Routledge
Taylor & Francis Group

Check for updates

Local item density modulates adaptation of learned contextual cues

Efsun Annac[a,b], Markus Conci[a], Hermann J. Müller[a,c] and Thomas Geyer[a]

[a]Department of Psychology, Ludwig-Maximilians-Universität München, Munich, Germany; [b]Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität München, Munich, Germany; [c]Department of Psychological Sciences, Birkbeck College, London, UK

ABSTRACT

In everyday scenes, searched-for targets do not appear in isolation, but are embedded within configurations of non-target or distractor items. If the position of the target relative to the distractors is invariant, such spatial contingencies are implicitly learned and come to guide visual scanning ("contextual cueing"). However, the effectiveness of contextual cueing depends heavily on the consistency between bottom-up perceptual input and context memory: following configural learning, re-locating targets to an unexpected location within an unchanged distractor context completely abolishes contextual cueing, and gains deriving from the invariant context recover only very slowly with increasing exposure to the changed displays. The current study induces variations of the local target context, i.e., item density, to investigate the relation between this factor and contextual adaptation. The results showed that learned contextual cues can be adapted quickly if the target is re-positioned to a sparse local distractor context (consisting of one neighbouring non-target item), as compared to no adaptation with a dense context (with three surrounding non-targets). This suggests that contextual adaptation is modulated by spatial factors and is not per se limited by order effects in the learning process.

Scanning a complex environment is a highly demanding process. Given this, having consistent context information can help to guide visual search and object recognition. One approach to studying memory-based visual search is provided by the contextual cueing paradigm (Chun & Jiang, 1998). In the standard paradigm, observers perform a relatively difficult search for a target letter T amongst letter L distractors. Unbeknown to them, half of the trials contain stable, i.e., repeated target–distractor (spatial) arrangements. The key finding is that reaction times (RTs) are faster to repeated than non-repeated displays, indicating that contextual regularities are learned and come to guide or "cue" visual search. While the importance of contextual memory to perception and attention is now widely acknowledged (see Cheung & Bar, 2012), recent work has shown that context memory is also severely limited. For example, Conci and collaborators reported that once observers have acquired a memory representation for a given target–distractor context, changes of the target location – within a repeated, i.e., unchanged, distractor layout – are difficult to incorporate in the existing configural memory representation (Zellin,

von Mühlenen, Müller, & Conci, 2014, see Figure 1). The present study induces target position changes in learned distractor contexts under conditions that carefully control local item density, in order to examine how this factor affects the adaptation of contextual cueing.

Lack-of-adaption of contextual cueing

In their seminal study, Chun and Jiang (1998; see also Jiang, Song, & Rigas, 2005) proposed that the memory underlying contextual cueing is of high capacity. However, while learning of contextual information is efficient to acquire spatial regularities, it has nevertheless turned out to be remarkably inflexible to adapt learned regularities subsequent to environmental changes. For example, a number of experiments suggest that contextual cueing does not recover easily after target location changes within a learned display (Conci, Sun, & Müller, 2011; Makovski & Jiang, 2010; Manginelli & Pollmann, 2009; Zellin, Conci, von Mühlenen, & Müller, 2013). These studies typically consisted of two phases: learning and test. The key manipulation was a change of the target location at

---

CONTACT Efsun Annac ✉ Efsun.Annac@psy.lmu.de 💬 Department of Psychology, Ludwig-Maximilians-Universität München, Munich, Germany Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität München, Munich, Germany

the transition from the learning to the test phases while keeping the distractor locations unchanged (see Figure 1 for an illustration of the basic procedure). The main findings were: (1) contextual cueing was substantially reduced immediately after the target location change; and (2) the effect recovered only with massive amounts of training on the changed displays (see Zellin et al., 2014, who had their observers perform 3600 trials on the changed displays across several days, with contextual cueing recovering only after some 1500 trials following the presentation of the "relocated" displays). It should be noted that the problem to re-instantiate contextual cueing after a change of the target location is unlikely to result from limitations in memory capacity. For instance, Jiang et al. (2005) reported reliable contextual cueing across five days of training, where observers

Figure 1. Examples of repeated search displays used in the present experiments. During learning, observers are presented with invariant target–distractor layouts to promote contextual learning of the repeated displays. In the subsequent test phase, the location of the target is changed whereas the context of distractors remains as before. Experiment 1 was a baseline, introducing a change of the target location at the transition from the learning to the test phase while keeping the distractor arrangement constant. In Experiment 2, in addition to the changes of the target location, displays in the learning phase (left panel) always contained two distractors in the immediate surround of the target location. In the subsequent test phase (right panel), targets were presented in local contexts of one or three distractors (sparse or dense display conditions), respectively. Experiment 3 consisted of only a learning phase to examine the impact of sparse and dense display conditions on the initial learning of context cues. Experiment 4 tested the effect of density on new learning using the same procedure as Experiment 2, except that a novel set of distractor-dense and distractor-sparse displays was shown in the test phase. The grey squares, of approximately $7.5° \times 7.5°$, mark the target's local context and are provided here for the purpose of illustration only.

were presented with variable sets of repeated displays on each day. Thus, while the capacity of the memory underlying contextual cueing is quite large, the adaptation of the cueing effect subsequent to a change of the target location (in an otherwise invariant layout) appears to be much less efficient.

The lack of adaption may at least in part suggest a primacy effect in contextual cueing (see Jungé, Scholl, & Chun, 2007), with learning being largely confined to early phases of the experiment. However, such a "turn-off" of learning would be highly maladaptive in the real world since statistical structure – which is usually present in the environment – would go undetected.

Arguably, though, temporal constraints on learning may not be the only conceivable reason for the failure to adapt contextual cues. One other, as yet unexplored, cause may relate to bottom-up, spatial grouping processes. For instance, it has been shown that visual search is aided by perceptual, i.e., spatial, grouping, effectively making it easier to detect and/or recognize a target within grouped regions compared to the processing of individual items (e.g., Conci, Müller, & Elliott, 2007; see also Han, Humphreys, & Chen, 1999, for a comparison of the effectiveness of different grouping mechanisms in vision). Such grouping and segmentation processes may reveal a crucial influence on paradigms that examine the adaptability of contextual cueing. For instance, it is possible that, in changed displays, the target and the distractors may be grouped together, thus forming a segmentable region (Conci & von Mühlenen, 2009) or a "salient" cluster of items that is prioritized for the assignment of bottom-up attention. Thus, because the search display is removed upon the response, insufficient time may be available for the system to encode, i.e., (re-)learn, the spatial target–distractor relations (e.g., Ogawa & Kumada, 2008) or there may be no incentive for (re-)learning as the target is found efficiently. Note that these accounts are neutral as to whether bottom-up attention affects the processes of the build-up, i.e., acquisition of a new association between the changed target position and the old distractor context or the expression, i.e., retrieval of this (adapted) representation from context memory.

Alternatively, the rich local context provided by (attended) grouped regions might facilitate the
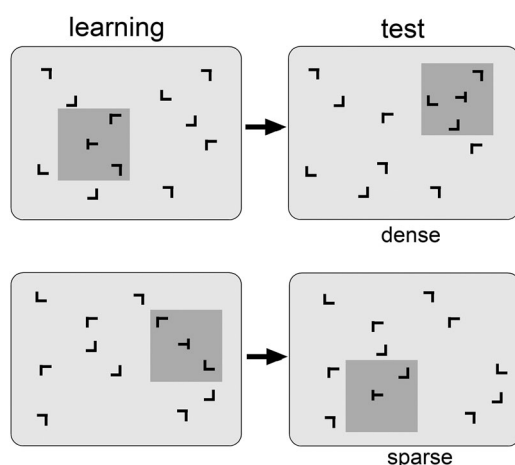
acquisition of the new target–distractor associations – thus promoting the adaptation of contextual cueing. It has, in fact, been shown that contextual cueing is particularly effective within segmented regions of an invariant display (Conci, Müller, & von Mühlenen, 2013; Geyer, Shi, & Müller, 2010; Hodsoll & Humphreys, 2005). In this view, presenting a changed target in a dense cluster of distractors may, in fact, facilitate the adaptation of contextual cueing because the dense local context surrounding the relocated target would allow a learned context to be effectively associated with the novel target location.

