Awareness in contextual cueing of visual search as measured with concurrent access- and phenomenal-consciousness tasks

Bernhard Schlagbauer
Department Psychologie, Ludwig-Maximilians-Universität München, München, Germany

Hermann J. Müller
Department Psychologie, Ludwig-Maximilians-Universität München, München, Germany
School of Psychology, Birkbeck College, University of London, UK

Michael Zehetleitner
Department Psychologie, Ludwig-Maximilians-Universität München, München, Germany

Thomas Geyer
Department Psychologie, Ludwig-Maximilians-Universität München, München, Germany

In visual search, context information can serve as a cue to guide attention to the target location. When observers repeatedly encounter displays with identical target-distractor arrangements, reaction times (RTs) are faster for repeated relative to nonrepeated displays, the latter containing novel configurations. This effect has been termed “contextual cueing.” The present study asked whether information about the target location in repeated displays is “explicit” (or “conscious”) in nature. To examine this issue, observers performed a test session (after an initial training phase in which RTs to repeated and nonrepeated displays were measured) in which the search stimuli were presented briefly and terminated by visual masks; following this, observers had to make a target localization response (with accuracy as the dependent measure) and indicate their visual experience and confidence associated with the localization response. The data were examined at the level of individual displays, i.e., in terms of whether or not a repeated display actually produced contextual cueing. The results were that (a) contextual cueing was driven by only a very small number of about four actually learned configurations; (b) localization accuracy was increased for learned relative to nonrepeated displays; and (c) both consciousness measures were enhanced for learned compared to nonrepeated displays. It is concluded that contextual cueing is driven by only a few repeated displays and the ability to locate the target in these displays is associated with increased visual experience.

Keywords: attention, visual search, implicit memory, awareness, contextual cueing


Introduction

Attention and learning interact in many ways: selective attention is an important prerequisite for both successful explicit (e.g., Voss, Baym, & Paller, 2008) and implicit learning (e.g., Nissen & Bullemer, 1987); conversely, memory for past events can guide our expectations, helping us to determine what we perceive in a scene (e.g., Biederman, Mezzanotte, & Rabinowitz, 1982) or where we look (Chun, 2000). The current study is concerned with the latter question, in particular: (a) whether guidance of focal attention in visual search is aided by perceptual memory, that is, contextual cueing (Chun & Jiang, 1998), and (b) how the memory underlying the cueing effect has to be characterized, as implicit or explicit in nature.

In visual search, human performance relies on several different cues that can aid attention. One of
These cues are contextual information, that is, the consistent arrangement of the target relative to the distractors, which is learned throughout the course of an experiment and can facilitate search performance (e.g., Chun & Jiang, 1998, 1999; Chun, 2000). In a standard contextual cueing search task, observers look for the target letter “T” among distractor letters “L” and report its orientation (left- vs. right-oriented). In the “repeated” condition, the target-distractor arrangement is repeated across trials; in the “nonrepeated” (baseline) condition, only target, but not distractor, locations are repeated across trials so as to equate target location repetition effects between the two conditions. Typically, in the repeated condition, there is a set of 12 different displays with a certain target-distractor configuration that is presented repeatedly throughout the course of the experiment. The standard finding is that RTs are faster for repeated than for nonrepeated displays—an effect that emerges after about 100 to 150 experimental trials (i.e., after four to six repetitions of each of the repeated displays; Chun & Jiang, 1998). The conclusion from these findings is that memory for learned target-distractor arrangements facilitates “early” processes of target selection (the argument put forward by, e.g., Chun & Jiang, 1998; Johnson, Woodman, Braun, & Luck, 2007; or Geyer, Zehetleitner, & Müller, 2010)—though the context effect might also aid “late” processes of response selection (an argument put forward by, e.g., Kunar, Flusberg, Horowitz, & Wolfe, 2007).

Usually, contextual cueing is regarded as an implicit effect (e.g., Chun & Jiang, 1998). This is tested by recognition tests in which observers are presented with both repeated and nonrepeated displays and are asked to indicate their feeling of familiarity with a given display (e.g., whether they believe, having seen a given display already during the earlier performance of the search task, the method employed by, e.g., Chun & Jiang, 1998) or to indicate the location of the target (the method employed by, e.g., Jiang & Chun, 2003; note that in to-be-judged displays, the target is replaced by a distractor item). The common finding is that participants’ recognition performance is only at chance level, which is taken as evidence that observers have no explicit knowledge about repeated displays (e.g., Chun & Jiang, 1998, 1999). However, there are at least two problems associated with these types of recognition tests. First, results from these tests are ambiguous as to whether observers’ conscious reports about display repetitions reflect knowledge of a particular distractor context or the precise target location within this context. While this limitation applies mostly to tests involving familiarity measures (e.g., Chun & Jiang, 1998), more recent assessments (e.g., Chun & Jiang, 2003) have attempted to be more precise by (only) examining explicit memory of the target location. Second, almost all recognition tests—whether using familiarity or target location measures—have only low statistical power: observers usually perform only very few recognition trials (e.g., 24 trials with 12 repeated and 12 nonrepeated displays, as in Chun & Jiang, 2003), which compares with hundreds of search trials (e.g., 576 trials, including 288 repeated and 288 nonrepeated displays, in Chun & Jiang, 2003). Given this small number of memory trials, the effect of explicit memory on these measurements would have to be massive to be detectable statistically. This also implies that the results (and conclusions) from recognition tests could change dramatically when the power of these tests is increased. For example, Smyth and Shanks (2008; Experiment 2) used 48 instead of 24 recognition trials and found that observers were able to consciously access the memory underlying contextual cueing: target location judgments were more accurate for repeated than for nonrepeated displays. It is, however, questionable whether the recognition task of Smyth and Shanks (2008) really assessed explicit knowledge. The argument here is that inferences based on recognition data require a clear definition of the term awareness or consciousness.

