attended Context × Ignored Context ANOVA on RTs in Color Stay trials (see **Table S3**) revealed a main effect of Ignored set size, $F(1, 97) = 17.68$, $p < .001$, $\eta_p^2 = .154$, and Attended Context, $F(1, 97) = 14.78$, $p < .001$, $\eta_p^2 = .132$. As in the training stage, the main effect of Ignored Context failed to reach statistical significance, $F(1, 97) = 2.81$, $p = .097$, $\eta_p^2 = .028$. None of the interactions reached statistical significance. The same analysis on Color Switch trials returned a significant main effect of Ignored set size, $F(1, 97) = 611.53$, $p < .001$, $\eta_p^2 = .863$. No other main effects or interactions reached statistical significance (see **Table S4** in the Supplementary Materials).

Taken collectively, these results confirm that contextual cueing proceeded as expected for distractors in the attended context. In contrast, for the distractors presented in the irrelevant color contextual cueing was weak and failed to reach statistical significance, both during the training stage and on Color Stay trials of the transfer stage. Contextual cueing for irrelevant distractors was also absent in Color Switch trials, suggesting that swapping the colors of previously relevant and irrelevant distractors did not uncover any evidence of "latent learning" (Jiang & Leung, 2005). Perhaps most importantly, we found robust evidence that participants did not completely ignore distractors presented in the irrelevant color, as search times were noticeably longer when the search displays included 16 irrelevant distractors than when they included only four.

Experiment 2

It is possible that the fact that participants did not completely ignore the irrelevant distractors in Experiment 1 is due to the fact that the visual search task was relatively simple. In particular, the L-shaped distractors were depicted in such a way that it was easy for participants to discriminate them from the T-shaped targets. The top row in **Figure 5** shows the distractors used in Experiment 1. Jiang and Chun (2001) hypothesized that simple visual search tasks like this might not be sufficiently demanding to exhaust participants' attentional resources, resulting in a surplus of attention that participants can devote to processing the irrelevant distractors. Following their example, in Experiment 2 we increased the difficulty of the task by making the distractors more similar to the target. The bottom row of **Figure 5** shows the shape of the distractors used in Experiment 2.

Method

Following the same power analysis of Experiment 1, we recruited 100 participants in Experiment 2. They all performed the task in identical conditions to Experiment 1. On average, they were 19.34 years old (*SD* = 1.25) and 87% were females. The method of Experiment 2 was identical to Experiment 1, except for the shape of the distractors. While the distractors employed in Experiment 1 were easily discriminable from the target, the distractors used in Experiment 2 were relatively similar to the target, rendering the visual search task more complicated for participants (see **Figure 5**).

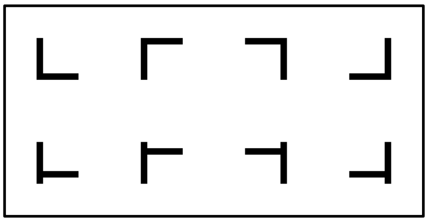

Figure 5. Set of distractors used in Experiments 1 and 3 (top row) and in Experiment 2 (bottom row).

Results and Discussion

Visual Search Performance during the Training Stage

All participants met the inclusion criteria. RTs were filtered following the same procedure as in Experiment 1. The middle row of **Figure 2** shows mean RTs during the training stage of Experiment 2. As can be seen, RTs were substantially slower than in Experiment 1, confirming that the change in the shape of the visual distractors affected the difficulty of the visual search task as expected. As in Experiment 1, RTs from the training stage were analyzed with a 2 (Attended Context: repeated vs. new) × 2 (Ignored Context: repeated vs. new) × 2 (Ignored set size: 4 vs. 16) × 12 (Epoch) repeated-measures ANOVA. The results are reported in detail in **Table S5** of the Supplementary Material. The ANOVA detected significant main effects of Attended Context, $F(1, 99) = 35.16$, $p < .001$, $\eta_p^2 = .262$, and Epoch, $F(6.39, 632.78) = 119.35$, $p < .001$, $\eta_p^2 = .547$, qualified by a significant Attended Context × Epoch interaction, $F(9.28, 918.82) = 5.46$, $p < .001$, $\eta_p^2 = .052$. These results suggest that contextual cueing for attended distractors developed gradually over the course of the training stage. As in Experiment 1, the main effect of Ignored Context was non-significant $F(1, 99) = 0.00$, $p < .967$, $\eta_p^2 < .001$. But, given its relevance for the present study, the main effect of Ignored Context in this experiment is also summarized in the central panel of **Figure 3**. As can be seen, there was no clear advantage for search displays were the irrelevant distractors were predictive of the target location and those were they were non-predictive. In other words, we did not find a contextual cueing effect for irrelevant distractors. Finally, the main effect of Ignored Set Size was significant, $F(1, 99) = 6.64$, $p = .011$, $\eta_p^2 = .063$, confirming that participants were slower at finding the target when the search display included many distractors in the irrelevant color. No other effects reached significance.

Visual Search Performance during the Transfer Stage

RTs during the transfer stage are depicted in the central panel of **Figure 4**. The full ANOVA revealed significant main effects of Ignored Set Size, $F(1, 99) = 926.26$, $p < .001$, $\eta_p^2 = .903$, Attended Context, $F(1, 99) = 6.23$, $p = .014$, $\eta_p^2 = .059$, and Switch, $F(1, 99) = 123.12$, $p < .001$, $\eta_p^2 = .554$. It also detected significant Ignored Set Size × Switch, $F(1, 99) = 672.54$, $p < .001$, $\eta_p^2 = .872$, and Attended Context × Switch interactions, $F(1, 99) = 21.81$, $p < .001$, $\eta_p^2 = .181$. All other main effects and interactions were non-significant (see **Table S6** in the Supplementary Materials). As in Experiment 1, we followed up these analyses with separate ANOVAs in Color Stay and Color Switch trials. In Color Stay trials, only the main effect of Attended Context was significant, $F(1, 99) = 29.86$, $p < .001$, $\eta_p^2 = .232$ (see **Table S7** in the Supplementary Materials). In Color Switch trials, only the main effect of Ignored Set Size was significant, $F(1, 99) = 1036.13$, $p < .001$, $\eta_p^2 = .913$ (see **Table S8** in the Supplementary Materials).

Overall, these analyses reveal a strong contextual cueing effect for attended distractors, both during the training stage and in Color Stay trials of the transfer stage. In contrast, there was no evidence whatsoever of contextual cueing for irrelevant distractors in either stage. Interestingly, despite the fact that the visual search task was more difficult in Experiment 2 than in Experiment 1, we nevertheless found that participants still paid some attention to the irrelevant distractors, as shown by the fact that the number of distractors in the irrelevant color made a significant difference in search times. Therefore, attention to irrelevant distractors does not seem to depend on having a surplus of attentional resources, since it also occurs when the task is very difficult.

Experiment 3

As in Experiment 2, in Experiment 3 we also tried to increase the difficulty of the task relative to Experiment 1, but instead of changing the shape of the distractors, we increased the number of distractors in the task-relevant color, from eight to 16. In the terminology of Reddy and VanRullen (2007), while Experiment 2 tried to exhaust participants' "attention for recognition" (i.e., their ability to recognize distractors and targets), Experiment 3 puts the stress on "attention against competition" (i.e., the ability to suppress interference from distractors appearing in the same receptive field as the target).

Method

Following the same power analysis of Experiment 1, we recruited 100 participants in Experiment 3. They performed the task in identical conditions to Experiments 1 and 2. On average, they were 19.36 years old (*SD* = 1.41) and 82% were females. The method of Experiment 3 was identical to Experiment 1, except that the visual search displays included 16 (instead of eight) distractors in the relevant color, increasing the total number of distractors on the screen (to 20 or 32).

Results and Discussion

Visual Search Performance during the Training Stage

All participants met the inclusion criteria. RTs were filtered following the same procedure as in Experiments 1 and 2. The bottom row of **Figure 2** shows mean RTs during the training stage of Experiment 3. Again, RTs were substantially slower than in Experiment 1, suggesting that, as intended, the increase in the number of distractors made the task more difficult. Note, however, that RTs were not as slow as in Experiment 2. As in Experiments 1 and 2, RTs from the training stage were analyzed with a 2 (Attended Context: repeated vs. new) × 2 (Ignored Context: repeated vs. new) × 2 (Ignored Set Size: 4 vs. 16) × 12 (Epoch) repeated-measures ANOVA. The results are reported in detail in **Table S9** of the Supplementary Material. The ANOVA revealed significant main effects of Attended Context, $F(1, 99) = 15.83$, $p < .001$, $\eta_p^2 = .138$, and Epoch, $F(2.64, 260.95) = 152.94$, $p < .001$, $\eta_p^2 = .607$, confirming that, as expected, contextual cueing was observed for distractors in the attended color and that RTs decreased over the course of the training stage. Unlike in Experiments 1 and 2, the main effect of Ignored Context was statistically significant, $F(1, 99) = 4.79$, $p = .031$, $\eta_p^2 = .046$. As can be seen in the right-most panel of **Figure 3**, RTs were slightly faster when the irrelevant distractors were predictive of the target location, although the effect is small and barely noticeable. Finally, the main effect of Ignored Set Size was significant as well, $F(1, 99) = 54.33$, $p < .001$, $\eta_p^2 = .354$, confirming that participants took more time to find the target when the search display included many distractors in the task-irrelevant color, even if they had been instructed to ignore those distractors. Among the interactions, only the Ignored Set Size × Epoch interaction approached significance.

Visual Search Performance during the Transfer Stage

The lower panel of **Figure 4** shows mean RTs during the transfer stage. After data collection was over, we discovered an error in the program of Experiment 3 that affected conditions *Both-New* and *Ignored-Old* during the transfer stage. Specifically, search displays in these conditions included only eight distractors in the attended color (instead of 16). Therefore, although the analyses reported below follow the preregistered protocol, the reader must bear in mind that the main effect of Attended Context (and any interaction involving this factor) is essentially uninterpretable in the transfer stage of Experiment 3, because the experimental conditions where the attended distractors did not repeat over trials (i.e., conditions *Both-New* and *Ignored-Old*) included fewer distractors than originally intended. Note, in any case, that this error does not affect the interpretation of the (crucial) effects of repeating the ignored context: conditions *Both-New* and *Ignored-Old* were still identical in all respects, except for the fact that the ignored distractors repeated in the latter but not in the former; and, similarly, conditions *Attended-Old* and *Both-Old* were identical except for the repetition of ignored stimuli in the latter. In other words, despite the programming error, the second stage of Experiment 3 still provides a valid test of the effects of Ignored Context, Ignored Set Size and Color Switch and, as such, we decided to retain the analyses in the present article.

The full ANOVA on RTs from the transfer stage revealed significant main effects of Attended Context, $F(1, 99) = 81.85$, $p < .001$, $\eta_p^2 = .453$, Ignored Set Size, $F(1, 99) = 215.15$, $p < .001$, $\eta_p^2 = .685$, and Switch, $F(1, 99) = 8.38$, $p < .001$, $\eta_p^2 = .078$. Most importantly, the Attended Context × Switch, $F(1, 99) = 41.36$, $p < .001$, $\eta_p^2 = .295$, and Ignored Set size × Switch interactions, $F(1, 99) = 195.90$, $p < .001$, $\eta_p^2 = .664$, were significant as well. No other main effects or interactions were significant, although the main effect of Ignored Context approached significance, $F(1, 99) = 3.30$, $p = .072$, $\eta_p^2 = .032$. As in Experiments 1 and 2, we followed up these analyses with separate ANOVAs in Color Stay and Color Switch trials. In Color Stay trials, the main effects of Attended

Context, $F(1, 99) = 104.67$, $p < .001$, $\eta_p^2 = .514$, Ignored Context, $F(1, 99) = 7.67$, $p = .007$, $\eta_p^2 = .072$, and Ignored Set Size, $F(1, 99) = 11.56$, $p < .001$, $\eta_p^2 = .105$, were significant, replicating the results observed in the training stage. In contrast, none of the interactions were significant. In Color Switch trials, only the main effects of Attended Context, $F(1, 99) = 4.89$, $p = .029$, $\eta_p^2 = .047$, and Ignored Set Size, $F(1, 99) = 368.15$, $p < .001$, $\eta_p^2 = .788$, were significant.