On this background, the present study was designed to test the effects of spatial grouping on contextual adaptation by systematically varying the number of distractors in the (relocated) target's immediate context. This manipulation is based on findings showing that contextual cueing is almost entirely supported by memory of individual target–distractor associations formed in the local vicinity, i.e., quadrant, of the target (Brady & Chun, 2007). Applied to the adaptation of contextual cueing, a manipulation of local target density should facilitate shifts of attention towards the grouped region either curtailing the time available for encoding individual distractor locations and thus updating the contextual memory representation (*local-grouping hypothesis*) or promoting the acquisition of new local target–distractor associations in the grouped (contextually-rich) region, engendering the rapid adaptation of contextual cueing (*local-learning hypothesis*). On both accounts, fast (bottom-up) stimulus-based and comparatively slower (top-down) memory-based processes would provide separable sources of information for search guidance, where contextual cueing acts in the way of memory-based automaticity (e.g., Chun & Jiang, 1998; Logan, 1988). Thus, although cueing operates automatically, it requires the activation of memory representations by the search array, which provides the guidance signals.

Rationale of the present study

In a departure from previous adaptation studies, which did not control for local density effects, the present study was designed to examine for the effects of local item density on contextual cueing by contrasting dense versus sparse distractor contexts. The study consisted of four experiments. Experiment 1 was a baseline experiment that investigated the "default" processes of

contextual learning and adaptation using arrangements of letter stimuli with the aim to replicate the basic pattern revealed in previous experiments. This experiment was divided into a learning and test phase, measuring contextual learning and adaptation, respectively. The critical manipulation was a change of the target position at the transition from the learning to the test phase (while keeping distractor locations unchanged). In Experiment 2, the placement of the target in the learning and test sessions was carefully controlled (unlike in Experiment 1 and in previous studies, where item locations were selected randomly), allowing the effects of spatial grouping on contextual cueing adaptation to be investigated. Specifically, in the initial learning phase, the target was positioned in such a way that it was always surrounded by two distractor items (see Figure 1). In the subsequent test phase, distractor-dense and distractor-sparse contexts were introduced, consisting of three and one distractor in the vicinity of the re-positioned target. In Experiment 3, distractor-dense and distractor-sparse contexts were introduced already in the initial learning phase in order to investigate how this factor affects the initial build-up of contextual cues (over and above the effects of density on the adaptation of learned contextual cues). Finally, Experiment 4 investigated the effects of spatial grouping on learning of a new set of repeated displays subsequent to initial learning of a different set of displays. In this experiment, distractor-dense and distractor-sparse contexts were novel in a sense that the respective arrangements were not shown in the previous learning phase.

Experiment 1 (baseline)

The purpose of Experiment 1 was to show, in the first instance, that with our stimuli we can actually replicate previous findings showing a general lack of adaptation in contextual cueing to relocated targets (e.g., Zellin et al., 2014).

Method

Participants and setup

For Experiment 1, we tested a sample of $N = 20$ observers (Experiments 2 and 3: $N = 14$ observers each; Experiment 4: $N = 13$ observers). But following previous investigations of contextual cueing (e.g., Conci & Müller, 2012; Kunar & Wolfe, 2011; Olson, Chun, & Allison, 2001;

Zellin et al., 2013), of these 20 participants we took into account only those who displayed an above-zero contextual cueing effect (RT [non-repeated display] minus RT [repeated display] > 0) in the learning phase. The rationale of this was that observers who fail to display contextual cueing in the learning phase are likely to acquire memory of the target in relation to the stable distractor configuration only later on, that is, they only show successful learning of relocated targets in the test phase (see Zellin et al., 2013, who, in an analysis of N = 38 excluded observers, showed this effect of late contextual learning). Conceivably, such (late) contextual learning of relocated targets may be preceded by some (target-independent) configural learning initially, with the target-context association being formed only later, during the test phase (cf. Beesley, Vadillo, Pearson, & Shanks, 2015). Thus, in order to examine true adaptation of previously learned target–distractor associations, rather than late learning, observers with negative contextual cueing scores in the learning session were excluded from the study (N = 7 observers in Experiment 1, N = 1 observer each in Experiments 2 and 3, and N = 0 observers in Experiment 4). These nine excluded observers showed a mean negative contextual cueing effect of 24 ms in the learning phase and a mean positive effect of 84 ms in the test phase (see Appendix).

Amongst the 13 participants in Experiment 1 with positive contextual cueing effects in the initial learning phase (six female; mean age 27.1 years, SD 3.83), all reported normal or corrected-to-normal vision and were naïve as to the purpose of the study. Prior to the experiment, participants provided written informed consent and were compensated with either course credit or monetary payment (8 Euro ∼ 9 USD). Note that the sample sizes of our remaining, selected group of observers (N = 13 observers in each experiment) are comparable to the sample sizes in previous studies that investigated multiple target location learning in contextual cueing, including the pioneering study of Chun and Jiang (1998).

Stimuli and design

The experiment was programmed in Matlab (version 7.5.0.342 R2007b), in combination with the OpenGL-Psychtoolbox extension (Brainard, 1997), and run on an Intel computer. Search displays used in both the learning and test phases contained one target and 11 distractor stimuli. The target was a T-shape,

rotated by either 90° or 270°, distractors were L-shaped and were rotated by either 0°, 90°, 180° or 270°. All stimuli subtended 0.7° × 0.7° of visual angle and were presented in white colour (67.0 cd/m^2) on a grey background (33.1 cd/m^2). The items of a search display were presented in the cells of an invisible 6 × 8 matrix. Cell size was 2.5° × 2.5°. The items were slightly jittered (0.1° × 0.1°) to prevent collinearities in the display.

Learning (phase 1)

The learning phase consisted of 384 trials divided into 24 blocks of 16 trials. Each block contained two different types of search displays: eight repeated and eight non-repeated displays. In repeated displays, the location of the target and the location and identities of distractors were held constant across trials. Non-repeated displays were generated anew on each trial. In order to equate target location repetition effects across the two conditions, the target was presented equally often at a fixed set of 16 locations across learning: eight locations were used for repeated eight (different) locations for non-repeated displays.

Test (phase 2)

The test phase was almost identical to the learning phase. There were 384 trials divided into 24 blocks of 16 trials. The major change consisted of the placement of the target at a novel location, while keeping the distractor locations unchanged (in the repeated condition). A total of 16 new target locations were selected at the beginning of the test phase. Eight locations were used in repeated and eight (different) locations in non-repeated displays. These new target locations were chosen in a way to equate the distances between old and new locations across the two types of displays (6.9° and 7.0° in repeated and non-repeated displays, respectively).

