cortex 154 (2022) 287–298

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex


Registered Report

# The role of working memory in contextual cueing of visual attention

*Francisco Vicente-Conesa* [a], *Tamara Giménez-Fernández* [a], *David R. Shanks* [b] and *Miguel A. Vadillo* [a,*]

[a] *Departamento de Psicología Básica, Universidad Autónoma de Madrid, Spain*
[b] *Department of Experimental Psychology, University College London, UK*

A R T I C L E I N F O

*Article history:*
Protocol received 3 November 2020
Protocol approved 6 July 2021
Received 9 March 2022
Reviewed 11 April 2022
Revised 23 May 2022
Accepted 24 May 2022
Action editor Chris Chambers
Published online 13 June 2022

*Keywords:*
Contextual cueing
Implicit learning
Working memory

A B S T R A C T

It is usually easier to find objects in familiar contexts that we have seen before. The type of learning that underlies this facilitation, known as *contextual cueing*, has been understood as an incidental and automatic process given that, among other reasons, it seems to be independent of working memory (WM) resources. This claim has found support in previous research showing that contextual cueing can be acquired latently, while participants perform a demanding WM task. However, previous studies have not always found this pattern of results and, in general, the available evidence is far from conclusive. The aim of the present study was to clarify the role of WM in contextual learning with two large-sample, confirmatory experiments. Our results show a robust contextual cueing effect even when visuospatial working memory resources were recruited by a demanding secondary task.


Regularities in our environment can be exploited to improve visual search. Contextual information allows us to focus our attention on the most important regions of a given scene. In the laboratory, this process has been studied extensively using an experimental paradigm known as *contextual cueing* (Chun & Jiang, 1998). In a typical experiment, a target (usually a tilted 'T') is presented among several distractors (usually tilted 'L' letters) and participants are instructed to find the target as fast as possible and respond to its orientation. Unknown to them, half of the search displays are repeated in each block of trials, whereas the other half are generated randomly (novel displays). The results usually reveal a steeper decrease in search times for repeated than for novel displays.

This contextual cueing effect has been understood as a form of incidental and automatic learning. One of the arguments supporting this claim is that contextual cueing seems to take place in the absence of any explicit knowledge about the repeated displays (Chun & Jiang, 2003). For instance, many

* *Corresponding author*. Departamento de Psicología Básica, Facultad de Psicología, Universidad Autónoma de Madrid, 28049, Madrid, Spain.
E-mail address: miguel.vadillo@uam.es (M.A. Vadillo).
https://doi.org/10.1016/j.cortex.2022.05.019

studies have reported that if participants are presented with a recognition test at the end of the experiment, they are unable to recognize repeated patterns, although they have seen them many times during the experiment.

However, studies using more fine-grained and high-powered measures of explicit knowledge usually reveal a certain level of awareness. For example, Smyth and Shanks (2008) applied awareness tests not only after the contextual cueing task, but also during the course of the experiment. Their results suggested that contextual cueing and explicit knowledge develop simultaneously. Furthermore, the absence of evidence of awareness about repeated displays reported in the literature may be due to a combination of low power and noisy measures of awareness (Vadillo et al., 2016; Smyth & Shanks, 2008). At present, whether or not contextual cueing can be considered an unconscious learning effect is the subject of vigorous debate (see, e.g., Colagiuri & Livesey, 2016; Geyer et al., 2020; Vadillo et al., 2022).

A second argument supporting the view that contextual cueing arises from automatic processes is that learning seems to demand few cognitive resources. Apparently, learning can occur even for stimuli that have been actively ignored by participants, although this learning is only expressed once the display becomes fully attended (Jiang & Leung, 2005; but see Vadillo et al., 2020).

Following the same reasoning, if contextual cueing is a purely automatic process, it should not depend on the availability of working memory (WM) resources. In a seminal series of experiments, Vickery et al. (2010) explored the role of WM in contextual cueing. During the first stage of their experiments (i.e., the training phase), participants had to perform a secondary WM task, such as maintaining a set of four colours in memory, concurrently with the visual search task. Then, contextual cueing was measured in a final test phase in which participants only performed the visual search task. Since contextual cueing was observed in the test phase, Vickery et al. concluded that the acquisition of contextual information must be independent of WM constraints. Since then, many experiments have followed up on this result manipulating whether the concurrent WM task is presented during the learning or the test stage (Annac et al., 2013; Chen et al., 2019; Manginelli et al., 2013b; Travis et al., 2013). The general consensus arising from this research is that WM load can interfere with the expression of contextual cueing, but not with learning itself (Goujon et al., 2015; Pollman, 2019; Sisk et al., 2019).

Some of the experiments supporting this view employ what we will call 'Type-1' designs. In these studies, the secondary task is performed in the initial training phase, but not in the final test phase. If WM load interferes with the expression of contextual cueing but not with learning, then these studies should find no or only weak evidence of contextual cueing during the first stage, but a sudden onset of contextual cueing in the test phase. The left panel of Fig. 1 shows the ideal pattern of results that would be expected following this logic. Other studies rely on what we call 'Type-2' designs. In these studies, the initial training phase is performed under single-task conditions and the concurrent WM task is administered only in the test phase. Again, if WM load interferes with the expression of contextual cueing but not with learning, one would expect to find a strong contextual cueing effect during the first phase, which would immediately disappear or at least be attenuated under dual-task conditions. The right panel of Fig. 1 shows the predicted pattern of results for these experiments. A simple summary of the predictions in Fig. 1 is that reaction times should differ more between the novel and repeated conditions under No Load (blue) than under WM Load (brown).

The different panels in Fig. 2 summarize the results of the empirical studies published so far with both types of designs.[1] The first surprising feature of this figure is the scarcity of studies, especially for Type-2 experiments. The second is that the observed results do not always fit the expected patterns.

Among Type-1 studies, two of the four experiments (Chen et al., 2019, Experiment 3; Travis et al., 2013, Experiment 3) clearly fail to show the expected pattern of results, as they find no evidence of contextual cueing in the test stage, after removing the concurrent WM task. This is particularly interesting with regard to Travis et al. (2013, Experiment 3). In this experiment, participants showed a small but significant cueing effect during the training stage, even in the condition with the highest WM load, but, if anything, removing the concurrent WM task reduced the size of the effect at test.

The results reported by Chen et al. (2019, Experiment 3) are also remarkable because they contradict previous evidence that only visuospatial loading can have an impact on contextual cueing (Manginelli et al., 2012, 2013b). The WM task used in this study required participants to store and manipulate numerical information. The results depicted in Fig. 2 show that this secondary task interfered with both the acquisition and expression of contextual cueing, even more strongly than in the other studies represented in the figure, which used visuospatial tasks. In any case, and most importantly, these results do not support the conclusion that WM load interferes only with the expression, but not the acquisition, of contextual cueing.

Experiment 1 by Manginelli et al. (2013b) is also a Type-1 study. Their results show clear evidence of contextual cueing in early stages of the experiment even under dual-task conditions, indicating that learning may be expressed despite a high WM load. Furthermore, once the WM load was removed, the observed contextual cueing effect remained significant but did not increase; it actually decreased slightly. Removing the WM load did not seem to facilitate the expression of contextual cueing.

To the best of our knowledge, the only Type-1 experiment fitting the expected pattern is that by Annac et al. (2013, Experiment 2D). The results show clear evidence of contextual cueing in the test phase (i.e., under single-task conditions), compared with the immediately preceding dual-task phase. This led the authors to conclude that WM load impairs the expression of implicit learning, not learning itself. However, this difference could also be due to an intriguing irregularity in the data. Apparently, contextual cueing emerged in the first stage of the

[1] A handful of studies using Type-1 and Type-2 designs were not included in this figure because their authors did not predict any interference of the secondary WM task on contextual cueing. For instance, many authors predicted that contextual cueing would be unaffected by WM tasks not involving executive processes (Chen et al., 2019) or visuospatial representations (Manginelli et al., 2013b).

