Vision Research 230 (2025) 108591

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Attentional misguidance from contextual learning after target location changes in natural scenes☆

Markus Conci *, Feifei Zhao

Department of Psychology, Ludwig-Maximilians-Universität München, München, Germany
Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität München, München, Germany

ARTICLE INFO

Keywords:
Visual attention
Visual search
Statistical learning
Contextual cueing
Natural scenes
Context adaptation

ABSTRACT

Attentional orienting in complex visual environments is supported by statistical learning of regularities. For instance, visual search for a target is faster when a distractor layout is repeatedly encountered, illustrating that learned contextual invariances improve attentional guidance (contextual cueing). Although contextual learning is usually relatively efficient, relocating the target (within an otherwise unchanged layout) typically abolishes contextual cueing, while revealing only a slow recovery of learning. However, such a "lack-of-adaptation" was usually only shown with artificial displays with target/distractor letters. The current study in turn used more realistic natural scene images to determine whether a comparable cost would also be evident in real-life contexts. Two experiments compared initial contextual cueing and the subsequent updating after a change in displays that either presented artificial letters, or natural scenes as contexts. With letter displays, an initial cueing effect was found that was associated with non-explicit, incidental learning, which vanished after the change. Natural scene displays either revealed a rather large cueing effect that was related to explicit memory (Experiment 1), or cueing was less strong and based on incidental learning (Experiment 2), with the size of cueing and the explicitness of the memory representation depending on the variability of the presented scene images. However, these variable initial benefits in scene displays always led to a substantial reduction after the change, comparable to the pattern in letter displays. Together, these findings show that the "richness" of natural scene contexts does not facilitate flexible contextual updating.

1. Introduction

The human visual system constantly extracts regularities from our environment in order to generate predictions about upcoming events. For example, when searching for an item (e.g. a loaf of bread) in your local supermarket, the layout of the shop as encountered on previous instances would typically help you to find the desired target object more quickly, as compared to a situation where search for the same item is performed in an unknown environment. Such examples illustrate that processing of information in a visual scene is not solely based on the analysis of the currently perceived features and objects in a bottom-up manner at any given moment, but the analysis of the visual environment is also biased by statistical learning, thereby facilitating attentional guidance to task-relevant objects at their expected locations based on predictions derived from previous encounters (see Goujon, Didierjean, & Thorpe, 2015; Nobre & Stokes, 2019; Oliva & Torralba, 2007; Võ, Boettcher, & Draschkow, 2019, for reviews).

One popular example that illustrates how statistical learning may improve the guidance of attention is the contextual cueing paradigm (Chun & Jiang, 1998), where observers are asked to perform a visual search task that requires them to detect a target T in an array of distractor L's (see Fig. 1A for example displays) and to indicate the target's left/right orientation. Unknown to the observers, half of the search displays are repeatedly presented with invariant target-distractor configurations (old contexts) throughout the experiment. The other half of displays, by contrast, presents repeated target locations embedded in layouts of randomly generated distractor arrangements (new contexts). This setup usually results in a contextual cueing effect, namely faster response times (RTs) to old context displays, as compared to new contexts. Moreover, observers are mostly unable to discern the repeated

☆ This article is part of a special issue entitled: 'Scene Processing' published in Vision Research.

* Corresponding author at: Allgemeine und Experimentelle Psychologie, Department Psychologie, Ludwig-Maximilians-Universität München, Leopoldstr. 13, D-80802 München, Germany.

E-mail address: conci@psy.lmu.de (M. Conci).

https://doi.org/10.1016/j.visres.2025.108591
Received 4 December 2024; Received in revised form 18 March 2025; Accepted 20 March 2025
Available online 26 March 2025

search layouts from novel item arrangements above chance level in a final recognition test (Chun & Jiang, 1998; but see Vadillo, Konstantinidis & Shanks, 2016). Together, this pattern of results thus suggests that the repeated spatial layouts of invariant distractors are incidentally learned, that is, without the explicit instruction to learn and without explicit recognition of the repeated layouts. Nevertheless, these learned contextual layouts bias attention to the target more efficiently (Zinchenko, Conci, Töllner, Müller, & Geyer, 2020; see also Zhao, & Ren, 2020) by generating predictions from the repeated contexts that infer the likely location of the target (Zinchenko, Conci, Müller, & Geyer, 2018; 2024).

Contextual cueing has usually been explored in "artificial" search displays that present randomly generated arrangements of letters (1 target T and 11 distractor L's, see Fig. 1A) in order to provide a controlled perceptual input for learning. A comparable benefit was also reported in search displays that presented natural scene contexts. For example, Brockmole and Henderson (2006a; 2006b) presented pictures of indoor or outdoor environments with a target object (e.g. a letter T) embedded in the scene context (e.g. a bedroom). In old contexts, the same scenes would be repeatedly presented with the target appearing at the same location on every trial, whereas in new contexts, new scenes would be presented on every trial. The results mirrored the findings from artificial letter displays and showed a large and reliable contextual cueing effect, that is, search benefited substantially from the repeated scene contexts. Moreover, observers were clearly able to identify the repeated contexts and distinguish them from novel scene images in a final recognition test, thus indicating (in contrast to the usual finding with letter displays) that the rather strong cueing effect with natural scenes is clearly associated with reliable, above-chance level explicit learning (see also Rosenbaum & Jiang, 2013). Various other studies with real-life images also reported a similar benefit with frequently co-occurring objects (Mack & Eckstein, 2011), meaningful scene configurations (Conci & Müller, 2014) and in predictable scenes that facilitate attentional orienting (Summerfield, Lepsien, Gitelman, Mesulam & Nobre, 2006). Together, these results thus indicate that contextual learning facilitates attentional guidance not only in lab-based, artificial search tasks but also in more realistic, natural environments.

An important, ecologically relevant aspect of statistical learning in natural environments concerns the flexibility to update previously learned regularities after a change, which happen to occur frequently in daily life. For instance, the items in your local supermarket might be rearranged from time to time, such as a permanent relocation of the bread to a different shelf. Such changes would require that an existing memory representation is updated quickly in order to include the changed location, thus allowing to maintain efficient attentional guidance by the otherwise invariant context. However, studies that explored this type of relearning in contextual cueing with "standard" letter search arrays reported that the ability to incorporate changes in already-established contextual memory representations is severely limited: Efficient contextual cueing is typically found after few encounters with the repeated spatial item arrangements during initial learning, as compared to a rather inefficient and time consuming relearning of contextual cueing when the target in a given repeated context consistently changed its location (e.g., Conci & Zellin, 2022; Conci, Sun, & Müller, 2011; Makovski & Jiang, 2010; Manginelli & Pollmann, 2009; Yang, Coutinho, Greene, & Hannula, 2021; Zellin, Conci, von Mühlenen,