Procedure

Observers received written instructions at the beginning of the experiment. The experiment started with a practice session (16 trials) to familiarize observers with the task. The practice session was immediately followed by the learning and test sessions. Each search trial started with the presentation of a black fixation cross at the centre of the screen for 500 ms. The search array was presented subsequently and

remained until a response was issued. Participants were to respond as fast and as accurate as possible to the orientation of the target stimulus. If the target was tilted to the left (right) they pressed the left (right) key of the computer mouse with their left (right) index finger. After an erroneous response, a red horizontal line was presented for 1000 ms. The inter-trial interval was 500 ms.

At the end of the search experiment, observers performed a yes–no recognition test, intended for the assessment of explicit memory of the repeated configurations (cf. Chun & Jiang, 1998). To this end, eight repeated displays from the initial learning session and eight newly composed displays were shown and observers were asked to indicate whether they have seen a given display previously (by pressing the left and right mouse keys; unspeeded responses).

Results and discussion

Data analysis was performed using R (R Core Team, 2014). Both frequentist and Bayes analyses were performed. Bayes Factors were calculated using the package BayesFactor (Morey & Rouder, 2015). The natural logarithm of a Bayes factor (BF_{10}) is reported, with values greater than 1 [$= \log(3)$] providing substantial, and values greater than 2.3 [$= \log(10)$] strong evidence for the alternative hypothesis (Wetzels et al., 2011). Error trials and trials with extreme RTs (outside 2.5 SD from the individual mean) were discarded.

Accuracy

Overall response accuracy was 98.8%. A repeated-measures analysis of variance (ANOVA) on the error rates with the factors phase (learning vs. test) and display type (repeated vs. non-repeated displays) revealed neither of the main effects nor their interaction to be significant (all $Fs < 1$). If anything, error rates were slightly lower in the learning than in the test session (1.2% vs. 1.3%).

RT performance

RTs in the learning phase were collapsed into four epochs, with each epoch representing an average of six consecutive blocks to obtain reasonably stable RT estimates over time. For the learning phase, a 2 (display type: repeated, non-repeated) × 4 (epoch: 1–4) repeated-measures ANOVA revealed main effects

of display type, $F(1,12) = 15.38$, $p < .01$, $BF_{10} = 12.87$, and epoch, $F(3,36) = 16.16$, $p < .001$, $BF_{10} = 5.48$. For the test phase, by contrast, a 2 × 4 repeated-measures ANOVA failed to yield a significant main effect for either display type, $F(1,12) = 0.11$, $p = .73$, $BF_{10} = 0.15$, or epoch, $F(3,36) = 2.80$, $p > .05$, $BF_{10} = 0.02$. As shown in Figure 2 (panels A1 and A2), the contextual cueing effect (i.e., RTs for non-repeated minus repeated displays) dropped substantially between the learning and test phases (71 vs. 6 ms) and did not recover even with extended practice on the changed displays. This pattern clearly replicates previous (own) results relating to the lack of adaptation in contextual cueing (see Figure 3). Specifically, in our previous work, we conducted six other experiments on the adaptation of contextual cueing. All experiments revealed reliable differences in contextual cueing obtained between the learning and test sessions (all $ps < .01$). An overview across all (85) observers in the present Experiment 1 and the previous experiments revealed a mean contextual cueing effect of 131 ms in the learning phase, which compares to a contextual cueing effect of only 4 ms in the test phase (reduction: 97%). This pattern of results suggests that, while contextual cueing can provide a (potentially) powerful mechanism of search guidance, this mechanism is severely limited in its ability to adapt to target location changes.

Recognition test

Observer's ability to recognize repeated displays was estimated by using the signal detection measure d prime [$d' = Z$ (hit rate) $- Z$ (false-alarm rate); Green & Swets, 1966]. A hit means that observers correctly classified a repeated display as having seen this layout in the previous search task, while a false alarm means that they incorrectly judged a non-repeated layout as a repeated display. Mean d-prime score was .11, which was not significantly different from zero, $t(12) = 0.49$, $p = .63$, $BF_{10} = 0.46$. The low BF_{10} value supports the null hypothesis of no positive relationship between cueing and recognition. Thus, there was little evidence that observers had explicit knowledge of the repeated displays. One caveat here is, of course, the low power of the current recognition test (with only eight trials with repeated and eight with non-repeated displays), which limits any conclusions as regards the involvement of explicit
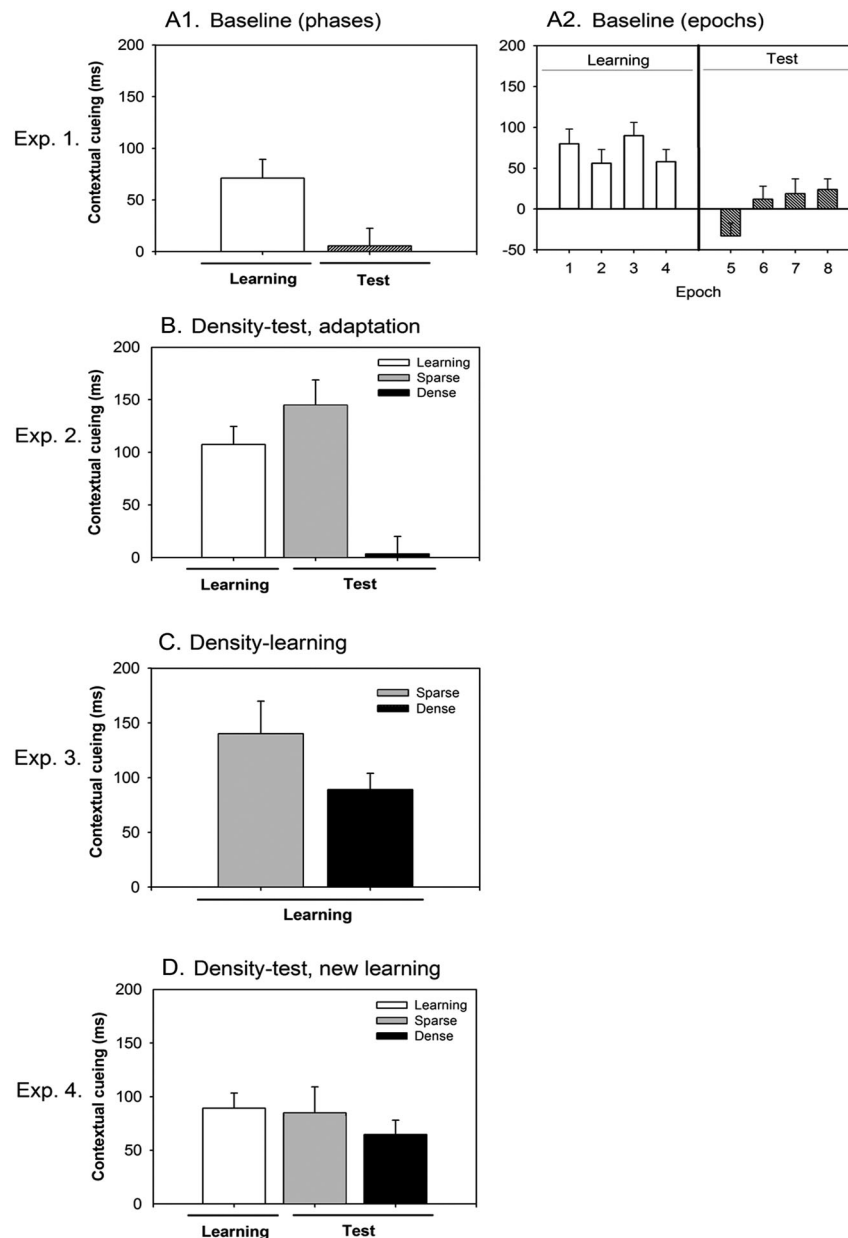
Figure 2. Results of Experiments 1–4. Panels A1 and A2: Contextual cueing as a function of phase (panel A1) and epoch (panel A2) in the baseline Experiment 1. Panel B: Experiment 2. Contextual cueing in learning and test, where the test phase presented a target location change in otherwise unchanged distractor-sparse and distractor-dense displays. Panel C: Experiment 3. Contextual cueing arising from dense and sparse displays in the initial learning session. Panel D: Experiment 4. Contextual cueing in learning and test, where the learning phase was identical to Experiment 1 and the test used a novel set of previously unseen distractor-dense and distractor-sparse displays. Contextual cueing is computed by subtracting reaction times to repeated displays from reaction times to non-repeated displays. Error bars represent the standard error of the mean.

knowledge in contextual cueing (see Vadillo, Konstantinidis, & Shanks, 2016).