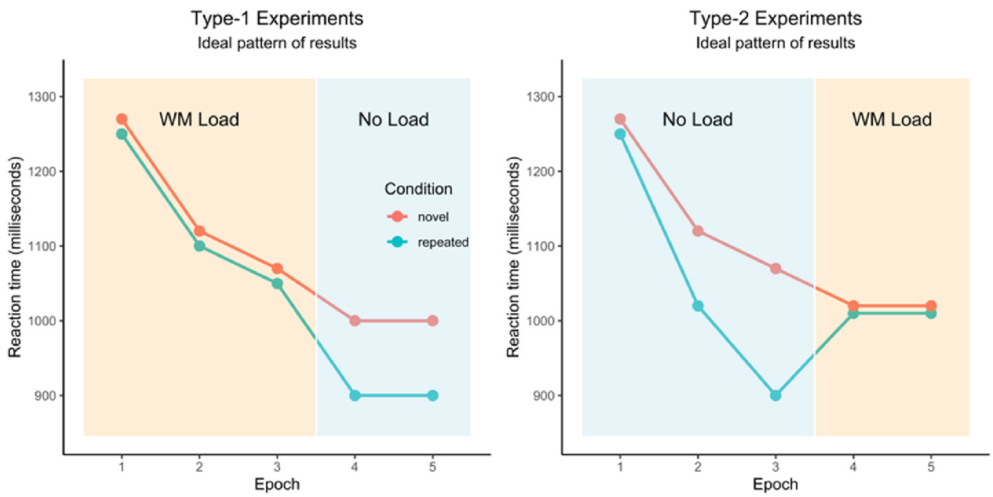


**Fig. 1 – Ideal pattern of results for Type-1 experiments (left panel) and Type-2 experiments (right panel).**

experiment (in epoch 2), but then disappeared in the last epoch of the training phase, perhaps resulting from sampling error.

To sum up, the four Type-1 experiments conducted so far have yielded variable results, possibly due to the different methods and manipulations used in each of those experiments, but perhaps also due to the typical amount of random noise and between-participants variability observed in contextual cueing experiments. Most importantly, only one of the four experiments using Type-1 designs lends clear support to the hypothesis that WM load interferes with the expression but not the acquisition of contextual cueing. In contrast, two of the experiments (Manginelli et al., 2013b, Experiment 1; Travis et al., 2013, Experiment 3) show a numerical decrease in the size of contextual cueing after removing the concurrent WM task. This suggests that, at least under some conditions, an abrupt change in the dynamics of the experimental task may, on its own, have an impact on the expression of learning. We return to this issue later on. To make things more complicated, although two studies found significant evidence of cueing at test with no WM load (Annac et al., 2013, Experiment 2D; Manginelli et al., 2013b, Experiment 1), the design of those experiments makes it unclear whether these results reveal any "latent" learning from previous stages of the experiment, or simply new learning arising during the test stage itself. Since the test phase contained a relatively large number of trials in both cases (120), we cannot know if the learning occurred during the training phase, test phase, or both. In the present study, we tried to discriminate between these two explanations.

We are aware of only two studies (Manginelli et al., 2013b, Experiment 3; Annac et al., 2013, Experiment 2C) relying on what we have called Type-2 designs, both of them showing the expected pattern of results. However, we find oddities in them too. Firstly, the size of contextual cueing during the first stage, conducted in the absence of any WM load, is surprisingly small. In addition to the experiments shown in Fig. 2, both Annac et al. (2013) and Manginelli et al. (2013b) included "baseline" studies where participants completed all the stages of the contextual cueing task in the absence of any working memory load. These control experiments yielded systematically larger cueing effects than the two Type-2 experiments depicted in Fig. 2. For instance, in Experiment 2C of Annac et al. (2013) the size of the contextual cueing effect (that is, the difference between mean RTs to novel *vs* repeated displays) by the end of the training stage was 84 msec. However, in the corresponding baseline Experiment 1, where training took place under identical conditions, the size of the cueing effect was 114 msec, almost 36% larger. Similarly, in Experiment 3 of Manginelli et al. (2013b) the size of the contextual cueing effect at the end of the training stage had an effect size of Cohen's $d_z = .47$, while in the baseline Experiment 5 the comparable size was $d_z = 1.13$, i.e., 140% larger. These differences did not reach statistical significance, but with the typically small samples of these experiments ($N = 19.83$ on average for the six experiments included in Fig. 2), null results are difficult to interpret.

This variability across experiments cannot be easily attributed to methodological differences, as the crucial experiments in Annac et al. (2013) and Manginelli et al. (2013b) were procedurally identical to their corresponding baseline studies. This suggests a considerable level of imprecision in the measurement of contextual cueing. Although all these experiments were sufficiently powered to detect the basic contextual cueing effect, the amount of noise or between-participants variability seen in these data suggests that larger samples may be needed to obtain precise estimates of contextual cueing with and without WM load. Given the small sample sizes of most contextual cueing experiments, the difference in cueing between stages 1 and 2 could simply be due to sampling error. A Bayesian reanalysis of the crucial results of Annac et al. (2013, Experiment 2C) and Manginelli et al. (2013b, Experiment 3) confirms that the evidence in favour of a reduction in cueing from stage 1 to stage 2 is indeed weak. In Annac et al. (2013), this reduction was tested by means of a 2 (Experiment: 1 "baseline" *vs* 2C) × 2 (Epoch: 3 *vs* 4) ANOVA on the size of contextual cueing, which yielded a significant interaction, $F(1, 32) = 5.61$. We computed a Bayes Factor, BF for this contrast assuming that this analysis is equivalent to an independent samples *t*-test comparing the evolution of