Two concepts of consciousness

When arguing about the implicit or explicit nature of contextual cueing, the term awareness (or consciousness) is often used without a precise conceptual definition. It is important to consider that “being conscious of something” can have different meanings. Following the reasoning of Block (2002, 2005), there are at least two distinguishable concepts of consciousness: phenomenal and access consciousness. Phenomenal consciousness refers to the phenomenal experience of a stimulus, or what it is like to be in a particular state, and thus can be referred to as what is usually meant when talking about having a visual experience or being aware of something. It is, however, easily confused with, but not identical to, the second concept of consciousness. Access-conscious contents of experience are not necessarily phenomenally conscious, but available to, in Block’s (2002, 2005) terms, the cognitive global workspace; that is, access-conscious contents are available to various cognitive systems, such as memory and learning, attention, or executive processes (e.g., decision making). It is important to note that phenomenal consciousness is not a necessary precondition for access consciousness. Perhaps, the clearest example of a distinction between the two terms of “conscious” is blindsight: patients with a (unilateral) lesion in striate cortex may report that they are unaware of visual stimuli in the affected hemifield (i.e., they do not have phenomenal consciousness);
however, they may be able to discriminate visual stimuli presented in the critical hemifield above chance level, while also reporting a feeling (or confidence) that something was there (i.e., they do have access consciousness; see e.g., Sahraie et al., 1997). Applied to Smyth and Shanks (2008), it is quite possible that participants showed above-chance recognition performance because the learned target-distractor arrangements accessed the cognitive global workspace, rather than because observers had phenomenal experience of the repeated search displays. In other words, observers may not have been aware of contextual memory representations, which would “reduce” contextual cueing to an implicit effect.

For determining the extent to which observers are aware of a particular stimulus, there are various measures or criteria that can be applied. For example, Sandberg, Timmermans, Overgaard, and Cleeremans (2010) suggested three measures of subjective awareness: the perceptual awareness scale (PAS), confidence ratings (CR), and post-decision wagering (PDW). The PAS is a rating of the clarity of a visual experience and can thus be understood as an index of phenomenal consciousness. CRs measure the confidence of observers in their identification response, and PDW measures refer to how much money observers would bet on their identification decision. The last two measures can be taken as indicators of access consciousness, because the information used for performing the visual discrimination does not need to be phenomenally conscious, but must be available in the global workspace for perceptual decision making. The validity of these measures has recently been re-evaluated by Zehetleitner and Rausch (2011) who used a novel approach, in a visual masking paradigm, of asking observers to rate the visual clarity (PAS) of the stimulus, rather than visual experience in general, on each trial; in addition, the CR and PDW measures were recorded on each trial. They found the PAS on the one hand and the CR and PDW on the other to form two different classes of consciousness measures, with one relating to the stimulus and the other to the response decision. On this basis, they proposed to use two measures (e.g., PAS and CR) concurrently on each trial in order to capture differential aspects of consciousness.

Rationale of the present study

For resolving the issue whether contextual cueing is implicit or explicit in nature,—more precisely, whether observers are aware of the target location in repeated distractor contexts, it is important to distinguish between (implicit) learning of repeated displays and (implicit) access to learned displays. The common view is that access is implicit if performance measures (RTs) reveal a contextual cueing effect, while subjective (recognition) scores do not differ from chance level (Chun & Jiang, 2003). And even if test power is raised by increasing the number of recognition trials (e.g., Smyth & Shanks, 2008), given the distinction between access and phenomenal consciousness, it would still not be possible to tell whether above-chance recognition scores index explicit knowledge about display repetitions.