To sum up, as in Experiments 1 and 2, we observed a robust contextual cueing effect for distractors presented in the attended color. Unlike in Experiments 1 and 2, in Experiment 3 we found significant contextual cueing for irrelevant distractors as well. This was found both in the training stage and in Color Stay trials of the transfer stage. Swapping the color of distractors at test did not uncover any evidence of “latent” learning, though. Finally, as in Experiments 1 and 2, Experiment 3 detected a strong effect of the number of irrelevant distractors on RTs, suggesting that participants do not completely ignore these distractors.

Combined analysis of Experiments 1-3

The preregistered protocol also included combined analyses of Experiments 1-3, with experiment as an additional factor. **Tables S13-S16** in the Supplementary Materials report the details of the ANOVAs for the training stage, transfer stage, Color Stay trials and Color Switch trials, respectively. Given the programming error discussed above, Experiment 3 was excluded from all the analyses conducted on data from the transfer stage. The main goal behind these analyses was to confirm that the two manipulations implemented in Experiments 2-3 to increase the difficulty of the task made a significant difference in RTs. Consistent with our expectations, the main effect of Experiment was significant in all the analyses. Interestingly, although Experiment interacted with other factors in different analyses, it did not interact with Ignored Set Size either in the training stage or in Color Stay trials of the transfer stage. This suggests that increasing the number of irrelevant distractors in the search display had a similar effect across experiments, despite the fact that visual search was more difficult in Experiments 2 and 3 than in Experiment 1. Therefore, the effect of set size for irrelevant distractors does not seem to depend on the availability of attentional resources.

Experiment 4

Experiment 4 sought to find converging evidence that despite the experimental instructions to focus only on stimuli presented in the relevant color, participants nevertheless paid some attention to irrelevant distractors as well. In this case, we tracked participants' eye movements while they were completing the task. This allowed us to measure the average number of fixations in each experimental condition and to obtain a measure of direct fixations on the irrelevant distractors. To minimize participants' discomfort during the experiment (and also reduce calibration problems), we reduced the length of the task to 12 blocks (instead of 24), we removed the transfer stage, and we also spared conditions *Attended-Old* and *Ignored-Old*. The experimental task was otherwise identical to Experiment 1. Our predictions were that we would find a significant cueing effect (i.e., a search advantage for *Both-Old* displays compared to *Both-New* displays) in RTs and also in the number of fixations and, most importantly, that the set size of the ignored context would affect not only RTs (as in Experiments 1-3) but also the number of fixations. Unlike Experiments 1-3, Experiment 4 was conceived as an exploratory study and was not formally preregistered. The sample size was also considerably smaller, since data acquisition had to be performed individually.

Method

Participants, Stimuli, and Design

Twenty-four participants were tested for Experiment 4. This sample size was not based on a formal power analysis and is substantially smaller than the sample size of Experiments 1-3. Note, however, that it is still 50% larger than the typical contextual cueing experiment (median $N = 16$ in the systematic review conducted by Vadillo et al., 2016). Furthermore, this experiment was based on the procedure of Experiment 1, where the size of the main effect of Ignored Set Size (the crucial effect for this experiment) was a Cohen's d_z of 1.06. The power to detect an effect of this size with 24 participants with $\alpha = .05$ is .99. Even assuming a gross reduction in the size of this effect to, for instance, $d_z = 0.60$, 24 participants still afford .80 power. As explained below, some analyses relied on data from only 20 participants, due to calibration problems in four participants. The smallest effect that can be detected with .80 in this reduced sample is $d_z = 0.66$.

Participants were psychology students from UAM, who were rewarded with course credit for their contribution. One of the participants misreported their age as 97. Among the remaining participants the average age was 20.46 (*SD* = 1.50). 56% of participants were female. All participants had normal or corrected-to-normal vision. Participants were not excluded if they used glasses or contact lenses and performed the task wearing their eye correction. All of them conducted the experiment individually. The study was approved by the UAM ethics committee (ref. CEI-94-1724) and all participants provided informed consent.

Unless noted otherwise, all aspects of the stimuli, procedure and design were identical to those of Experiment 1. As explained above, the main differences were that the experiment comprised only 12 blocks of trials and did not include a transfer stage.

To reduce the number of trials even further, we also removed conditions *Attended-Old* and *Ignored-Old*, retaining only conditions *Both-New* and *Both-Old*. This yielded a total of 192 trials, but it implies that the present experiment does not allow us to measure contextual cueing separately for relevant and irrelevant distractors. We did preserve the manipulation of the Ignored Set Size as in Experiment 1, so that one half of the trials included four distractors in the irrelevant color, while the other half included 16 distractors in the irrelevant color.

Apparatus and Procedure

Eye movements were recorded using a monocular (left eye) infrared eye-tracker (ASL 6000, Applied Science Laboratories). This eye tracker has a frequency of 60Hz and a spatial error of less than 1 deg, according to manufacturers. Stimuli were displayed on a computer monitor (1280 × 1024 LCD), located 78 cm from the participant. From this viewing distance, each cell of the 12 × 12 grid containing the search display subtended approximately 1.3° of visual angle, the distractors and targets presented in the center of each cell subtended 0.53º, and the minimum distance between two adjacent stimuli was 0.80º. Head was stabilized during the experiment with a chin rest. As in Experiments 1-3, stimulus presentation and eye tracker operations were controlled with the Psychophysics Toolbox extension for Matlab. A second, standard PC computer controlled the eye tracker, and both computers communicated through a parallel port connection.

The experiment started with the calibration of the eye-tracker. For that a 9-point calibration screen was presented. Once calibrated, participants performed all the trials of the experiment in one continuous session, with just one short break after 100 trials. Since the experiment was short (each recording session lasted about 15 minutes) recalibration was not performed during the experiment. Participants with bad eye tracker data (4 cases) were not considered for the analysis of eye fixations.

Eye data analysis

Raw eye scan paths from all trials and participants were visually inspected to determine data quality. In this step data from four participants was removed due to low quality of the eye recordings. Two of these participants showed a high degree of noise and multiple artefacts (very high velocity peaks) in their recordings, while the other two had moved during the experiment, so eye position was lost for half of the trials. Eye movement data was further processed with Matlab custom made scripts. Fixation locations were determined using a simple, velocity based algorithm, based on the I-VT method (Salvucci & Goldberg, 2000). A velocity threshold of 1,272 pixels per second was used to define saccades. This approximately corresponds to a velocity of 25º/sec. Using such low value is justified in this task, since participants perform many saccades of short amplitude (Holmqvist et al., 2011). Blinks and missed values were also marked as saccades and the remaining points were marked as fixations. Groups of consecutive fixation points with a minimum of three data points were collapsed and the centroid, or geometrical mean of all the points, was taken as the location for that fixation. Fixations with fewer than three data points (shorter than 50 msec) and those that occurred outside of the computer screen were eliminated.

Visual inspection of the scan-paths also showed that in some cases there were small and systematic offsets, so a custom-made algorithm was used to correct drifts in eye position. Our algorithm was based on the idea of implicitly required fixation locations, suggested by Hornof and Halverson (2002). Given the characteristics of our task, in each trial we expected participants to look at the center of the screen at the beginning of the trial and also at the target letter in the display at the end of the trial (once the target was found). Our algorithm used the eye position detected when looking at the center of the screen right before the search screen was presented to realign the complete scan path for that trial. In a second step the algorithm checked whether this realignment allowed for a fixation in the target to be now detected. To avoid wrong drift corrections the algorithm was applied only in those trials that meet the following conditions: (a) The initial fixation position was more than 25 pixels away from the center of the screen, but not further than 200 pixels (these values are based on the range of systematic drifts previously reported by Hornof & Halverson, 2002); (b) a fixation in the target for that trial was not initially detected in the scan path but was detected after the realignment. These limitations were included to maintain drift correction to a minimum.

After drift correction was applied, fixation locations were used to estimate, for each trial, how many fixations were made in total (fixation count, including initial fixation at the center of the screen) and how many of these were on the target and on each of the relevant and the irrelevant distractors. For that, a square area of 60 pixels was defined around the center of each element and fixations falling inside that area were computed as fixations on that element. This area was used as reference because, as in Experiments 1-3, distractor and targets appeared in an invisible 12 × 12 grid where each cell was 60 pixels square. We did not eliminate those trials in which a fixation in the target was not found, because there were multiple cases in which gaze location at the time of the response was in the proximity of the target, but not inside the square area that we defined and the response was still correct.

Results and Discussion

None of the 24 participants failed to meet the accuracy criterion used in Experiment 1-3. RTs were filtered following the same procedure as in Experiments 1-3. The top panel of **Figure 6** shows average RTs in each experimental condition across blocks of trials. As can be seen, in general RTs tended to be faster in condition *Both-Old* than in condition *Both-New*, suggesting that participants showed a contextual cueing effect. A 2 (Condition: *Both-New* vs. *Both-Old*) × 2 (Ignored Set Size: 4 vs. 16) × 12 (Epoch) ANOVA, showed that the main effect of Condition was only marginally significant, though, $F(1, 23) = 4.24$, $p = .051$, $\eta_p^2 = .156$, possibly due to the lower power of the present experiment. The ANOVA did reveal a main effect of Epoch, $F(5, 115) = 31.42$, $p < .001$, $\eta_p^2 = .577$, and most importantly, a significant main effect of Ignored Set Size, $F(1, 23) = 11.56$, $p = .002$, $\eta_p^2 = .335$, confirming that participants were slower when the search display included many distractors in the irrelevant color. The Epoch × Ignored Set Size × Condition interaction was marginally significant, $F(3.49, 80.26) = 2.40$, $p = .064$, $\eta_p^2 = .095$, suggesting that the contextual cueing effect increased gradually with training, but only in search displays with the lowest number of distractors. No other effects were significant or approached significance.

We repeated the same analyses on the number of fixations. For this analysis, we had to remove data from four participants due to the poor quality of their eye-tracking data, as explained above. The results are shown in the middle panels of **Figure 6**. As can be seen, the results observed in the number of fixations closely mirror those found in RTs. The ANOVA returned a main effect of Epoch, $F(3.16, 60.05) = 10.15$, $p < .001$, $\eta_p^2 = .348$, showing a decline in the number of fixations over training, and, most importantly, a main effect of Ignored Set Size, $F(1, 19) = 8.62$, $p = .008$, $\eta_p^2 = .312$, confirming that, as expected, participants needed to make more fixations to find the target when the search display included more distractors in the irrelevant color. Again, the main effect of Condition failed to reach statistical significance, $F(1, 19) = 1.86$, $p = .187$, $\eta_p^2 = .090$, possibly due to the low power of this test. All other effects were non-significant, p's > .264.