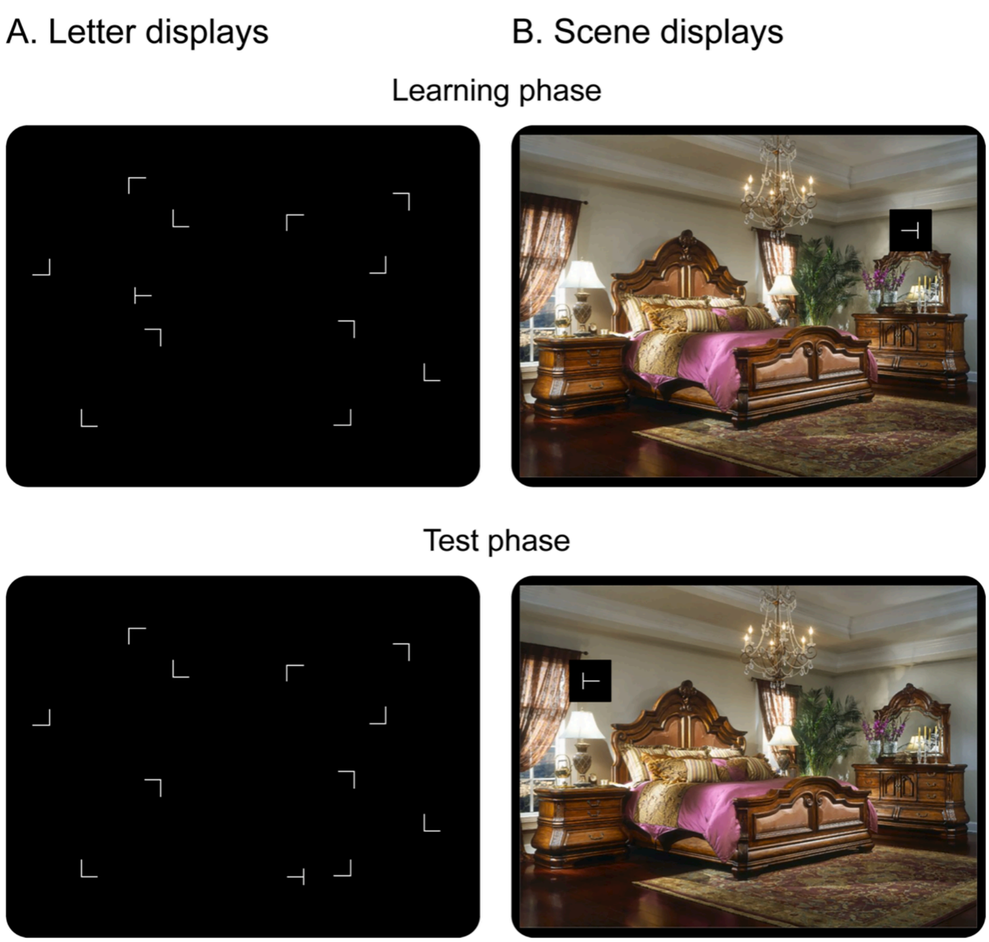

Fig. 1. Example of old-context letter (A) and scene (B) displays in the learning phase (top) and test phase (bottom) of the experiments. Search displays were initially paired with fixed, unique target locations in the learning phase. Subsequently, in the test phase, the targets were relocated and repeatedly presented at novel positions in the letter or scene contexts. Note that the displays are not drawn to scale.

& Müller, 2013; Zellin, von Mühlenen, Müller, & Conci, 2014; Zinchenko, Conci, Taylor, Müller, & Geyer, 2019; Zinchenko et al., 2020). For example, Zellin et al. (2014), reported that contextual cueing for the relocated targets only reemerged after some 80 repetitions of each repeated display arrangement after several days of training. Context-based learning thus appears to be efficient to initially register statistical regularities from our environment to guide search, but it seems to be surprisingly insensitive to task-relevant changes once an updating of the context–target associations is required. Importantly, this lack of efficient adaptation in contextual cueing is unlikely to have resulted from restrictions in overall memory capacity (Jiang, Song, & Rigas, 2005), or from a general lack of flexible learning of contextual information (Brockmole & Henderson, 2006a; Brockmole & Võ, 2010; Jiang & Wagner, 2004; Zang, Shi, Müller, & Conci, 2017), which, in natural environments, may be supported in particular by the availability of semantic (Goujon et al., 2012) and categorical (Bahle, Kershner, & Hollingworth, 2021) information and/or by global scene information (Brockmole et al., 2006; Brooks, Rasmussen, & Hollingworth, 2010).

One potential reason for the lack of efficient contextual adaptation may relate to the limited availability of rich environmental information in the usual laboratory-based contextual cueing experiments, as these would typically only present rather limited perceptual variations across all search displays (i.e. all displays simply consist of white lines that form letter shapes on a black background). Such rather simple arrays are usually employed to control potential confounding factors (e.g. scene semantics and/or positional constraints of real-world objects) that might in itself influence statistical learning (see Võ, 2021 for review). However, the rich variability of the stimuli in natural, real-world environments might conversely also support the flexible updating of statistical learning. One might therefore expect that the availability of "rich" environmental variation in natural scene displays not only boosts initial statistical learning and leads to explicit memory representations (Brockmole & Henderson, 2006b; Rosenbaum & Jiang, 2013), but these rich, real-world contexts could also facilitate flexible adaptation after an unexpected change of the target.

In light of these considerations, the current study reports two experiments that directly measured contextual cueing in natural scenes and artificial letter displays in order to test how different types of search context influence initial learning and the subsequent ability to update previously learned memories after a change of the target location.

2. Experiment 1

Experiment 1 employed a variant of the contextual cueing paradigm that used a learning-/test-phase design (e.g. Manginelli and Pollmann, 2009) to assess initial contextual cueing (in a learning phase) and the subsequent adaptation after a change of the target location (in a test phase) in search displays that (i) consist of typical letter arrangements (i.e., a target T among various distractor L's) with (ii) search displays that present the very same target T in a natural scene context (see Fig. 1 for example displays). While previous studies with letter displays typically revealed efficient contextual cueing during initial learning, followed by a rather inefficient contextual adaptation during the test phase (see above), it remains to be seen whether a comparable modulation of contextual learning and its updating is also evident in natural scene displays.

2.1. Methods

Participants. Experiment 1 tested a sample of 24 adults (7 men, 5 left-handed, mean age = 24.25 years, SD = 3.00 years). All participants reported normal or corrected-to-normal vision and were compensated with 10 Euros or course credits for participating in the experiment. The study was approved by the ethics committee of the Department of Psychology at LMU Munich, and all participants provided written informed consent prior to their participation.

Previous experiments that compared contextual cueing before and after a target location change across sequential phases typically tested between 12 and 16 participants (e.g., Zellin et al., 2013, 2014; Zinchenko et al., 2020). Moreover, a power analysis, which was based on the effect sizes reported in Zinchenko et al. (2020) revealed that a sample of only 8 participants would be required to detect a "lack-of-adaptation" effect in contextual cueing with an effect size f of 1.22 with a power of 95 % at an alpha level of 0.05. That is, on the basis of these previous studies, one would expect that the cueing effect should vanish after the target location change, which would be evidenced by a significant two-way context by phase interaction (alongside with a corresponding reduction in contextual cueing). We further increased (i.e., almost doubled) our sample relative to these previous studies to ensure sufficient statistical power to additionally detect a potential difference in contextual adaptation across letter and scene displays.

Apparatus and stimuli. Participants were seated in a dimly lit, experimental room. Stimulus presentation and data collection were controlled by a Windows 7 PC using Matlab and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). A standard mouse was used as the response device. The viewing distance was approximately 57 cm. Stimuli were presented on a 24-inch LCD monitor (with 1920 × 1080 pixel resolution and 60 Hz refresh rate).

There were two types of search displays that either presented letters or real-world scenes (Fig. 1A and 1B, respectively). Letter displays consisted of 12 items, one T-shaped target rotated 90° to the left or right and 11 L-shaped distractors, which were rotated randomly by 0°, 90°, 180°, or 270°. All items subtended 0.7° x 0.7° of visual angle and were presented in gray (8.5 cd/m^2) against a black background (0.2 cd/m^2). Search displays were generated by placing all items randomly within the cells of an invisible 6 × 8 matrix (cell size 2.5° × 2.5°), with the constraint that the target would neither appear in the four most central matrix cells nor in the two most lateral cells in the four corners of the display matrix. The target could thus appear in 36 out of 48 possible cells. The placement of the distractors was random. Moreover, distractors were jittered horizontally and vertically in steps of 0.1°, with a range of ± 0.6° within each cell to avoid collinearities between neighbouring items.

Scene displays presented one of 160 pictures of natural scenes (800 x 600 pixels) that were preselected from the "scene understanding of 397 categories" (SUN397) image database (Xiao, Hays, Ehinger, Oliva, & Torralba, 2010), and which depicted full-color photographs of conventional indoor rooms (e.g. living rooms, dining rooms, kitchens, bathrooms, bedrooms, dressing rooms and offices). A given search display would present one of the photographs within the boundaries of the 6 × 8 display matrix (15° × 20°), and the T-shaped target was added on top of the scene picture at one of the 36 possible matrix locations that was allowed to contain a target (see description above). To ensure that the gray target was clearly visible in front of the scene pictures, a small black patch, that subtended 0.9° in width and height, was presented behind the target (see example displays in Fig. 1B; see also Rosenbaum & Jiang, 2013, for a comparable approach). With these measures, the overall size of the search display and the physical properties of the target and its possible locations were identical for both letter and scene displays.