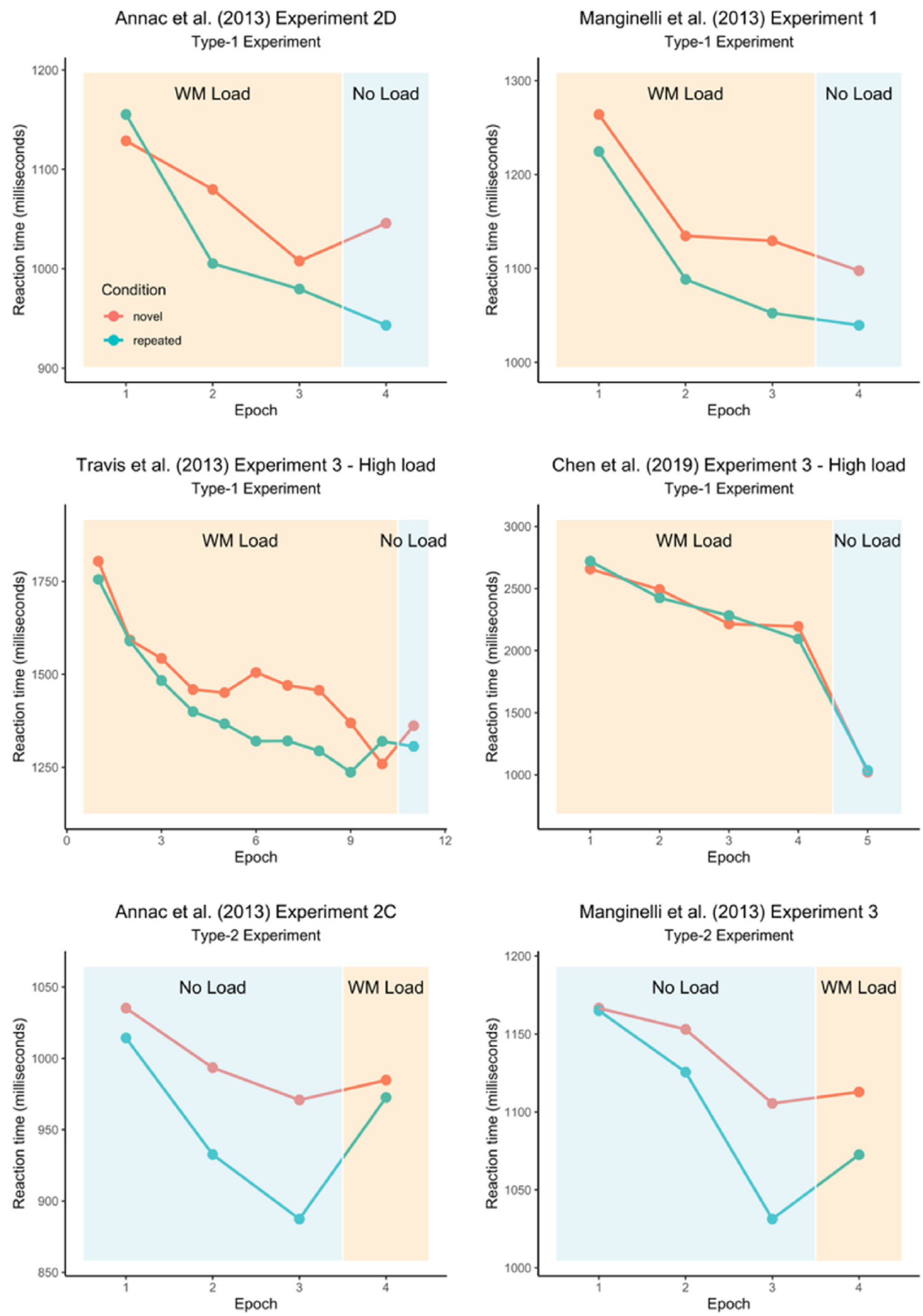


**Fig. 2 – Empirical results of previous experiments using Type-1 and Type-2 designs.**

contextual cueing from epoch 3 to epoch 4 in Experiments 1 and 2C, yielding $t(32) = 2.37$. Entering this value into the 'BayesFactor' package for R with the default settings results in a $BF_{10} = 2.57$, suggesting only anecdotal evidence in favour of the alternative hypothesis. A similar analysis can be applied to Manginelli et al. (2013b, Experiment 3). In this case, the inference that contextual cueing was reduced during stage 2 is based on a significant Epoch × Configuration interaction on reaction times, $F(1, 33) = 5.23$. Again, this is mathematically equivalent to a paired-samples $t$-test with the result $t(34) = 2.29$, which yields a $BF_{10} = 1.77$, once more just anecdotal evidence in favour of the alternative hypothesis. These analyses suggest that the crucial results of Annac et al. (2013, Experiment 2C) and Manginelli et al. (2013b, Experiment 3) should be confirmed in new research, ideally with much larger sample sizes to ameliorate the impact of sampling error in this kind of study.

Perhaps more importantly, there are good reasons to suspect that the impairment of contextual cueing at test, if real, may be caused or at least influenced by the mere lack of familiarity with the dual task. Some of the experiments shown in Fig. 2 suggest that some amount of practice may be essential to familiarize participants with the procedure of the test stage. For instance, two Type-1 experiments in Fig. 2 (Manginelli et al., 2013b, Experiment 1; Travis et al., 2013, Experiment 3) show a decline in contextual cueing after removing the concurrent WM task. If removing the secondary task disrupts contextual cueing, one can only speculate that introducing it suddenly after an initial stage with just the visual search task might be even more disruptive. Therefore, the impairment of contextual cueing in the test phase could be explained simply by the sudden addition of a demanding WM task. Again, the experiments reported here took measures to discount this possibility.

It is important to acknowledge that the idea that the acquisition of contextual cueing is independent of WM resources has also received some support from neuroimaging studies. Manginelli et al., 2013a conducted separately a spatial WM task and a visual search task and compared the common brain areas activated during them. They found that the size of the contextual cueing effect correlated with activation in an area along the descending segment of the left intraparietal sulcus, which was also involved in the WM task. Most importantly, this correlation was only observed after substantial training with the contextual cueing task. During early stages, the size of the contextual cueing effect was unrelated to the areas recruited by the WM task. According to the authors, this result dovetails with the idea that the expression but not the acquisition of contextual cueing depends on working memory resources. Despite these promising results, more empirical evidence is needed to reveal the neural basis of the interaction between implicit context learning and WM. Neuroimaging studies like this will necessarily lag behind the development of robust behavioural models (Niv, 2021). As we have argued above, the extant behavioural evidence is less than compelling.

In the present series of experiments, we aimed to replicate the two ideal patterns of results shown in Fig. 1, addressing some of the concerns highlighted above. Some Type-1 studies show contextual cueing during the training phase even under dual-task conditions. This could mean that the secondary tasks used in these studies are not sufficiently demanding to have a noticeable impact on contextual cueing (Annac et al., 2013; Manginelli et al., 2013b). In the two experiments reported below, the visual search task was combined with a working memory task that has been shown to effectively impair contextual cueing (Travis et al., 2013).

As noted above, one problem with previous Type-1 experiments is that learning can occur during the test phase. In the experiments reported below, the test phase included a new set of repeated displays, from now on called *control displays*. This allowed us to compare the magnitude of the facilitation of the repeated displays versus the novel and the control displays to check whether contextual cueing was acquired during the training or the test phase, respectively. Moreover, to reduce noise in the data due to lack of familiarity with the procedure used during the test stage, both experiments included a set of practice trials before the test stage. Annac et al. (2013) and Manginelli et al. (2013b) included practice trials with and without the WM task at the beginning of their experiments. While this might suffice to familiarize participants with the dual task, we think that presenting these practice trials between the training and test stages might be more effective at stabilizing participants' performance (and therefore reducing noise) at test. Finally, we argue that the variability observed in Fig. 2 across experiments and even across epochs within each study is a compelling reason to suspect that the data collected in these studies may not be sufficiently precise, probably due to sampling error. To solve this issue, the present experiments relied on considerably larger samples than previous experiments.

## 1. Experiment 1

The goal of Experiment 1 was to replicate the expected pattern of results for Type-1 designs (see the left panel of Fig. 1) while overcoming the methodological shortcomings identified above. Specifically, the aim of the experiment was to elucidate whether an initial training phase performed under dual-task conditions resulted in any latent learning, as revealed by evidence of contextual cueing when the WM load was removed in the test phase. In this experiment, one group of participants (the load group) performed the training phase under WM load and then performed the test without the secondary WM task. In addition, a separate group of participants (the no-load group) performed the experiment under the exact same conditions, except that they completed both stages without WM load. The latter thus served as a baseline condition to measure what is the size of the contextual cueing effect that one could expect to find at test in the complete absence of working memory load during the learning stage.

### 1.1. *Method*

#### 1.1.1. *Participants and apparatus*

There were two crucial contrasts for Experiment 1. On the one hand, one of the main goals of the study was to test whether RTs for repeated displays (trained during the first phase under WM load) were significantly faster than RTs for control displays at test in the load group. It is difficult to estimate the potential size of this within-participants effect. In our previous experiments, the size of the basic contextual cueing effect has usually been quite large, often larger than $d_z = 1.00$ (see Vadillo et al., 2021). It is reasonable to expect that after training with a demanding secondary task this effect might become much smaller. We thought it was reasonable to plan the sample size of the load group assuming a potential effect of roughly one third, $d_z = .33$. A sample of $N = 100$ would achieve 90% power to detect an effect of this size. On the other hand, we also wanted to test for potential differences in the size of cueing at test between the load (with training under WM load) and no-load groups (without WM load). If the sample size of the no-load group was also set to $N = 100$, this granted 90% power to detect a small-to-medium difference of $d_s = .46$ between the groups, which we thought was reasonable for a study of these characteristics. To put this figure in context, note that with the average sample size of the experiments included in Fig. 2 ($N = 19.83$), those experiments only achieve 90% power to detect effect sizes larger than $d_s = 1.05$.