Regarding fixation locations, we found that a fixation in the target (square area defined around it, see above) was detected in about 85% of the trials. After removing initial fixations on the center of the screen, we found that most fixations (between 48 and 53% of all detected, depending on the experimental condition, or 1.8-1.9 fixations per trial) landed on empty areas of the screen. This might correspond to “centre-of-gravity” fixations: fixations in empty regions between items that often occur in difficult visual search tasks (van der Stigchel & Nijboer, 2011; Venini, Remington, Horstmann & Becker, 2014). 17% of detected fixations (0.6 fixations per trial on average) fell on relevant distractors, and this value was similar for all conditions. Interestingly, we also found that around 5% of detected fixations fell on irrelevant distractors when ignored set size was equal to 4 (0.2 fixations per trial), and that this value went up to 11% when ignored set size was equal to 16 (0.45 fixations per trial). The bottom panel of Figure 6 shows how the percentage of fixations on irrelevant distractors evolved over the course of the task. A Condition × Ignored Set Size × Epoch ANOVA on these percentages yielded only a significant main effect of Ignored Set Size, $F(1, 19) = 88.67$, $p < .001$, $\eta_p^2 = .824$, confirming that ignored distractors received direct attention to some degree. In fact, it is worth noting that the manipulation of set size had a substantially larger effect on this dependent variable than on RTs or the total number of fixations. The remaining main effects and interactions were non-significant, p's > .153, except for an unexpected double interaction between Epoch, Ignored Set Size and Condition, $F(4.02, 76.29) = 2.52$, $p = .048$, $\eta_p^2 = .117$. We do not have any explanation for such complex interaction and can only attribute it to sampling error and random variability in display configurations. Since displays are randomly generated for each participant, it is possible that relevant and irrelevant distractors showed different grouping patterns that might have affected gaze paths differently, even for the same experimental condition.

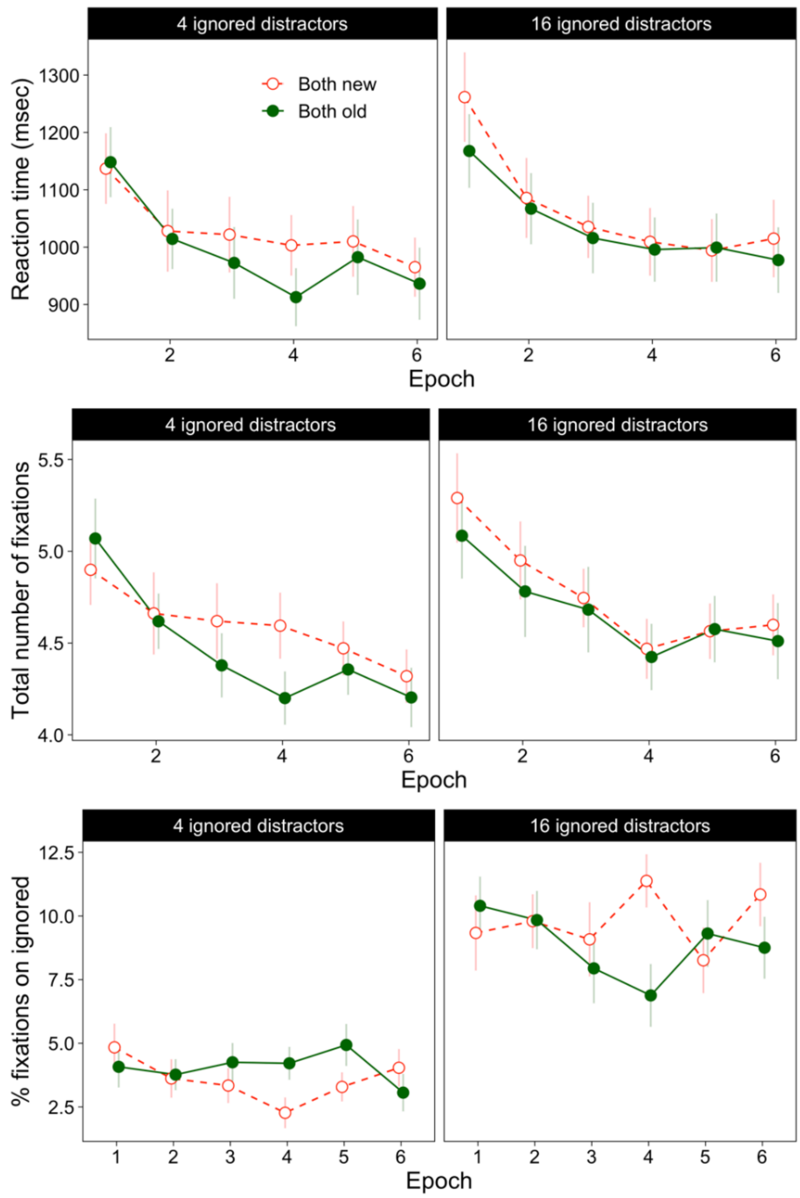

Figure 6. Reaction times (top panel), total number of fixations (middle panel), and percentage of fixations on irrelevant distractors (bottom panel) in Experiments 4. Error bars denote the standard error of the mean.

Although Experiment 4 failed to find robust evidence of contextual cueing in either dependent measure, the results replicated the main effect of Ignored Set Size detected in Experiments 1-3 and showed that the same effect can be found in eye movements, with participants systematically making more fixations (including quite a few on irrelevant items) when the target is embedded in search displays with many irrelevant distractors. This pattern of results confirms that participants do not completely filter these stimuli out, even though they are completely task irrelevant.

Meta-analysis

The results reported so far show quite well that participants do not completely ignore distractors presented in the irrelevant color. Paradoxically, though, not all the experiments included in this study show clear evidence of learning for irrelevant stimuli. For instance, the main effect of Ignored Context during the training stage was significant in Experiment 3, but non-significant in Experiments 1 and 2 (see **Figure 3**). This variability could reflect genuine differences across studies. But it cannot be discarded that the effect is simply too small to be detected reliably in our experiments. To further explore this possibility, we run a series of (non-registered) meta-analyses aimed at estimating the average effect size of learning for irrelevant distractors. These meta-analyses included not only evidence from Experiments 1-3 in the present study, but also from previous experiments addressing this question with similar designs and procedures. The data and scripts used for all the analyses presented in this section are publicly available at the public repository of the project (Vadillo et al., 2023).

The first of our meta-analyses focused on the size of contextual cueing for irrelevant distractors during the training stage. Except for the manipulation of Ignored Set Size, the procedure and design of the present series of experiments was virtually identical to the four experiments reported in Vadillo et al. (2020). Therefore, these four experiments were also included in the meta-analysis. The training stage in Experiments 1-3 was also very similar to Experiments 3 and 4 in Jiang and Chun (2001) and the main experiment in Jiang and Leung (2005). However, only the first of these studies reported sufficient statistical information to compute an effect size estimate. Experiment 2 from Jiang and Chun (2001) also measured contextual cueing for irrelevant distractors, in this case using only conditions *Both-New* and *Ignored-Old*. However, it had to be excluded from the meta-analysis because it also failed to report sufficient information.

The upper half of **Figure 7** shows a forest plot and meta-analysis of the size of the main effect of repeating the irrelevant stimuli during the training stage in those experiments. In addition to effect sizes and confidence intervals, **Figure 7** also reports Bayes Factors (BF_{10}), which quantify to what extent each contrast favors the alternative hypothesis (modelled as a Cauchy distribution with the scaling factor set to .707) over the null hypothesis of no effect. As can be seen, the effect does not reach statistical significance in every single study and, in fact, Bayes Factors often provide substantial support for the null hypothesis (i.e., $BF_{10} < 1$). However, the random-effects meta-analysis reveals a small but significant effect, $d_z = 0.14$, 95% CI [0.05, 0.22]. Furthermore, the meta-analytic Bayes Factor (Rouder & Morey, 2011) also provides moderate support for the alternative hypothesis, $BF_{10} = 6.97$. Although some individual effects are significant and some are not, the effect sizes are actually quite consistent across experiments, as shown by the fact that heterogeneity was small and non-significant, $I^2 = 8.55\%$, $Q(7) = 10.63$, $p = .156$.

The bottom half of **Figure 7** shows an otherwise identical meta-analysis, but this time based on Color Stay trials from the transfer stage. Note that these trials are identical to those of the training stage, except for the fact that they are randomly intermixed with Color Switch trials. In addition to Vadillo et al. (2020), Jiang and Leung (2005) also included a transfer stage with Color Stay trials that in principle could be collated in this meta-analysis. However, as in previous cases, the authors did not report sufficient information to compute an effect size estimate. As can be seen, the results of this meta-analysis are in almost perfect agreement with the meta-analysis of the training stage. Although most individual studies fail to detect a significant effect of repeating the irrelevant stimuli and Bayes Factors are often lower than 1, at the meta-analytic level there is a small but significant effect, $d_z = 0.14$, 95% CI [0.04, 0.24], and the amount of heterogeneity across studies is again small and non-significant, $I^2 = 30.31\%$, $Q(6) = 8.65$, $p = .194$.

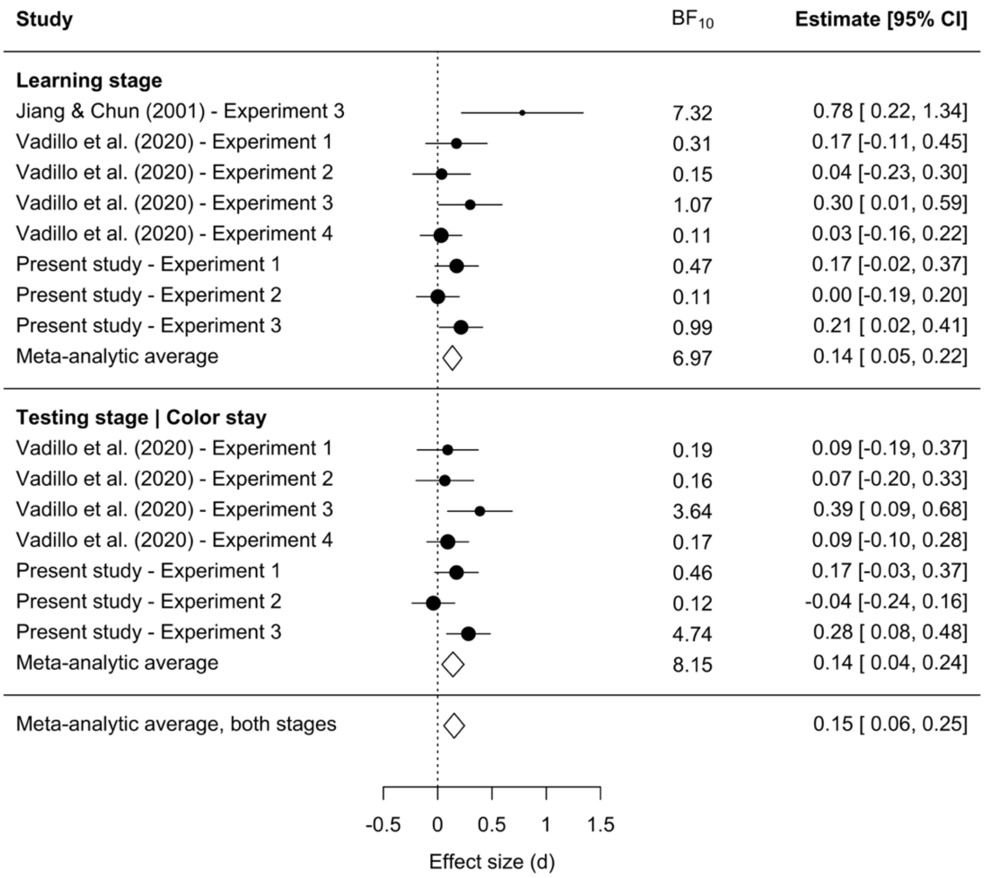

Figure 7. Meta-analytic evidence of contextual cueing for irrelevant distractors.