To address the latter question, of the conscious accessibility of target location information in repeated displays, the present experiment was divided into a training and a test session.1 In the training session, observers performed the standard contextual cueing task (e.g., Chun & Jiang, 1998), to foster contextual learning of the repeated displays. In the test session (comprising the same repeated configurations already encountered in the training session), search displays were presented only briefly and terminated by visual masks. Observers had to localize the target stimulus (similar to the “generation task” used by Smyth & Shanks, 2008), with the dependent variable being response accuracy. Each search display was followed by a visual experience (VE) as well as a confidence (CO) rating. The ratings were similar to those used by Sandberg et al. (2010), the difference being that both ratings referred to the target stimulus in a given display rather than to visual experience in general. The predictions for the test session were as follows: First, regarding response accuracy, we hypothesized an effect of repeated displays on the deployment of attention which should be expressed by higher accuracy scores for repeated compared to non-repeated displays (see Geyer et al., 2010, for the same rationale and positive findings, however, for a pop-out search task). Second, regarding the VE and CR measures, if the target location in repeated displays cannot be accessed explicitly, then neither of the two consciousness ratings should differ between repeated and nonrepeated displays; restated, contextual cueing should not be associated with improved “consciousness” of the target, whether in terms of access or phenomenal consciousness measures (i.e., “implicit” hypothesis). By contrast, contextual cueing might lead to higher ratings of confidence for the target in repeated over nonrepeated configurations, while not being associated with any improved phenomenal percept. We refer to this as the “access consciousness” hypothesis. Finally, both—access and phenomenal—consciousness measures might be revealed to be higher for repeated than for nonrepeated displays. Such a pattern would indicate that observers could access the location of the target in repeated displays explicitly (i.e., “phenomenal consciousness” hypothesis).
Analyses of contextual cueing at the level of individual displays

Prior studies suggest that contextual cueing manifests for only a subset of repeated displays (e.g., Peterson & Kramer, 2001; Johnson, Woodman, Braun, & Luck, 2007). For example, Peterson and Kramer (2001) observed in an oculomotor study a larger proportion of initial target fixations for repeated compared to nonrepeated displays. However, when trials in which the eyes that went directly to the target were excluded from analysis, there was no trend for eye movements (fixations) to land closer to the target location in repeated displays. This led Peterson and Kramer (2001) to propose that contextual cueing is highly accurate in guiding attention, but that guidance is limited to repeated and, importantly, “recognized” displays. The idea of contextual cueing being due to only a subset of (in Peterson and Kramer’s terms: “recognized”) displays was reinforced by Smyth and Shanks (2008), who introduced a formal procedure for identifying displays that gave rise to contextual cueing. Smyth and Shanks (2008) compared RTs to each of the 12 repeated displays with the mean RTs to nonrepeated displays (i.e., single-display analysis). An individual repeated display was classified as a “recognized”—or “cueing”—display if its associated RT value was below the 99% confidence interval (CI) of the mean RTs to nonrepeated displays. Using this criterion, Smyth and Shanks (2008) found that, on average, only 2–4 repeated configurations (of the typical 12) produced a cueing effect.

On this background, the most stringent test of the above hypotheses as to the relation between contextual cueing and awareness would involve a comparison of visual consciousness measures between displays actually learned by observers and nonrepeated displays.

In the current study, we were interested in two types of displays: those repeated displays that produced a “positive” contextual cueing effect (i.e., cueing displays) and those that yielded a “negative” cueing effect (noncueing displays). The two types of display were identified on the basis of the same 99% CI criterion, depending on whether RTs to an individual repeated display were below the lower or above the upper 99% CI boundary (cueing vs. noncueing displays, respectively). Furthermore, participants were classified according to whether they did or did not show contextual cueing (see Lleras & von Mühlener, 2004, who introduced the distinction between positive and negative cueing observers). In order to assess whether variations in the size of contextual cueing, for a given observer, are due to variations in the size of the RT gains derived by a relatively constant set of learned displays, or whether contextual cueing scales with the number of learned and nonlearned displays (assuming that RT gains derived from contextual cueing are relatively constant), the number of cueing and noncueing displays was assessed for each (positive and, respectively, negative) observer.

Method

Participants

Twenty observers took part in the experiment (four male; mean age: 27.1 years). All participants reported normal or corrected-to-normal vision and were naïve as to the purpose of the study. Subjects received 12€ (~$16 USD) or course credit for their participation.

Stimuli and procedure

Figure 1 illustrates the sequence of events in both the training and the test session. In the training session, each trial started with a 500-ms presentation of a black fixation cross (size: 0.80° × 0.80°; viewing distance: 57 cm; luminance: 0.5 cd/m²) in the center of the computer screen (a 19-inch CRT monitor; the display resolution was set to 1024 × 768 pixels; the refresh rate was 85 Hz [AOC, Amsterdam, The Netherlands]). After a blank interval of 200 ms, the search display appeared. Search displays consisted of one black “T” and 11 black “L”s (0.46° × 0.46°; 0.5 cd/m²) presented on a white background (30 cd/m²); the “T” target was rotated by 90° or 270°, and the “L” distractors by 0°, 90°, 180°, or 270° from the vertical. The 12 display stimuli were scattered randomly over an invisible 8 × 6 matrix (total matrix size: 24° × 18°). The target was never presented at the four most inner positions, leaving a total of 44 possible target locations. From these 44 locations, 24 target locations were randomly selected at the start of
the experiment, for each observer: 12 locations were used for repeated and 12 for nonrepeated displays. Note that targets appeared equally likely in each of the four display quadrants. Further, target eccentricity was equated across the two (repeated, nonrepeated) conditions. Observers had to detect the target letter “T” (present on each trial) and discriminate its orientation (left- vs. right-oriented) by pressing a corresponding key on a computer keyboard (left and right arrow keys) with their left and right index fingers, respectively. The stimuli were visible until observers produced their search task response. Observers were instructed to respond as fast and as accurately as possible. Error feedback was given by presenting the word “Fehler” (German word for “error”) in black letters in the center of the screen. The intertrial interval was 500 ms (increased to 1000 ms following a response error). The training phase consisted of 288 trials, divided into 12 blocks of 24 trials each. Each block contained 12 displays with identical target-distractor arrangements (i.e., repeated displays; note that the repeated displays had a different arrangement relative to each other) and 12 newly generated arrangements (i.e., nonrepeated displays), constituting the independent variable display type (repeated vs. nonrepeated display). For analysis, four blocks of trials were aggregated into one epoch in order to obtain a reasonable estimate of the contextual cueing effect.