In other words, our planned sample size would allow us to detect effects less than half the size of those detectable with the usual samples.

Participants were Psychology students at Autonomous University of Madrid (UAM), and they participated in exchange for course credit. They performed the experimental task in groups of 4–6 participants in a laboratory with individual cubicles. All participants provided informed consent. The UAM ethics committee approved the experimental protocol.

### 1.1.2. *Stimuli*

1.1.2.1. VISUAL SEARCH TASK. Each visual search display comprised 15 L-shaped distractors rotated 0°, 90°, 180° or 270° and a T-shaped target rotated clockwise or anticlockwise by 90°. All the figures were coloured in blue, red, green or yellow. The stimuli were presented on a grey background and positioned in the cells of a 12 × 12 grid invisible to participants.

1.1.2.2. WM TASK. For the load group, four black dots were presented sequentially on a grey background in the cells of a 4 × 4 grid invisible to participants.

### 1.1.3. *Procedure and design*

For the load group, the task consisted of 3 phases: (1) a training phase, in which participants performed the visual search task in combination with a concurrent visuospatial WM task; (2) a practice phase, in which they completed practice trials to become familiar with the single-task condition; and (3) a test phase, in which participants only performed the visual search task.

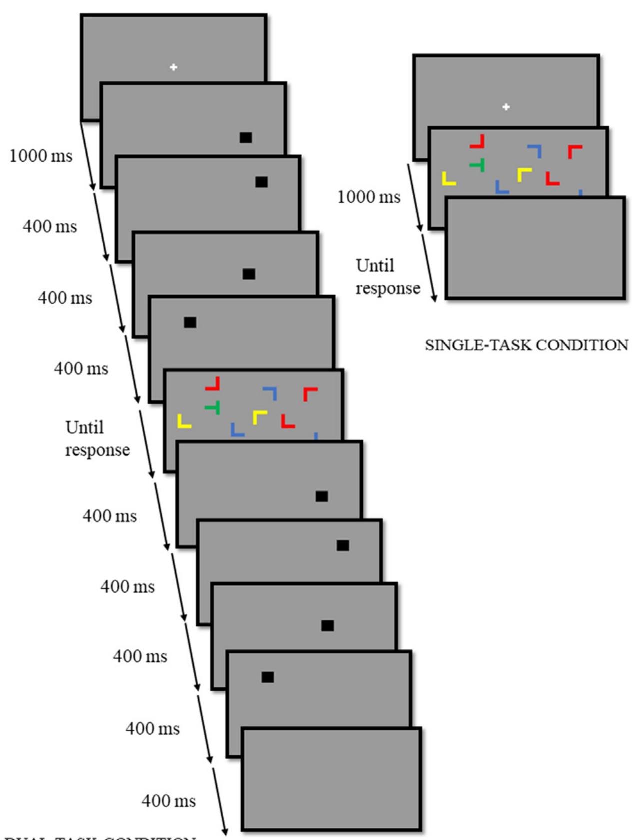


**Fig. 3 – Sequence of events in single-task (right panel) and dual-task (left panel) trials in Experiments 1 and 2.**

Each trial in the training phase included a search task and a WM task (see Fig. 3). Each trial began with a fixation cross lasting 1000 msec. Then, a sequence of four dots was presented on the screen. Each dot was displayed for 400 msec and followed by a 100 msec blank screen. Participants were asked to remember that sequence until the end of the trial. In the visual search task, a search display was presented on the screen until the target response was made. Participants were instructed to respond to the orientation of the stem of the T-shaped target as fast as possible pressing the <left arrow> key if the stem of the T pointed to the left or the <right arrow > key if it pointed to the right. Negative feedback (the word 'Error') was presented on the screen for 800 msec if they made an error. When the search task was completed, the four dots of the sequence presented previously appeared on the screen in the same or different order than the first time. These probe displays were identical to the initial WM encoding displays on half of the trials and, in the remaining half of the trials, differed from the encoding display in that two dots swapped their positions in the sequence. When the sequence was completed, a blank screen was presented until the participant responded to indicate whether the sequence was the same as or different from the initial sequence by pressing the <a> or <z> keys. As in the visual search task, incorrect responses were followed by an 800 msec error message. The intertrial interval was 1000 msec.

The training phase consisted of 16 blocks, each comprising 24 trials. This training length is similar to other Type-1 experiments in Fig. 2 (e.g., Annac et al., 2013, Experiment 2D; Manginelli et al., 2013b, Experiment 1). Each block contained 12 search displays that were generated at the beginning of the task and belonged to the *repeated* condition throughout the entire experiment. The orientation of the target was chosen randomly on each trial, so that any facilitation in responding was a location-based effect rather than orientation-based. The remaining 12 displays in each block were generated randomly in each block (*novel* condition) and appeared only once throughout the experiment. At the beginning of the experiment, 12 locations roughly equidistant from the screen centre were preselected to contain the target. These locations were used both for repeated and novel search displays.

Trials in the practice and test phases only included the visual search task. The practice phase consisted of two blocks, each with 12 novel search displays. This amount of practice is consistent with the number of practice trials that other studies have included at the beginning of the experiment to familiarize participants with the task (see, e.g., Manginelli et al., 2013b; Travis et al., 2013). The test phase comprised six blocks, each with 36 trials. At the beginning of the test phase, 12 new search displays were generated for the *control* condition. These displays were repeated once per block throughout the test phase. Therefore, each block of the test phase contained 12 repeated, 12 novel and 12 control trials. The length of the test stage is slightly longer than in other Type-1 studies included in Fig. 2 (e.g., Annac et al., 2013; Manginelli et al., 2013b). This is due to the fact that the present study included 12 control displays per block. This feature of the design is important, because comparing the facilitation to repeated versus control displays allowed us to determine if contextual cueing was acquired during the test phase, instead of (or in addition to) during the training phase. Participants could take a break after every 100 trials.

The procedure was identical for the no-load group in all respects, except that participants performed all the stages under single task conditions.

### 1.2. Results and discussion

#### 1.2.1. Data preprocessing

Due to a computer problem, four participants failed to complete the task and their data were removed from the analyses. Participants with accuracies below 95% in the search task or 3 standard deviations (*SD*) below the group mean in the WM task were also removed from the analyses. Two participants failed to meet these selection criteria. This resulted in a final sample of 96 (84 female) and 98 (81 female) participants in the load and no-load groups, respectively, with a mean age of 19.6 ($SD = 1.96$). For the analyses of reaction times (RT) in the search task, only trials with correct responses were considered. RTs from trials immediately following a rest break and RTs longer than 10 sec in the search task were also removed. Then, for each participant we calculated the mean and *SD* of RTs and removed trials three or more *SD*s higher or lower than the mean. To reduce noise in the data, adjacent blocks were binned in two-block epochs.

