

Meaning in learning: Contextual cueing relies on objects' visual features and not on objects' meaning

Tal Makovski¹

Published online: 2 August 2017
© Psychonomic Society, Inc. 2017

Abstract People easily learn regularities embedded in the environment and utilize them to facilitate visual search. Using images of real-world objects, it has been recently shown that this learning, termed contextual cueing (CC), occurs even in complex, heterogeneous environments, but only when the same distractors are repeated at the same locations. Yet it is not clear what exactly is being learned under these conditions: the visual features of the objects or their meaning. In this study, Experiment 1 demonstrated that meaning is not necessary for this type of learning, as a similar pattern of results was found even when the objects' meaning was largely removed. Experiments 2 and 3 showed that after learning meaningful objects, CC was not diminished by a manipulation that distorted the objects' meaning but preserved most of their visual properties. By contrast, CC was eliminated when the learned objects were replaced with different category exemplars that preserved the objects' meaning but altered their visual properties. Together, these data strongly suggest that the acquired context that facilitates real-world objects search relies primarily on the visual properties and the spatial locations of the objects, but not on their meaning.

Keywords Contextual cueing · Visual search · Semantics · Visual learning

People are highly efficient in extracting statistical regularities embedded in the environment. Consequently, visual search is

facilitated when the target location is repeated within the same context. That is, observers are able to learn the association between the context and the target location, and they can use it as a cue of where the target is going to appear (Chun & Jiang, 1998; Makovski & Jiang, 2010). This effect, termed contextual cueing (CC), has traditionally been shown in search tasks that examined the repetition of spatial configurations of simple, meaningless items, in the absence of any semantic context (e.g., finding a *T* among *Ls*). Naturally, there are excellent reasons to test CC using meaningless stimuli, and this line of research has been proven fruitful, yielding countless informative insights (e.g., Goujon, Didierjean, & Thorpe, 2015). Nevertheless, the world we live in is heterogeneous and filled with meaningful objects, and it is therefore imperative to test CC under more realistic settings that involve meaningful, complex objects.

When meaningful objects were tested in previous CC studies, they were typically embedded within the context a meaningful scene. That is, the role of semantics in CC was usually tested by using natural scenes, rather than arbitrary spatial arrangements, as the repeated-search context. This line of research revealed that subjects readily learn to associate the global properties of the scenes with the target locations (e.g., Brockmole, Castelano, & Henderson, 2006; Brockmole & Henderson, 2006b). These findings are in accord with the notion that the meaning of the scene (or gist) is by itself a powerful cue that guides scene processing and visual search (Bar, 2004; Vo & Wolfe, 2013). Further, it was found that predictive scenes are such powerful cues that they actually preclude learning from predictive spatial configurations (Rosenbaum & Jiang, 2013; see also Brooks, Rasmussen, & Hollingworth, 2010). Thus, while there is an agreement that scenes are easily learned, powerful, contextual cues, it is still unknown whether the meaning of the objects themselves is part of the context that is being learned in CC.

✉ Tal Makovski
tal.makovski@gmail.com

¹ Department of Education and Psychology, The Open University of Israel, Ra'anana, Israel

The issue of whether the meaning of objects is learned in CC, in the absence of a coherent scene, is further important because many researchers agree that different mechanisms might underlie scene-based CC versus CC when there is no coherent structure and learning is based on arbitrary configurations of unrelated items (henceforth, array-based CC). For instance, scene-based CC is believed to rely on explicit memory and the global properties of the display (Brockmole et al., 2006; Brockmole & Henderson, 2006a, 2006b; Brockmole & Vo, 2010), whereas array-based CC relies on implicit memory (Chun & Jiang, 2003; Colagiuri & Livesey, 2016; but see Vadillo, Konstantinidis, & Shanks, 2016) and the local elements of the display (e.g., Brady & Chun, 2007). Array-based CC is also considered to be a fundamental type of learning, as it was observed in infants (Bertels, San Anton, Gebuis, & Destrebecqz, 2016), nonhuman primates (Goujon & Fagot, 2013), and even birds (Gibson, Leber, & Mehlman, 2015). Thus, it is important to examine whether semantics (in this case, the meaning of the distractors) is a key factor not only in scene-based CC but also in array-based CC.

Most CC models have emphasized the spatial domain in learning (Brady & Chun, 2007; Jiang & Wagner, 2004; Olson & Chun, 2002), whereas the identities of the objects received only little attention. However, it was recently reported that when real-world objects are used in arbitrary displays, then learning is found only when the same distractors are repeated at the same locations (Makovski, 2016, 2017). That is, the repetition of spatial (*where*) information was insufficient to facilitate search when item identities (*what*) varied across repetitions. Similarly, no benefit was found when the *what* information was repeated, and CC was found only when both *what* and *where* information were preserved throughout the experiment (see Fig. 1 for an illustration of these conditions). That observers were able to take advantage of the repetition of distractors' identities and locations, even when a scene gist was absent, was found to be robust as it was not modulated by set size or memory load manipulations (Makovski, 2016, 2017). Furthermore, the repetition of both identities and locations did not benefit search when the two were not bound together, suggesting that CC critically depends on what and where binding (namely, the same repeated object must be at the same repeated location) rather than on the two types of information being learned independently (Makovski, 2017).

These results raise several important questions. First, is it possible that the meaning of the objects prevented the learning of “pure” configural regularities (i.e., that there was no benefit when only the spatial information was repeated)? Indeed, this finding seems inconsistent with the notion that the spatial dimension is special for CC. Thus, it is important to clarify what it is about real-world objects that impede the learning of configuration regularities: is it their meaning, or their visual complexity and heterogeneity? Second, the finding that

identities do play a part in CC (see also Chun & Jiang, 1999; Endo & Takeda, 2004) highlights the question of what constitutes the context of an array-based CC. Specifically, do people extract the visual properties of the objects (there was a green circle here and a silver rectangle there) or their meaning (there was an apple here and a phone there)?