Trial sequence. Each trial started with the presentation of a fixation cross in the center of the screen for 500 ms. The fixation cross was then replaced by a (letter or scene) search display, which was presented until the participants responded. Participants were instructed to react as quickly and as accurately as possible by pressing the left or right mouse key to indicate the (left/right) pointing direction of the target T. In case of an erroneous response, a minus sign appeared in the center of the screen for 1000 ms. An inter-trial interval of 1000 ms separated one trial from the next. Different experimental blocks were separated from each other by a short break of 5000 ms.

Design and procedure. Experiment 1 used a repeated-measures design with the (within-subject) factors Display type (letter, scene), Context (old, new) and Phase (learning, test). Letter displays presented search layouts that are usually employed in "standard" contextual cueing experiments (e.g. Chun & Jiang, 1998) and which presented a T-shaped target among L-shaped distractors (Fig. 1A). Scene displays, in turn, presented the very same target T together with a natural scene picture (Fig. 1B; see also Brockmole & Henderson, 2006b; Rosenbaum & Jiang, 2013 for comparable search task variants). Old-context displays would present a repeated search layout throughout the experiment, while the new-context displays presented random contexts. A set of 6 old-context letter displays with an invariant arrangement of the distractors was generated for each observer and repeated throughout the experiment. Moreover, a second set of 6 old-context scene displays would repeatedly present the same six scene pictures (which were randomly selected for each observer) throughout the experiment. For new contexts, the configuration of distractor items in letter displays was generated randomly on each trial, while scene displays would, on each trial, present one randomly selected scene picture (from the 154 pictures that were not assigned to the old-context scenes). Each old-context display was paired with two different, randomly selected, target locations: One of these two target locations was presented in the first half of the experiment in the initial learning phase, while the second target was presented during the subsequent test phase. For new-context displays, another set of unique target locations was presented, which did, however, not change from learning to test. Thus, altogether 36 distinct target locations were presented, with 12 locations each assigned to the old-context displays during learning and test, and 12 to new-context displays. While target locations were fixed, the (left/right) orientation of the target was random on each trial to avoid that observers learn to associate a specific response with a given repeating (old) context.

The experiment started with a practice block of 24 randomly generated trials that presented an equal number of letter and scene search displays in random order. The subsequent, formal experiment consisted of 30 blocks of 24 trials each (yielding 720 trials in total), which were also presented in randomized order within each block. Each phase (learning, test) consisted of 15 consecutive blocks, with an equal number of old- and new-context, letter and scene search displays.

Recognition test. After the search task, a recognition test was administered to assess whether observers had formed any explicit memory of the repeated search displays. To this end, the 12 old-context (letter and scene) displays from the search task and 12 randomly generated/selected new-context (letter and scene) displays were shown. Participants were instructed with a text that was displayed on the monitor after the end of the main (search) experiment, and which remained on the screen until a button press was issued. They were asked to indicate whether or not they had repeatedly seen a given display previously, that is, participants were asked to recognize the picture –, or search display identity. There were six old- and new-context scene and letter displays each (i.e., 24 trials in total), which were presented in random order. Displays were presented with the target at the initial learning-phase location because explicit recognition of a given old context would be expected to manifest in particular for the initial, more reliably learned target-context association. The recognition responses were non-speeded, and no error feedback was provided.

2.2. Results

Search task. Overall, accuracy was rather high (98 %), ranging from 93 % to 100 % across observers. Given that there were only few errors, no further statistical analyses were performed.

Next, mean RTs were calculated for each factorial combination. Trials with response errors and RTs above 10 s and below 20 ms as well as trials with RTs above or below 3 standard deviations of each participant's mean for each condition were discarded (2 % of all trials in total). The mean RTs were then subjected to a repeated-measures analysis of variance (ANOVA) with the factors Display type (letter, scene), Context (old, new) and Phase (learning, test). This analysis yielded significant main effects of Context, $F(1, 23) = 23.69$, $p < 0.001$, $\eta_p^2 = 0.51$, and Phase, $F(1, 23) = 28.30$, $p < 0.001$, $\eta_p^2 = 0.55$, showing faster RTs to old- as compared to new-context displays (yielding a mean contextual cueing effect of 86 ms) as well as a reduction of the mean RTs (by 77 ms) from the initial learning to the subsequent test phase (see Fig. 2). The main effect of Display type was not significant ($p = 0.53$). In addition, the Context by Phase interaction was also significant, $F(1, 23) = 7.58$, $p = 0.01$, $\eta_p^2 = 0.25$. To decompose this interaction, Bonferroni-corrected post-hoc comparisons were performed, which revealed a reliable contextual cueing effect in the learning phase (118 ms, $t(23) = 5.58$, $p < 0.001$, Cohen's d = 0.66) as compared to a smaller and non-significant difference in the test phase (54 ms, $t(23) = 2.57$, $p = 0.08$, Cohen's d = 0.30). There was also a reliable reduction of contextual cueing (by 64 ms, or 54 %) from learning to test, $t(23) = 2.75$, $p = 0.01$, Cohen's d = 0.56, thus indicating that the change of the target location indeed hampered contextual cueing. Moreover, the Display type by Context interaction was also significant, $F(1, 23) = 9.83$, $p < 0.01$, $\eta_p^2 = 0.3$, which showed that contextual cueing with scene displays was substantially (i.e. almost two times) larger than for letter displays (126 ms vs. 47 ms, respectively), $t(23) = 3.14$, $p < 0.01$, Cohen's d = 0.64, indicating that scene search arrays indeed boosted contextual cueing. There were no other significant effects (all p's > 0.53), also including the (theoretically interesting) 3-way interaction, ($p = 0.12$).

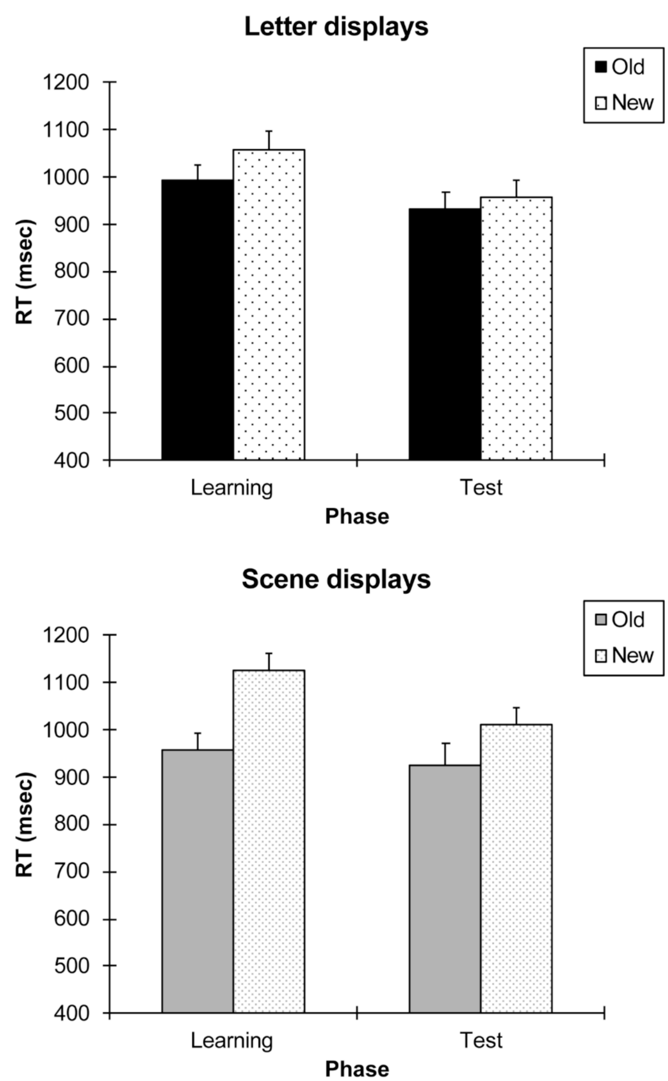

Fig. 2. Mean reaction times (RTs, in milliseconds) for old and new contexts (solid and dotted bars, respectively) in the learning and test phase for letter (top) and scene (bottom) displays in Experiment 1. Error bars show + 1 standard error of the mean.