In the test session, a trial started with the presentation of a fixation cross for 1500 ms followed by a blank interval of 200 ms. Then, the search array (consisting of one “T” plus 11 “L” letter stimuli) was presented for a certain (across observers variable) time and followed by a visual mask of figure-8 placeholders presented at each of the 8 × 6 matrix locations. The placeholders had the same size as the search items and were displayed until observers responded to the location (quadrant) of the “T” target letter. Note that this location-based response is different to that required in the training session, in which observers had to detect the target and subsequently discriminate its orientation. Responses were given with the right hand on the numeric keypad section (on the far right) of the computer keyboard. The response-relevant numerical keys were chosen in order to spatially correspond to the likely target quadrant. Observers pressed the “1” key if they believed the target was located in the bottom-left quadrant, “3” for the bottom-right quadrant, “7” for the top-left quadrant, and “9” for the top-right quadrant.

The stimulus onset asynchronies (SOAs) between the search displays and masks (i.e., the effective display exposure durations) were determined individually in a pre-experimental session using an adaptive staircase procedure that aimed at a level of 75% correct localization responses. The staircase started with an SOA value of 400 ms and was adjusted stepwise until the first reversal point was reached (i.e., an erroneous trial after a correct trial or vice versa). During the first six reversals, SOA values were modified by step sizes of six frame durations (~72 ms; at the screen refresh rate of 85 Hz one frame lasted approximately 12 ms) to step sizes of one frame duration for the last six reversals. In order to achieve the 75%-correct threshold, the SOA step size was doubled (increased) following an error response trial. A total of 12 reversal points was determined. The 75%-threshold was taken as the average SOA value across the last six reversal points. The mean SOA established in this way was 581 ms (standard deviation: 133 ms). Search displays presented in the staircase procedure contained novel target-distractor configurations. None of them was shown later in the training or test sessions. In the test session of the search task (as well as in the staircase task), observers were instructed to respond as accurately as possible without stress on response speed. Following their localization responses (and a blank interval of 500 ms), observers had to rate their visual experience of the target and their confidence in their quadrant response on a four-point scale by pressing the corresponding keys on the computer keyboard with their left hand (“1,” “2,” “3,” or “4” key). The visual experience (VE) question was: “Wie klar haben Sie das T gesehen?” (German for “How clearly did you see the T?”). The label for the left side (anchor) of the scale was “sehr unklar” (“very unclear”); the label for the right side was “sehr klar” (“very clear”). The confidence rating (CO) question was: “Wie sicher waren Sie sich in der Ortsentscheidung?” (“How sure were you with your localization decision?”) The label for the left side of the scale was “sehr unsicher” (“very unsure”); the label for the right side was “sehr sicher” (“very sure”). No labels were presented for each point on the rating scale. The order of the ratings was randomly determined at the beginning of each block, but kept constant throughout a given block. At the beginning of the test session, participants were informed about the variable order of the rating questions, providing them with an incentive to carefully read the queries on each trial. To preview the results, the VE and CO ratings were statistically different, in particular for observers who showed a contextual cueing effect (see Results section). 3.00 versus 3.11, \( t(11) = 8.52, p < 0.01 \). This suggests that they responded to the two questions (at least to some extent) differently, rather than treating them as one and the same. Both questions were presented on the screen, in black letters, including the scale from 1 to 4 with the corresponding labels at the left and right ends of the scale (see Figure 1). There was a blank interval of 200 ms between the two ratings, followed by a 200-ms intertrial interval. The test session consisted of 96 trials, divided into four blocks of 24 trials each.
Results

Training session

Data analysis was performed using R (R Development Core Team, 2007). RTs in the training session outside the range ± 2.5 standard deviations of the mean were discarded as outliers (overall 2.23% of trials). Further, incorrect responses (in the training session) were also discarded from analysis. The mean error rate was 2.4%. A 2 × 3 (display type × epoch) ANOVA on the error rates revealed no reliable effects (all ps > 0.10).