Because these two meta-analyses are essentially exploring the same effect, simply measured at different times of the experiments, we collated the 15 effect sizes in a single multi-level meta-analysis, adding a random intercept at the study level to account for statistical dependencies among effect sizes computed from the same sample. The meta-analytic average of this model is also shown in the figure. Again, the average effect was statistically significant, $d_z = 0.15$, 95% CI [0.06, 0.25], and heterogeneity was non-significant, $Q(14) = 19.29$, $p = .154$. Taken collectively, the previous meta-analyses show that participants *do learn* about ignored distractors, despite the fact that they have been instructed to ignore them. However, the effect is too small to be reliably detected with the small samples typically included in contextual cueing experiments. Even the relatively large samples of the present experiments (each with a planned $N = 100$) grant only .32 power to detect an effect of $d_z = 0.15$. In light of this, it is unsurprising that the effect does not reach statistical significance in every single study.

Jiang and Chun (2001) speculated that contextual cueing for irrelevant distractors might perhaps be due to the fact that the experimental task is not always sufficiently demanding for participants, and this leaves them with sufficient attentional resources to process even task-irrelevant stimuli. Consistent with this hypothesis, they observed that irrelevant stimuli no longer supported contextual cueing when the difficulty of the visual search task was increased. Specifically, the last experiment reported by Jiang and Chun (2001) employed visual distractors that were highly similar to the target. In response to this manipulation, participants became much slower at finding the target and, most importantly, any evidence of learning about irrelevant distractors disappeared. In Vadillo et al. (2020) we failed to find any effect of target-distractor discriminability in the size of contextual cueing for irrelevant distractors. However, the present meta-analysis provides another opportunity to put this hypothesis to the test in a larger sample of studies and participants. Among the studies included in the previous multi-level meta-analysis, the target was easily discriminable from the distractors in Experiment 3 by Jiang and Chun (2001), Experiments 2 and 3 from Vadillo et al. (2020), and Experiments 1 and 3 in the present study. In the remaining cases, the horizontal line of the L-shape distractors was slightly offset with respect to the vertical line, rendering distractors perceptually similar to the target and hindering visual search. A moderator analysis showed that the size of cueing for irrelevant distractors was significantly different between these two sets of studies, $Q(1) = 6.47$, $p = .011$. To assess the size of the effect in each subgroup, we repeated the moderator analysis, but removing the intercept from the model. This analysis revealed that the effect size of cueing for irrelevant distractors was significant among the studies using the easiest version of the search task, $d_z = 0.22$, 95% CI [0.13, 30], but not in the studies using the difficult version, $d_z = 0.05$, 95% CI [-0.05, 0.14]. Therefore, at the meta-analytic level there is some evidence that contextual cueing for irrelevant distractors is only found among experiments where targets and distractors are easily discriminable. Note, however, that once again the effect is too small to be reliably detected in a single experiment with the typical sample sizes. In fact, our previous analysis of this question (in Vadillo et al., 2020) failed to detect the moderating role of target-distractor similarity.

As explained in the introduction, Jiang and Leung (2005) argued that the fact that participants show no clear evidence of contextual cueing for irrelevant distractors need not mean that they failed to learn about these stimuli. It could simply mean that attention is necessary for the behavioral expression of contextual cueing, which might still be learned latently. In support for this view, they found that swapping the color of distractors during the transfer stage gave rise to a sudden search advantage for *Ignored-Old* trials than for *Both-New* trials, showing that participants must have learned something about them during the early stages of the experiment, even if the effect was not manifest at that time. This latent learning effect was later replicated in a semantic contextual cueing experiment by Goujon et al. (2009, Experiment 4). However, neither Vadillo et al. (2020) or Experiments 1-3 in the present study replicated this effect. With all the evidence available so far, it is possible to run a high-powered meta-analytic test of the latent-learning effect. For this purpose, we estimated the standardized effect size of the difference between RTs in conditions *Ignored-Old* trials and *Both-New* trials in the Color Switch trials of the transfer stage in Jiang and Leung (2005), Goujon et al. (2009, Experiment 4), the four experiments in Vadillo et al. (2020) and the three new empirical studies reported in the present study. Despite the large number of participants and studies, the meta-analysis failed to find a significant latent learning effect, $d_z = 0.06$, 95% CI [-0.02, 0.14], and the meta-analytic Bayes Factor provided moderate support for the null hypothesis, $BF_{10} = 0.14$. Heterogeneity was again low, $I^2 = 0.01\%$, and non-significant, $Q(8) = 11.85$, $p = .158$. Therefore, taken collectively, the body of the evidence available so far *does not support* the hypothesis that swapping the colors of distractors at test uncovers latent learning for previously ignored distractors.

Incidentally, the distribution of the effect sizes included in the meta-analyses reported in this section also reveal potential evidence of publication bias. **Figure S1** in the Supplementary Material shows clear evidence of funnel plot asymmetry both in the meta-analysis of direct learning of ignored context (i.e., main effect of Ignored Context during the training stage and in Color Stay trials of the transfer stage) and in the meta-analysis of latent learning (i.e., condition *Both-New* vs. *Ignored-Old* in Color Switch trials). Funnel plot asymmetry is a suggestive but imperfect indicator of publication bias (Sterne et al., 2011) and, therefore, these analyses should be interpreted with caution. But in any case, these results in combination with the relatively small size of the effects under investigation suggest that this is an area of research that would benefit enormously from the adoption of open research practices and high-powered experiments, possibly in the form of multi-site collaborative projects (e.g., Klein et al., 2018).

In sum, at the meta-analytic level we found that there is a small but reliable amount of learning for ignored distractors during the training stage and in Color Stay trials of the transfer stage (with potential evidence of publication bias). The discriminability of distractors and targets seems to modulate this effect. In fact, it seems to be restricted to conditions where distractors and targets are easily discriminable (consistent with Jiang & Chun, 2001). In contrast, all evidence of learning for ignored distractors disappears in Color Switch trials, that is, when previously ignored stimuli become attended and vice versa. Changing the colors of stimuli only seems to disrupt any evidence of contextual cueing observed during training. This is at odds with the idea that contextual cueing can develop latently for ignored items (Jiang & Leung, 2005).

General Discussion

The four experiments reported in this paper tried to clarify two issues: whether ignored elements could induce contextual cueing when they were predictive of target location, and whether instructions to ignore irrelevant elements were enough to do so. Both issues pinpoint to the role of selective attention in contextual cueing.

Research conducted with the contextual cueing paradigm has occasionally produced evidence suggesting that people can learn about aspects of the visual context they have been explicitly asked to ignore (Goujon et al., 2015; Jiang & Chun, 2001; Jiang & Leung, 2005; Vadillo et al., 2020). From these results, it is sometimes concluded that contextual cueing is possibly based on an automatic extraction of reliable visual information and can proceed without attention (e.g., Goujon et al., 2015; Jiang, 2018; Jiang & Leung, 2005). The present set of experiments replicates these findings: Participants were slightly faster to find visual targets when they were embedded in familiar search displays, even when what made those contexts "familiar" was the repetition of task-irrelevant distractors. However, as discussed above, the effect was small and only reliable at the meta-analytic level. In fact, the effect only reached statistical significance in Experiment 3 of the present study, when the task was especially difficult due to the increase in the number of attended distractors. On its own, the fact that the effect is so small and difficult to replicate should invite us to take with skepticism the prevailing view that contextual cueing is independent from selective attention. If anything, contextual cueing seems to become a marginal effect once predictive stimuli are presented as task-irrelevant and participants are explicitly instructed to ignore them.

The present experiments provide a second cause for concern. Even if contextual cueing for irrelevant distractors is considered a meaningful effect, however weak, there is little reason to assume that participants actually ignored those distractors just because they were instructed to do so. Contrary to this assumption, the evidence reported in the present study shows that participants do pay some attention to irrelevant distractors. This is shown by the fact that the number of irrelevant distractors in the search display always made a systematic difference in search times. Furthermore, eye-tracking data confirms that these stimuli were occasionally fixated. Overall, these findings confirm the hypothesis that participants are unable to completely filter out the task-irrelevant stimuli in the contextual cueing task, as originally suggested by Jiang and Chun (2001).

In light of this evidence, it seems premature to conclude that contextual cueing is independent from selective attention.

Of course, it is not our intention to imply that the instruction to ignore a specific set of stimuli made no difference in participants' performance. Quite on the contrary, everything suggests that participants paid more attention to task-relevant than to task-irrelevant distractors (for converging evidence, see Egeth et al., 1984; Kaptein et al., 1995). Perhaps the clearest evidence for this is provided by **Figure 4**. While the manipulation of set size only made a small difference in search times in Color Stay trials, the effect was large and robust in Color Switch trials. That is, the impact of set size was stronger when it affected task-relevant than task-irrelevant items. Eye-tracking data also confirms that participants were much more likely to fixate on task-relevant than task-irrelevant distractors. But even if attention to task-irrelevant distractors is diminished, the evidence that they are attended, in one way or another, is robust and reliable. In fact, as discussed above, it is more reliable than any evidence suggesting that contextual cueing can take place for irrelevant stimuli. It is difficult to see how we can make any strong inference based on the small contextual cueing effect detected in this paradigm without necessarily admitting that irrelevant-distractors were attended; if anything, the evidence for the latter effect is stronger and easily replicable.

Note that, strictly speaking, none of the experiments reported in this study prove that contextual cueing took place *because* participants paid attention to the irrelevant distractors. The design of our experiments allowed us to measure attention towards irrelevant distractors, but not to manipulate it experimentally. Consequently, we cannot make any strong causal inference based on our data. In fact, if anything, we actually found a mismatch between attention and learning: While we found very robust evidence of attention to irrelevant distractors, the amount of contextual cueing elicited by those distractors was small and only significant in one experiment. In other words, it is far easier to detect attention towards irrelevant distractors than learning about them.[2] Our results do not allow us to conclude that there is a causal relationship between both processes, but they do cast doubts on the conclusions of previous studies using this procedure: The fact that some studies find contextual cueing for stimuli that participants have been asked to ignore provides weak support, at best, for the idea that it is based on an automatic and efficient process.

Our concerns about the conclusions of previous studies with the contextual cueing paradigm dovetail with decades of research showing that task-irrelevant stimuli are not completely filtered out in visual search tasks, although their impact on visual search is logically modest compared to that of task-relevant stimuli. For instance, Green and Anderson (1956) observed that participants were significantly slower to complete a visual search task when it included relevant and irrelevant stimuli compared to an identical task that only included the relevant stimuli. Similarly, Carter (1982) found that a distractor presented in task-irrelevant colors affected visual search, although the effect disappeared if the distractor color was too different from the target's color. Other authors have shown that reaction times vary depending on the number of irrelevant items presented. For example, using small set sizes Kaptein et al. (1995) found that search times increased with the number of items presented in the task-irrelevant color, although the effect was comparatively small and restricted to target-absent trials. Using

[2] In fact, we have not found any systematic relationship between the amount of attention to irrelevant distractors (defined as the search cost induced by including 16, as opposed to 4, irrelevant distractors in the search display) and the amount of contextual cueing supported by those distractors (defined as the search advantage granted by the repetition of irrelevant distractors compared to search displays where the irrelevant distractors are non-predictive). Additional, non-registered analyses reported in the Supplementary Material show that these two variables do not correlate positively with each other. The correlation is even negative for Experiment 1. Note, however, that the lack of correlations is hardly surprising given the asymmetry between both effects (i.e., the limited range of variation in contextual cueing scores). Furthermore, the split-half reliability of both measures (also reported in the Supplementary Material) leaves much to be desired, obscuring the interpretation of these correlations even further (see Draheim et al., 2019; Loken & Gelman, 2017; Schmidt & Hunter, 2014; Vadillo et al., 2022).

set sizes including up to 600 items Benjamins et al. (2009) found that irrelevant distractors impaired search, although in their experimental paradigm the effect on reaction times was modulated by the *ratio* between relevant and irrelevant elements. Our results show a similar pattern to these studies: reaction times increased when the number of irrelevant distractors was higher and more fixations on irrelevant distractors were detected when more of them were present on the display (a result also reported by Benjamins et al., 2009).