There are good reasons to suggest that objects' meaning is not part of CC. For instance, it was recently reported that category-level information does not play a significant role in visual short-term memory tasks (Quinlan, & Cohen, 2016). Moreover, the original CC effect was, in fact, observed using meaningless stimuli (Chun & Jiang, 1998), suggesting that, at least in relatively simple, homogenous environments, meaning is not necessary for CC. Learning also occurred when search targets were embedded within visually complex yet meaningless images (Goujon, Brockmole, & Ehinger, 2012). Although it was not clear in advance whether such images fall under the category of scene-based or array-based CC, it was found that learning in this case depended on the global properties of the display (color scheme) and was associated with explicit memory, and thus might reflect a scene-based, rather than an array-based, CC.

On the other hand, there are also good reasons to suggest that even array-based CC involves meaning. Indeed, people are extremely efficient in extracting objects' meaning rapidly (e.g., Potter, 1976), and this categorical information is known to support visual long-term memory (Konkle, Brady, Alvarez, & Oliva, 2010). Furthermore, visual statistical learning, which is another form of implicit learning, does seem to involve category-level abstraction (Brady & Oliva, 2008; Otsuka, Nishiyama, & Kawaguchi, 2014; Otsuka, Nishiyama, Nakahara, & Kawaguchi, 2013). Nonetheless, in this procedure, each item is presented in isolation for relatively long duration, and it is not clear whether visual-search processes are sufficient to support such a categorical learning. Thus, the present study aims at investigating the extent to which the meaning of objects contribute to CC, particularly in arbitrary complex displays where learning is confined.

Experiment 1

The first experiment repeated Makovski's (2016, Experiment 1) procedure and logic with the exception that the meaning of the distractors was largely removed. Four display conditions were tested. The *location-repeat* condition mirrored typical array-based CC experiments in that only the locations, but not the identities, of the distractors, were repeated across blocks. In the *identity-repeat* condition, only the identities, but not the locations, of the distractors were repeated. In contrast, both the identities and the locations of the distractors were repeated in the *all-repeat* condition. These conditions

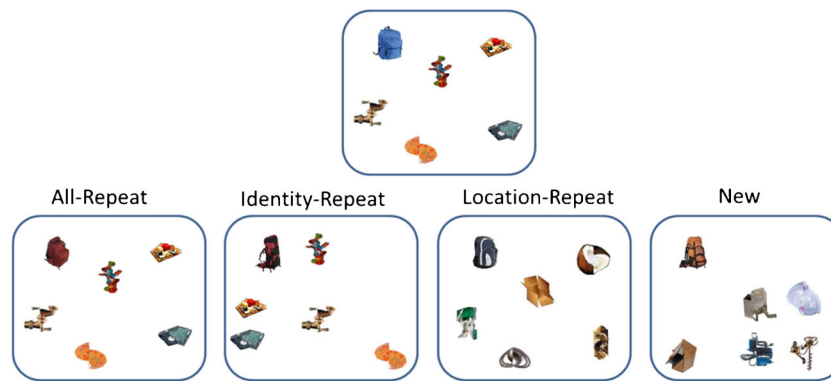


Fig. 1 Schematic illustration of the conditions tested in Experiment 1. Compared to the top display, both the identities (“what”) and the locations (“where”) of the distractors were repeated in the *all-repeat* condition. Only the “what” information was repeated in the *identity-repeat* condition, whereas only the spatial configuration (“where”) was

repeated in the *location-repeat* condition. In the *new* condition, both the identities and the locations of the distractors were randomly selected. Note that in all of the conditions, the repeated context was associated only with the location of the target, and not with a specific target image. (Color figure online)

were compared to *new* display trials, in which only the target locations were repeated, but the distractors’ locations and identities were randomly selected (see Fig. 1).

Testing these conditions with meaningless objects enables us to address two questions: First, why was there not the repetition of spatial configuration sufficient to produce CC in Makovski (2016)? Was it due to the use of meaningful distractors? One might argue, for example, that the processing of the distractors’ meaning came at the expense of the processing of the spatial information, or that it encouraged individual rather than configural processing. An alternative explanation is that the increased visual variability and display heterogeneity, regardless of meaning, diminished CC (Feldmann-Wustefeld & Schubo, 2014). Thus, if the object’s meaning underlies the lack of learning of configuration regularities, then the repetition of spatial configurations should facilitate search when the distractor’s meaning is removed. Alternatively, if the lack of learning is due to heterogeneity, and not semantics, then spatial configurations’ repetition should not elicit CC even when meaningless objects are used. The second goal of this experiment was to test whether meaning is necessary for the learning of complex, heterogeneous context or that array-based CC would still emerge when the locations of meaningless objects are repeated.

Method

Participants

All participants were students from the Open University of Israel who took part in the experiments for course credit. All reported having normal or corrected-to-normal visual acuity. Thirty-nine participants (10 males, age: $M = 25.7$ years) completed Experiment 1, and thus the experiment had power of more than 0.95 to detect small learning effects across epochs.