To further assess the time course of contextual cueing after the change of the target location in the test phase, an additional analysis was performed on separate epochs in the experiment. Each epoch aggregated the RT data from five consecutive experimental blocks (to increase the statistical power). Mean contextual cueing effects were computed from the last epoch in the learning phase (epoch 3, which combines blocks 10 to 15) onwards, where initial learning would be expected to reveal a strong benefit – throughout the test phase until the end of the experiment. A repeated-measures ANOVA with the factors Display type (letter, scene) and Epoch (3–6) was performed on these mean contextual cueing effects. This analysis resulted in significant main effects of Display type, $F(1, 23) = 7.43$, $p < 0.02$, $\eta_p^2 = 0.24$, and Epoch, $F(3, 69) = 12.01$, $p < 0.001$, $\eta_p^2 = 0.34$, but no significant interaction ($p = 0.28$). As depicted in Fig. 3, contextual cueing was again larger with scene (110 ms) as compared to letter (43 ms) displays. In addition, at the end of the learning phase in epoch 3, a large cueing effect of 143 ms was observed, that then dropped substantially after the target location change in epoch 4 (19 ms, $t(23) = 5.46$, $p < 0.001$, Cohen's d = 0.88) while recovering (at least numerically) in epoch 5 (45 ms) and epoch 6 (99 ms, albeit not revealing a significant recovery across sequential [Bonferroni-corrected] comparisons across epochs, p's > 0.10). Finally, a series of one-sample t-tests additionally revealed contextual cueing to be reliably different from zero in epochs 3, 5 and 6 ($t(23)$'s > 2.10, p's < 0.05, Cohen's d's > 0.43), but not in epoch 4 ($p = 0.47$), thus showing that the target location change resulted in a transient reduction and delayed recovery of contextual cueing.

Recognition test. The mean accuracy of recognizing old and new letter displays was 55 %, as compared to 91 % for scene displays. To further compare the hits (correct categorization of old-context displays as 'old') to the rate of false alarms (erroneous categorization of new-context displays as 'old'), we computed a repeated-measures ANOVA with the factors Display type (letter, scene) and Response type (hits, false alarms). This analysis revealed a significant main effect of Response type, $F(1, 23) = 185.28$, $p < 0.001$, $\eta_p^2 = 0.89$, that was additionally modulated by Display type – as evidenced by a significant interaction effect, $F(1, 23) = 154.33$, $p < 0.001$, $\eta_p^2 = 0.87$. The main effect of Display type was not significant (p = 0.53). Bonferroni-corrected post-hoc comparisons, performed to decompose the significant interaction, revealed that there was no difference between hits (53 %) and false alarms (44 %) in letter displays ($p = 0.38$), suggesting that context memory for the letter displays was not based on explicit recognition. By contrast, for scene displays the hits (92 %) were substantially higher than the corresponding false alarms (11 %), $t(23) = 18.43$, $p < 0.001$, Cohen's d = 4.38, which clearly indicates that participants developed strong explicit memories for the repeated scenes. This difference in recognition performance between letter and scene displays can also be quantified by means of the signal-detection sensitivity measure d-prime (d'), which was derived from the z-transformed hit and false alarm rates (see Macmillan & Creelman, 2004). The resulting mean d' scores were 0.22 for letter displays as compared to a much larger sensitivity d' score of 3.47 for scene displays, again illustrating that there was a large difference in memory sensitivity across the two display types, $t(23) = 10.91$, $p < 0.001$, Cohen's d = 3.11. Thus, scene displays not only differed from letter displays in terms of contextual cueing magnitude, but also with regards to the explicitness of the underlying memory representation.

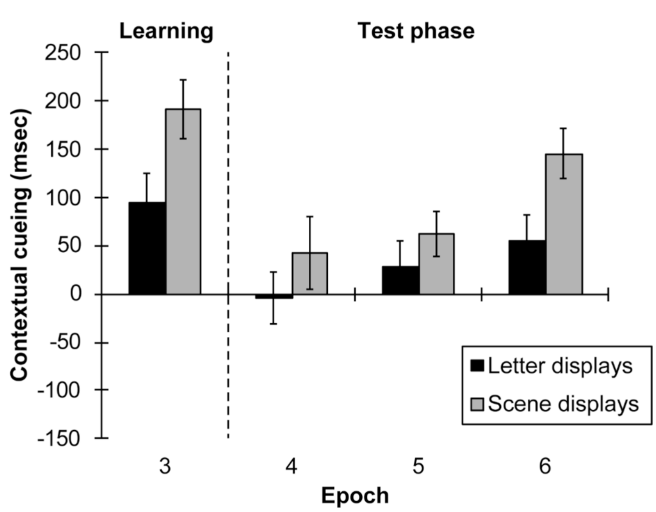

Fig. 3. Mean contextual cueing effects (in milliseconds) in letter and scene displays (black and gray bars, respectively) across epochs in Experiment 1. Each epoch presents the cueing effects from five consecutive experimental blocks. After epoch 3 (the last epoch in the learning phase), the target locations changed in epochs 4 to 6 (test phase). Error bars show +/- 1 standard error of the mean.

2.3. Discussion

In Experiment 1, contextual cueing was found to be substantially larger for scene (126 ms) than letter (47 ms) displays. This benefit of cueing for natural scene displays (as compared to letter displays) appears to relate to two (not mutually exclusive) sources, namely the (i) availability of "rich" contextual information in real-world environments and/or the (ii) explicit learning of the repeating scenes. For instance, while it has been shown that contextual cueing in letter searches essentially relies on the representation of a search-guiding spatial item layout (Chun & Jiang, 1998), in natural scenes, both spatial and object-related invariances as well as semantic and categorical associations might all contribute to the comparably large benefit for the repeating scene contexts (Brockmole & Henderson, 2006b). Moreover, contextual cueing with the repeating scenes not only boosted search but scene recognition performance was also very accurate (91 %), while revealing a high memory sensitivity (d' score: 3.47). This shows that observers formed explicit memory representations of the old scene contexts (see also Brockmole & Henderson, 2006b; Rosenbaum & Jiang, 2013). By contrast, for letter displays, the recognition accuracy for the repeating search layouts was at chance level (53 %) and memory sensitivity was poor (d' score: 0.22), which is again largely consistent with comparable studies that suggested search guidance in these displays to rely predominantly on incidental learning and implicit memory representations (Goujon et al., 2015, for review). Thus, both the amount of contextual detail and the awareness for repeating displays might contribute to the larger contextual benefit for real-life scenes.

Irrespective of these overall differences in contextual cueing between letter and scene displays, the sudden change of the target location (within otherwise unchanged contextual layouts) resulted in substantial costs (i.e. a reduction of cueing by some 57 %), with these costs being particularly strong directly after the change, thus essentially replicating previous findings. For instance, a comparable reduction of cueing after a target location change (within otherwise unchanged letter displays) was also reported in a recent *meta*-analysis (Annac, Conci, Müller & Geyer, 2017). Moreover, the reduction after the change was statistically (and numerically) comparable with both letter and scene displays. Together, these findings thus suggest that while initial contextual learning is rather fast and efficient, the updating of an already-existing context representation after a change is more effortful (Zellin et al., 2014). Importantly, the current experiment extends these previous findings by showing that this cost after the change appears to occur to a similar extent both in letter and scene displays.

3. Experiment 2

Experiment 2 was performed to dissociate the role of "rich" contextual details from concurrent effects of awareness about repeating search layouts during initial learning and the subsequent updating (after a change) in contextual cueing with scene and letter displays. To this end, slight changes to the experimental design were introduced to test contextual cueing in rich contexts (as given in natural scenes) before and after a change, when there is no (or only reduced) awareness about the repeating (letter and scene) search displays. For instance, in Experiment 1, only 6 scene pictures were consistently repeated throughout the entire experiment and presented in each block (in the old-context condition), whereas the new-context scene displays would present a randomly selected picture of the remaining 154 scenes on each trial throughout the experiment. It thus seems likely that the large amount of repetitions of the few old-context scenes (presented among many variants of the new-context scenes) resulted in explicit memory for these frequently repeating scenes, while a comparable effect was absent in the letter displays, where the old-context search layouts were also repeated to equal amounts, but they were supposedly harder to identify among the new-context displays, given that all displays comprised rather similar objects (i.e. one gray target T and eleven gray distractor L's) that were presented in every single (letter) display on a uniform black background.