Experiment 2 (density at test, adaptation)

Experiment 2 went on to re-examine the lack of contextual cueing found in Experiment 1 (and prior studies). In a departure from Experiment 1 (and

previous adaptation studies), which did not control for local density effects, Experiment 2 introduced a grouping manipulation, examining the effects of local item density on contextual adaptation. During initial learning, search layouts were presented with an "intermediate", i.e., baseline level of density, where two distractor items always surrounded the target. In the subsequent test phase, and following
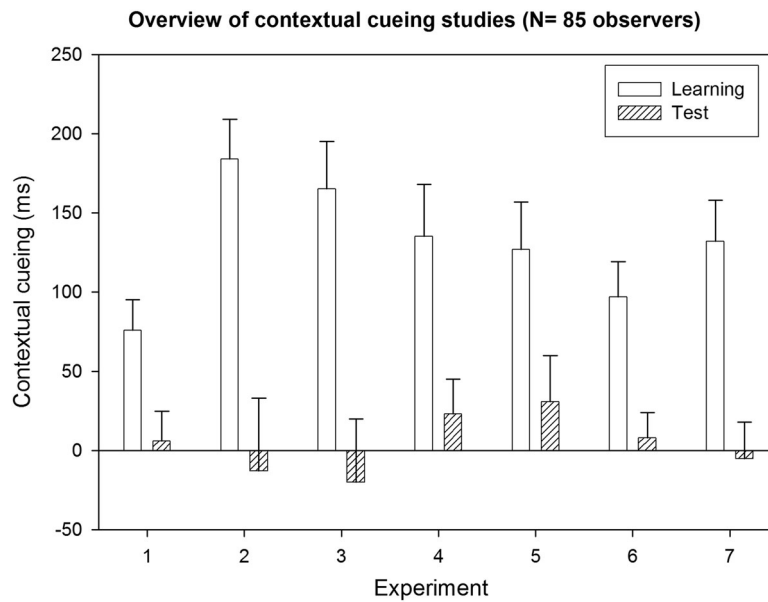
Overview of contextual cueing studies (N= 85 observers)

Figure 3. Overview of seven experiments (with $N = 85$ observers) showing that contextual cueing drops massively after target location changes at the transition from a learning to a test phase (white and dashed bars, respectively). Note that the distractor layout remained unchanged throughout the entire experiments. Experiments: (1) Experiment 1 of the current study; (2) Conci et al. (2011), Experiment 1; (3) Conci and Müller (2012), Experiment 2; (4) Zellin et al. (2013), Experiment 1A; (5) Zellin et al. (2013), Experiment 1B; (6) Zellin et al. (2013), Experiment 3; (7) Zellin et al. (2014), Experiment 1 (mean contextual cueing for test summarizes performance on day 1).

the target position changes, distractor-dense and distractor-sparse contexts were introduced, consisting of three and one distractor in the vicinity of the target (Figure 1). In total, participants were to learn eight repeated displays during the initial learning phase. During test, half of them became distractor-dense and the other half distractor-sparse displays.

Assuming that contextual cueing is supported by memory for local, i.e., individual, target–distractor pairs (Conci et al., 2013; Geyer et al., 2010; Hodsoll & Humphreys, 2005), re-positioning the target to a "rich" learning region – of many local distractor items – should facilitate the re-learning of target–distractor associations. The hypothesis of *local-learning* thus predicts rapid recovery of contextual cueing in the distractor-dense condition, which provides a contextually rich region for the re-learning of target–distractor relations. For the same reason, adaptation of contextual cueing should be less strong in the distractor-sparse condition. By contrast, the *local-grouping* hypothesis predicts no advantage of contextual cueing in the distractor-dense condition, but an advantage in the distractor-sparse condition. This is because with dense (but not sparse) displays, attention is guided efficiently in a bottom-up manner to the target region, thus decreasing the incentive for encoding the local target–distractor

relations, i.e., for relying on top-down contextual cues.

Method

The method of Experiment 2 was similar to Experiment 1, except for the details as follows.

Participants and setup

A total of 13 new participants took part in the experiment (eight female; mean age 24.0 years, *SD* 2.79).

Stimuli and design

The learning and test phases consisted of 384 trials each, divided into 24 blocks of 16 trials. Each block contained two different types of search displays: eight repeated and eight non-repeated displays. Repeated displays were generated prior to the experiment in order to implement different target densities across learning and test. This involved the creation of a set of eight search displays with manually chosen target and distractor locations. Note that the same repeated displays were shown to individual observers. For the learning phase, items were positioned to have exactly two distractors in the eight cells in the immediate surround of the target location (Figure 1). For the test phase, item locations were chosen to have three

or one distractors in the target's neighbouring cells, corresponding to the dense and sparse conditions, respectively. From the eight repeated displays in learning, four became dense and the other four sparse displays in the test phase. There were three further restrictions in determining target locations: (1) the target was never presented in the four central locations or in the corners of the display matrix; (2) targets were presented equally often in each of the four display quadrants in each experimental phase (learning, test) and display condition (repeated, non-repeated display); (3) the average distance of targets from the display centre was held constant at 3.5° across the four display type (repeated, non-repeated displays) x session (learning, test) conditions.

Results and discussion

Accuracy

Overall response accuracy was 98.2%. A repeated-measures ANOVA on the error rates with the factors phase (learning vs. test) and display type (repeated vs. non-repeated displays) revealed the main effect of phase to be significant, $F(1,12) = 4.89$, $p < .05$, $BF_{10} = 0.19$: error rates were slightly lower in the learning than in the test phase (1.6% vs. 2.0%).

RT performance

RTs in the learning phase were again collapsed into four epochs, with each epoch representing an average of six consecutive blocks to obtain reasonably stable RT estimates. For the test phase, contextual cueing was assessed by comparing RTs in dense and, respectively, sparse repeated displays against RTs from (in terms of density) comparable non-repeated displays, that is: for the non-repeated condition, only a subset of RTs was included in the analysis dependent on whether the target was embedded in a dense or a sparse local context (3 vs. 1 distractors). RTs were thus collapsed across all blocks of the test phase.

For the learning phase, a 2×4 repeated-measures ANOVA with the factors display type (repeated, non-repeated) and epoch (1–4) revealed the main effects of display type, $F(1,12) = 49.34$, $p < .01$, $BF_{10} = 3.93$, and epoch, $F(3,36) = 19.66$, $p < .01$, $BF_{10} = 2.44$, to be significant. The interaction was borderline-significant, $F(3,36) = 2.77$, $p = .05$, $BF_{10} = 0.80$, indicative of an increase in the gain from contextual cueing over the

course of the learning phase (from epoch 1: 73 ms to epoch 4: 128 ms). The mean cueing effect was 107 ms, demonstrating that participants formed robust contextual memories in the first phase of the experiment (Figure 2C).