Our failure to find complete filtering of task-irrelevant distractors also complements recent studies centered in trying to understand how we manage to ignore or suppress irrelevant information (see, e.g., Beck, Luck & Hollingworth, 2018; Gaspelin & Luck, 2018; Gaspelin & Vecera, 2019). The picture that is emerging suggests that the rejection or inhibition of irrelevant items in attentional tasks is a complex process that can be more or less effective depending on multiple variables. For example, aspects like the time course of attentional guidance (Cunningham & Egeth, 2016; Moher & Egeth, 2012; Palmer et al, 2018), previous knowledge about the location of the irrelevant items (Munneke et al., 2008) or whether the to-be-ignored feature is explicitly cued or not (Stilwell & Vecera, 2018) can modulate distractor rejection. These factors might partially explain why the repetition of irrelevant distractors leads, under some conditions, to contextual cueing.

One factor that might be particularly relevant in this context is search difficulty. More complex tasks might lead to improved attentional guidance, while, at the same time, reducing the availability of attentional resources. This could modulate how information about relevant and irrelevant items is used. For example, Conci et al. (2019) found that information about to-be-avoided distractors facilitated performance, but this effect only occurred when the visual search task was difficult. This is interesting, since difficulty is also relevant regarding contextual cueing effects. As explained above, Jiang and Chun (2001) found a small amount of contextual cueing for irrelevant distractors in one of their experiments, using a relatively simple visual search task, but the effect disappeared in a subsequent experiment using a more difficult display, with perceptually similar distractors and targets.

Our results point in this direction as well: Our meta-analysis shows that, taken collectively, the evidence collected so far only shows consistent evidence of contextual cueing for irrelevant distractors in experiments that use easily discriminable distractors and targets. When distractors are very similar to the target, as in the bottom row of **Figure 5**, the visual search task becomes relatively difficult and pushes participants' attention to the limit, leaving few or no attentional resources available for the processing of task-irrelevant stimuli. In contrast, when the visual search task is made easier by using dissimilar distractors and targets, as in the top row of **Figure 5**, participants might not need to be as selective in their allocation of attentional resources, which would result in some attention spilling over to the task-irrelevant stimuli. This hypothesis, based on perceptual load theory (Lavie, 1995), has received support in other experiments. For example, Lavie and Cox (1997) used a visual search task with different levels of difficulty (high or low perceptual load of the search display) and presented flankers as irrelevant stimuli. Their results showed that the flanker produced *less* interference under high-load conditions than under low-load conditions. A similar result was reported by Theeuwes et al. (2004), although in their case the effect depended on whether perceptual load was stable per block or varied at the trial level.

Although our results show that participants pay some attention to task-irrelevant distractors, they are relatively silent about the potential mechanisms how or why irrelevant distractors support contextual cueing. Some models of contextual cueing assume that the visual search advantage for repeated patterns is supported by learned associations between individual distractors and the location of the target (Brady & Chun, 2007). As the strength of these associations grows with training, it becomes increasingly more likely that fixating on any of these individual distractors will be followed by a fixation to the target. An alternative view is that, with repeated exposure to a search display, participants eventually develop a configural representation of that display that then guides attention towards the target location. Consistent with the later view, Chun and Jiang (1998) found that the identity of the distractors was relatively irrelevant for contextual cueing, as long as the general distribution of distractors in the display remained constant (Chun & Jiang, 1998). In the same vein, there is evidence that contextual cueing is weaker under conditions that prevent the development of consistent configural representations (Beesley et al., 2015, 2016). Following up on this idea, it is possible that the reason why the repetition of task-irrelevant distractors strengthens contextual cueing is that such repetitions facilitate the process of developing stable configural representations. It may be easier to learn that a particular configuration of distractors predicts the location of the target when all the elements in the display, relevant or irrelevant, appear in the exact same locations over blocks (i.e., in condition *Both-Old*) than when only the relevant distractors are presented in predictable configurations (i.e., in condition *Attended-Old*). Or, in other words, if the irrelevant distractors appear in random locations, this might prevent the development of a configural representation of the search display, which would oblige participants to rely on elemental, and probably less efficient, strategies. Perhaps this might explain why these irrelevant stimuli are attended at all: Participants may know that none of these stimuli can be the target, but paying some attention to them may contribute to develop richer and more stable representations of the search display.

It is worth noting that none of the present experiments yielded reliable evidence of "latent learning" of irrelevant contexts. Drawing from previous research with the serial reaction time task (Frensch et al., 1998, 1998), Jiang and Leung (2005) hypothesized that selective attention might be necessary for the expression, but not the acquisition, of contextual cueing. To put this idea to the test, they implemented a transfer stage where distractors that had previously been ignored were now presented in the attended color and vice versa. The results showed that ignored distractors, which during the initial stages of the experiment seemed to make no difference in visual search, elicited a significant contextual cueing effect during the transfer stage, suggesting that participants had indeed learned something about them, although this learning remained "silent" until participants were encouraged to pay attention to the previously irrelevant distractors. A similar latent learning effect was also reported by Goujon et al. (2009) in a semantic contextual cueing task. However, none of the experiments reported in the present study or in our previous preregistered experiments addressing this question (Vadillo et al., 2020) have been able to replicate this effect.

To some extent, our results add to a growing body of evidence suggesting that the acquisition and expression of contextual cueing might be less automatic than previously thought (see Giménez-Fernández et al., 2023, for a critical review). As explained in the introduction, the prevalent view is that contextual cueing is a form of implicit learning that can take place without participants becoming aware of the repetitions in the search displays or their impact on search performance (Chun & Jiang, 2003; Colagiuri & Livesey, 2016; Goujon et al., 2015). But the methods used to test participants' awareness in the task are often too underpowered to detect small but theoretically meaningful levels of awareness (Meyen et al., in press; Vadillo et al., 2016) and they are too unreliable to allow any inference about their relationship with performance in the contextual cueing task (Vadillo et al., 2022). In fact, the few experiments that have employed sufficiently sensitive measures have tended to find significant evidence of awareness and even significant correlations between awareness and contextual cueing (Geyer et al., 2020; Kroell et al., 2019). Similarly, it is unclear, at best, that contextual cueing possesses other features of automaticity. For instance, the attentional bias produced by contextual cueing seems to be perfectly controllable, in the sense that participants can be instructed to ignore their previous experience and find the target in new locations instead (Luque et al., 2017, 2021).

In sum, the present studies show that the most common strategy to manipulate selective attention in the contextual cueing task often fails to prevent participants from paying some attention to the task-irrelevant distractors they have been instructed to ignore. As a result, the conclusion that contextual cueing can be acquired or expressed without selective attention must be taken with caution. Ideally, future research should address this problem by devising new and more powerful methods to manipulate selective attention in the contextual cueing task.

References

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences,* **16**, 437–443.

Beck, V. M., Luck, S. J., & Hollingworth, A. (2018). Whatever you do, don't look at the...: Evaluating guidance by an exclusionary attentional template. *Journal of Experimental Psychology: Human Perception and Performance,* **44**, 645–662.

Beesley, T., Vadillo, M. A., Pearson, D., & Shanks, D. R. (2015). Pre-exposure of repeated search configurations facilitates subsequent contextual cuing of visual search. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **41**, 348–362.

Beesley, T., Vadillo, M. A., Pearson, D., & Shanks, D. R. (2016). Configural learning in contextual cuing of visual search. *Journal of Experimental Psychology: Human Perception and Performance,* **42**, 1173–1185.

Biederman, I. (1972). Perceiving real-world scenes. *Science,* **177**, 77–80.

Brady, T. F., & Chun, M. M. (2007). Spatial constraints on learning in visual search: Modeling contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance,* **33**, 798–815.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision,* **10**, 443-446.

Carter, R. C. (1982). Visual search with color. *Journal of Experimental Psychology: Human Perception and Performance,* **8**, 127–136.

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

Chun, M. M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition,* **29**, 224–234.

Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current Opinion in Neurobiology,* **17**, 177–184.

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 & Review,* **23**, 1996–2009.

Conci, M., Deichsel, C., Müller, H. J., & Töllner, T. (2019). Feature guidance by negative attentional templates depends on search difficulty. *Visual Cognition,* **27**, 317–326.

Cunningham, C. A., & Egeth, H. E. (2016). Taming the white bear: Initial costs and eventual benefits of distractor inhibition. *Psychological Science,* **27**, 476–485.

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin,* **145**, 508–535

Frensch, P. A., Lin, J., & Buchner, A. (1998). Learning versus behavioral expression of the learned: The effects of a secondary tone-counting task on implicit learning in the serial reaction task. *Psychological Research,* **61**, 83–98.

Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance,* **10**, 32–39.

Frensch, P. A., Wenke, D., & Rünger, D. (1999). A secondary tone-counting task suppresses expression of knowledge in the serial reaction task. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **25**, 260–274.

Gaspelin, N., & Luck, S. J. (2018). The role of inhibition in avoiding distraction by salient stimuli. *Trends in Cognitive Sciences,* **22**, 79–92.

Gaspelin, N., & Vecera, S. (2019). An introduction to the special issue on "Dealing with distractors in visual search". *Visual Cognition,* **27**, 183–184.

Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics,* **67**, 1252–1268.

Geyer, T., Rostami, P., Sogerer, L., Schlagbauer, B., & Müller, H. J. (2020). Task-based memory systems in contextual-cueing of visual search and explicit recognition. *Scientific Reports,* **10**: 16527.

Giménez-Fernández, T., Luque, D., Shanks, D. R., & Vadillo, M. A. (2023). Rethinking attentional habits. *Current Directions in Psychological Science,* **32**, 494–500.

Goujon, A., Didierjean, A., & Marmèche, E. (2009). Semantic contextual cuing and visual attention. *Journal of Experimental Psychology: Human Perception and Performance,* **35**, 50–71.

Goujon, A., Didierjean, A., & Thorpe, S. (2015). Investigating implicit statistical learning mechanisms through contextual cueing. *Trends in Cognitive Sciences,* **19**, 524e533.

Green, B. F., & Anderson, L. K. (1956). Color coding in a visual search task. *Journal of Experimental Psychology,* **51**, 19–24.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. & van der Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures.* Oxford: Oxford University Press.

Hornof, A. & Halverson, T. (2002). Cleaning up systematic error in eye tracking data by using required fixation locations. *Behavior Research Methods, Instruments, & Computers,* **34**, 592–604.

Jiang, Y. V. (2018). Habitual versus goal-driven attention. *Cortex,* **102**, 107–120.

Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning. *Quarterly Journal of Experimental Psychology,* **54A**, 1105–1124.

Jiang, Y., & Leung, A. W. (2005). Implicit learning of ignored visual context. *Psychonomic Bulletin & Review,* **12**, 100–106.

Jiang, Y. V., & Sisk, C. A. (2019). Habit-like attention. *Current Opinion in Psychology,* **29**, 65–70.

Jiang, Y. V., Swallow, K. M., & Rosenbaum, G. M. (2013). Guidance of spatial attention by incidental learning and endogenous cuing. *Journal of Experimental Psychology: Human Perception and Performance,* **39**, 285–297.

Kaptein, N. A., Theeuwes, J., & van der Heijden, A. H. C. (1995). Search for a conjunctively defined target can be selectively limited to a color-defined subset of elements. *Journal of Experimental Psychology: Human Perception and Performance,* **21**, 1053–1069.