Experiment 2 thus aimed to reduce the conspicuity of the few, repeating (old-context) scene pictures (to make the scenes in this regard more comparable to letter displays). To this end, a new baseline condition was introduced, which replaced the new-context condition used in Experiment 1 (see also Chaumon, Drouet, & Tallon-Baudry, 2008, for a comparable task variant). In order to reduce the explicit recognition of the 6 old scene displays, a second set of 6 scenes was used for this baseline measure. A similar change was also implemented in the letter displays, thus essentially presenting repeating scene and repeating letter displays on every trial in the experiment. Now, in order to induce learning of predictable target-context associations, in some displays (termed the "old-fixed" displays), each single repeated context would be presented with a repeating (fixed) target location throughout the learning phase of the experiment, and with a second repeating target location for a given repeated context in the subsequent test phase (i.e., identical to the procedure used for old-context displays in Experiment 1). By contrast, in "old-random" (scene and letter) contexts, randomly changing target-context pairings would be presented on every trial. That is, in old-random contexts six repeating scene and six repeating letter displays would each be randomly paired with six unique, yet constantly changing target locations. Given this, observers would not be able to associate a given, repeating context with a given, repeating target location. These old-random contexts thus replaced the new-context baseline as used in Experiment 1. Moreover, to further increase the similarity across the repeating scenes, in Experiment 2, scene displays were chosen from a reduced set of 24 scenes that all depicted images of living rooms (see examples in Fig. 4), thereby increasing scene similarity which should further reduce the ability to explicitly recognize consistent target-scene context associations.

3.1. Methods

Experiment 2 tested a new sample of 29 adults (8 men, all right-handed, mean age = 22.97 years, SD = 3.00 years). All of them reported normal or corrected-to-normal vision and they received payment (10 Euros) or course credits for their participation in the experiment.

Apparatus, stimuli, design and procedure were similar to Experiment 1, except that we now refer to the "old" context condition as the "old-fixed" context condition. In Experiment 2, these old-fixed contexts were compared to a new "old-random" context condition that was used as a baseline (instead of the "new" context condition as employed in Experiment 1). For old-random context displays, six scene pictures (randomly selected from 24 natural scene images that depicted living rooms) and six invariant letter displays would be presented in every block of the experiment (in randomized order together with the 12 trials depicting old-fixed, scene and letter, contexts). Each of these six old-random scene and letter displays was associated with six unique target locations that would be randomly assigned to the six displays (separately for each display type) on a given trial, thus preventing that observers could learn a consistent target-context association. Given this, an equal amount of 12 distinct target locations would be assigned to the old-fixed condition and to the old-random condition, but only the old-fixed

Fig. 4. Examples of natural scene displays that depict living rooms as used in Experiment 2. Note that the displays are not drawn to scale.

condition would present predictable, and thus learnable target-context associations (Note that these predictable associations would again change from the initial learning to the test phase). By contrast, the old-random condition would always present one of six possible target locations, selected randomly on a given trial. We expected that this new variant of contextual cueing (with either variable or predictive targets together with the repeating contexts that were presented in all trials of the experiment) should lead to somewhat smaller benefits of learning as in the standard contextual cueing variant with random contexts, because observers could potentially also benefit from learning the repeating contexts on their own – without relation to the repeating target locations (see Vadillo, Giménez-Fernández, Beesley, Shanks & Luque, 2021), thus possibly reducing the magnitude of the observable cueing effect. For this reason, in Experiment 2 we slightly increased both the sample size (from N = 24 in Experiment 1 to N = 28 in Experiment 2, see below) and the number of blocks (from 30 blocks in Experiment 1 to 36 blocks in Experiment 2, thus presenting 18 blocks in both the learning and test phase).

In the final recognition test, observers were then (after an instruction on the screen) presented with the 12 old-fixed (letter and scene) displays and with the 12 old-random (letter and scene) displays from the previous search task. Participants were now asked to indicate whether or not they had previously seen a given display *together with a repeating target location* (i.e., as depicted in the old-fixed displays). Thus, in contrast to Experiment 1, the task was not simply to recognize a given repeating context, but to identify the repeating contexts that were paired with a repeating target. That is, the recognition test in Experiment 2 asked observers for their memory of the presented contingencies between a given picture/search display and the corresponding target location. It should be noted that the recognition test variant used in Experiment 1 essentially coincides with the "standard" tasks that are typically presented after contextual cueing experiments (see e.g. Chun & Jiang, 1998; Brockmole & Henderson, 2006b), while in Experiment 2, this test had to be changed given the change of the paradigm (which now always presented old, i.e. repeated displays together with either fixed or random targets). There were six old-fixed and old-random scene and letter displays each (24 trials in total) presented in random trial order in the recognition test.

3.2. Results

Search task. One participant revealed a rather high error rate of 25 % and was excluded from the data proper, thus leaving 28 participants for the actual analysis. The overall accuracy for these remaining 28 participants was again rather high (97 %), ranging from 93 % to 100 % across observes. Accuracies were thus again not analyzed any further.

For the RT analysis, erroneous responses and trials with extreme RTs (with identical outlier removal criteria as in Experiment 1) were again discarded (5 % of all trials). As above, we performed a repeated-measures ANOVA with the factors Display type (letter, scene), Context (old-fixed, old-random) and Phase (learning, test) on the mean RTs. This analysis yielded a significant Context by Phase interaction, $F(1, 27) = 28.32$, $p < 0.001$, $\eta_p^2 = 0.51$. Subsequent (Bonferroni-corrected) posthoc comparisons revealed faster RTs (by 63 ms) for old-fixed as compared to old-random context displays in the initial learning phase, $t(27) = 3.41$, $p < 0.01$, Cohen's d = 0.26, thus, showing a reliable contextual cueing effect with our newly established, old-random baseline contexts in the initial part of the experiment. By contrast, in the subsequent test phase, a non-significant contextual cost (of −31 ms) was observed, which showed that responses were (at least numerically) slower to old-fixed contexts with relocated targets than to targets in old-random contexts ($p = 0.63$; see also Fig. 5). There were no other significant effects (all p's > 0.06; the 3-way interaction was again not significant, $p = 0.83$). Overall, this pattern of results thus shows comparable contextual cueing effects for both types of display during initial learning, which then vanishes (for both letter and scene displays) after the location change.

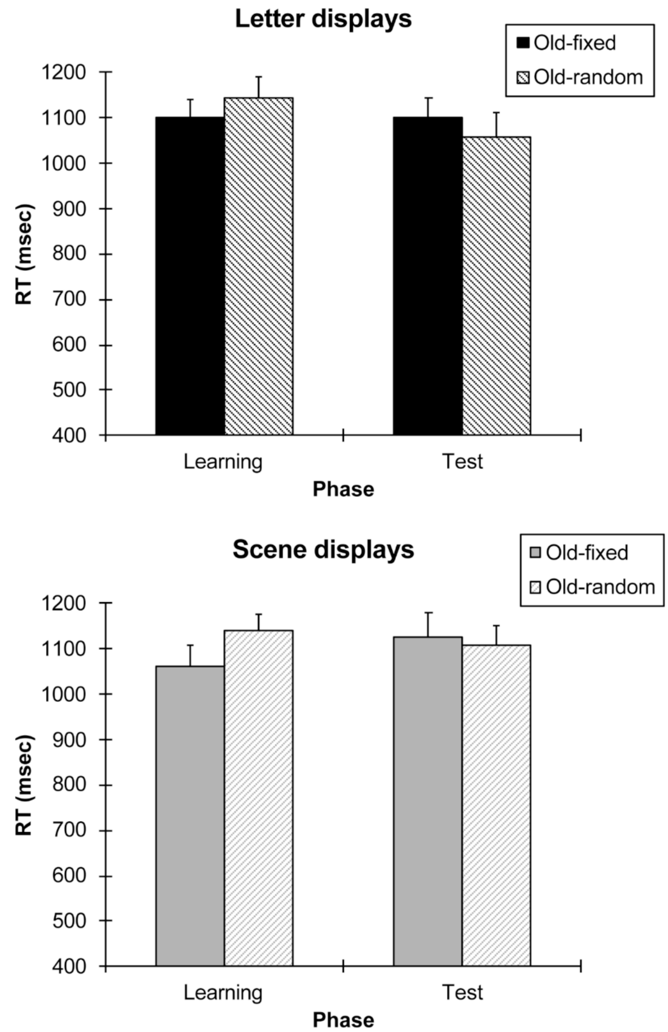

Fig. 5. Mean reaction times (RTs, in milliseconds) for old-fixed and old-random contexts (solid and striped bars, respectively) in the learning and test phase for letter (top) and scene (bottom) displays in Experiment 2. Error bars show + 1 standard error of the mean.