For the test phase, a 2×2 repeated-measures ANOVA with the factors local context (dense, sparse) and display type (repeated, non-repeated) revealed the main effect of local context to be significant: $F(1,12) = 47.92$, $p < .01$, $BF_{10} = 3.18$: RTs were overall faster for dense than for sparse displays: 812 vs. 946 ms. Furthermore, and most importantly, the display type x local context interaction was significant, $F(1,12) = 34.03$, $p < .01$, $BF_{10} = 3.40$, with contextual cueing being stronger for sparse than for dense displays (145 vs. 4 ms; see Figure 2C).

This result clearly supports the *local-grouping hypothesis*, assuming a competitive relationship of attention guidance from contextual cueing and spatial grouping. In the presence of local grouping cues (distractor-dense condition), visual search is relatively effectively guided towards the target location (region) by these cues, leaving relatively little time for the in-depth processing of and adaptation to relocated displays. By contrast, in the absence of local grouping cues (distractor-sparse condition), visual search is to a large extent driven by long-term memory of repeated search arrangements, importantly, also including guidance from adapted context representations.

However, the above analysis also showed that RTs were significantly faster in the dense condition, which may be taken to suggest that mere response speed, rather than spatial inter-element grouping, modulates contextual cueing adaptation. For instance, it is possible that RTs in the distractor-dense condition leave no room for the modification of context memory.

This idea presupposes that context-based guidance of visual search will never manifest in the presence of local grouping cues at fast RTs. To test this, in Experiment 3 we examined contextual cueing in two new conditions. Here, distractor-dense and distractor-sparse contexts were introduced already in the initial learning phase. Under these conditions, it was possible to examine the impact of response speed particularly on the initial acquisition of context memory. If contextual cues can be established only with relatively slow RTs, then we expected a reliable cueing effect

only in the distractor-sparse condition. If, however the build-up of cueing is relatively independent of response speed, and thus the time available for processing the repeated displays is critical only in the adaption of context cues, then RT advantages due to contextual cueing should be observed in both the distractor-dense and the distractor-sparse condition.

Recognition test

Mean d prime was .13, which was not significantly different from zero, $t(12) = 0.81$, $p = .43$, $BF_{10} < 0$, suggesting that observers were unable to explicitly recognize repeated displays.

Experiment 3 (density at learning)

Experiment 3 was carried out to examine the alternative explanation that the original build-up of context memory, rather than the adaptation of an existing context representation, may occur only when there is enough time for the visual search, and contextual cueing, to evolve. If this alternative hypothesis is correct, then distractor-dense displays should produce hardly any contextual cueing effect during the initial learning of target–distractor arrangements.

Method

The method of Experiment 3 was essentially similar to Experiment 2, except for the following differences.

Participants and setup

A total of 13 new observers took part in this experiment (seven female; mean age 26.6 years, SD 4.02).

Stimuli and design

Participants performed a single "learning" session of 24 blocks of 16 trials each, yielding a total of 384 trials. A given block contained eight repeated and eight non-repeated trials. Among the repeated trials, half were distractor-dense and half distractor-sparse displays. The learning session was followed by a recognition test of one block of 16 trials.

Results and discussion

Accuracy

Overall response accuracy was 98.7%. A repeated-measures ANOVA on the error rates with the factors

epoch (1–4) and display type (repeated, non-repeated displays) revealed no effects, all $Fs < .1$, $BF_{10} < 1$.

RT performance

A 2×2 repeated-measures ANOVA with the factors local context (dense, sparse) and display type (repeated, non-repeated display) revealed the main effect of local context to be significant, $F(1,12) = 85.14$, $p < .001$, $BF_{10} = 17.20$: RTs were again faster for dense than for sparse displays: 969 vs. 1219 ms. Furthermore, the main effect of display type was significant, $F(1,12) = 36.30$, $p < .001$, $BF_{10} = 2.38$, with repeated displays giving rise to faster responses than non-repeated displays (1034 and 1149 ms). The display type x local context interaction was only borderline-significant $F(1,12) = 3.28$, $p = .09$, $BF_{10} = 25.25$, though the large BF_{10} value provides support for the hypothesis that contextual cueing has a lesser effect in dense than in sparse displays: 90 vs. 140 ms, respectively. This result pattern indicates that contextual cueing is not per se limited by the presence of local grouping cues: the cueing effect of 90 ms with relatively fast RTs was reliably larger than zero, $t(12) = 6.01$, $p < .01$, $BF_{10} = 8.29$. On the other hand, the cueing effect was reduced to some extent with faster as compared to slower search performance (i.e., for the distractor-dense vs. the distractor-sparse condition). Overall, this pattern of effects may be taken to suggest that spatial grouping is detrimental to contextual cueing, with the effect being most marked in the adaptation of (already) learned contextual cues.

Recognition test

Mean d prime was .23, which was not significantly different from zero, $t(12) = 0.88$, $p = .39$, $BF_{10} = .38$, suggesting that observers could not reliably tell apart repeated from non-repeated layouts.

Experiment 4 (density at test, new learning)

Experiment 3 showed reaction time facilitation for distractor-dense displays suggesting that processing time per se is not the limiting factor in context learning, but may be of particular importance for successful adaptation of previously learned contextual cueing displays. The requirement to spend a sufficient amount of time with the changed display may arise because of the high similarity between the original

and relocated displays (see General discussion for a more in-depth discussion of this idea). A prediction that follows from this proposal is that dense displays should generate a reliable contextual cueing effect even if these displays are shown only later in the experiment. To this end, Experiment 4 implemented a variant of the contextual cueing task which presented two entirely different sets of repeated displays during the learning and test phase (see Jiang et al., 2005; Zellin et al., 2013). Under these conditions, no interference should arise between the original and the novel set of displays because these displays would differ in terms of their underlying memory representations. As a result, a reliable cueing effect should again emerge even for the distractor-dense displays in the test session. This hypothesis was tested in Experiment 4.

Method

The method of Experiment 4 was essentially similar to Experiment 2, except for the following differences.

Participants and setup

A total of 13 new observers took part in this experiment (eight female; mean age 27.1 years, *SD* 4.25).

Stimuli and design

Participants performed "learning" and "test" sessions of 24 blocks of 16 trials each, with a total of 768 trials. In the learning session (which was identical to Experiment 2), in each block, participants encountered eight repeated and eight non-repeated displays with the target item always being surrounded by two distractor items. Target positions were different for repeated and non-repeated displays. During test, again eight repeated and eight non-repeated displays were shown in a given block of trials. Half of these displays were distractor-dense and the other half distractor-sparse displays, with the target surrounded by either three or one distractors, respectively. Unlike Experiment 2, observers were presented with different sets of repeated displays (and target positions) in each phase. Participants received one set of repeated contexts in the training phase (density level: 2) and another set in the test phase (density levels: 3, 1). At the end of the experiment, participants performed a recognition test of one block of 16 trials, which

presented the eight repeated displays from the test session and eight newly generated random layouts.

Results and discussion

Accuracy

Overall response accuracy was 98.6%. A repeated-measures ANOVA on the error rates with the factors phase (learning vs. test) and display type (repeated, non-repeated displays) revealed no effects, all $Fs < 1$, $BF_{10} < 1$.

RT performance

For the learning phase, a 2×4 repeated-measures ANOVA with the factors context (repeated, non-repeated displays) an epoch (1–4) revealed main effects of display type, $F(1,12) = 42.97$, $p < .001$, $BF_{10} = 10.85$, and epoch, $F(3,36) = 7.36$, $p < .001$, $BF_{10} = 0.10$, suggesting an emerging contextual cueing effect in the learning phase. The display type main effect shows a reliable contextual cueing effect of 89 ms.