Equipment and stimuli

Participants were tested individually in a dimly lit room. They sat about 67 cm away from a 17-in. CRT monitor (resolution: 1024×768 , 85 HZ). The experiments were programmed using Psychophysics Toolbox (www.psychtoolbox.org), implemented in MATLAB (www.mathworks.com). Six hundred colored images ($1.89^\circ \times 1.89^\circ$) of real-world objects were taken from Brady, Konkle, Alvarez, and Oliva (2008; <http://timbrady.org/stimuli.html>). A distorted version of each image was created by flipping one half of the object (see Fig. 2a). This manipulation presumably preserved the “objecthood” of the items as well as most of their visual statistics (color, orientation, brightness), but largely removed their meaning. The latter was confirmed by several independent manipulation checks: (1) A group of 10 observers was substantially faster to verbally name the intact objects ($M = 2,270$ ms) than their distorted counterparts ($M = 3,417$ ms), $t(9) = 4.47$, $p < .002$, Cohen’s $d = 1.93$.¹ (2) When asked to rate how meaningful is each item on a scale of 0 (*no meaning*) to 5 (*meaningful*), 15 subjects rated the intact images as much more “meaningful” than the distorted images (3.73 vs. 2.26), $t(14) = 10.5$, $p < .001$, Cohen’s $d = 2.0$. Finally, 14 new observers repeated the last procedure, except that each item was presented briefly (250 ms) at one of the possible search positions, to somewhat simulate a search task in which distractors are only briefly scanned. Still, although overall meaning ratings were greatly reduced, similar pattern of results was found (intact: 2.94 vs. distorted: 1.88), $t(13) = 5.82$, $p < .001$, Cohen’s $d = 1.32$.

¹ Accuracy was not measured because errors in labeling the image do not necessarily entail lack of meaning. On the flip side, correct guessing of the item’s label might be achieved when enough time and visual cues are provided.



Fig. 2 **a** Examples of the intact (*top row*) and distorted (*bottom row*) distractor images used in the study. **b** Examples of two target categories. All seven target categories were consisted of 16 different exemplars. (Color figure online)

These manipulation checks demonstrate that, first, meaning is hard to define and there is no single, optimal way to measure it, as every image might contain some meaning for someone. Second, and more importantly, they confirm that it is considerably more difficult to extract meaning from the distorted distractor images (as can also be seen in Figs. 1, 2, and 4), and it is even more so when those are viewed for only short durations during search.

Design and procedure

The design and procedure were identical to Makovski's (2016) Experiment 1, except for the use of distorted images as distractors. Each subject was randomly assigned to one of seven target categories (guitars, backpacks, sofas, butterflies, gift-wrappers, shoes, horses) and to a random set of 350 distractors. On each trial, a target was randomly selected from 16 possible exemplars (see Fig. 2b) and thus there was never a consistent association between the target image and the repeated context. The (intact) target together with 14 distractor distorted objects were presented against a white background on an invisible 8×6 grid ($21.6^\circ \times 16.2^\circ$, with a random jitter of up to

0.54°). Subjects were instructed to press the space bar as fast as they could when they found the target. Afterwards, the items disappeared, and the digits 1–6 appeared at the positions of the target and five random distractors. Subjects were asked to insert the digit occupying the target's position. A green plus sign (+) was presented for 500 ms after correct responses, whereas a red minus sign (–) was displayed for 2,000 ms after errors.

Participants performed 20 blocks; each consisted of 32 trials (eight displays of the four experimental conditions) presented in a random order. In the *location-repeat* displays, only the distractor locations, but not the images, were repeated across blocks. Conversely, on *identity-repeat* trials only the distractor images, but not their locations, were repeated together with the target locations. On *all-repeat* trials, both the distractor images and their locations were repeated together with the target locations. All of the repeated displays were generated randomly for each participant and were compared to *new* trials, wherein target locations were repeated, yet the distractor images and locations were randomly selected on each block (see Fig. 1).

A surprise familiarity test was administered at the end of all experiments, and these data are reported and briefly discussed in the Appendix.

Results

Accuracy was high (>97.6%), and none of the repeated conditions significantly differed from *new* ($ps > .19$). Error trials as well as outliers—trials deviating 2.5 *SD* above and below each participant's mean of each cell (2.86% of the correct trials), were removed from the response time (RT) analyses (see Fig. 3).

Planned repeated-measures ANOVAs, with epoch (a bin of four consecutive blocks) and display condition (*new* vs. *repeated*) were conducted to assess learning separately for each condition. All of these analyses revealed robust effects of epoch, whereby RT became faster as the experiment progressed, $F(4, 152) > 14.9$, $ps < .001$, $\eta_p^2 > .28$. More importantly, neither the *location-repeat* nor the *identity-repeat* displays were different than *new* displays, $F(1, 38) < 1$, $\eta_p^2 = .01$; $F(1, 38) < 1$, $\eta_p^2 = .015$, respectively. There was also no significant Epoch \times Condition interaction for the *identity-repeat* condition, $F(4, 152) < 1$, $\eta_p^2 = .025$. However, such an interaction was found for the *location-repeat* condition, $F(4, 152) = 2.87$, $p = .03$, $\eta_p^2 = .07$. This interaction was driven by the first epoch, where *location-repeat* trials were exceptionally slow. Notably, *location-repeat* trials were not reliably faster than *new* trials in any of the other epochs, and thus it seems unlikely that this interaction reflects learning.

By contrast, a clear benefit was found for the *all-repeat* displays, which yielded faster responses than *new* displays did, $F(1, 38) = 4.57$, $p = .04$, $\eta_p^2 = .11$, and this facilitation increased as the experiment progressed, $F(4, 152) = 3.29$, $p = .01$, $\eta_p^2 = .08$. This interaction was accompanied by a significant linear trend, $F(1, 38) = 6.07$, $p = .02$, $\eta_p^2 = .14$, suggesting the difference between the conditions increased with epoch (Epoch 1: -30 ms, $p = .24$; Epoch 2: 33 ms, $p = .18$; Epoch 3: 59 ms, $p = .01$; Epoch 4: 44 ms, $p = .08$; Epoch 5: 60 ms, $p < .01$).