Table 1 presents the mean RTs for repeated and nonrepeated displays in the third epoch. For all observers, contextual cueing gains were calculated by subtracting, for the last epoch, mean RTs for repeated displays from mean RTs for nonrepeated displays. Figure 2 shows the individual contextual cueing effects in the third epoch. As can be seen, contextual cueing was manifest for 12 out of the 20 observers (“positive cueing group”); for the remaining eight observers, no contextual cueing effects were evident (“negative cueing group”).

Subsequent analyses were conducted for the positive cueing group only, as it is reasonable to assume that only observers in this group had learned the repeated displays to a significant extent. Figure 3 shows RTs for observers in the positive cueing group. Mean RTs were examined by a display type (repeated, nonrepeated) × epoch (1, 2, 3) repeated-measures ANOVA. The main effect of display type, F(1, 11) = 7.23, p = 0.02, was significant. No other effects were reliable: main effect of epoch, F(2, 22) = 2.22, p = 0.13, and display type × epoch interaction, F(2, 22) = 1.77, p = 0.19. Separate t tests (Holm-corrected) comparing old and new displays in each epoch revealed no reliable RT difference for the first two epochs (epoch 1: 1153 ms vs. 1180 ms, p = 0.39; epoch 2: 1108 ms vs. 1161 ms, p = 0.163), but a reliable effect for the third epoch (1087 ms vs. 1173 ms, p < 0.01; contextual cueing gain of 86 ms). This pattern indicates that contextual cueing developed over the course of the training session.

Test session

Observers made fewer localization errors with repeated than with nonrepeated displays; however, this difference was only borderline-significant [30.6% vs. 34.9%; one-tailed t(11) = 1.33, p = 0.10].² Further, VE and CO consciousness ratings did not differ significantly between repeated and nonrepeated displays [both ps > 0.30], though there was some numerical tendency for both consciousness ratings to be increased for repeated versus nonrepeated displays (VE: 3.00 vs. 2.99; CO: 3.14 vs. 3.09). Table 2 presents the results of the test session.

<table>
<thead>
<tr>
<th>Training</th>
<th>RT of epoch 3 (in ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Old</td>
</tr>
<tr>
<td>Positive group</td>
<td>1086</td>
</tr>
<tr>
<td>Negative group</td>
<td>1351</td>
</tr>
<tr>
<td>All</td>
<td>1185</td>
</tr>
</tbody>
</table>

Table 1. Training session: Mean RTs, in epoch 3 of the training session, to repeated and nonrepeated displays, separately for all observers (N = 20), observers showing a “positive” contextual cueing effect only, (N = 12), and observers showing a “negative” cueing effect only (N = 8). The contextual cueing effect was determined by subtracting RTs to repeated from RTs to nonrepeated displays (in epoch 3).

Figure 2. Training session: Contextual cueing scores (RT nonrepeated display minus RT repeated display) in the third epoch of the training session, separately for each observer. The dashed line separates the negative cueing group (RT nonrepeated display < RT repeated display) from the positive cueing group (RT nonrepeated display > RT repeated display).

Figure 3. RTs in the training session, separately for repeated and nonrepeated displays (positive cueing group only). Error bars indicate the standard error of the mean.

Figure 3. RTs in the training session, separately for repeated and nonrepeated displays (positive cueing group only). Error bars indicate the standard error of the mean.
Single-display analysis

To further investigate the difference between the negative and positive cueing groups and to obtain a finer-grained picture of the data in the test session, the dependent variables were re-investigated individually for each of the repeated displays (i.e., single-display analyses). To investigate the number of repeated displays that were learned by participants in the present experiment, RTs were computed separately for each of the repeated displays—and each participant, as the repeated configurations were different for each observer—and compared to the 99% CI of the participant’s mean RT for nonrepeated displays (cf. Smyth & Shanks, 2008). A repeated configuration was classified as a “cueing display” (i.e., a display generating a contextual-cueing effect) if RTs to this configuration fell below the left boundary of the 99% CI interval. On the contrary, a repeated configuration was classified as a “noncueing display” if RTs to this display fell above the right boundary of the 99% CI. Note that the number of “cueing” displays is not necessarily dependent on the number of “noncueing” displays (negative correlation), because both types of displays are linked by a third type of “neutral” displays, also repeated displays whose RT values are within the 99% CI.

Table 2. Test session: Mean response errors, visual experience, and confidence ratings for repeated and nonrepeated displays in the test session, separately for all observers (N = 20), observers in the positive cueing group (N = 12), and observers in the negative cueing group (N = 8).