Next, we again analyzed the time course of contextual cueing after the target location change in the test phase, in an epoch-wise analysis that aggregated the RTs from six consecutive experimental blocks into one epoch. Analyses, were again performed on mean contextual cueing effects from the end of the learning phase (epoch 3) onwards. A repeated-measures ANOVA with the factors Display type (letter, scene) and Epoch (3–6) only revealed a significant main effect of Epoch, $F(3, 81) = 11.22$, $p < 0.001$, $\eta_p^2 = 0.29$, but no other significant effects (all p's > 0.3), thus indicating that contextual cueing varied across epochs (see Fig. 6). At the end of the learning phase (epoch 3) a large cueing effect (82 ms) was observed, that then dropped substantially after the target location change in epochs 4, 5 and 6 (−63 ms, −5 ms and −25 ms, respectively), $t(27)$'s < 3.32, $p < 0.01$, Cohen's d's > 0.45) while revealing no difference across epochs in the test phase (all p's > 0.16, with Bonferroni correction). Moreover, a series of one-sample t-tests additionally showed that contextual cueing was reliably larger than zero in epoch 3 ($t(27) < 4.50$, $p < 0.001$, Cohen's d > 0.85), while revealing a substantial cost in epoch 4 ($t(27) < 2.36$, $p < 0.03$, Cohen's d > 0.45) and non-significant deviations from zero in epochs 5 and 6 (p's > 0.41). This outcome again shows that the change of the target location led to a rather sustained reduction of contextual cueing.

An additional analysis was performed to compare contextual cueing across the two experiments with either random, "new context" displays

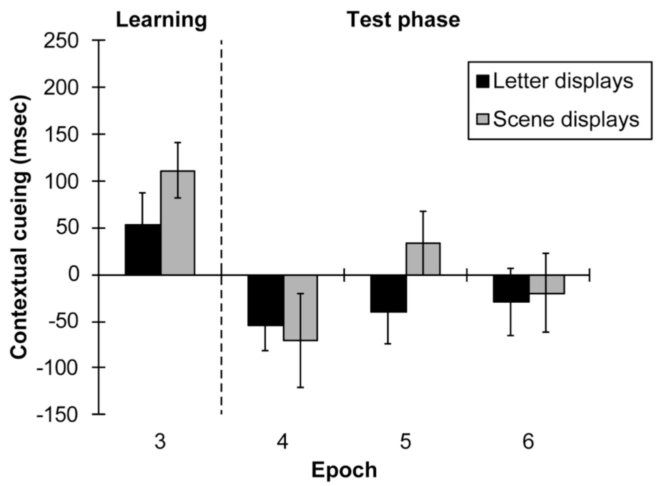

Fig. 6. Mean contextual cueing effects (in milliseconds) in letter and scene displays (black and gray bars, respectively) across epochs in Experiment 2. Each epoch presents the cueing effects from six consecutive experimental blocks. After epoch 3 (the last epoch in the learning phase), the target locations changed in epochs 4 to 6 (test phase). Error bars show +/- 1 standard error of the mean.

(Experiment 1), or with repeated displays and random target locations (i.e. "old-random" contexts; Experiment 2) as the baseline. A series of planned comparisons were performed using independent-samples t-tests, which showed that the cueing effect in the learning phase was significantly larger with scene displays in Experiment 1 (169 ms) than in Experiment 2 (80 ms), $t(50) = 2.18$, $p = 0.03$, Cohen's d = 0.61, while there was no significant difference across the two experiments in letter displays, despite revealing a certain numerical difference (68 ms and 46 ms in Experiments 1 and 2, respectively, $p = 0.61$). In the test phase, contextual cueing in scene displays was again larger in Experiment 1 (83 ms) than in Experiment 2 (−19 ms), $t(50) = 2.38$, $p = 0.02$, Cohen's d = 0.66, while there was again no significant difference in letter displays but only some numerical trend (26 ms and −43 ms in Experiments 1 and 2, respectively, $p = 0.07$). Together, these results thus show that contextual cueing is particularly pronounced with natural scene displays in Experiment 1, where the repeating scenes were compared to a large set of non-repeating, new-context displays.

Recognition test. In the final recognition test, the overall mean accuracy was 48 % for letter displays, as compared to 55 % for scene displays. It should be noted that the change of the baseline condition (from new-context displays in Experiment 1 to old-random contexts in Experiment 2) also changed the task in the recognition test (see also above). In Experiment 1, observers were required to recognize the repeating context displays among non-repeating, variable contexts. By contrast, in Experiment 2, all contexts were repeated, and observers were thus required to decide whether a given target location was repeated with a given context or not. The two task variants are therefore not directly comparable, but can nevertheless inform whether observers were aware of the repeating contexts (in Experiment 1) and/or of the repeating target-context contingencies (in Experiment 2).

In the analysis of Experiment 2, we compared the hits (correct categorization of the fixed target locations in repeating context displays as 'repeating') to the rate of false alarms (erroneous categorization of the randomly changing target locations in repeating context displays as 'repeating') for the two types of display by means of a repeated-measures ANOVA with the factors Display type (letter, scene) and Response type (hits, false alarms). This analysis revealed a significant main effect of Display type, $F(1, 27) = 11.20$, $p < 0.01$, $\eta_p^2 = 0.29$, which showed that observers tended to classify more targets in scene than letter displays as repeating (64 % vs. 46 %, respectively). This may thus be taken to indicate that the repeating scenes were partly recognized (and erroneously labelled as repeated) *without* relation to the repeated vs. random target locations. However, importantly, there were no significant (main or interaction) effects that involved the factor Response type (all p's > 0.12). For letter displays, the hits (44 %) were comparable to the false alarms (48 %, $p = 0.51$). Moreover, for scene displays, the hits (69 %) were also comparable to the false alarms (59 %, $p = 0.14$). This lack of a difference was also mirrored in the d' scores, that were again computed for Experiment 2, and which revealed a statistically comparable memory sensitivity for the repeating target locations in letter and scene displays (d' scores of −0.1 and 0.45, respectively, $p = 0.31$). Together, these findings thus show that recognition memory for target-context associations was incidental for both types of display in Experiment 2.

3.3. Discussion

Experiment 2 revealed a reliable contextual cueing effect (of 63 ms) that was comparable across letter and scene displays in the initial learning phase. After the target location change, the cueing effect then vanished and reversed into a numerical cost (of −31 ms) that was again comparable for both display types in the test phase (see Fig. 5), without showing major signs of recovery across sequential test-phase epochs (see Fig. 6). Overall, this substantial reduction of cueing after an unexpected change thus again replicates previous, comparable findings (e.g., Annac et al., 2017; Geyer, Zinchenko, Seitz, Balik, Müller, & Conci, 2024) and indicates that contextual learning is rather inflexible. Moreover, performance in the final recognition test suggested that observer's memory for the repeating target-context associations in both letter and scene displays was mostly implicit in nature. Together, these results thus indicate that initial contextual learning and its subsequent adaptation after a change is essentially comparable in both artificial letter – and "rich" natural scene displays, while statistical learning in both types of displays was based on incidental, non-explicit memories. The "richness" of a given search layout therefore does not appear to facilitate the updating of learning after an unexpected target location change both when learning of the context is explicit (in scene displays in Experiment 1) or context learning is implicit (in letter displays in Experiment 1 and in both letter and scene displays in Experiment 2).

4. General Discussion

The current experiments investigated contextual learning and its adaptation after a change in natural scenes, relative to more abstract letter arrays. In Experiment 1, contextual cueing was found to be substantially larger in magnitude in natural scene than in letter displays, while revealing clear evidence for explicit memory of the repeating scenes as opposed to incidental learning in letter displays (see also Brockmole & Henderson, 2006b; Rosenbaum & Jiang, 2013 for comparable findings with a similar task variant). However, despite these differences, the change of the target location led to comparable costs of contextual cueing in the test phase for both display types (relative to the initial learning phase), thus essentially replicating previous studies that used a comparable learning-/test-phase design with letter displays (see Annac et al., 2017, for a metanalysis). Next, in Experiment 2, the paradigm was changed slightly, now presenting *only* repeating search layouts with either fixed (i.e., predictable) or random (i.e. unpredictable) target locations. This task variant was implemented to test whether a comparable pattern of results would also be evident when the difference between predictable and unpredictable scene contexts would be less pronounced. The results again showed reliable cueing during initial learning for both types of display. However, in Experiment 2, the cueing effect for scenes was smaller than in Experiment 1 and statistically comparable to learning in letter displays, while now also revealing incidental/implicit learning with both display types. This shows that the change of the paradigm was successful in establishing non-explicit, comparable context memories across both display types. Moreover, after the change, cueing was reduced substantially and to equal amounts for both types of display, now even revealing a numerical cost. Together, these findings thus show that awareness about the repeating contexts may boost overall learning of the invariant target-context associations. However, unexpected changes lead to attentional misguidance with essentially comparable costs in artificial letter and natural scene displays.