Interestingly, a direct comparison of this benefit with the *all-repeat* advantage observed in Makovski's (2016) Experiment 1 (where images were intact), revealed that

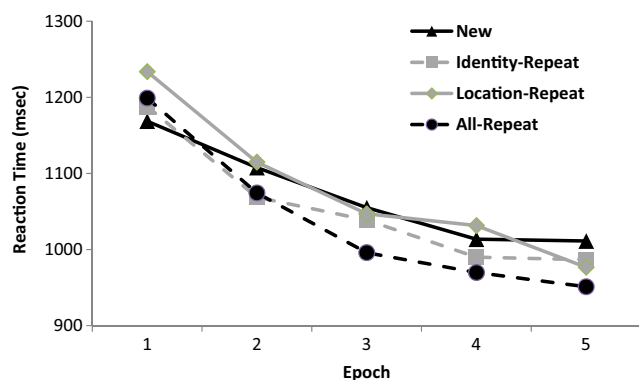


Fig. 3 Experiment 1's results: mean RT as a function of epoch and display condition

although search latencies were slower within distorted images (*all-repeat* = 1,038 ms; *new* = 1,071 ms) than intact images (*all-repeat* = 868 ms; *new* = 922 ms), $F(1, 67) = 14.07$, $p < .001$, $\eta_p^2 = .17$, there was no interaction between CC and experiment, $F(1, 67) < 1$, $p = .38$, $\eta_p^2 = .01$. Note, however, that these results should be interpreted with caution, and a dedicated experiment is needed in order to closely examine the differences between searching through meaningless and meaningful objects.

Discussion

Experiment 1 showed that relative to the *new* condition, there was no benefit for either the *identity-repeat* or the *location-repeat* conditions, and only the *all-repeat* condition facilitated search. These findings differ from previous studies showing that CC tolerates some identity and spatial variability (Chun & Jiang, 1999; Endo & Takeda, 2004; van Asselen, Sampaio, Pina, & Castelo-Branco, 2011). In contrast, they are in full agreement with a recent study that tested CC with real-world objects and found a similar pattern of results (Makovski, 2016). That study further ruled out several possible explanations for the apparent inconsistency regarding CC tolerance to variability. Specifically, the lack of learning in the *identity-repeat* and the *location-repeat* condition could not be explained by insufficient statistical power, or by overshadowing by the *all-repeat* condition, because no learning was found even when these conditions were tested separately. Other methodological differences were further rejected, as no learning was found even when the display was less crowded (set size was reduced to 12), targets were defined by a single exemplar, and more displays (12) and more repetitions (28) were used. The results of Experiment 1 (as well as the results of the *identity-repeat* condition in the next experiments) further corroborate the conclusion that at least for heterogeneous, complex environments, both what and where repetitions are required for CC, and that the repetition of only one type of information is insufficient to facilitate search.

In addition, the present findings imply that the lack of learning of configuration regularities reported in Makovski (2016) is not the result of using meaningful distractors, because no such learning was observed here as well, when meaningless objects were used. Instead, these results support the notion that identity variability interferes with the learning of spatial configurations. Finally, and more importantly for the current purposes, the results of Experiment 1 clearly indicate that meaning is not critical for learning even in heterogeneous, complex displays (Goujon et al., 2012), and array-based CC can be found as long as the same distractors, meaningful or not, are repeated at the same locations.

Experiments 2 and 3

The first experiment showed that CC occurs even when there is no coherent scene and the context consists of arbitrary distorted images of real-world objects. Nonetheless, the finding that CC does involve distractor identities (as well as locations) highlights the question of what observers actually learn in an array-based CC. Namely, when subjects learn to associate the context of the display with the target location, do they extract the meaning of the objects or is learning strictly visual?

To address this question, subjects performed CC tasks using intact real-world objects. To assess what subjects learned during the training phase, all of the distractor items were replaced during the transfer phase of the experiments. In Experiment 2, each distractor item was replaced with its distorted, meaningless version, whereas in Experiment 3 each distractor item was replaced with a different exemplar from the same category (e.g., a different picture of a ladder; see Fig. 4). The latter manipulation of using different exemplars alters the basic visual features of the objects while keeping the meaning intact. Conversely, in Experiment 2, the meaning is distorted, but most of the basic visual properties (e.g., color, orientation, brightness) are largely preserved. The logic is straightforward: If CC mainly relies on the visual properties, then learning should transfer only when these are preserved (Experiment 2). Similarly, if CC involves category-level, abstract representations, then a transfer of learning should be found when the meaning is preserved (Experiment 3).

Experiment 2

Method

The training phase was identical to Experiment 1, except for the following changes. First, the distractor stimuli were the original, intact images of the objects. Second, the *location-repeat* displays were excluded and each block consisted of ten displays of the remaining three conditions. After completing the 20 blocks of training, subjects immediately started the transfer phase that was composed of four blocks in which all distractor items were replaced with their distorted counterparts. Thirty-one subjects (eight males, age: $M = 25.9$ years) participated in Experiment 2.²

² Relative to Experiment 1, fewer subjects were tested in Experiments 2 and 3 because only three conditions were used in the training phase. Still, more subjects were tested here than in other experiments that included a transfer phase (e.g., Brockmole & Vo, 2010; Goujon et al., 2012; Jiang & Wagner, 2004; Makovski & Jiang, 2010, 2011). A power analysis showed that a sample size of 30 subjects has power of more than 0.88 to detect medium size effects (0.3–0.5) in the transfer phase.