<table>
<thead>
<tr>
<th>Test</th>
<th>Response errors</th>
<th>Visual experience</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Old</td>
<td>New</td>
<td>Old</td>
</tr>
<tr>
<td>Positive group</td>
<td>30.56</td>
<td>34.90</td>
<td>3.00</td>
</tr>
<tr>
<td>Negative group</td>
<td>30.70</td>
<td>29.20</td>
<td>2.93</td>
</tr>
<tr>
<td>All</td>
<td>31.98</td>
<td>33.75</td>
<td>2.97</td>
</tr>
</tbody>
</table>

Figure 4 shows the results of the single-display analysis. Interestingly, the number of cueing displays was larger for the positive than for the negative group [5.1 vs. 3.0; t(17) = 3.97, p < 0.01]. By contrast, the number of noncueing displays was larger for the negative than for the positive group [5.0 vs. 2.3; t(15) = 5.46, p < 0.01]. And, within the positive group, the number of cueing displays was significantly larger than that of noncueing displays [5.1 vs. 2.3; t(11) = 4.21, p < 0.01]. The opposite was found for the negative group [noncueing vs. cueing displays: 5.0 vs. 3.0; t(7) = 4.32, p < 0.01]. These results suggest that differences in the magnitude of the contextual cueing (positive vs. negative group) can be attributed to differences in both the number of cueing and that of noncueing displays. In other words, the pattern of RTs observed in contextual cueing result from a mixture distribution of fast and slow RTs (considered further below).

Further, a comparison of the accuracy performance in the test session between repeated displays that generated contextual cueing and nonrepeated displays yielded a significant difference: target location responses were more accurate for cueing compared to nonrepeated displays [21.2% vs. 34.9%; one-tailed t(11) = 3.88, p < 0.01]. By contrast, response accuracy was comparable between repeated displays that did not yield contextual cueing and nonrepeated displays [37.4% and 34.9%, t(11) = 0.79, p = 0.44].

Most importantly, consciousness ratings were higher for cueing than for nonrepeated displays (see Figure 5): VE-rating, 3.43 vs. 2.99 [one-tailed t(11) = 5.14, p < 0.001], and CO-rating, 3.50 vs. 3.09 [one-tailed t(11) = 4.27, p < 0.001]. By contrast, VE-ratings were lower for noncueing compared with nonrepeated displays: 2.11 and 2.99 [(t(11) = 3.63, p = 0.004); the same was observed for CO-ratings: 2.27 vs. 3.09 [(t(11) = 3.03, p = 0.01]. This pattern of results suggests that observers had explicit knowledge of displays that generated contextual cueing—in line with the phenomenal consciousness hypothesis.

Figure 4. Training session: Number of repeated cueing displays (RT < left boundary of 99% confidence interval of the new mean RT) and noncueing displays (RT > right boundary of 99% confidence of new mean RT). The dashed line separates the negative cueing group from the positive cueing group.

Role of perceptual fluency?

An alternative way to explain the above findings, in particular, the enhanced consciousness ratings (along...
with the enhanced response accuracy) for cueing displays, is to assume that these ratings reflect perceptual fluency—or “meta-memory” of the speed with which the target localization responses were made.3 In other words, the consciousness ratings might reflect the (perceived) ease or speed of the localization responses, rather than perception of the target stimuli. Consistent with this, RTs in the present experiment were faster for cueing than for nonrepeated displays (main effect of display type in the training session). Thus, assuming that this speed advantage transfers to the test session, the target would also be more likely to be localized rapidly in the test phase, which might, in turn, increase the visual consciousness ratings in addition to response accuracy. In an attempt to test this alternative account, we first examined RTs in the test phase (even though participants had not been instructed that response speed was critical on these trials), which did indeed reveal response times to be overall faster for cueing than for nonrepeated displays (main effect of display type in the training session). Accordingly, response speed might have had an effect on the dependent measures in the test session. Next, for each participant, based on an RT median split, test trials were partitioned into “fast” and “slow” response trials, respectively, and the dependent measures were re-analyzed as a function of response speed (see Table 3). Technically, observers’ VE ratings, CO ratings, and accuracy performance were examined by separate 2 × 2 ANOVAs, each with display type (cueing, nonrepeated) and response speed (fast vs. slow) as factors. For the visual experience ratings, the ANOVA revealed significant main effects of display type, F(1, 11) = 29.49, p < 0.001, and response speed, F(1, 11) = 27.97, p < 0.001, but, importantly, the interaction was nonsignificant: F(1, 11) = 0.29, p = 0.60. The same pattern was observed for the confidence ratings, display type: F(1, 11) = 14.46, p < .001; response speed: F(1, 11) = 14.46, p < 0.01; interaction: F(1, 11) = 2.05, p = .18; and for response accuracy, display type: F(1, 11) = 76.91, p < 0.01; display type: F(1, 11) = 20.70, p < 0.01; interaction: F(1, 11) = 1.03, p = 0.33. Importantly, after the median-split RTs were not reliably different for fast cueing as compared to fast nonrepeated displays [1316 ms vs. 1321 ms; t(11) = 0.23, p = 0.826] (the same was true for slow cueing versus slow nonrepeated displays [1932 ms vs. 1969 ms; t(11) = 0.61, p = 0.552]), suggesting that fast cueing and nonrepeated displays were effectively equated in terms of perceptual fluency. Nevertheless, visual consciousness ratings, along with response accuracy, were still higher for fast cueing relative to fast nonrepeated displays [VE: 3.79 vs. 3.41; t(11) = 3.49, p = 0.01; CO: 3.85 vs. 3.56; t(11) = 2.40, p = 0.035; response errors: 7.9 vs. 17.2; t(11) = 2.77, p = 0.01 (Holm-corrected)]. This argues that the consciousness ratings are not simply a function of perceptual fluency; instead, contextual cueing is associated with enhanced visual consciousness, as well as enhanced localization performance (this association is further considered below).