In the test phase, RT performance was analysed my means of a 2×2 repeated-measures ANOVA with the factors local context (dense, sparse) and display type (repeated, non-repeated display). This ANOVA revealed the main effects of local context, $F(1,12) = 80.01$, $p < .001$, $BF_{10} = 2.80$, and display type, $F(1,12) = 28.92$, $p < .001$, $BF_{10} = 0.67$, to be significant. Further, the local context x display type interaction was borderline significant: $F(1,12) = 3.75$, $p = .07$, $BF_{10} = 0.71$. Concerning the main effect of local context, RTs were faster for distractor-dense than distractor-sparse displays: 912 vs. 1030 ms. The main effect of display type indicated that there was a reliable cueing effect of 75 ms overall in the test session. The (borderline significant) interaction was due to the fact that contextual cueing scores were reduced in distractor-dense compared to distractor-sparse displays: 64 and 85 ms, respectively. However, and as confirmed by direct *t*-tests, the RT difference between repeated and non-repeated contexts was highly reliable for both types of display ($ps < .01$). The results of Experiment 4 thus further confirm Experiment 3 in showing that distractor-dense displays do not per se hinder that contextual cueing emerges, even if these displays are shown only during later phases of the search experiment.

Recognition test

Mean *d* prime in Experiment 4 was .13, which was not significantly different from zero, $t(12) = 0.47$, $p = .64$, $BF_{10} = .30$, suggesting that observers could not reliably tell apart repeated from non-repeated arrangements.

General discussion

The present experiments investigated the adaptation of long-term memory for target–distractor associations (Chun & Jiang, 1998). Previous research showed that while contextual cueing is of high capacity (Jiang et al., 2005) and is considered a genuine form of effortless, nonconscious, learning (Colagiuri & Livesey, 2016), the memory underlying the cueing effect is quite resistant to forming an association between a new target location and an existing distractor representation. This "lack of adaptation" was replicated in Experiment 1 and in various other experiments conducted in the course of our research (see Figure 3). In Experiment 2, variations of local contexts were induced by carefully controlling the placement of the target before and after target location changes (see Figure 1). We observed that distractor-sparse contexts gave rise to robust adaptation of contextual cueing, while distractor-dense contexts did not. Further, mean RTs were faster overall for dense displays, suggesting that bottom-up search towards a dense target region conflicts with adaptation of learned contextual cues in these regions. Experiments 3 and 4 revealed that the lack of adaptation with dense displays was not due to a floor effect, that is, the overall faster RTs to displays with distractor-dense contexts. Moreover, across all three experiments, explicit recognition of repeated displays was not different from chance level, suggesting that observers had little explicit knowledge of the repeated displays.

Lack of adaptation of contextual cueing revisited

These findings provide evidence for the view that contextual cueing is an adaptive mechanism that is not per se restricted by order effects in context learning (with an overall advantage of early relative to late trials; see Jungé et al., 2007). Further, they appear to be at odds with Zellin et al. (2014), who found adaptation of contextual cueing only after massive amounts of training with the "relocated" displays.

However, Zellin et al. did not control for local item density – a manipulation that was central to the current experiments. It is thus possible that (fast) adaptation did take place in prior studies but was statistically reduced by local distractor density effects.

In the present study, local item density interacted with contextual cueing in two ways: first, sparse contexts produced a reliable contextual adaptation; second, dense contexts interfered with the adaptation of contextual cues. The fact that RTs were overall faster for dense displays demonstrates that the interference is likely due to attention capture by dense regions, facilitating the detection of the target when it is present in such a region, as was the case in the present experiment. Cueing then is effectively bypassed by stimulus-based, bottom-up factors. By contrast, memory-based cueing can come into play when search is more effortful and time-consuming, that is, when search is not immediately summoned to the target region based on display density factors. In this case, associations may be retrieved from memory and come to guide search. But again, for adaptation of a relocated target, this is only effective when the repositioned target is not located within a dense region.

Evidence for this proposal is provided by the findings that contextual cueing was smaller with dense than with sparse displays (though still reliable for the former) when these displays were presented in the initial learning phase or when a new set of repeated displays was shown in the test phase (after initial learning; Experiments 3 and 4, respectively). In addition, contextual cueing was absent in dense displays during test, when context memory about these displays had already been established during the learning phase (Experiment 2). Given these findings, it appears that the differences in contextual cueing with dense displays across learning and test are quantitative, rather than qualitative, in nature. In both phases, dense regions attract attention and thus shorten the time for in-depth processing of the display arrangements, including the detection and subsequent learning of target–distractor relations in these displays. Such temporal restrictions may be particularly harmful for the adaptation of contextual cueing, given that the re-learning of a changed, but relative to the original display highly similar, display would require extended processing time to encode the new target–distractor relations (recall that only

the target position differs between the original and the changed displays). However, since processing times are relatively short for dense displays, due to the effective capture of attention by the dense region, this counteracts the build-up of a new memory representation of the changed target position in relation to the previous (and unchanged) distractor positions.

Overall, this pattern resembles relatively recent work on contextual cueing in feature-singleton search (e.g., Geyer, Zehetleitner, & Müller, 2010; Ogawa & Kumada, 2008), showing that contextual cueing can also be observed for these relatively fast (pop-out) searches, but requires time to become effective. In these studies, the feature-singleton displays were preceded by placeholder displays (presented for some 700 ms), with the placeholders marking the locations of the subsequent search items, without providing information as to their identities (as distractors or target). The cueing effect was observed only under these "pre-view" conditions.

Adaptation, recognition and crowding

Although it is theoretically appealing to argue that spatial grouping plays a significant role in contextual cueing adaptation, there are other possible accounts that need to be considered: pattern recognition and crowding.

Concerning the former, it is possible that re-positioning the target across learning and test may have differentially changed the appearance of the display, making the original context either more or less recognizable. Given that the context-based guidance of visual search requires the activation of corresponding memory representations, it is conceivable that a small change in the position of one item may impact the display appearance and thus the recognition of changed displays. Assuming that dense contexts provide stronger cues for retrieving the originally learned configuration (e.g., because attending to a dense region, of three distractors, may more robustly trigger an originally learned spatial memory representations than attending to a sparse context, of only one distractor), this could explain why adaptation of contextual cueing failed with dense displays. However, this idea is from our point of view potentially unlikely because it is not clear whether targets surrounded by two distractors in the originally learned configurations

are actually better recognizable targets in dense displays (with three surrounding distractors). Instead, re-locating the target to a dense region may make the changed display less similar to the original display than when relocating the target to a sparse region. This may occur because the target location change modifies the display summary statistics such as the display items' centre of mass (e.g., Alvarez & Oliva, 2008) towards the dense region. The latter would predict the opposite outcome, namely that contextual cueing should recover for dense displays after the target location changes. Of interest in this regard, Manginelli and Pollmann (2009) found that, following target location changes in a contextual cueing adaptation task, RTs can actually be even slower to repeated than to non-repeated displays – which they attributed to a "misguidance" of attention towards the originally learned target location. If a dense context was more "reminiscent" of the original configuration, this would imply that focal attention tends to be misguided to the "wrong" – the originally learned – target location, where this misguidance would require a corrective, time-consuming shift of attention to the changed target position and thus increase RTs. But this is not what we observed: mean RTs were overall faster for dense than for sparse displays, suggesting that attention was effectively captured by the dense region. For this reason, we argue that our findings are better explained by a spatial grouping account

A second alternative account of the present findings may be in terms of "crowding" that is, a deterioration of performance for targets that appear in a dense (i.e., cluttered or "crowded') display region (e.g., Whitney & Levi, 2011) due to a loss of visual resolution in peripheral vision for such items. Applied to contextual cueing, for dense local contexts, crowding may impede target detection, which in turn could interfere with the adaptation of contextual cues. For instance, crowding may require search to operate with a narrow, item-based focus of attention, impeding the encoding and thus the re-learning of local target–distractor relations (cf. Lleras & Von Mühlenen, 2004). However, this account too encounters a difficulty: Crowding should not only reduce contextual cueing but also slow search, resulting longer RTs. However, RTs were faster, rather than slower, for dense displays in Experiments 2–4 of the present study.