Klein, R. A., Vianello, M., Hasselman, F., et al. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science,* **1**, 443–490.

Kleiner, M., Brainard, D. H. & Pelli, D. G. (2007). What's new in Psychtoolbox-3? *Perception,* **36**, 14.

Kroell, L. M., Schlagbauer, B., Zinchenko, A., Müller, H. J., & Geyer, T. (2019). Behavioural evidence for a single memory system in con- textual cuing. *Visual Cognition,* **27**, 551–562.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance,* **21**, 451–468.

Lavie, N., & Cox, S. (1997). On the efficiency of visual selective attention: Efficient visual search leads to inefficient distractor rejection. *Psychological Science,* **8**, 395–398.

Loken, E., & Gelman, A. (2017). Measurement error and the replication crisis. *Science,* **355**(6325), 584–585.

Luque, D., Beesley, T., Molinero, S., & Vadillo, M. A. (2021). Contextual cuing of visual search does not guide attention automatically in the presence of top-down goals. *Journal of Experimental Psychology: Human Perception and Performance,* **47**, 1080–1090.

Luque, D., Vadillo, M. A., López, F. J., Alonso, R., & Shanks, D. R. (2017). Testing the controllability of contextual cuing of visual search. *Scientific Reports,* **7**:39645.

Meyen, S., Vadillo, M. A., von Luxburg, U., & Franz, V. H. (in press). No evidence for contextual cueing beyond explicit recognition. *Psychonomic Bulletin & Review.*

Moher, J., & Egeth, H. E. (2012). The ignoring paradox: Cueing distractor features leads first to selection, then to inhibition of to-be-ignored items. *Attention, Perception and Psychophysics,* **74**, 1590–1605.

Munneke, J., van der Stigchel, S., & Theeuwes, J. (2008). Cueing the location of a distractor: An inhibitory mechanism of spatial attention? *Acta Psychologica,* **129**, 101–107.

Palmen, E. M., Van Wert, M. J., Horowitz, T. S., & Wolfe, J. M. (2018). Measuring the time course of selection during visual search. *Attention, Perception and Psychophysics,* **81**, 47–60.

Pearson, D., Watson, P., Albertella, L., & Le Pelley, M. E. (2022). Attentional economics links value-modulated attentional capture and decision-making. *Nature Psychology Reviews,* **1**, 320–333.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision,* **10**, 437–442.

Pollmann, S. (2019). Working memory dependence of spatial contextual cueing for visual search. *British Journal of Psychology,* **110**, 372–380.

Reddy, L., & VanRullen, R. (2007). Spacing affects some but not all visual searches: Implications for theories of attention and crowding. *Journal of Vision,* **7**, 3.

Rouder, J. N., & Morey, R. D. (2011). A Bayes factor meta-analysis of Bem's ESP claim. *Psychonomic Bulletin & Review,* 18, 682–689.

Salvucci, D. D., & Goldberg, J. H. (2000). *Identifying fixations and saccades in eye-tracking protocols.* Paper presented at the Eye Tracking Research and Applications Symposium., Palm Beach Gardens, FL, USA.

Schmidt, F. L., & Hunter, J. E. (2014). *Methods of meta-analysis: Correcting error and bias in research findings (3rd ed.).* SAGE.

Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rücker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., & Higgins, J. P. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ,* **343**, d4002.

Stilwell, B. T., & Vecera, S. P. (2018). Learned and cued distractor rejection for multiple features in visual search. *Attention, Perception and Psychophysics,* **81**, 359–376.

Tanaka, D., Kiyokawa, S., Yamada, A., Dienes, Z., & Shigemasu, K. (2008). Role of selective attention in artificial grammar learning. *Psychonomic Bulletin & Review,* **15**, 1154–1159.

Theeuwes, J. (2018). Visual selection: Usually fast and automatic; seldom slow and volitional. *Journal of Cognition,* **1**, 29.

Theeuwes, J., Bogaerts, L., & van Moorselaar, D. (2022). What to expect where and when: How statistical learning drives visual selection. *Trends in Cognitive Sciences,* **26**, 860–872.

Theeuwes, J., Kramer, A. F., & Belopolsky, A. V. (2004). Attentional sets interacts with perceptual load in visual search. *Psychonomic Bulletin & Review,* **11**, 697–702.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General,* **134**, 552–564.

Vadillo, M. A., Giménez-Fernández, T., Aivar, M. P., & Cubillas, C. P. (2020). Ignored visual context does not induce latent learning. *Psychonomic Bulletin & Review,* **27**, 512–519.

Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review,* **23**, 87–102.

Vadillo, M. A., Malejka, S., Lee, D. Y. H., Dienes, Z., & Shanks, D. R. (2022). Raising awareness about measurement error in research on unconscious mental processes. *Psychonomic Bulletin & Review,* **29**, 21–43.

Vadillo, M. A., Aniento, P., Hernández-Gutiérrez, D., Saini, L., & Aivar, M. P. (2023). Data and materials for "Can we learn ignored visual information? Figuring out is harder than you think!". Retrieved from osf.io/697cg

van der Stigchel, S., & Nijboer, T. C. W. (2011). The global effect: What determines where the eyes land? *Journal of Eye Movement Research,* **4**, 3.

Venini, D., Remington, R. W., Horstmann, G. & Becker, S. I. (2014). Center-of-gravity fixations in visual search: When looking at nothing helps to find something. *Journal of Ophthalmology*, **2014**, 237812.

Vicente-Conesa, F., Giménez-Fernández, T., Shanks, D. R., & Vadillo, M. A. (2022). The role of working memory in contextual cueing of visual attention. *Cortex,* **154**, 287–298.

Vickery, T. J., Sussman, R. S., & Jiang, Y. V. (2010). Spatial context learning survives interference from working memory load. *Journal of Experimental Psychology: Human Perception and Performance,* **36**, 1358–1371.

Wang, B., & Theeuwes, J. (2018). Statistical regularities modulate attentional capture. *Journal of Experimental Psychology: Human Perception and Performance,* **44**, 13–17.

SUPPLEMENTARY MATERIAL

Measuring attention to irrelevant distractors in contextual cueing

Miguel A. Vadillo, Patricia Aniento, David Hernández-Gutiérrez,
Luca Saini, & M. Pilar Aivar

Table S1. Analysis of Reaction Times During the Learning Stage in Experiment 1

Effect	**df**	***MSE***	***F***	$\boldsymbol{\eta_p^2}$	***p*-value**
Epoch	3.75, 363.52	127878.92	108.24 ***	.527	<.001
Set size	1, 97	38879.48	110.70 ***	.533	<.001
Ignored	1, 97	35661.08	3.04 †	.030	.084
Attended	1, 97	32164.00	73.22 ***	.430	<.001
Epoch × Set size	7.54, 731.47	17153.68	1.11	.011	.357
Epoch × Ignored	8.68, 841.56	13531.15	1.07	.011	.380
Set size × Ignored	1, 97	40743.71	1.06	.011	.305
Epoch × Attended	8.28, 802.78	14258.95	1.47	.015	.162
Set size × Attended	1, 97	37684.31	0.05	<.001	.826
Ignored × Attended	1, 97	39037.52	0.60	.006	.440
Epoch × Set size × Ignored	7.88, 764.72	17079.05	0.65	.007	.733
Epoch × Set size × Attended	7.09, 688.12	18461.43	1.06	.011	.388
Epoch × Ignored × Attended	9.27, 899.60	13424.70	1.11	.011	.354
Set size × Ignored × Attended	1, 97	37025.80	2.82 †	.028	.096
Epoch × Set size × Ignored × Attended	9.50, 921.97	11288.74	1.27	.013	.249

Note. In this and all subsequent ANOVA tables, **df** = degrees of freedom; ***MSE*** = mean squared error for each effect in the ANOVA; ***F*** = *F*-value for the contrast of each effect in the ANOVA; ***p*** = *p*-value associated with the *F* and df in the same row; $\boldsymbol{\eta_p^2}$ = partial eta squared. ***$p < .001$, **$p < .01$, *$p < .05$, †$p < .10$.

Table S2. Analysis of Reaction Times During the Testing Stage in Experiment 1

Effect	**df**	***MSE***	***F***	η_p^2	***p*-value**
Set size	1, 97	18369.09	448.49 ***	.822	<.001
Ignored	1, 97	14158.49	5.00 *	.049	.028
Attended	1, 97	12769.69	5.57 *	.054	.020
Switch	1, 97	9808.02	67.26 ***	.409	<.001
Set size × Ignored	1, 97	11348.29	1.19	.012	.278
Set size × Attended	1, 97	12622.08	0.41	.004	.523
Ignored × Attended	1, 97	12625.32	4.28 *	.042	.041
Set size × Switch	1, 97	14816.33	332.18 ***	.774	<.001
Ignored × Switch	1, 97	10214.58	0.00	<.001	.995
Attended × Switch	1, 97	8428.72	11.56 ***	.106	<.001
Set size × Ignored × Attended	1, 97	11179.39	0.22	.002	.639
Set size × Ignored × Switch	1, 97	8845.43	0.02	<.001	.891
Set size × Attended × Switch	1, 97	7055.26	2.51	.025	.116
Ignored × Attended × Switch	1, 97	7534.10	0.09	<.001	.761
Set size × Ignored × Attended × Switch	1, 97	9464.96	1.54	.016	.218

Table S3. Analysis of Reaction Times in Color Stay Trials During the Testing Stage in Experiment 1

Effect	df	*MSE*	*F*	η_p^2	*p*-value
Set size	1, 97	12013.07	17.68 ***	.154	<.001
Ignored	1, 97	12514.38	2.81 †	.028	.097
Attended	1, 97	11335.92	14.78 ***	.132	<.001
Set size × Ignored	1, 97	10798.90	0.49	.005	.484
Set size × Attended	1, 97	9677.23	0.19	.002	.661
Ignored × Attended	1, 97	11203.53	2.99 †	.030	.087
Set size × Ignored × Attended	1, 97	10222.02	1.42	.014	.236

Table S4. Analysis of Reaction Times in Color Switch Trials During the Testing Stage in Experiment 1

Effect	df	*MSE*	*F*	η_p^2	*p*-value
Set size	1, 97	21172.35	611.53 ***	.863	<.001
Ignored	1, 97	11858.69	3.00 †	.030	.087
Attended	1, 97	9862.49	0.11	.001	.746
Set size × Ignored	1, 97	9394.82	0.89	.009	.348
Set size × Attended	1, 97	10000.11	2.10	.021	.150
Ignored × Attended	1, 97	8955.90	2.37	.024	.127
Set size × Ignored × Attended	1, 97	10422.33	0.24	.002	.625

Table S5. Analysis of Reaction Times During the Learning Stage in Experiment 2

Effect	df	*MSE*	*F*	η_p^2	*p*-value
Epoch	6.39, 632.78	304755.27	119.35 ***	.547	<.001
Set size	1, 99	253654.22	6.64 *	.063	.011
Ignored	1, 99	179166.42	0.00	<.001	.967
Attended	1, 99	190711.17	35.16 ***	.262	<.001
Epoch × Set size	9.10, 900.67	100106.31	1.42	.014	.176
Epoch × Ignored	9.37, 927.87	94394.96	1.06	.011	.390
Set size × Ignored	1, 99	238581.57	0.14	.001	.707
Epoch × Attended	9.28, 918.82	94017.48	5.46 ***	.052	<.001
Set size × Attended	1, 99	269112.89	1.54	.015	.217
Ignored × Attended	1, 99	194002.42	1.76	.017	.187
Epoch × Set size × Ignored	8.84, 875.28	96058.33	0.51	.005	.867
Epoch × Set size × Attended	9.25, 915.44	93159.05	1.21	.012	.283
Epoch × Ignored × Attended	9.55, 945.39	88977.52	0.55	.006	.847
Set size × Ignored × Attended	1, 99	151626.66	0.17	.002	.680
Epoch × Set size × Ignored × Attended	9.19, 909.90	101722.23	1.05	.010	.400