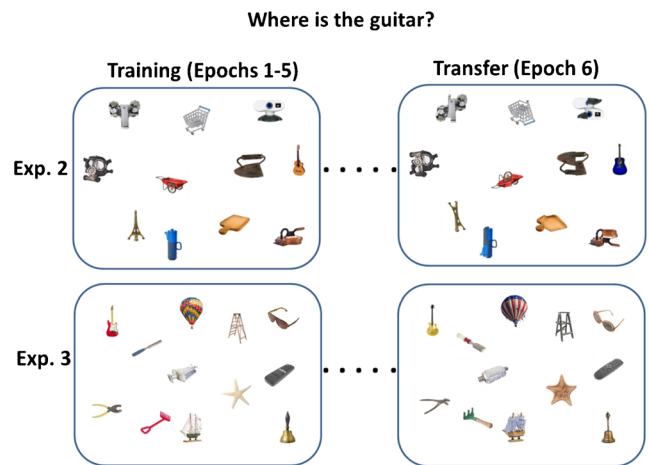


Fig. 4 Schematic illustrations of all-repeat displays in Experiment 2 (transfer to distorted images) and Experiment 3 (transfer to different category exemplars). (Color figure online)

Results and discussion

Error trials as well as outliers (2.76% of the correct trials) were removed from the RT analyses (see Fig. 5).

Training phase

Accuracy was high (>98.2%), and neither the All-Repeat nor the *identity-repeat* conditions differed from *new*, $F(1, 30) < 1$, $p = .36$, $\eta_p^2 = .03$; $F(1, 30) < 1$, $p = .74$, $\eta_p^2 = .004$, respectively. Replicating previous results (Makovski, 2016), there was no difference in RT between the *identity-repeat* and *new* displays, $F(1, 30) < 1$, $p = .78$, $\eta_p^2 = .003$, and no interaction with epoch, $F(4, 120) < 1$, $p = .49$, $\eta_p^2 = .03$. In contrast, *all-repeat* trials were faster than *new* trials, $F(1, 30) = 4.5$, $p = .04$, $\eta_p^2 = .13$, and this effect did not interact with epoch, $F(4, 120) < 1$, $p = .65$, $\eta_p^2 = .02$.

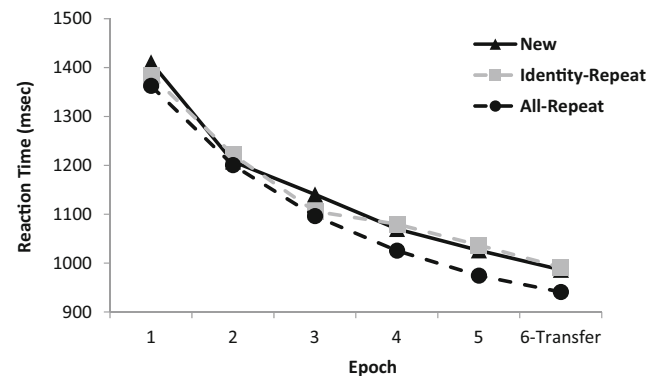


Fig. 5 Experiment 2’s results: mean RT as a function of epoch and display condition. Epochs 1–5 = training phase; Epoch 6 = transfer to distorted images

Transfer phase

Accuracy was high again (>98.5%), with no significant difference between the conditions ($ps > .44$). Importantly, while *identity-repeat* trials did not differ from *new* trials, ($t < 1$), responses were faster in the *all-repeat* condition (941 ms) than in the *new* condition (986 ms), $t(30) = 2.78, p = .009, d = 0.24$. Moreover, there was no interaction between block (21–24) and display, $F(3, 90) < 1, p = .82, \eta_p^2 = .01$, confirming that this facilitation was not the result of new learning acquired during the transfer phase. Finally, an analysis comparing the last epoch of training with the transfer epoch revealed a main effect of display, $F(1, 30) = 7.39, p = .01, \eta_p^2 = .20$, that was not modulated by phase, $F(1, 30) < 1, p = .74, \eta_p^2 = .004$. Thus, it seems safe to conclude that CC was hardly affected by the distortion of the meaning of the objects, and the benefit of learning was transferred in full.³

Experiment 3

Method

Experiment 3 followed the same logic and design of Experiment 2, only now during the transfer phase each distractor image was replaced with a picture of a different exemplar of the object's category. To that end, the distractor items were sampled from a smaller set of 93 pairs of objects taken from <http://timbrady.org/stimuli.html>. Thirty subjects (seven males, age: $M = 25.1$ years) participated in Experiment 3.

Results and discussion

Error trials and outliers (2.7% of the correct trials) were removed from the RT analyses (see Fig. 6).

Training phase

Accuracy was higher in the *all-repeat* condition (98.9%) than in the *new* condition (98.6%, $p = .036$), which did not differ from the *identity-repeat* condition (98.7%, $p = .51$). As before, there was no difference in RT between *identity-repeat* and *new* displays, $F(1, 29) < 1, p = .51, \eta_p^2 = .015$, and no interaction between epoch and display, $F(4, 116) < 1, p = .56, \eta_p^2 = .025$.

Although a reliable learning effect was observed in the *all-repeat* accuracy data, it was less pronounced in RT, perhaps

³ The same pattern of results (full transfer of learning) was obtained in an additional experiment that was identical to Experiment 2, except that during the transfer phase the distractor items were flipped upside down. This manipulation also impaired subjects' ability to verbally name the items, but to a lesser extent, and thus it is not clear how effective it was in removing the meaning of the objects.