Grouping and contextual cueing

We propose that the absence of contextual cueing in combination with fast RTs for dense displays is better explained by spatial grouping and its role in the guidance of attention, importantly also including perceptual long-term learning. This view complements previous ideas that propose a central role for grouping processes in visual search. For instance, Duncan and Humphreys ([1989](), [1992](); see also Humphreys & Müller, [1993]()), in their attentional engagement theory, propose that the output representations of an early, preattentive coding stage (so-called "structural units') are formed through the operation of basic grouping processes (notably grouping based on similarity, but also proximity). Structural units then compete for access to visual short-term memory, which is thought to be equivalent to the deployment of attention. Attentional engagement theory thus equates the efficiency of visual search with the efficiency of visual grouping (and segmentation) processes: the better the inter-element grouping – of the target and the distractors into separable target and non-target groups – the faster the selection of the target group. Besides bottom-up grouping processes, Duncan and Humphreys ([1989](), [1992]()) allowed for top-down – template-based – influences on visual selection, i.e., a top-down enhancement of items sharing features specified in the target template. The present study investigated one variant of such a top-down effect: long-term memory of (consistently encountered) target–distractor spatial relations, which may act as a specific form of "search template" guiding attention more effectively to the target location. Our results show that both stimulus-based (grouping) and memory-based (contextual cue) factors are available to guide visual search, where the contribution of context memory however is dependent on the efficiency of spatial grouping processes (see Conci & von Mühlenen, [2011](); Feldmann-Wüstefeld & Schubö, [2014](), for a related proposal, albeit investigating feature-based grouping in contextual cueing).

This relationship is best explained by theories assuming a reciprocal relationship between bottom-up attention and memory (e.g., Logan's, [1988](), instance theory of automaticity; see also Jiang & Leung, [2005](); or Annac et al., [2013](), for evidence of the effects of attention – both in terms of selectivity and processing resources deployed – on contextual cueing). At the heart of these proposals is the idea that repeated encounters of visual search displays lead to the build-up of memory about these displays, with the memory representation, in turn, facilitating detection performance. However, the present results qualify this relationship in that the effects of bottom-up attention on memory formation turned out to be highly specific, with a crucial difference depending on whether memory is established initially or whether an existing memory representation is updated to incorporate a changed target position into it. Only in the latter case did fast, bottom-up search processes interfere substantially with contextual cueing.

Note that our suggestion that context memory requires time to be adapted is essentially consistent with other studies on the effects of grouping on contextual cueing. For example, Feldmann-Wüstefeld and Schubö ([2014]()) manipulated the featural (colour, orientation) identity of the background distractors and found that an increase in similarity (or in their terms homogeneity) of the distractors can enhance contextual cueing (see also Conci & von Mühlenen, [2011]()). While this facilitatory effect of grouping on contextual cueing may in the first instance be difficult to reconcile with the current findings, a striking difference between Feldmann-Wüstefeld and Schubö ([2014]()) and the present investigation is that of the time, or experimental phase, when the grouping manipulation was applied. While Feldmann-Wüstefeld and Schubö ([2014]()) investigated grouping effects in the (initial) learning of context cues, here we addressed the issue of grouping effects in the adaptation of contextual cueing. In fact, spatial grouping during initial learning (Experiment 3) did not interfere with contextual cueing (though it was smaller with dense compared to sparse displays). Furthermore, it is possible that different forms of grouping vary in their effects upon contextual cueing (Conci & von Mühlenen, [2011]()). Assuming that grouping expedites visual search (Duncan & Humphreys, [1989]()), for grouping by proximity this may come along with the disadvantage of observers having less time to encode and subsequently learn target–distractor associations. This could have reduced the contextual cueing effect with dense displays in learning and effectively interfered (abolished) the adaptation of the cueing effect in the test phase. By contrast, feature-based grouping (investigated by Feldmann-Wüstefeld & Schubö,

2014), though effectively boosting visual search, may reinforce the impression of a coherent ensemble of distractor items and thus effectively enhance target–distractor learning in repeated displays. Whatever the explanation, together these studies suggest a role of grouping in contextual cueing. However, whether grouping facilitates or interferes with contextual cueing may be dependent on the type of grouping (spatial proximity versus featural grouping) and/or the experimental phase when grouping is administered (learning versus adaptation of contextual cueing). This might be an avenue for future research.

Conclusion

In conclusion, the present results show that contextual adaptation occurs more readily when the number of distractors surrounding the target is low. This rules out factors such as order effects in perceptual learning or a temporally sluggish re-learning process in the adaptation of learned target–distractor arrangements. Instead, the results suggest that one important factor for the lack of adaptation found in previous contextual cueing studies is spatial grouping, effectively guiding attention towards grouped regions, and thus limiting the time for processing the changed displays in order to bring into play existing target–distractor representations. The more general implication is that Gestalt factors relating to display design are crucial in investigations of contextual cueing in visual search. Future work should take this into consideration when interpreting context-based learning in visual search.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

References

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science, 19*, 392–398.

Annac, E., Manginelli, A. A., Pollmann, S., Shi, Z., Müller, H. J., & Geyer, T. (2013). Memory under pressure: Secondary-task effects on contextual cueing of visual search. *Journal of Vision, 13*(13), 6. doi:10.1167/13.13.6

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.

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*, 433–436.

Cheung, O. S., & Bar, M. (2012). Visual prediction and perceptual expertise. *International Journal of Psychophysiology, 83*(2), 156–163.

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

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

Conci, M., & Müller, H. J. (2012). Contextual learning of multiple target locations in visual search. *Visual Cognition, 20*, 746–770.

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., Müller, H. J., & von Mühlenen, A. (2013). Object-based implicit learning in visual search: Perceptual segmentation constrains contextual cueing. *Journal of Vision, 13*(3), 15, 1–17.

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

Conci, M., & von Mühlenen, A. (2009). Region segmentation and contextual cuing in visual search. *Attention, Perception, & Psychophysics, 71*, 1514–1524.

Conci, M., & von Mühlenen, A. (2011). Limitations of perceptual segmentation on contextual cueing in visual search. *Visual Cognition, 19*, 203–233.

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

Duncan, J., & Humphreys, G. W. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 578–588.

Feldmann-Wüstefeld, T., & Schubö, A. (2014). Stimulus homogeneity enhances implicit learning: Evidence from contextual cueing. *Vision Research, 97*, 108–116.

Geyer, T., Shi, Z., & Müller, H. J. (2010). Contextual cueing in multiconjunction visual search is dependent on color- and configuration-based intertrial contingencies. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 515–532.

Geyer, T., Zehetleitner, M., & Müller, H. J. (2010). Contextual cueing of pop-out visual search: When context guides the deployment of attention. *Journal of Vision*, *10*(5), 1–11.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Oxford: Wiley.

Han, S., Humphreys, G. W., & Chen, L. (1999). Uniform connectedness and classical gestalt principles of perceptual grouping. *Perception & Psychophysics*, *61*, 661–674.