Table S6. Analysis of Reaction Times During the Testing Stage in Experiment 2

Effect	df	*MSE*	*F*	η_p^2	*p*-value
Set size	1, 99	144916.08	926.26 ***	.903	<.001
Ignored	1, 99	104324.04	0.69	.007	.407
Attended	1, 99	107553.55	6.23 *	.059	.014
Switch	1, 99	81373.52	123.12 ***	.554	<.001
Set size × Ignored	1, 99	80029.93	0.26	.003	.610
Set size × Attended	1, 99	90481.80	0.26	.003	.614
Ignored × Attended	1, 99	95149.43	0.20	.002	.653
Set size × Switch	1, 99	180031.80	672.54 ***	.872	<.001
Ignored × Switch	1, 99	85606.65	2.39	.024	.125
Attended × Switch	1, 99	65465.52	21.81 ***	.181	<.001
Set size × Ignored × Attended	1, 99	78751.81	2.31	.023	.132
Set size × Ignored × Switch	1, 99	76745.72	0.61	.006	.435
Set size × Attended × Switch	1, 99	78939.41	0.27	.003	.603
Ignored × Attended × Switch	1, 99	74829.81	0.25	.003	.619
Set size × Ignored × Attended × Switch	1, 99	87807.14	0.05	<.001	.818

Table S7. Analysis of Reaction Times in Color Stay Trials During the Testing Stage in Experiment 2

Effect	**df**	***MSE***	***F***	η_p^2	***p*-value**
Set size	1, 99	78705.29	2.15	.021	.145
Ignored	1, 99	77420.24	0.22	.002	.642
Attended	1, 99	67878.31	29.86 ***	.232	<.001
Set size × Ignored	1, 99	61127.18	1.07	.011	.303
Set size × Attended	1, 99	52718.41	0.00	<.001	.987
Ignored × Attended	1, 99	76159.36	0.50	.005	.482
Set size × Ignored × Attended	1, 99	69725.69	1.75	.017	.189

Table S8. Analysis of Reaction Times in Color Switch Trials During the Testing Stage in Experiment 2

Effect	**df**	***MSE***	***F***	η_p^2	***p*-value**
Set size	1, 99	246242.59	1036.13 ***	.913	<.001
Ignored	1, 99	112510.45	2.31	.023	.131
Attended	1, 99	105140.76	0.67	.007	.414
Set size × Ignored	1, 99	95648.46	0.03	<.001	.869
Set size × Attended	1, 99	116702.80	0.38	.004	.538
Ignored × Attended	1, 99	93819.89	0.00	<.001	.995
Set size × Ignored × Attended	1, 99	96833.26	0.66	.007	.418

Table S9. Analysis of Reaction Times During the Learning Stage in Experiment 3

Effect	df	*MSE*	*F*	η_p^2	*p*-value
Epoch	2.64, 260.95	390289.60	152.94 ***	.607	<.001
Set size	1, 99	64018.84	54.33 ***	.354	<.001
Ignored	1, 99	75669.25	4.79 *	.046	.031
Attended	1, 99	81497.48	15.83 ***	.138	<.001
Epoch × Set size	8.91, 881.60	24355.50	1.67 †	.017	.092
Epoch × Ignored	9.34, 924.58	22889.30	0.77	.008	.651
Set size × Ignored	1, 99	92436.85	0.00	<.001	.979
Epoch × Attended	9.00, 890.63	25191.02	1.11	.011	.352
Set size × Attended	1, 99	74369.44	0.32	.003	.572
Ignored × Attended	1, 99	72106.52	1.20	.012	.276
Epoch × Set size × Ignored	9.43, 933.57	21155.37	0.71	.007	.709
Epoch × Set size × Attended	9.88, 977.74	21110.53	1.08	.011	.374
Epoch × Ignored × Attended	9.21, 912.10	24991.81	0.96	.010	.477
Set size × Ignored × Attended	1, 99	67609.06	0.01	<.001	.915
Epoch × Set size × Ignored × Attended	9.40, 930.45	23686.15	0.97	.010	.469

Table S10. Analysis of Reaction Times During the Testing Stage in Experiment 3

Effect	df	*MSE*	*F*	η_p^2	*p*-value
Set size	1, 99	25104.14	215.15 ***	.685	<.001
Ignored	1, 99	18598.21	3.30 †	.032	.072
Attended	1, 99	12678.72	81.85 ***	.453	<.001
Switch	1, 99	10359.06	8.38 **	.078	.005
Set size × Ignored	1, 99	23647.01	0.01	<.001	.923
Set size × Attended	1, 99	15946.67	0.07	<.001	.788
Ignored × Attended	1, 99	14240.23	0.01	<.001	.929
Set size × Switch	1, 99	14465.44	195.90 ***	.664	<.001
Ignored × Switch	1, 99	12369.86	1.93	.019	.167
Attended × Switch	1, 99	11838.30	41.36 ***	.295	<.001
Set size × Ignored × Attended	1, 99	15387.24	0.01	<.001	.915
Set size × Ignored × Switch	1, 99	13113.69	0.51	.005	.478
Set size × Attended × Switch	1, 99	9699.53	0.54	.005	.465
Ignored × Attended × Switch	1, 99	9456.97	1.48	.015	.227
Set size × Ignored × Attended × Switch	1, 99	12921.43	0.00	<.001	.989

Table S11. Analysis of Reaction Times in Color Stay Trials During the Testing Stage in Experiment 3

Effect	**df**	***MSE***	***F***	η_p^2	***p*-value**
Set size	1, 99	17758.77	11.56 ***	.105	<.001
Ignored	1, 99	10566.28	7.67 **	.072	.007
Attended	1, 99	14106.41	104.67 ***	.514	<.001
Set size × Ignored	1, 99	18372.01	0.12	.001	.728
Set size × Attended	1, 99	13511.56	0.05	<.001	.817
Ignored × Attended	1, 99	12575.57	0.66	.007	.419
Set size × Ignored × Attended	1, 99	15414.27	0.00	<.001	.946

Table S12. Analysis of Reaction Times in Color Switch Trials During the Testing Stage in Experiment 3

Effect	**df**	***MSE***	***F***	η_p^2	***p*-value**
Set size	1, 99	21810.81	368.15 ***	.788	<.001
Ignored	1, 99	20401.78	0.21	.002	.646
Attended	1, 99	10410.61	4.89 *	.047	.029
Set size × Ignored	1, 99	18388.69	0.25	.003	.616
Set size × Attended	1, 99	12134.64	0.47	.005	.496
Ignored × Attended	1, 99	11121.62	0.52	.005	.472
Set size × Ignored × Attended	1, 99	12894.40	0.01	<.001	.927

Table S13. Analysis of Reaction Times During the Learning Stage in Experiments 1, 2, and 3

Effect	**df**	***MSE***	***F***	η_p^2	***p*-value**
Experiment	2, 295	4896821.95	541.25 ***	.786	<.001
Epoch	4.70, 1385.85	246183.50	348.37 ***	.541	<.001
Experiment × Epoch	9.40, 1385.85	246183.50	15.86 ***	.097	<.001
Set size	1, 295	119393.03	76.68 ***	.206	<.001
Experiment × Set size	2, 295	119393.03	1.40	.009	.248
Ignored	1, 295	97246.97	3.09 †	.010	.080
Experiment × Ignored	2, 295	97246.97	0.88	.006	.416
Attended	1, 295	101927.34	90.40 ***	.235	<.001
Experiment × Attended	2, 295	101927.34	5.50 **	.036	.005
Epoch × Set size	9.60, 2831.16	43862.68	1.40	.005	.176
Experiment × Epoch × Set size	19.19, 2831.16	43862.68	1.42	.010	.104
Epoch × Ignored	9.87, 2912.74	41245.27	0.69	.002	.736
Experiment × Epoch × Ignored	19.75, 2912.74	41245.27	1.16	.008	.279
Set size × Ignored	1, 295	124484.62	0.43	.001	.513
Experiment × Set size × Ignored	2, 295	124484.62	0.10	<.001	.908
Epoch × Attended	9.77, 2882.40	41724.86	6.56 ***	.022	<.001
Experiment × Epoch × Attended	19.54, 2882.40	41724.86	3.06 ***	.020	<.001
Set size × Attended	1, 295	127661.45	1.84	.006	.176
Experiment × Set size × Attended	2, 295	127661.45	0.80	.005	.452
Ignored × Attended	1, 295	102140.42	0.06	<.001	.807
Experiment × Ignored × Attended	2, 295	102140.42	2.18	.015	.115
Epoch × Set size × Ignored	9.53, 2812.37	41562.22	0.69	.002	.727
Experiment × Epoch × Set size × Ignored	19.07, 2812.37	41562.22	0.48	.003	.970
Epoch × Set size × Attended	9.85, 2906.78	40810.02	0.59	.002	.823
Experiment × Epoch × Set size × Attended	19.71, 2906.78	40810.02	1.46 †	.010	.088
Epoch × Ignored × Attended	10.01, 2952.50	40301.59	0.62	.002	.796
Experiment × Epoch × Ignored × Attended	20.02, 2952.50	40301.59	0.71	.005	.821
Set size × Ignored × Attended	1, 295	85748.61	0.07	<.001	.787
Experiment × Set size × Ignored × Attended	2, 295	85748.61	0.73	.005	.482
Epoch × Set size × Ignored × Attended	9.81, 2894.49	43186.92	1.13	.004	.334
Experiment × Epoch × Set size × Ignored × Attended	19.62, 2894.49	43186.92	1.00	.007	.456

Table S14. Analysis of Reaction Times During the Testing Stage in Experiments 1 and 2

Effect	**df**	***MSE***	***F***	η_p^2	***p*-value**
Experiment	1, 196	900425.14	757.26 ***	.794	<.001
Set size	1, 196	82288.23	1262.03 ***	.866	<.001
Experiment × Set size	1, 196	82288.23	453.84 ***	.698	<.001
Ignored	1, 196	59701.29	2.40	.012	.123
Experiment × Ignored	1, 196	59701.29	0.00	<.001	.999
Attended	1, 196	60645.21	9.66 **	.047	.002
Experiment × Attended	1, 196	60645.21	2.46	.012	.118
Switch	1, 196	45955.90	171.09 ***	.466	<.001
Experiment × Switch	1, 196	45955.90	59.22 ***	.232	<.001
Set size × Ignored	1, 196	46039.53	0.01	<.001	.929
Experiment × Set size × Ignored	1, 196	46039.53	0.74	.004	.391
Set size × Attended	1, 196	51949.19	0.48	.002	.489
Experiment × Set size × Attended	1, 196	51949.19	0.06	<.001	.807
Ignored × Attended	1, 196	54308.42	0.08	<.001	.773
Experiment × Ignored × Attended	1, 196	54308.42	1.27	.006	.261
Set size × Switch	1, 196	98267.00	883.55 ***	.818	<.001
Experiment × Set size × Switch	1, 196	98267.00	386.74 ***	.664	<.001
Ignored × Switch	1, 196	48295.27	2.10	.011	.149
Experiment × Ignored × Switch	1, 196	48295.27	2.09	.011	.150
Attended × Switch	1, 196	37238.12	30.32 ***	.134	<.001
Experiment × Attended × Switch	1, 196	37238.12	10.29 **	.050	.002
Set size × Ignored × Attended	1, 196	45310.35	1.54	.008	.216
Experiment × Set size × Ignored × Attended	1, 196	45310.35	2.48	.012	.117
Set size × Ignored × Switch	1, 196	43142.00	0.61	.003	.437
Experiment × Set size × Ignored × Switch	1, 196	43142.00	0.48	.002	.490
Set size × Attended × Switch	1, 196	43364.09	0.90	.005	.343
Experiment × Set size × Attended × Switch	1, 196	43364.09	0.00	<.001	.967
Ignored × Attended × Switch	1, 196	41525.30	0.14	<.001	.706
Experiment × Ignored × Attended × Switch	1, 196	41525.30	0.32	.002	.574
Set size × Ignored × Attended × Switch	1, 196	49035.76	0.03	<.001	.865
Experiment × Set size × Ignored × Attended × Switch	1, 196	49035.76	0.36	.002	.547