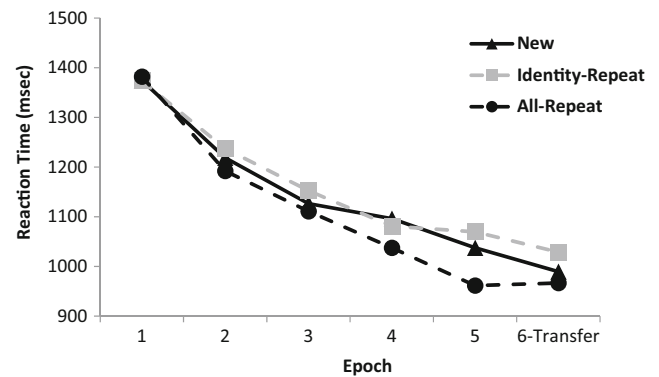


Fig. 6 Experiment 3's results: mean RT as a function of epoch and display condition. Epochs 1–5 = training phase; Epoch 6 = transfer to new exemplars

because of the limited set of objects used in this experiment. *All-repeat* trials were not overall faster than *new* trials, $F(1, 29) = 2.13, p = .16, \eta_p^2 = .07$, yet there was a borderline interaction between display and epoch, $F(4, 116) = 2.14, p = .08, \eta_p^2 = .07$. Importantly, there was a significant linear trend in the interaction, $F(1, 29) = 7.5, p = .01, \eta_p^2 = .21$, indicating that the difference between the conditions increased as the experiment progressed. This was reflected by the lack of a significant difference between the conditions in the first three epochs, a marginal effect in the fourth ($p = .058, \eta_p^2 = .23$), and by the fifth epoch responses were markedly faster in the *all-repeat* condition (962ms) than in the *new* condition (1,038 ms), $F(1, 29) = 8.76, p = .006, \eta_p^2 = .23$.

Transfer phase

Accuracy was above 98.2%, with no significant difference between the conditions ($ps > .26$). RT did not differ between *identity-repeat* trials and *new* trials, $F(1, 29) = 2.83, p = .10, \eta_p^2 = .089$. Of greater interest and in contrast to Experiment 2, responses were not significantly faster in the *all-repeat* condition (967 ms) than in the *new* condition (989 ms), $F(1, 29) = 0.99, p = .33, \eta_p^2 = .033$, and no difference was found between these conditions in any of the four transfer blocks (all $ps > .16$). Moreover, a direct comparison between the last training epoch and the transfer epoch revealed a significant interaction between display and phase, $F(1, 29) = 6.61, p = .016, \eta_p^2 = .19$. It is worth noting that performance in the *new* and *identity-repeat* conditions improved in the transfer epoch relative to the final training epoch, $F(1, 29) = 4.12, p = .05, \eta_p^2 = .12$. However, this improvement was not found in the *all-repeat* condition, $F(1, 29) < 1, p = .73, \eta_p^2 = .004$, probably because the general improvement due to practice was counteracted by the disappearance of the learning advantage.

Finally, to directly compare between the results of Experiments 2 and 3 and to overcome baseline differences, a

percentage of measurement of learning was created, and benefit scores were calculated for Epochs 5 and 6 using the following formula (Makovski, 2017):

$$\text{Benefit score} = \frac{(\text{rt}[\text{New}] - \text{rt}[\text{Repeat}]) \times 100}{\text{rt}[\text{New}]}$$

For Experiment 2, this score was significantly different from zero (indicating a reliable learning advantage) in both Epochs 5 (final training epoch, 4.3%, $p = .05$) and 6 (transfer epoch, 4.5%, $p = .005$). Importantly, while there was a significant learning effect in Experiment 3's Epoch 5 (6.2%, $p = .01$), it completely disappeared in Epoch 6 (1.4%, $p = .54$). A direct comparison between the two experiments showed a marginal significant interaction between experiment (2 vs. 3) and epoch (5 vs. 6), $F(1, 59) = 3.19$, $p = .079$, $\eta_p^2 = .05$.

Taken together, these results indicate that in contrast to the meaning manipulation that had little effect on learning, CC did not survive the transfer to new category exemplars, even though the items in the transfer phase of Experiment 3 were visually similar to the items used in the training (e.g., in their general shape). In effect, CC was eliminated when the distractors preserved their meaning and their spatial locations but not other basic visual properties (e.g., color, brightness). Moreover, that some of the exemplars were visually similar to one another only strengthens the conclusion that intact meaning is insufficient for learning to transfer. This is in contrast to the meaning removal manipulation of Experiment 2, which kept most of the summary statistics of the low visual features, and had no effect on learning.

General discussion

People extract the meaning of objects rapidly and efficiently (e.g., Potter, 1976), but is this information part of the context that facilitates search? Indeed, it has been recently shown that distractor identities (and locations) are part of CC (Makovski, 2016, 2017); however, it was unknown whether these identities include the visual properties of the objects, their meaning, or both. The present findings strongly propose that what people actually learn in array-based CC is to associate the position of the target with the visual properties, and not the meaning, of the distractors.

Experiment 1 demonstrated that meaning is not necessary for learning even in complex, heterogeneous search displays (Goujon et al., 2012), as an array-based CC effect was found even with distorted, meaningless distractors. This experiment further revealed that the lack of configural learning (no learning in the *location-repeat* condition) when object identities vary, cannot be explained by the use of meaningful real-world objects, because configural learning was also absent when the meaning of the distractors was largely removed.

This conclusion is in line with the finding of no configural learning, even when identity processing was discouraged, and subjects looked for a single target exemplar rather than for category-defined targets (Makovski, 2016). Taken together, these data confirm that both *what* and *where* repetitions are necessary for array-based CC and thus challenge the notion that the spatial domain is special for CC, and visual cognition in general (e.g., Treisman, 1988; Tsal & Lavie, 1993).