While the dynamics of contextual learning and updating seem comparable across the two types of display, search in the letter and scene contexts nevertheless appears to be rather different from each other and it might thus be difficult to directly compare these two forms of search. For instance, letter displays consist of configurations that are all made up of horizontal and vertical line segments, and these high similarities across target and distractor stimuli tend to make the search task rather difficult (e.g. Duncan & Humphreys, 1989; Conci, Müller, & Elliott, 2007), thus leading to serial scanning of the display (e.g. Wolfe, 2021). By contrast, scene displays are more variable, and search in such real-world environments is not necessarily governed solely by a serial inspection of candidate target objects (see e.g. Evans & Treisman, 2005; Martin, Davis, Riesenhuber & Thorpe, 2018), while other properties of the scene (e.g. the overall scene gist) may additionally impact search (Wolfe, Võ, Evans, & Greene, 2011, for review). Despite these obvious differences in how search is performed in different arrays, the overall search RTs were nevertheless rather comparable across letter and scene displays (e.g. there was no main effect of Display type in both experiments). A direct comparison of search in different search displays may nevertheless be problematic, but it should be noted that the current study rather aimed to track the dynamics of initial learning and updating after a change *within* artificial search displays and real-world scenes.

In scene displays, the target not only consists of a T-shape, but it was always presented on top of a black square (see methods), which could additionally be used to guide attention more efficiently (at least in some displays). However, context-based learning seems not to be influenced by this background square since a large and explicit contextual cueing effect with real-world scenes has been reported both with and without an additional square underneath the target (see Experiment 1 in the current study and Brockmole & Henderson, 2006b, respectively). Similar findings were also reported in experiments that combined real-world scenes with letter search arrays, which were placed on background circles (Rosenbaum & Jiang, 2013) thereby eliminating any potential benefit of attentional guidance by the additional (square/circle) objects. Contextual cueing in real-world scenes thus appears to reveal a strong benefit that relies on explicit memory, but this benefit seems to be largely independent from variations of how the target was presented in these images.

Basic search may thus vary to some extent across display types, but our results additionally show that the type of memory representation (i.e. explicit vs. incidental/implicit) determines the overall strength of cueing: Contextual cueing in scene displays was much larger and the repeated contexts were clearly associated with explicit memory representations in Experiment 1, relative to a smaller cueing effect and implicit target-scene memory in Experiment 2. This pattern of results thus suggests that the explicit awareness about repeating scenes enhances the magnitude of cueing (see Brockmole & Henderson, 2006b). However, after the unexpected change, the costs were always comparable suggesting that awareness does not benefit the flexible updating of a previously learned context. Of note, both experiments presented scene displays that depicted a "rich" source of contextual information (e.g. in terms of object and scene semantics, which were available in addition to the invariant spatial relations of the search items). However, these rich contextual details did apparently not enhance contextual adaptation – both in Experiments 1 and 2. This may suggest that awareness about display repetitions improves contextual learning, but neither awareness about the repeating displays, nor the richness of the underlying contextual memory representation supports flexible contextual updating.

While explicit memory enhances overall contextual learning in the first place, inflexible contextual updating (after the change) may, at least partly, be due to the target location change that is typically not noticed. For instance, a lack of attention to the change seems to be associated with overly strong cue automatization which is counterproductive for search when the target is not at the expected location (Zinchenko et al., 2020). Thus, when a given repeated display layout is reencountered, the initially acquired context automatically biases attention to the initially learned target location, thereby facilitating search initially but also inducing a persistent erroneous attentional misguidance signal when the target has changed in the meantime. This misguidance signal is triggered both by abstract, geometric contexts and rich scene contexts to a similar extent, and it occurs irrespective of whether the target-context representation is explicit or implicit. However, a persistent misguidance signal may also be "corrected" (i.e. updated) by salient attentional cues to the changed target location (Conci & Zellin, 2022). The flexibility of contextual cueing may also be facilitated by additional global regularities (Zinchenko et al., 2024), thus illustrating that further information that helps to predict a change can be used to overcome the context-induced attentional misguidance that occurs after a task-critical change.

It should be noted that the difficulty to adapt a previously acquired context to an unexpected target location change does not imply that contextual learning is, per se, inflexible. Several types of changes to the search context, such as mirror-, or depth reversals, or a rescaling of the search displays were found not to harm contextual cueing (Brockmole & Henderson, 2006a; Jiang & Wagner, 2004; Zang et al., 2017). Moreover, the availability of (global) semantic and categorical cues have also been shown to facilitate memory flexibility (Bahle, Kershner, & Hollingworth, 2021; Brockmole et al., 2006; Brockmole & Võ, 2010; Brooks, Rasmussen, & Hollingworth, 2010; Goujon, Brockmole, & Ehinger, 2012). These findings thus indicate that some changes can in fact be incorporated into a given contextual memory representation as long as the previously learned target-context associations are not altered, or when additional cues are available to guide search. However, unexpected and non-explicit changes that disrupt previously learned associations seem to be rather hard to compensate.

Various previous studies debated about the overall role of incidental/implicit vs. explicit memory for the learning of a repeated context. For instance, contextual cueing has typically been attributed to rely upon incidental, implicit long-term memory (Chun & Jiang, 1998; 2003). However, other studies have questioned the implicit nature of contextual cueing and have suggested that observers are in fact aware of the repeating contexts (Smyth & Shanks, 2008) – with evidence for explicit memory typically depending on a sufficiently powered recognition test (Vadillo et al., 2016). Moreover, it has also been proposed that implicit and explicit memory measures may actually originate from a single memory system (Kroell, Schlagbauer, Zinchenko, Müller, & Geyer, 2019). It thus seems that contextual memory in "standard" cueing experiments could relate both to implicit/incidental memory, or to some weak, yet above chance-level recognition performance that would index at least some explicit memory for the repeating search layouts. However, compared to such "borderline" cases, our results with natural scenes (in Experiment 1) nevertheless reveal a qualitative difference as we show that context memory was clearly above chance level and explicit (even in a recognition test that is considered to be underpowered, Vadillo et al., 2016), while the memory for specific target-context associations (in Experiment 2) was mostly incidental. This finding thus indicates that memory representations in contextual cueing may in fact vary in strength, but this graded strength of the contextual memory representation primarily affects initial learning but does not impact the flexibility of contextual updating.

In addition to this differentiation in terms of memory strength, the current results also reveal *what* information is stored in a given contextual memory representation. For instance, our results show that contextual cueing is largely reflecting learning of an association between a fixed target and its invariant context. For instance, Experiment 2 showed a somewhat reduced, yet reliable contextual cueing effect during learning (relative to Experiment 1). We indeed expected that contextual cueing would be reduced to a certain extent in Experiment 2, because we compared repeated search layouts with fixed targets to a baseline that also presented repeated search layouts but with random targets. One might, thus assume that the baseline condition itself incurs some residual learning of the repeated layouts (irrespective of the repeated, i.e., fixed targets; see Vadillo et al., 2021), thus, potentially reducing the magnitude of cueing relative to Experiment 1 where random, i.e. "non-learnable" search layouts served as the baseline measure. However, with the repeated-context/random-target baseline (in Experiment 2), contextual cueing was nevertheless reliable and only showed a rather small, numerical reduction of cueing relative to the random-context baseline (in Experiment 1), which suggests that the (incidental) learning of specific target-context associations drives contextual cueing to a major extent relative to a smaller benefit that might result from learning of the context on its own (Beesley, Vadillo, Pearson, & Shanks, 2016; Vadillo et al., 2021).