Hodsoll, J. P., & Humphreys, G. W. (2005). Preview search and contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 1346–1358.

Humphreys, G. W., & Müller, H. J. (1993). Search via recursive rejection (SERR): A connectionist model of visual search. *Cognitive Psychology*, *25*, 43–110.

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

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

Jungé, J. A., Scholl, B. J., & Chun, M. M. (2007). How is spatial context learning integrated over signal versus noise? A primacy effect in contextual cueing. *Visual Cognition*, *15*, 1–11.

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

Lleras, A., & Von Mühlenen, A. (2004). Spatial context and top-down strategies in visual search. *Spatial Vision*, *17*, 465–482.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.

Makovski, T., & Jiang, Y. V. (2010). Contextual cost: When a visual-search target is not where it should be. *The 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.

Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes factors for common designs. R package version 0.9.10-1. Retrieved from http://cran.r-project.org/package=BayesFactor

Ogawa, H., & Kumada, T. (2008). The encoding process of non-configural information in contextual cuing. *Perception and Psychophysics*, *70*, 329–336.

Olson, I. R., Chun, M. M., & Allison, T. (2001). Contextual guidance of attention: Human intracranial event-related potential evidence for feedback modulation in anatomically early temporally late stages of visual processing. *Brain*, *124*(7), 1417–1425.

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. 2012: ISBN 3-900051-07-0.

Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review*, *23*(1), 87–102.

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical evidence in experimental psychology an empirical comparison using 855 t tests. *Perspectives on Psychological Science*, *6*, 291–298.

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, *15*, 160–168.

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), e59466. doi:10.1371/journal.pone.0059466

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.

Appendix

Additional analyses were performed to investigate the development of contextual cueing in the learning and test phase while also taking into account observers who displayed a negative contextual cueing effect initially and effective learning of re-located target in the later test session. To this end, our complete samples of $N = 20$, 14 and 14 observers in Experiments 1, 2 and 3, respectively, were included in the subsequent analysis. (Note that in Experiment 4 no observers were excluded and, hence, no additional analyses are reported.)

For experiment 1 (with 7 "late" learners), a 2 (display type: repeated, non-repeated) × 8 (epoch: 1–4 in learning; 5–8 in test) × 2 (phase: learning, test) repeated-measures ANOVA revealed main effects of display type, $F(1,19) = 12.28$, $p < .01$, $BF_{10} = 0.23$, and epoch, $F(3,57) = 32.58$, $p < .01$, $BF_{10} = 0.36$, in addition to a significant epoch x phase interaction, $F(3,57) = 5.90$, $p < .01$, $BF_{10} = 0.08$. The context x epoch x phase interaction did not reach significance, $F(3,57) = 1.71$, $p = .17$, $BF_{10} = 0.00$. Separate 2 × 4 repeated-measures for the learning and test phase revealed a borderline-significant effect of display type, $F(1,19) = 3.79$, $p = .06$, $BF_{10} = 0.18$, and a significant main effect of epoch, $F(3,57) = 22.53$, $p < .001$, $BF_{10} = 0.15$. The epoch and context interaction was not significant, $F(3,57) = 0.25$, $p = .85$, $BF_{10} = 0.02$. For the test phase, the 2 × 4 repeated-measures ANOVA yielded a significant main effect of display type, $F(1,19) = 5.26$, $p = .03$, $BF_{10} = 0.02$, and epoch, $F(3,57) = 7.04$, $p < .001$, $BF_{10} = 0.22$. Interestingly, the epoch and context interaction was also significant, $F(3,57) = 2.79$, $p < .05$, $BF_{10} = 0.00$, suggesting that the $N = 7$ observers with initial negative cueing effects display very high learning of re-located targets, leading to a reliable contextual cueing effect in the test session at the level of the entire sample.

In an attempt to further examine the hypothesis that observers who lack a contextual cueing effect initially develop a reliable effect later on, we performed a correlation analysis, comparing individual contextual cueing scores in the learning and test session (across the 20 observers in Experiment 1, given that this experiment contained the largest number of observers with initial negative cueing). Specifically, we predicted a significant negative correlation between contextual cueing values for two reasons: (1) contextual cueing, established during learning, should lead to relearning deficits at test (i.e., lack of adaptation effect; $N = 13$ observers in Experiment 1); (2) a failure to develop a contextual cueing effect during learning should be followed by a solid cueing effect in the subsequent test session (i.e., a late learning effect; $N = 7$ observers in Experiment 1). The results confirmed this prediction. Quantifying the relationship between contextual cueing scores across the training and test phases of Experiment 1 revealed a significant negative correlation: $r(20) = -.44$, $p = .05$, confidence interval: $-.74$ to .01.

For Experiment 2, a 2 × 4 repeated-measures ANOVA performed on the entire sample of $N = 14$ observers with the factors display type (repeated, non-repeated) and epoch (1–4) revealed a reliable

contextual cueing effect in the training phase (74 ms-effect; main effects of display type, $F(1,13) = 33.76$, $p < .001$, $BF_{10} = 15.39$), in addition to a reliable main effect of epoch, $F(3,39) = 19.98$, $p < .001$, $BF_{10} = 0.16$. Further, the interaction was significant, $F(3,39) = 4.19$, $p < .05$, $BF_{10} = 3.07$. For the test phase, contextual cueing was reliable only for sparse, but not dense, displays (153 and −16 ms-effects, respectively), as indicated by a 2×2 repeated-measures ANOVA with the factors local context (dense, sparse) and display type (repeated, non-repeated), which revealed the display type x local context interaction significant, $F(1,13) = 22.64$, $p < .01$, $BF_{10} = 18.66$. Further, RTs were faster for dense than for sparse displays: 829 vs. 954 ms (main effect of local context: $F(1,13) = 17.27$, $p < .01$, $BF_{10} = 0.62$).

For Experiment 3, the 2 (local context) $\times$ 2 (display type) repeated-measures ANOVA performed in the learning phase revealed a main effect of display type, $F(1,13) = 31.23$, $p < .001$, $BF_{10} = 0.56$, indicating that mean RTs were on average faster for repeated than non-repeated displays (1067 vs. 1146 ms) when analysing the entire sample of $N = 14$ observers. Further, RTs were faster for dense than for sparse displays: 998 vs. 1215 ms and the display type x local context interaction was significant $F(1,13) = 47.39$, $p < .001$, $BF_{10} = 7.03$, illustrating that contextual cueing was less pronounced in dense than in sparse displays: 88 vs. 141 ms, respectively.

In a final analysis, we analysed only the "excluded" observers and compared their mean RTs for repeated and non-repeated contexts by means of paired-sample t-tests for each experimental phase (learning, test). This analysis was limited to Experiment 1, as this was the experiment with the largest number of ($N = 7$) observers with negative contextual cueing scores in the learning session (and associated with this: differences in the results when either the data of a subset of observers or the entire sample was analysed). For the $N = 7$ "excluded" observers in Experiment 1, RTs were slower, by 57 ms for repeated compared to non-repeated contexts in the initial learning phases, $t(6) = 4.04$, $p < .01$, $BF_{10} = 2.20$. Subsequently, however, significant contextual cueing, of 88 ms was observed for relocated targets in the test phase, $t(6) = 5.00$, $p < .01$, $BF_{10} = 3.00$. This pattern shows that contextual cueing can develop rather late and thus massively confound data regarding the adaptation of initially acquired target–distractor associations. Thus, in order to examine the "true" adaptation of contextual cues, it is important to exclude observers with negative contextual cueing scores in the learning session.