Experiments 2 and 3 directly tested whether category-level information is acquired during repetitions. The results showed that CC was resistant to a manipulation that distorted the meaning of the objects but preserved their visual properties. In sharp contrast, CC was eliminated when the objects kept their category-level meaning but the visual properties were altered. That people can learn to associate the position of the target with the visual properties of the distractors is consistent with the finding that color scheme changes diminished learning effects (Goujon et al., 2012). It is also in accord with the notion that items scanned briefly during search leave visual memory traces, regardless of intentions to remember (Castelhano & Henderson, 2005; Williams, Henderson, & Zacks, 2005). Of greater interest, the current results entail that the context that is being used to facilitate search, when scene meaning is not available, relies primarily on the visual properties and the spatial locations of the items, whereas the meaning of the objects play little role in this type of learning. However, further examination is still needed in order to isolate the critical visual features that are specifically important for array-based CC. For instance, it is possible that the shape of the objects is less important for this type of learning than, for example, color information, as CC completely transferred in Experiment 2, even though it involved some distortion of the objects' shapes (while other basic visual properties, such as color, were less affected by this manipulation).

It is worth noting that these findings were obtained in spite of the fact that the search targets were defined categorically. Presumably, this manipulation should have encouraged subjects to rely more on semantic processing than on visual processing because they could not search for specific visual features. That is, forming a target template was more difficult here than in typical search tasks, where targets are defined by a single exemplar, because on each trial in the present experiments the target could appear in a different brightness, shape or color (see Fig. 2b). Nevertheless, even under these conditions, where search cannot be guided by specific target features, subjects acquired the visual properties of the items and not their abstract meaning.

From a broader perspective, it is important to emphasize that the conclusion that objects' meaning is not part of array-based CC does not entail that semantics plays no role in CC in general. In fact, there is strong evidence that scene meaning is a key factor in CC (Brockmole et al., 2006; Brockmole & Henderson, 2006a, 2006b; Rosenbaum & Jiang, 2013).

Furthermore, associations between scene meaning and target positions can be learned in spite of large variability in the visual properties of the display. For instance, learning that the search target was on a pillow occurred despite the use of multiple bedroom images, and this learning even transferred to a semantically related context (pillows presented in living rooms; Brockmole & Vo, 2010). This finding seems inconsistent with the present findings that show that learning was more sensitive to changes to the visual properties of the objects than to changes in their meaning. Nevertheless, several noticeable differences between the studies that show involvement of semantics in scene-based CC, and the present one that shows no involvement of objects' meaning in array-based CC. First, it is likely easier to extract a single scene, or a target, meaning than the meaning of multiple distractors. Second and as discussed above, different mechanisms might underlie a scene-based CC and an array-based CC (e.g., Rosenbaum & Jiang, 2013). In line with this distinction, the set of studies that tested CC with real-world objects, but without a coherent scene, shows that learning under these conditions is more specific than scene-based CC. That is, an array-based CC relies on the specific visual properties of the objects and on the binding of these properties to specific locations (Makovski, 2017) and, unlike scene-based CC, does not tolerate large variability in the distractors locations or visual features (Makovski, 2016).

That CC with real-world objects does not involve semantics is somewhat unexpected given that objects' categories are easily extracted and are known to affect visual search and attentional control (e.g., Nako, Wu, Smith, & Eimer, 2014). Moreover, semantic flexibility was found in visual statistical learning tasks (Brady & Oliva, 2008; Otsuka, et al., 2013; 2014), whereas here there was no evidence for category-level generalization. A possible explanation for this difference is that category-level learning is overshadowed in CC by the dominant learning of the visual features. This is consistent with recent evidence suggesting that category-level learning in visual statistical learning occurs mainly when the regularities of the simpler features are absent (Emberson & Rubinstein, 2016). Nevertheless, there are other important methodological differences between the two procedures, particularly concerning the presentation mode of the objects (e.g., short vs. long exposures, central vs. peripheral vision) that might be related to the difficulty to extract their meaning, which can in turn explain this discrepancy. Additional investigation is therefore needed to elucidate whether and when distractors categorical information is acquired during CC.

In sum, people are able to utilize the repetition of both *what* and *where* information to facilitate search, even when those are embedded in complex heterogeneous, arbitrary displays. The present study further revealed that the context that

facilitates search does not involve the distractors meaning and relies instead on precategorical representations of visual and spatial information.

Acknowledgements I thank Ayelet Golestani, Tomer Sahar, and Meital Nechmadi for help in data collection, and Ido Liviatan and James Brockmole for helpful comments and suggestions. Correspondence address: Tal Makovski, Department of Psychology, The Open University of Israel, The Dorothy de Rothschild Campus, 1 University Road, P. O. Box 808, Ra'anana. 43107. E-mail: talmak@openu.ac.il; Phone: +972-97781714

Appendix

Table 1 Mean familiarity scores (and *SEs*) as a function of display condition and experiment

	New	All-repeat	Identity-repeat	Location-repeat
Exp. 1	2.12 (0.13)	2.46 (0.14) <i>p</i> < .001	2.33 (0.15) <i>p</i> = .02	2.29 (0.14) <i>p</i> = .08
Exp. 2	2.49 (0.17)	2.72 (0.19) <i>p</i> = .05	2.48 (0.17) <i>p</i> = .96	
Exp. 3	3.02 (0.19)	3.24 (0.18) <i>p</i> = .09	3.14 (0.19) <i>p</i> = .15	

At the end of each experiment, subjects were asked to rate the familiarity of all the repeated and new displays presented during training on a scale of 1 (*low familiarity*) to 5 (*high familiarity*). Appendix Table 1 shows mean familiarity scores as a function of display condition and experiment. Standard error of the means are presented in parentheses. The *p* values indicate the results of the *t* test comparisons of the repeated displays with the *new* displays.