In summary, the present results show that a previously learned context that is represented in long-term memory induces an attentional bias towards an expected target location, and this bias is hard to overcome both in artificial letter search displays and in "richer" natural scene displays, and even when the context memory is explicit. Importantly, this pattern of learning and relearning seems to be rather comparable in controlled laboratory settings and in more realistic real-world scenarios. One practical recommendation that follows from these findings would thus be that supermarkets should not rearrange their goods too often (see our real-world example presented in the introductory section and Croxton, 2012).

Author notes

This work was supported by a project grant from the German Research Foundation (DFG; FOR 2293). Feifei Zhao was supported by a Chongqing Social Science Planning Fund (2021NDQN70) and a scholarship from the Chinese Scholarship Council. We thank Dejan Draschkow for his help with the selection of the scene stimuli and for his valuable comments on an earlier draft of the manuscript.

All data, data-analysis scripts, and (scene) stimuli for the reported experiments have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/nf2mg/. The design and the analysis plans for the experiments were not preregistered.

CRediT authorship contribution statement

Markus Conci: Writing – review & editing, Writing – original draft, Visualization, Software, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Feifei Zhao:** Writing – review & editing, Methodology, Investigation, Funding acquisition, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data, data-analysis scripts, and (scene) stimuli for the reported experiments have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/nf2mg/.

References

Annac, E., Conci, M., Müller, H. J., & Geyer, T. (2017). Local item density modulates adaptation of learned contextual cues. *Visual Cognition, 25*, 262–277.

Bahle, B., Kershner, A. M., & Hollingworth, A. (2021). Categorical cuing: Object categories structure the acquisition of statistical regularities to guide visual search. *Journal of Experimental Psychology: General, 150*, 2552–2566.

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.

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

Brockmole, J. R., Castelhano, M. S., & Henderson, J. M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 699–706.

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

Brockmole, J. R., & Henderson, J. M. (2006b). Using real-world scenes as contextual cues for search. *Visual Cognition, 13*, 99–108.

Brockmole, J. R., & Võ, M.-L.-H. (2010). Semantic memory for contextual regularities within and across scene categories: Evidence from eye movements. *Attention, Perception, & Psychophysics, 72*, 1803–1813.

Brooks, D. I., Rasmussen, I. P., & Hollingworth, A. (2010). The nesting of search contexts within natural scenes: Evidence from contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1406–1418.

Chaumon, M., Drouet, V., & Tallon-Baudry, C. (2008). Unconscious associative memory affects visual processing before 100 ms. *Journal of Vision, 8(3):10*, 1–10.

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.

Conci, M., & Müller, H. J. (2014). Global scene layout modulates contextual learning in change detection. *Frontiers in Psychology, 5*, 89.

Conci, M., & Zellin, M. (2022). Stimulus-driven updating of long-term context memories in visual search. *Psychological Research, 86*, 252–267.

Conci, M., Müller, H. J., & Elliott, M. A. (2007). Closure of salient regions determines search for a collinear target configuration. *Perception & Psychophysics, 69*, 32–47.

Conci, M., Sun, L., & Müller, H. J. (2011). Contextual remapping in visual search after predictable target-location changes. *Psychological Research, 75*, 279–289.

Croxton, M. (2012). Older shoppers irritated by supermarket layout changes. Retail-Wire (May 22). https://retailwire.com/discussion/older-shoppers-irritated-by-supermarket-layout-changes/.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*, 433–458.

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: Is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1476–1492.

Geyer, T., Zinchenko, A., Seitz, W., Balik, M., Müller, H. J., & Conci, M. (2024). Mission impossible? Spatial context relearning following a target relocation event depends on cue predictiveness. *Psychonomic Bulletin & Review, 31*, 148–155.

Goujon, A., Brockmole, J. R., & Ehinger, K. A. (2012). How visual and semantic information influence learning in familiar contexts. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 1315–1327.

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

Jiang, Y., & Wagner, L. C. (2004). What is learned in spatial contextual cuing - Configuration or individual locations? *Attention, Perception & Psychophysics, 66*, 454–463.

Jiang, Y., Song, J. H., & Rigas, A. (2005). High-capacity spatial contextual memory. *Psychonomic Bulletin & Review, 12*(3), 524–529.

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

Mack, S. C., & Eckstein, M. P. (2011). Object co-occurrence serves as a contextual cue to guide and facilitate visual search in a natural viewing environment. *Journal of Vision, 11(9):9*, 1–16.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide* (2nd ed.). Cambridge, UK: Cambridge University Press.

Makovski, T., & Jiang, Y. V. (2010). Contextual cost: When a visual- search target is not where it should be. *Quarterly Journal of Experimental Psychology, 63*, 216–225.

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

Martin, J. G., Davis, C. E., Riesenhuber, M., & Thorpe, S. J. (2018). Zapping 500 faces in less than 100 seconds: Evidence for extremely fast and sustained continuous visual search. *Scientific Reports, 8*, 12482.

Nobre, A. C., & Stokes, M. G. (2019). Premembering experience: A hierarchy of time-scales for proactive attention. *Neuron, 104*, 132–146.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences, 11*, 520–827.

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

Rosenbaum, G. M., & Jiang, Y. V. (2013). Interaction between scene-based and array-based contextual cueing. *Attention, Perception, & Psychophysics, 75*, 888–899.

Smyth, A. C., & Shanks, D. R. (2008). Awareness in contextual cuing with extended and concurrent explicit tests. *Memory & Cognition, 36*, 403–415.

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

Vadillo, M. A., Giménez-Fernández, T., Beesley, T., Shanks, D. R., & Luque, D. (2021). There is more to contextual cuing than meets the eye: Improving visual search without attentional guidance toward predictable target locations. *Journal of Experimental Psychology: Human Perception and Performance, 47*, 116–120.

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

Võ, M.-L.-H. (2021). The meaning and structure of scenes. *Vision Research, 181*, 10–20.

Võ, M.-L.-H., Boettcher, S. E. P., & Draschkow, D. (2019). Reading scenes: How scene grammar guides attention and aids perception in real-world environments. *Current Opinion in Psychology, 29*, 205–210.

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review, 28*, 1060–1092.

Wolfe, J. M., Võ, M. L., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences, 15*, 77–84.

Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). SUN database: Large-scale scene recognition from abbey to zoo. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (San Francisco, CA, USA)*, 3485–3492.

Yang, Y., Coutinho, M. V., Greene, A. J., & Hannula, D. E. (2021). Contextual cueing is not flexible. *Consciousness and Cognition, 93*, Article 103164.

Zang, X., Shi, Z., Müller, H. J., & Conci, M. (2017). Contextual cueing in 3D visual search depends on representations in planar-, not depth-defined space. *Journal of Vision, 17*(5), 1–11.

Zellin, M., Conci, M., von Mühlenen, A., & Müller, H. J. (2013). Here today, gone tomorrow – Adaptation to change in memory-guided visual search. *PLoS ONE, 8*(3), Article e59466.

Zellin, M., von Mühlenen, A., Müller, H. J., & Conci, M. (2014). Long-term adaptation to change in implicit contextual learning. *Psychonomic Bulletin & Review, 21*, 1073–1079.

Zhao, F., & Ren, Y. (2020). Revisiting contextual cueing effects: The role of perceptual processing. *Attention, Perception, & Psychophysics, 82*, 1695–1709.

Zinchenko, A., Conci, M., Müller, H. J., & Geyer, T. (2018). Predictive visual search: Role of environmental regularities in the learning of context cues. *Attention, Perception, & Psychophysics, 80*(5), 1096–1109.

Zinchenko, A., Conci, M., Müller, H. J., & Geyer, T. (2024). Environmental regularities mitigate attentional misguidance in contextual cueing of visual search. *Journal of Experimental Psychology: Learning, Memory and Cognition, 50*, 699–711.

Zinchenko, A., Conci, M., Taylor, P. C. J., Müller, H. J., & Geyer, T. (2019). Taking attention out of context: Frontopolar transcranial magnetic stimulation abolishes the formation of new context memories in visual search. *Journal of Cognitive Neuroscience, 31*, 442–452.

Zinchenko, A., Conci, M., Töllner, T., Müller, H. J., & Geyer, T. (2020). Automatic guidance (and misguidance) of visuospatial attention by acquired scene memory: Evidence from an N1pc polarity reversal. *Psychological Science, 31*, 1531–1543.