Table S15. Analysis of Reaction Times in Color Stay Trials During the Testing Stage in Experiments 1 and 2

Effect	**df**	***MSE***	***F***	η_p^2	***p*-value**
Experiment	1, 196	415770.41	719.65 ***	.786	<.001
Set size	1, 196	45699.44	8.33 **	.041	.004
Experiment × Set size	1, 196	45699.44	0.03	<.001	.859
Ignored	1, 196	45298.46	0.04	<.001	.843
Experiment × Ignored	1, 196	45298.46	1.11	.006	.293
Attended	1, 196	39895.60	41.87 ***	.176	<.001
Experiment × Attended	1, 196	39895.60	12.66 ***	.061	<.001
Set size × Ignored	1, 196	36219.82	0.45	.002	.502
Experiment × Set size × Ignored	1, 196	36219.82	1.48	.008	.224
Set size × Attended	1, 196	31417.42	0.03	<.001	.874
Experiment × Set size × Attended	1, 196	31417.42	0.04	<.001	.851
Ignored × Attended	1, 196	44012.85	0.00	<.001	.974
Experiment × Ignored × Attended	1, 196	44012.85	1.62	.008	.205
Set size × Ignored × Attended	1, 196	40277.44	0.64	.003	.425
Experiment × Set size × Ignored × Attended	1, 196	40277.44	2.73	.014	.100

Table S16. Analysis of Reaction Times in Color Switch Trials During the Testing Stage in Experiments 1 and 2

Effect	**df**	***MSE***	***F***	η_p^2	***p*-value**
Experiment	1, 196	530610.64	726.26 ***	.787	<.001
Set size	1, 196	134855.79	1411.09 ***	.878	<.001
Experiment × Set size	1, 196	134855.79	558.73 ***	.740	<.001
Ignored	1, 196	62698.10	3.87 †	.019	.050
Experiment × Ignored	1, 196	62698.10	0.81	.004	.370
Attended	1, 196	57987.74	0.76	.004	.384
Experiment × Attended	1, 196	57987.74	0.47	.002	.496
Set size × Ignored	1, 196	52961.71	0.19	<.001	.662
Experiment × Set size × Ignored	1, 196	52961.71	0.02	<.001	.900
Set size × Attended	1, 196	63895.86	0.99	.005	.321
Experiment × Set size × Attended	1, 196	63895.86	0.03	<.001	.857
Ignored × Attended	1, 196	51820.87	0.20	.001	.654
Experiment × Ignored × Attended	1, 196	51820.87	0.21	.001	.646
Set size × Ignored × Attended	1, 196	54068.67	0.84	.004	.360
Experiment × Set size × Ignored × Attended	1, 196	54068.67	0.38	.002	.541

Exploring publication biases

As mentioned in the main article, the distribution of effect sizes suggests that this literature might suffer from publication bias or selective reporting of analyses with significant results. **Figure S1** plots the effect sizes (Cohen's d_z) against their precision (standard error). The effect sizes included in the left panel refer to the main effect of Ignored Context during the training stage and also in Color Stay trials of the transfer stage. In the absence of bias, effect sizes and precision should be unrelated (Sterne et al., 2001). However, the funnel plot shows that, overall, studies with higher precision (i.e., with larger samples) tend to yield smaller effect sizes. Furthermore. Although funnel-plot asymmetry can arise for reasons unrelated to publication bias (Sterne et al., 2011), it is often taken as evidence that small studies with non-significant results (or alternative analyses with non-significant results) may be missing in the literature, especially when the effects that reach significance are close to the border of significance (i.e., just beside the grey contour denoting the area of non-significance).

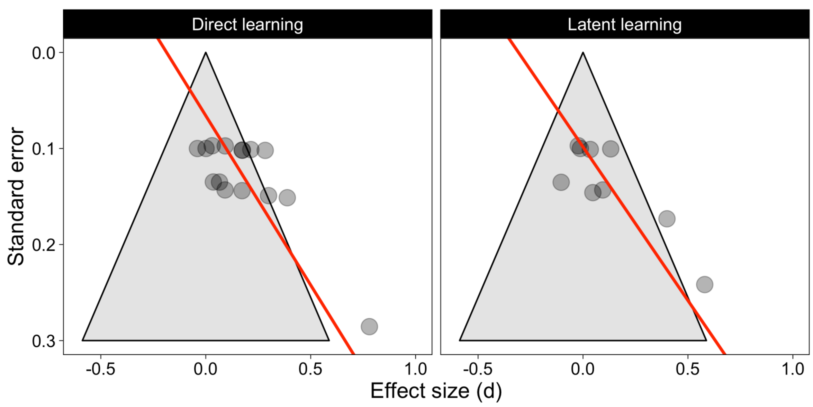

Figure S1. Funnel plots for the effect sizes included in the meta-analysis. The left panel refers to the main effect of Ignored Context during the training stage and also in Color Stay trials of the transfer stage. The right panel refers to the size of "latent learning", i.e., reaction times to *Both-New* vs. *Ignored-Old* trials during Color Switch trials of the transfer stage. The red line denotes Egger's test for funnel-plot asymmetry.

To test whether this asymmetry is significant, we fitted a multi-level meta-analytic model with a random intercept at the study level entering the standard error as a continuous moderator. This is essentially a multi-level version of the popular Egger test for funnel plot asymmetry (Rodgers & Pustejovsky, 2021). The slope of the model was significant, $b_1 = 2.84$, $SE = 1.16$, $z = 2.44$, $p = .015$. The right panel of **Figure S1** shows the same information for the effect sizes of "latent learning", defined as the difference in search times in conditions *Ignored-Old* vs. *Both-New* in color switch trials of the testing stage. Egger's test yielded significant evidence of funnel-plot asymmetry in this case as well, $b_1 = 3.13$, $SE = 1.31$, $z = 2.39$, $p = .017$.

Correlation between contextual cueing and attention to irrelevant distractors

On the assumption that participants learn about irrelevant distractors because they actually pay attention to them, one would expect to find a positive correlation between learning and attention to irrelevant distractors. To put this hypothesis to the test, we computed the correlation between the amount of contextual cueing for irrelevant distractors showed by each participant and the amount of attention that he/she paid to irrelevant distractors. In these analyses, contextual cueing for irrelevant distractors was defined as the reaction-time (RT) difference between experimental conditions where the ignored context was not predictive of the target location (i.e., conditions *Both-New* and *Attended-Old*) and the experimental conditions where the ignored context was predictive of the target location (i.e., conditions *Ignored-Old* and *Both-Old*). Attention to irrelevant distractors was defined as the RTs difference between trials with 16 distractors in the irrelevant color and trials with 4 distractors in the irrelevant color. Given that RTs were substantially slower in Experiments 2 and 3 compared to Experiment 1, we rescaled these measures by each participants average RT across all experimental conditions. Only trials from the training stage were included in these analyses. Trials from the first epoch were excluded because contextual cueing was expected to arise after a few repetitions of each search display. The scatterplot in the left panel of **Figure S2** plots both dependent variables against each other. As can be seen, the correlation was negative in Experiment 1, r = .23, p = .020, and fairly close to zero in Experiment 2, r = .03, p = .784, and Experiment 3, r = .01, p = .949. Across the three experiments, the correlation was small and non-significant, r = .02, p = .669. These negative and null correlations should be interpreted with caution, though. Firstly, because our experimental procedure proved to be much more sensitive to attention to irrelevant distractors than to contextual cueing by them. And, secondly, because the reliabilities of these dependent measures as far from ideal. The right panel of **Figure S2** shows the distribution of split-half reliabilities for each dependent variable and experiment across 1,000 iterations with random splits. As can be seen, the average reliability is only slightly above .60. Although this estimate is noticeably higher than previous analyses of the reliability of contextual cueing (e.g., Vadillo et al., 2022), it is nevertheless sufficiently low to cast doubts on the interpretation of correlations involving these dependent variables (Draheim et al., 2019; Loken & Gelman, 2017; Schmidt & Hunter, 2014).

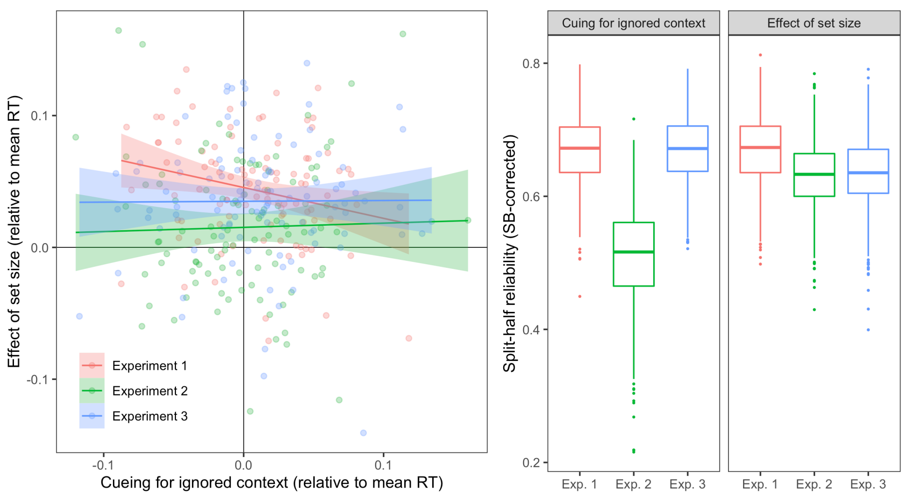

Figure S2. The left panel plots the size of the contextual cueing effect for irrelevant distractors (i.e., the main effect of Ignored Context in the training stage) against the search cost of having 16 instead of 4 irrelevant distractors in the search display. To facilitate the comparison of these effects across experiments, the reaction times in each condition have been divided by the average reaction time across conditions. The right panel shows the split-half reliability (Spearman-Brown corrected) of the two dependent variables analyzed in the left panel, across 1,000 iterations with random splits.

References

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin,* **145**, 508–535

Loken, E., & Gelman, A. (2017). Measurement error and the replication crisis. *Science,* **355**(6325), 584–585.

Rodgers, M. A. & Pustejovsky, J. E. (2021). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods,* **26**, 141–160.

Schmidt, F. L., & Hunter, J. E. (2014). *Methods of meta-analysis: Correcting error and bias in research findings (3rd ed.).* SAGE.

Sterne, J. A. C., Egger, M., & Smith, G. D. (2001). Investigating and dealing with publication and other biases in meta-analysis. *BMJ,* **323**, 101–105.

Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rücker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., & Higgins, J. P. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ,* **343**, d4002.

Vadillo, M. A., Malejka, S., Lee, D. Y. H., Dienes, Z., & Shanks, D. R. (2022). Raising awareness about measurement error in research on unconscious mental processes. *Psychonomic Bulletin & Review,* **29**, 21–43.