#### 1.2.2. Analytic approach

All the inferential analyses described below were conducted twice, using frequentist and Bayesian statistics. For Bayesian *t*-tests, we report Bayes factors ($BF_{10}$) comparing the evidence in favour of a two-sided alternative hypothesis (modelled as a Cauchy distribution scaled at .707) over the null hypothesis of no difference. For Bayesian analyses of variance, we report the model showing the best fit to the data and then Bayes factors comparing the fit of the remaining models to the best-fitting model. In all cases, *BF*s between 1/3 and 3 are considered anecdotal evidence, *BF*s between 3 and 10 or between 1/3 and 1/10 are considered moderate evidence, and *BF*s larger than 10 or lower than 1/10 are considered strong evidence.

#### 1.2.3. WM task

On average, participants in the load group responded correctly to the WM task in 84.16% of the trials, 95% CI [82.52%, 85.78%]. This level of performance was significantly greater than chance, $t(95) = 41.55$, $p < .001$, $d_z = 4.24$, $BF_{10} = 1.17 \times 10^{59}$.

#### 1.2.4. Search task: training stage

Fig. 4 depicts mean RTs over the course of the experiment. RTs from the training phase were analysed with a 2 (Group: load *vs* no-load) × 2 (Context: repeated *vs* novel) × 8 (Epochs) analysis of variance (ANOVA). The main effects of Group, $F(1, 192) = 12.93$, $p < .001$, $\eta_p^2 = .06$, Context, $F(1, 192) = 131.89$, $p < .001$, $\eta_p^2 = .41$, and Epoch, $F(7, 1344) = 208.49$, $p < .001$, $\eta_p^2 = .52$, were statistically significant. The Group × Epoch, $F(7, 1344) = 10.22$, $p < .001$, $\eta_p^2 = .05$, Context × Epoch, $F(7, 1344) = 10.90$, $p < .001$, $\eta_p^2 = .05$, and Group × Context × Epoch, $F(7, 1344) = 2.71$, $p = .009$, $\eta_p^2 = .01$, interactions were also significant. Only the Group × Context interaction failed to reach statistical significance, $F(1, 192) = 1.29$, $p = .256$, $\eta_p^2 = .006$. In a Bayesian ANOVA with the same design and dependent variable, the best-fitting model included the three main effects and the Group × Epoch and Context × Epoch interactions, with a *BF* of 11.76 over the next best-fitting model.

To follow up on the significant Group × Context × Epoch interaction, we conducted separate Context × Epoch ANOVAs on each group. In the load group, the main effects of Context, $F(1, 95) = 95.23$, $p < .001$, $\eta_p^2 = .50$, Epoch, $F(7, 665) = 104.59$,

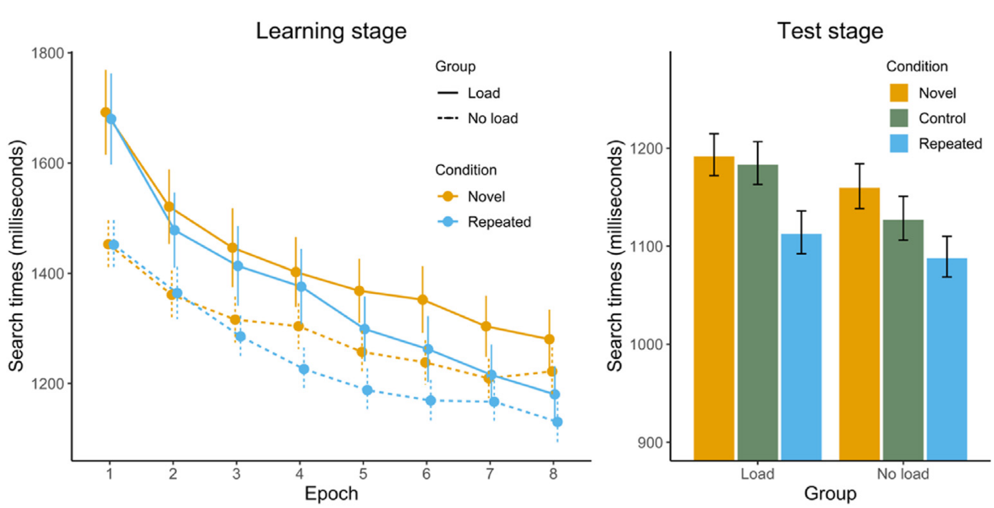


**Fig. 4 – Search times during the training and test stages in Experiment 1. Error bars denote 95% confidence intervals.**

$p < .001$, $\eta_p^2 = .52$, and the Context × Epoch interaction, $F(7, 665) = 5.93$, $p < .001$, $\eta_p^2 = .06$, were significant. In the Bayesian ANOVA, the best model included both main effects but not the interaction, with a *BF* of 7.39 over the next model. Similarly, in the no-load group, both the main effect of Context, $F(1, 97) = 46.28$, $p < .001$, $\eta_p^2 = .32$, and Epoch, $F(7, 679) = 118.41$, $p < .001$, $\eta_p^2 = .55$, reached statistical significance, and so did their interaction, $F(7, 679) = 7.80$, $p < .001$, $\eta_p^2 = .07$. The best fitting model in the Bayesian ANOVA included all three factors, with a *BF* of 650.56 over the next model.

Overall, these results suggest that contextual cueing was established successfully in both groups of participants and that the size of the effect was roughly similar in both cases, although visual search was generally slower for participants who were performing a concurrent WM task. In the Supplementary Material we report the results of (non-registered) exploratory analyses comparing the evolution of contextual cueing in the load and no-load groups. All the analyses converge to the conclusion that the WM task had no noticeable effect on the size of contextual cueing.

#### 1.2.5. *Search task: test stage*

RTs from the test phase were analysed with a 2 (Group: load *vs* no-load) × 3 (Context: repeated *vs* novel *vs* control) × 3 (Epochs) ANOVA. Only the main effects of Context, $F(2, 384) = 76.81$, $p < .001$, $\eta_p^2 = .29$, and Epoch, $F(2, 384) = 17.34$, $p < .001$, $\eta_p^2 = .08$, and the Group × Context interaction, $F(2, 384) = 3.43$, $p = .033$, $\eta_p^2 = .02$ were statistically significant. In the Bayesian ANOVA, the best-fitting model included only the main effects of Context and Epoch, with a *BF* of 1.41 over the model that also included the main effect of Group, a *BF* of 2.26 over the model that included the three main effects and the Group × Context interaction, and *BFs* > 9.69 over the remaining models.

A separate Context × Epoch ANOVA taking only data from the load group revealed significant main effects of Context, $F(2, 190) = 51.79$, $p < .001$, $\eta_p^2 = .35$, and Epoch, $F(2, 190) = 7.61$, $p < .001$, $\eta_p^2 = .07$, but no interaction, $F(4, 380) = .28$, $p = .892$, $\eta_p^2 = .003$. In the Bayesian ANOVA, the model with just the two main effects outperformed all other models, with a *BF* of 186.40 over the next best-fitting model. For the no-load group, again, the main effects of Context, $F(2, 194) = 30.30$, $p < .001$, $\eta_p^2 = .24$, and Epoch, $F(2, 194) = 11.59$, $p < .001$, $\eta_p^2 = .11$, reached statistical significance, but the interaction was non-significant, $F(4, 388) = .78$, $p = .540$, $\eta_p^2 = .008$. Similarly, the model with the two main effects outperformed the remaining models, with *BFs* > 118.65.

Pairwise comparisons showed that, in the load group, RTs were faster for repeated displays than for control displays, $t(95) = 7.44$, $p < .001$, $d_z = .76$, $BF_{10} = 1.97 \times 10^8$, but responses to control displays were not significantly faster than responses to novel displays, $t(95) = 1.15$, $p = .254$, $d_z = .11$, $BF_{01} = 4.69$. In the no-load group, RTs were also faster for repeated than for control displays, $t(97) = 4.16$, $p < .001$, $d_z = .42$, $BF_{10} = 258.89$, and they were also faster for control displays than for novel displays, $t(97) = 4.07$, $p < .001$, $d_z = .41$, $BF_{10} = 189.79$.