Importantly, the CC effect does not seem to follow familiarity as subjects were able to distinguish between repeated and new displays even when no CC was found, and conversely, CC was found even without familiarity (for similar findings, see Makovski, 2016, 2017). Furthermore, collapsing across all three experiments, there was no positive correlation between explicit familiarity (*all-repeat* minus *new*) and CC on Epoch 5, $r(100) = -0.15$, $p = .14$, and subjects who rated the *all-repeat* displays as more familiar than *new* displays did not show larger CC effect ($N = 50$, $M = 3.7\%$) than subjects who did not ($N = 50$, $M = 6.1\%$), $t(98) < 1$.

References

- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, 5, 617–629.

- Bertels, J., San Anton, E., Gebuis, T., & Destrebecqz, A. (2016). Learning the association between a context and a target location in infancy. *Developmental Science*, 20(4). doi:10.1111/desc.12397
- Brady, T. F., & Chun, M. M. (2007). Spatial constraints on learning in visual search: Modeling contextual cuing. *Journal of Experimental Psychology: Human Perception & Performance*, 33, 798–815.
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, USA*, 105(38), 14325–14329.
- Brady, T. F., & Oliva, A. (2008). Statistical learning using real-world scenes: Extracting categorical regularities without conscious intent. *Psychological Science*, 19, 678–685.
- Brockmole, J. R., Castelhan, M. S., & Henderson, J. M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, & 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(1), 99–108.
- Brockmole, J. R., & Vo, M. L. H. (2010). Semantic memory for contextual regularities within and across scene categories: Evidence from eye movements. *Attention, Perception & Psychophysics*, 72(7), 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(6), 1406–1418.
- Castelhan, M. S., & Henderson, J. M. (2005). Incidental memory for objects in scenes. *Visual Cognition*, 12, 1017–1040.
- 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. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science*, 10, 360–365.
- Chun, M. M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 224–234.
- Colagiuri, B., & Livesey, E. J. (2016). Contextual cuing as a form of nonconscious learning: Theoretical and empirical analysis in large and very large samples. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-016-1063-0. **Advance online publication.**
- Emberson, L. L., & Rubinsten, D. Y. (2016). Statistical learning is constrained to less abstract patterns in complex sensory input (but not the least). *Cognition*, 153, 63–78.
- Endo, N., & Takeda, Y. (2004). Selective learning of spatial configuration and object identity in visual search. *Perception & Psychophysics*, 66, 293–302.
- Feldmann-Wustefeld, T., & Schubo, A. (2014). Stimulus homogeneity enhances implicit learning: Evidence from contextual cueing. *Vision Research*, 97, 108–116.
- Gibson, B. M., Leber, A. B., & Mehlman, M. L. (2015). Spatial context learning in pigeons (*Columba livia*). *Journal of Experimental Psychology: Animal Learning and Cognition*, 41(4), 336–342.
- 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(5), 1315–1327.
- Goujon, A., Didierjean, A., & Thorpe, S. (2015). Investigating implicit statistical learning mechanisms through contextual cueing. *Trends in Cognitive Sciences*, 19, 524–533.
- Goujon, A., & Fagot, J. (2013). Learning of spatial statistics in nonhuman primates: Contextual cueing in baboons (*Papio papio*). *Behavioural Brain Research*, 247, 101–109.
- Jiang, Y., & Wagner, L. C. (2004). What is learned in spatial contextual cueing: Configuration or individual locations? *Perception & Psychophysics*, 66(3), 454–463.
- Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General*, 139, 558–578.
- Makovski, T. (2016). What is the context of contextual cueing? *Psychonomic Bulletin & Review*, 23, 1982–1988.
- Makovski, T. (2017). Learning “what” and “where” in visual search [Special issue]. *Japanese Psychological Research*, 59, 133–143.
- 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(2), 216–225.
- Makovski, T., & Jiang, Y. V. (2011). Investigating the role of response in spatial context learning. *Quarterly Journal of Experimental Psychology*, 64(8), 1563–1579.
- Nako, R., Wu, R., Smith, T. J., & Eimer, M. (2014). Item and category-based attentional control during search for real-world objects: Can you find the pants among the pans? *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1283–1288.
- Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition*, 9, 273–302.
- Otsuka, S., Nishiyama, M., & Kawaguchi, J. (2014). Constraint on the semantic flexibility in visual statistical learning. *Visual Cognition*, 22(7), 865–880.
- Otsuka, S., Nishiyama, M., Nakahara, F., & Kawaguchi, J. (2013). Visual statistical learning based on the perceptual and semantic information of objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 196–207.
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 2(5), 509–522.
- Quinlan, P. T., & Cohen, D. J. (2016). The precategorical nature of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(11), 1694–1712.
- Rosenbaum, G. M., & Jiang, Y. V. (2013). Interaction between scene-based and array-based contextual cueing. *Attention, Perception & Psychophysics*, 75, 888–899.
- Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 40A, 201–237.
- Tsal, Y., & Lavie, N. (1993). Location dominance in attending to color and shape. *Journal of Experimental Psychology: Human Perception & Performance*, 19, 131–139.
- Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review*, 12, 87–102.
- van Asselen, M., Sampaio, J., Pina, A., & Castelo-Branco, M. (2011). Object based implicit contextual learning: A study of eye movements. *Attention, Perception & Psychophysics*, 73, 297–302.
- Vo, M. L.-H., & Wolfe, J. M. (2013). The interplay of episodic and semantic memory in guiding repeated search in scenes. *Cognition*, 126, 198–212.
- Williams, C. C., Henderson, J. M., & Zacks, R. T. (2005). Incidental visual memory for targets and distractors in visual search. *Perception and Psychophysics*, 67, 816–827.