### Discussion

The aim of the present study was to investigate conscious access to the target element in contextual cueing of visual search. To examine this, observers performed a training session meant to establish contextual memory representations, and a test session intended to measure conscious accessibility of the
target location in repeated configurations. The results showed that the presentation of repeated displays led to performance gains in both RT (training session) and accuracy (test session) measures. That is, contextual associations acquired in a discrimination task (dependent variable: RT) can carry over to a localization task (dependent variable: accuracy), indicating that a common learned representation is generating the performance benefit in both tasks. Given that contextual cueing also facilitated the accuracy of nonspeeded (localization) responses, it is likely that the memory of repeated displays facilitates “early” processes of target selection, rather than “late” processes of response selection (see also Geyer et al., 2010). This conclusion is predicated on the assumption (e.g., Santee & Egeth, 1982; Pashler, 1989) that accuracy measures only index perceptual stages of processing, whereas RT measures reflect processing at both perceptual and postperceptual stages.

Contextual cueing and the global workspace

When comparing (the entire set of 12) repeated and nonrepeated displays, there was no difference in the two consciousness ratings thought to measure phenomenal and, respectively, access consciousness. This is in line with previous results from “standard” recognition tests, which appeared to show observers to be unable to consciously access the memory underlying contextual cueing (e.g., Chun, 2000; Chun & Jiang, 2003). The present approach, however, overcomes a severe problem of standard recognition tests in which participants are presented with only a very limited number (typically 24) of to-be-judged displays. In contrast, in the test session of the current experiment, participants performed 96 search and, importantly, 96 recognition trials. Under these conditions, it was possible to derive more powerful conclusions regarding explicit memory effects as well as to examine recognition performance (along with contextual cueing) for individual displays. These analyses revealed that (a) only some repeated displays produced contextual cueing and (b) observers showed increased awareness of these “cueing” displays in terms of both subjective clarity of the target’s visual experience and decision confidence. That is, the fact that participants exhibited increased confidence in their localization responses for cueing displays can be taken to indicate that learned target-distractor configurations enhance the accessibility of the target stimulus in the “global workspace” (Block, 2002, 2005). Furthermore, the enhanced visual experience ratings show that the visual stimulus is not only available functionally for solving the task at hand, but also in terms of the phenomenal experience of the target item. Contextual cueing can thus be interpreted as a mechanism that guides attention faster to the target location and makes visual information more readily available in the global workspace. This interpretation is in line with theories of visual attention such as that put forward by Bundesen (1990, 2005): visual stimuli that are attentionally selected, in a competitive (race) process, enter a limited-capacity visual short-term memory (VSTM), where they become available to higher-order cognitive processes (including response selection) and explicit report. In the case of contextual cueing, the memory of repeated distractor contexts modulates focal-attentional selection by conferring a competitive bias to (i.e., increasing the selection weight for) the object at the target location. Consequently, this location or object gains prior entry to VSTM, which translates into both improved target detection performance (i.e., faster RTs, higher perceptual accuracy) and increased availability or, respectively, accessibility in the global work space.5

Contextual cueing is due to only a few repeated displays

But how can it be explained that only a subset of observers showed a positive contextual cueing effect, while the others showed no or even a negative effect? Lleras and von Mühlener (2004) reported a very similar result pattern. In an attempt to replicate the original experiment of Chun and Jiang (1998), Lleras and von Mühlener failed to find consistent evidence for contextual cueing, which they attributed to differences in cueing effects among observers: while some observers did show contextual cueing (= positive group), others did not or even showed negative cueing effects (= negative group). Postexperimental debriefings suggested that different search strategies could be the reason for the positive and negative cueing effects. This was examined in their Experiment 3, in which they varied task instructions. Observers were told either to search actively for the target by deliberately shifting their attention (= active strategy), or to be as receptive as possible, letting the unique item “pop into their minds” (= passive strategy; Lleras & von Mühlener, 2004, p. 467). The results were that of significant contextual cueing for the passive search group only. Interestingly, the size of positive contextual cueing (passive group) was comparable to that of negative cueing (active group), a finding that led Lleras and von Mühlener to conclude that a passive search strategy leads to contextual cueing, while an active search inhibits the cueing effect. However, this conclusion is more descriptive rather than explanatory. In particular, it remains unclear what processes are responsible for positive and negative cueing effects and, once identified, whether these processes modulate
the learning or the expression of memory for repeated displays. To address this issue, the present study analyzed RTs at the level of individual repeated displays. These analyses revealed contextual cueing to scale with the number of cueing and noncueing displays: the effect was larger (and positive) with larger numbers of cueing and smaller numbers of noncueing displays (see Figure 4). This suggests that the search strategy adopted modulates the number of repeated displays that are learned by observers (rather than modulating the magnitude of the RT gains derived from contextual cueing). For example, using an active search strategy may pose pressure on the attention system (e.g., deliberate programming of attention shifts and/or voluntary tracking of already inspected locations, etc.), reducing attentional capacity for the item arrangement and, with it, contextual cueing. In a sense, active search would be a condition of high load (e.g., Lavie, Hirst, de Fockert, & Viding, 2004), making observers ignore the repeated displays. This is supported by the results of Manginelli et al. (under revision), who combined the visual search task with a second, spatial working memory (WM) task, the latter intended for reducing attentional capacity in the search task. It was found that the addition of the WM task abolished contextual cueing, particularly when the WM task was administered in late experimental trials. While this may suggest that the WM task hampers the retrieval, rather than the acquisition, of memory for repeated configurations (see also Jiang & Leung, 2005, or Conci & von Mühlener, 2009, for similar findings and conclusions, albeit based on different paradigms), the important point here is that a secondary WM task (Manginelli et al., under revision) or an active search task per se (Lleras & von Mühlener, 2004) take away attentional capacity for the processing of repeated displays.