A non-registered 2 (Group: load *vs* no-load) × 2 (Context: repeated *vs* control) ANOVA yielded a significant main effect of Context, $F(1, 192) = 67.76$, $p < .001$, $\eta_p^2 = .26$, and a significant interaction, $F(1, 192) = 5.95$, $p = .016$, $\eta_p^2 = .03$, but no main effect of Group, $F(1, 192) = 2.98$, $p = .086$, $\eta_p^2 = .02$. The full model outperformed all other models, with a *BF* of 2.29 over the model including just the two main effects, a *BF* of 2.36 over the model including just the main effect of Context, and a *BF* of $2 \times 10^{12}$ over the model including just the main effect of Group. The significant interaction suggests that the size of contextual cueing, defined as the difference between repeated and control conditions, was, if anything, larger for the load group.

We also conducted a second non-registered Group × Context ANOVA comparing RTs to control and novel displays. Again the main effect of Context, $F(1, 192) = 14.38$, $p < .001$, $\eta_p^2 = .97$, and the interaction, $F(1, 192) = 5.14$, $p = .025$, $\eta_p^2 = .03$, were significant. The main effect of Group approached but failed to reach full statistical significance, $F(1, 192) = 3.40$, $p = .067$, $\eta_p^2 = .02$. The full model outperformed all other models, with a *BF* of 1.76 over the model including just the two main effects, a *BF* of 1.87 over the model including only the main effect of Context, and a *BF* of 127.92 over the model including only the main effect of Group. These results suggest that, unlike participants in the load group, participants in the no-load group developed some contextual cueing for control displays during the test stage.

#### 1.2.6. *Search task: training versus test stage*

A key prediction of the latent learning hypothesis is that the size of contextual cueing should increase from the training phase to the test phase for the load group but not for the no-load group. To put this hypothesis to the test, we computed for each participant the size of contextual cueing (i.e., RTs to novel displays minus RTs to repeated displays) during the six blocks of the test phase and during the last six blocks of the training phase. Then we submitted these contextual cueing scores to a 2 (Group: load *vs* no-load) × 2 (Stage: training *vs* test) ANOVA. The main effect of Group approached but failed to reach statistical significance, $F(1, 192) = 2.92$, $p = .089$, $\eta_p^2 = .01$. The main effect of Stage, and the interaction were far from significance, largest $F(1, 192) = 1.42$. The model including only the main effect of Group outperformed all other models, with *BFs* > 4.

Paired-samples *t*-tests comparing the size of cueing in both stages, conducted separately for each group, showed that contextual cueing was not significantly larger in the test stage than in the training stage for either the load group, $t(95) = 1.23$, $p = .221$, $d_z = .12$, $BF_{01} = 4.26$, or the no-load group, $t(97) = -.41$, $p = .682$, $d_z = .04$, $BF_{01} = 8.24$.

#### 1.2.7. *Sensitivity analysis*

Following the preregistered protocol, to check that results are not biased due to performance in the WM task, we repeated the previous analyses eliminating trials with incorrect responses in the WM task. The results were qualitatively similar to those reported above, leading only to small numerical differences that never affected the pattern of statistical significance or the preferred models in Bayesian analyses.

## 2. Experiment 2

The aim of Experiment 2 was to replicate a 'Type-2' design (see the right panel of Fig. 1). As in Experiment 1, there was a training, practice and test phase. There was only one group of participants in Experiment 2, which was equivalent to the load group in Experiment 1, except for the time of administration of the WM task. Whereas in Experiment 1 the WM task was administered only in the training phase, in Experiment 2 it was present only in the practice and test phases. Except for this change in the design, all the methodological aspects of Experiment 2 were identical to those of the load group in Experiment 1. The main objective of Experiment 2 was to check whether any contextual learning acquired during the training phase under single task conditions could be expressed when the search task is combined with a demanding visuospatial WM task at test.

### 2.1. Method

#### 2.1.1. Participants, apparatus, stimuli, procedure and design

The participant population, apparatus, and stimuli, were exactly the same as in the load group of Experiment 1, though none of the participants from Experiment 1 took part in Experiment 2. The procedure and design were also exactly as in the load group of Experiment 1, with the only exception that there was a concurrent WM task in the practice and test phases, but not in the training phase.

### 2.2. Results and discussion

#### 2.2.1. Data pre-processing

One participant could not complete the experimental task due to a computer failure. Data were pre-processed following the same procedure described in Experiment 1, resulting in the exclusion of three additional participants. Therefore, the final sample included 96 participants (76 female), with a mean age of 19.4 ($SD = 1.59$).

#### 2.2.2. WM task

On average, participants responded correctly to the WM task in 88.16% of the trials, 95% CI [86.90%, 89.42%]. This level of performance was significantly greater than chance, $t(95) = 60.07$, $p < .001$, $d_z = 6.13$, $BF_{10} = 3.51 \times 10^{73}$.

#### 2.2.3. Search task: training stage

Fig. 5 shows the evolution of average RTs over the course of Experiment 2. RTs from the training phase were analysed with a 2 (Context: repeated *vs* novel) × 8 (Epochs) ANOVA. The main effects of Context, $F(1, 95) = 56.80$, $p < .001$, $\eta_p^2 = .37$, and Epoch, $F(7, 665) = 85.08$, $p < .001$, $\eta_p^2 = .47$, and their interaction, $F(7, 665) = 3.96$, $p < .001$, $\eta_p^2 = .04$, were statistically significant. In the Bayesian ANOVA, the best-fitting model included only main effects of Context and Epoch, with a *BF* of 4.89 over the next model, which also included their interaction.

#### 2.2.4. Search task: test stage

RTs from the test phase were also analysed with a 3 (Context: repeated *vs* novel *vs* control) × 3 (Epoch) ANOVA. The main effects of Context, $F(2, 190) = 44.55$, $p < .001$, $\eta_p^2 = .32$, and Epoch, $F(2, 190) = 13.33$, $p < .001$, $\eta_p^2 = .12$, were significant, but their interaction was not, $F(4, 380) = .99$, $p = .414$, $\eta_p^2 = .01$. The model including just the two main effects achieved the best performance, with a *BF* of 87.92 over the model including also the interaction. Collapsing across epochs, RTs to the repeated displays were significantly faster than RTs to control displays, $t(95) = 7.59$, $p < .001$, $d_z = .77$, $BF_{10} = 4 \times 10^9$, but the latter were not significantly faster than RTs to novel displays, $t(95) = 1.31$, $p = .195$, $d_z = .13$, $BF_{01} = 3.90$.

#### 2.2.5. Search task: training versus test stage

The contextual-cueing effect, defined as the difference in RTs to repeated displays and novel displays, was not significantly larger at the end of the training stage than during the test stage, $t(95) = -.28$, $p = .777$, $d_z = -.02$, $BF_{01} = 8.52$.

#### 2.2.6. Sensitivity analysis

As in Experiment 1, the pattern of results reported above was unaffected by the exclusion of trials with incorrect responses in the WM task.

## 3. General discussion

The aim of the present Registered Report was to investigate the role of WM in the acquisition and expression of contextual cueing in two high-powered experiments. Although this issue has been extensively investigated in the past (Annac et al., 2013; Chen et al., 2019; Manginelli et al., 2013b; Travis et al., 2013), previous studies have yielded inconsistent conclusions. A popular hypothesis arising from this literature is that contextual cueing can be learned independently from WM resources, but WM load can nevertheless prevent the expression of this learning in visual search performance. The two experiments reported in the present article were designed as a purely confirmatory, pre-registered test of this hypothesis. To this aim, both experiments combined a demanding secondary visuospatial WM task with a visual search task either in the training stage (Experiment 1) or in the test stage (Experiment 2) to test whether the expression or acquisition of contextual cueing, respectively, is hindered by WM load. Overall, our findings show that the reduction in the availability of WM resources did not make a noticeable difference in the acquisition or expression of contextual cueing.

In Experiment 1, we tried to replicate the ideal pattern of results of a typical Type-1 experiment (see Fig. 1). An experimental group of participants (the load group) performed the visual search task concurrently with a secondary task in the training phase, but not in the test phase, whereas a control (no-load) group performed only the visual search task in the training and test phases. A robust contextual cueing effect was observed in the load group both in the training and test stages, without any meaningful difference compared to the no-load group. These results contrast with those of other Type-1 experiments that found a diminished contextual

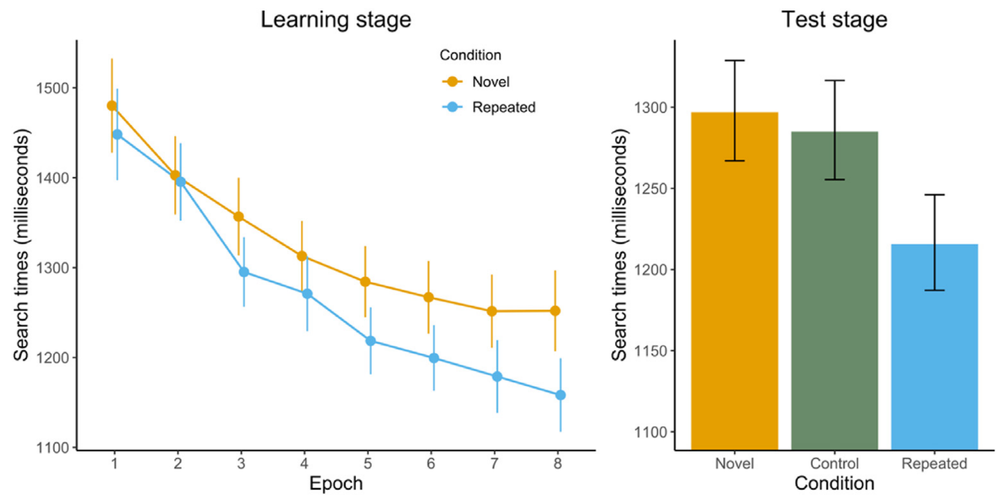


**Fig. 5 – Search times during the training and test stages in Experiment 2. Error bars denote 95% confidence intervals.**

cueing effect in the training phase but not in the test phase (Annac et al., 2013, Experiment 2D). On the basis of such results it has been argued that a secondary task may interfere with the expression, but not with the acquisition of this implicit learning. Our results do not support this conclusion and are, instead, consistent with the findings reported by Manginelli et al. (2012, Experiment 1).

The results of Experiment 1 also contrast dramatically with those of Chen et al. (2019, Experiment 3). Contrary to us, their results showed a complete absence of contextual cueing both in the training and test phases, suggesting that learning did not occur at all. Those discrepancies may be explained by some differences in the design of the experiments. The clearest one is possibly the type of secondary task used in each study. Whereas our participants performed a visuospatial task, Chen et al. used a numerical executive task. It is tempting to think that perhaps our secondary task was simply not difficult enough to impair contextual learning. However, a closer look at the results suggests that overall difficulty is probably not the main variable explaining the discrepancy between the two studies. If anything, accuracy in the WM task was slightly lower in our Experiment 1 (84.2%) than in the conditions where Chen et al. found an interference with contextual cueing in the 4 items condition (88.3%), but not in the 3 items condition (92.3%), suggesting that our WM task was not less demanding. Most likely, the discrepancy is due to differences in the nature of the information held in WM (numerical *vs* visuospatial) or in the operations performed on that information (simple storage *vs* active manipulation). More research about this topic is needed before any firm conclusion can be drawn.

The pattern of results observed in Experiment 2 led to similar conclusions. In this case, the secondary task was performed during the test, but not in the training stage, following what we have called a Type-2 design. The aim of the experiment was to investigate if adding a demanding secondary task in the test phase would interfere with the expression of the learning acquired during the training phase. Again, our results failed to support this hypothesis: we observed a strong contextual cueing effect in the test phase, even when a demanding and concurrent visuospatial WM task was performed concurrently with the visual search task. Consistent with the results of Experiment 1, this pattern of results also fails to support the latent learning hypothesis. WM load does not seem to hinder the expression of an already established contextual cueing effect.

This result stands in contrast with those reported by Manginelli et al. (2013b, Experiment 3) and Annac et al. (2013, Experiment 2C), who found a significant reduction in contextual cueing after adding a secondary task at test. The discrepancy between our results and those of Manginelli et al. and Annac et al. may be explained in different ways. Firstly, there may be crucial differences among the experimental designs. Whereas our participants completed a practice stage to familiarize themselves with the dual task before starting the test stage, this was not the case for previous experiments, which either did not include practice trials at all or included too few of them to make a meaningful difference. Furthermore, because our testing stage included three experimental conditions (novel, control, and repeated), the total number of trials was slightly larger than in previous experiments using a Type-2 design. In other words, our experiment provided participants with many opportunities to adapt their performance to dual-task conditions. Perhaps by providing additional practice with the dual task, we actually eliminated the crucial factor that was responsible for the deleterious effects of WM-load in previous experiments. From this point of view, the absence of contextual cueing in the test phase of Annac et al. and Manginelli et al. could perhaps be entirely due the mere lack of familiarity with the secondary task. This might suggest that it is not WM load itself with interferes with (either the learning or expression of) contextual cueing, but participants' transient inability to perform the WM and visual-search tasks concurrently. Note, however, that additional (non-registered) analyses reported in the Supplementary Material show that contextual cueing was already robust in the first block of the test stage and did not become significantly larger in subsequent blocks, suggesting that it did not depend on extensive experience with the WM task.

Another possible explanation is statistical power. Although the studies by Annac et al. (2013) and Manginelli et al. (2013b) were sufficiently powered to detect the basic contextual cueing effect under standard conditions (i.e., without a concurrent working-memory task), we hypothesized that these studies might have been underpowered to detect contextual cueing when the visual search task is combined with a WM task or to detect any difference in the size of contextual cueing with and without WM load. In fact, Bayesian analyses reported in the Introduction show that the diminution of the contextual cueing effect from the training to the test phase was weak. However, with our results at hand, the hypothesis that the original studies by Annac et al. and Manginelli et al. were underpowered seems unlikely. In fact, in our Experiment 1 the size of the contextual cueing effect did not differ between the load and no-load groups. If anything, it was slightly larger for the load group. And in Experiment 2, contextual cueing did not decrease significantly from the end of training to the test stage. With sample sizes of 17 and 35, the power of the experiments by Annac et al. (2013, Experiment 2C) and Manginelli et al. (2013b, Experiment 3) to detect an effect like the one found in the test stage of our Experiment 2, i.e., $d_z = .77$, in a two-tailed $t$-test with $\alpha = .05$, is equal to .85 and .99, respectively.

In general terms, we find no influence of WM in the acquisition or the expression of contextual cueing. This result dovetails with the idea that this type of learning can take place even when cognitive resources are scarce and is consistent with the common assumption that contextual cueing is based on automatic mental processes. Alternatively, it is possible that the visuospatial WM task used in the present series of experiments simply fails to exhaust participants' cognitive resources, allowing successful performance in the visual search task. This is, in fact, a possible explanation for the variety of results found in previous studies. Individual differences in visuospatial WM capacity may lead to varying results in similar experimental designs. Overall accuracy in the WM task was relatively good (84.2 and 88.2% in Experiments 1 and 2), suggesting that most participants have no problem to remember a four-item sequence while they perform the visual search task. This amount of difficulty might only exhaust the visuospatial WM span of some of participants. In studies with few participants, sampling variance in the proportion of participants who struggle with this version of the task could make a difference on whether the secondary task interferes with contextual cueing or not. This might explain the contradictory pattern of results in previous studies.