Conclusion

In summary, the current study provides evidence that repeated displays in contextual cueing do not only expedite search RTs under unlimited viewing conditions but also lead to improved response accuracy when displays are presented only briefly and followed by visual masks. Further, single display analyses showed that contextual cueing was due to only a very small number (2–4) of repeated displays and that for these “cueing” displays, observers exhibited increased awareness of the target location, as indexed by two consciousness measures: visual experience of the target element and confidence of the localization decision. These findings are in line with the idea that contextual cueing leads to faster allocation of attention towards the target location, and thus faster availability of the target in the global workspace—with this information being then available for further processing as well as phenomenal experience. We propose one promising direction for future research: the analysis of the dependent variables in contextual cueing of search at the level of individual repeated displays.

Acknowledgments

This research was supported by Deutsche Forschungsgemeinschaft (DFG) Research Group (FOR 480), Research Project (GE 1889/1-1) and CoTeSys Excellence Cluster (142/435).

Commercial relationships: none.
Corresponding author: Thomas Geyer.
Email: geyer@lmu.de.
Address: Department Psychologie, Lehrstuhl für Allgemeine und Experimentelle Psychologie, Ludwig-Maximilians-Universität München, München, Germany.

Footnotes

1 The aim of the current experiment was to study (conscious) processes associated with the accessing of learned information, rather than the learning of repeated search displays; therefore, measures of visual consciousness were obtained after learning had taken place. Of course, the division of the experiment into a training and test phase does not entirely rule out that learning was restricted to the former phase. However, in order to minimize new learning in the test phase, the test session was limited to one epoch only (as opposed to three epochs in the training session). This compares with earlier investigations of the contextual cueing effect that have addressed the issue of access to, rather than the learning of, repeated configurations (e.g., Jiang & Leung, 2005, or Manginelli, Langer, & Pollmann, in revision). In addition, observers did not receive feedback about the accuracy of their manual (i.e., localization) response in the test session (but did so in the training session). This should also have decreased new learning in the test session. Further, in “standard” contextual cueing tasks, learning of repeated displays reaches a plateau after some six to eight repetitions of individual repeated displays (see, e.g., Chun & Jiang, 1998; Experiment 1). Applied to the current experiment, this means that configural learning would have been saturated in epochs 2–3 of the training session.
Two participants in the positive group showed a relatively high proportion of response errors in the localization task, 47.9%, which compares with a mean error rate of only 29.1% (SD: 10.2%). If these two observers are excluded from analysis, the effect of contextual cueing on response accuracy was significant [repeated vs. nonrepeated displays: 26.9% vs. 33.1%; one-tailed t(9) = 2.11, p < 0.05].

We thank Todd Horowitz for alerting us to this possibility.

This null-finding is only seemingly incompatible with the above result of the RTs, in the test session, being overall faster for cueing than nonrepeated displays. The reason for this is that for cueing displays, the relative amount of fast trials, in terms of our classification, is larger than that for nonrepeated displays (59% vs. 47%); therefore also mean RTs are faster for the former type of display. However, when performing a median-split on both RTs to cueing and nonrepeated displays, this imbalance in the relative proportion of fast responses is “corrected,” which also leads to comparable RTs between the two types of displays.

Note that learning of the target-distractor configuration would not only require that the target is selected, but also that the context is encoded. Assuming that attentional selection of the target would involve suppression of the distractor surround (see, e.g., the literature on inhibitory positional intertrial effect: e.g., Maljkovic & Nakayama, 1996), this would hamper encoding of the (suppressed) distractor context. However, along the lines of Lavie’s (1995) “load” account, the more efficient target selection becomes over the course of learning, the more attentional capacity (required for suppression) is released for encoding the distractor context.

References