It is important to note, though, that in a series of exploratory analyses, presented in the Supplementary Material, we failed to find any significant correlation between participants' performance in the WM task and the size of contextual cueing. But given the low reliability of these dependent measures, the absence of significant correlations is not particularly surprising or informative (see Table S1 in the Supplementary Material and Vadillo et al., 2022). In any case, an interesting idea for future research is to adjust the difficulty of the secondary task to the WM span of each subject to ensure that all participants perform the visual search task under similar levels of WM load or, alternatively, implement changes in the secondary WM-task aimed at draining participants' WM resources more effectively. Interestingly, previous experiments have sometimes combined visuospatial WM tasks like the one employed in the present experiments with articulatory suppression. While at present there is little evidence that experiments using articulatory suppression (e.g., Manginelli et al., 2013b, Experiment 1) find a stronger impact of WM-load on contextual cueing than otherwise similar experiments without articulatory suppression (e.g., Manginelli et al., 2012, Experiment 1; Travis et al., 2013), it is not impossible that by combining different types of WM tasks future experiments will be more effective at exhausting participants cognitive resources. In any case, retrospectively, we suspect that future studies would benefit from including a manipulation check ensuring that the secondary task drained participants' WM-resources as expected. In our experiments, we can be sure that WM load made a noticeable difference in participants' performance, because their reaction times became substantially longer under dual-task conditions, but this does not necessarily mean that the manipulation was as powerful as intended.

In conclusion, our study failed to replicate the ideal result of Type-1 and -2 experiments, and does not support the hypothesis that WM load impairs the expression but not the acquisition of contextual learning. Conversely, our results suggest that the acquisition and expression of contextual cueing do not rely on visuospatial WM resources or, at least, that the secondary task used in this experiment is not sufficiently demanding to make a difference.

## CRediT authorship contribution statement

**Francisco Vicente-Conesa**: Conceptualization, Investigation, Methodology, Roles/Writing - original draft, Writing - review & editing.

**Tamara Giménez-Fernández**: Conceptualization, Investigation, Methodology, Roles/Writing - original draft, Writing - review & editing.

**David R. Shanks**: Conceptualization, Funding acquisition, Investigation, Methodology, Roles/Writing - original draft, Writing - review & editing.

**Miguel A. Vadillo**: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Roles/Writing - original draft, Writing - review & editing.

## Author note

FVC, TGF, and MV were supported by grants PSI2017-85159-P and PID2020-118583 GB-I00 from Agencia Estatal de Investigación and FEDER, UE. MAV and TGF were supported by grants 2016-T1/SOC-1395 and 2020-5A/SOC-19723 from Comunidad de Madrid (Programa de Atracción de Talento Investigador). DS and MAV were supported by grant ES/P009522/1 from the Economic and Social Research Council, United Kingdom. FVC was supported by a PhD fellowship (PRE2018-085148) from Ministerio de Ciencia e Innovación.

Correspondence concerning this article should be addressed to Miguel A. Vadillo, Departamento de Psicología Básica, Facultad de Psicología, Universidad Autónoma de Madrid, 28049 Madrid, Spain. E-mail: miguel.vadillo@uam.esReferences.

## Open practices

The experiments in this article earned Open Data, Open Materials and Preregistered badges for transparent practices. Materials and data for the study are available at https://osf.io/97f5d/. The preregistered protocol is available at https://osf.io/8ypxt.

## Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2022.05.019.

## REFERENCES

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

Chen, M., Wang, C., Sclodnick, B., Zhao, G., & Liu, X. (2019). Executive working memory involved in the learning of contextual cueing effect. *Experimental Brain Research, 237*, 3059–3070. https://doi.org/10.1007/s00221-019-05643-7

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, 36, 28–71. https://doi.org/10.1006/cogp.1998.0681

Chun, M. M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 29, 224–234. https://doi.org/10.1037/0278-7393.29.2.224

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

Geyer, T., Rostami, P., Sogerer, L., Schlagbauer, B., & Müller, H. J. (2020). Task-based memory systems in contextual-cueing of visual search and explicit recognition. *Scientific Reports, 10*, Article 16527. https://doi.org/10.1038/s41598-020-71632-4

Goujon, A., Didierjean, A., & Thorpe, S. (2015). Investigating implicit statistical learning mechanisms through contextual cueing. *Trends in Cognitive Sciences, 19*, 524–533. https://doi.org/10.1016/j.tics.2015.07.009

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

Manginelli, A. A., Baumgartner, F., & Pollmann, S. (2013a). Dorsal and ventral working memory-related brain areas support distinct processes in contextual cueing. *Neuroimage, 67*, 363–374. https://doi.org/10.1016/j.neuroimage.2012.11.025

Manginelli, A. A., Geringswald, F., & Pollmann, S. (2012). Visual search facilitation in repeated displays depends on visuospatial working memory. *Experimental Psychology*, 59, 47–54. https://doi.org/10.1027/1618-3169/a000125

Manginelli, A. A., Langer, N., Klose, D., & Pollmann, S. (2013b). Contextual cueing under working memory load: Selective interference of visuospatial load with expression of learning. *Attention, Perception, and Psychophysics, 75*, 1103–1117. https://doi.org/10.3758/s13414-013-0466-5

Niv, Y. (2021). The primacy of behavioral research for understanding the brain. *Behavioral Neuroscience*, 135, 601–609. https://doi.org/10.1037/bne0000471

Pollmann, S. (2019). Working memory dependence of spatial contextual cueing for visual search. *British Journal of Psychology, 110*, 372–380. https://doi.org/10.1111/bjop.12311

Sisk, C. A., Remington, R. W., & Jiang, Y. V. (2019). Mechanisms of contextual cueing: A tutorial review. *Attention, Perception, & Psychophysics, 81*, 2571–2589. https://doi.org/10.3758/s13414-019-01832-2

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

Travis, S. L., Mattingley, J. B., & Dux, P. E. (2013). On the role of working memory in spatial contextual cueing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 208–219. https://doi.org/10.1037/a0028644

Vadillo, M. A., Giménez-Fernández, T., Aivar, M. P., & Cubillas, C. P. (2020). Ignored visual context does not induce latent learning. *Psychonomic Bulletin & Review, 27*, 512–519. https://doi.org/10.3758/s13423-020-01722-x

Vadillo, M. A., Giménez-Fernández, T., Beesley, T., Shanks, D. R., & Luque, D. (2021). There is more to contextual cuing than meets the eye: Improving visual search without attentional guidance toward predictable target locations. *Journal of Experimental Psychology: Human Perception and Performance, 47*, 116–120. https://doi.org/10.1037/xhp0000780

Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review, 23*, 87–102. https://doi.org/10.3758/s13423-015-0892-6

Vadillo, M. A., Malejka, S., Lee, D. Y., Dienes, Z., & Shanks, D. R. (2022). Raising awareness about measurement error in research on unconscious mental processes. *Psychonomic Bulletin & Review, 29*, 21–43. https://doi.org/10.3758/s13423-021-01923-y

Vickery, T. J., Sussman, R. S., & Jiang, Y. V. (2010). Spatial context learning survives interference from working memory load. *Journal of Experimental Psychology: Human Perception & Performance, 36*, 1358–1371. https://doi.org/10.1037/a0020558