



## Special Issue “The Brain’s Brake”: Research Report

# Statistical learning in the absence of explicit top-down attention



Dock Duncan <sup>a,b,\*</sup> and Jan Theeuwes <sup>a,b</sup>

<sup>a</sup> Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

<sup>b</sup> Institute Brain and Behavior Amsterdam (iBBA), the Netherlands

## ARTICLE INFO

## Article history:

Received 7 April 2020

Reviewed 11 May 2020

Revised 21 May 2020

Accepted 10 July 2020

Published online 27 July 2020

## Keywords:

Attention

Statistical learning

Distractor suppression

Top-down/bottom-up

Habituation

## ABSTRACT

Recently it has been shown that statistical learning of regularities presented in a display can bias attentional selection, such that attentional capture by salient objects is reduced by suppressing the location where these distractors are likely to appear. The role of attention in learning these contingencies is not immediately clear. Specifically, it is not known whether attention needs to be directed to the contingencies present in the display for learning to occur. In the current study we investigated whether participants can learn statistical regularities present in the environment even when these regularities are not relevant for the participant and are not part of their top-down goals. We used the additional singleton paradigm in which a color singleton was presented much more often in one location than in all other locations. We show that after being exposed to these regularities regarding the location of the color singleton during an unrelated task in which there are no targets nor distractors, participants showed a suppression effect from the previously learned contingencies when switching to a task in which they search for a target and suppress a distractor. We conclude that visual statistical learning can occur in the absence of top-down attention.

© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

When we look upon a scene, the first things we are consciously aware of are the things to which we attend (Posner, 1980). In an everyday example that may be a billboard, a street sign or someone running across the street; a case which highlights how understanding attention is essential for designing the sort of environment that ensures the correct target will be selected by attention, thereby avoiding potentially dangerous situations.

Traditionally, attentional selection is considered to be the result of the interaction between the goals of the observers (the top-down attentional set) and the physical properties of the environment (the saliency of objects) (Egeth & Yantis, 1997; Itti & Koch, 2001; Theeuwes, 2010, 2018). On one end of the spectrum, it has been shown that objects that stand out from the environment are selected regardless of the task set. For example, in the so-called additional singleton paradigm (Theeuwes, 1991, 1992) participants need to search for a unique shape (a diamond among circles or a circle among diamonds)

\* Corresponding author. Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands.

E-mail address: [d.h.duncan@vu.nl](mailto:d.h.duncan@vu.nl) (D. Duncan).

<https://doi.org/10.1016/j.cortex.2020.07.006>

0010-9452/© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

while on some trials a unique color singleton distractor is present (a red item among green items). Typically, participants are slower when a unique color singleton is present than when it is absent indicating that the color singleton captured attention in a bottom-up, exogenous way. On the other end of the spectrum, it has been claimed that only objects that are in line with the goals and intentions of the observer are selected (Bacon & Egeth, 1994; Folk, Remington & Johnson, 1992; Leber & Egeth, 2006). The idea being that an observer's attention only gets captured by a particular stimulus feature if it matches their top-down attentional set (e.g., a red singleton captures attention when the observer is looking for red items). In other words, when a top-down set is adopted, only elements that fit that top-down set receive any attentional priority.

Recently, it was argued that selection is often neither controlled by top-down nor by bottom-up factors but instead is driven by lingering biases of previous selection episodes (Awh, Belopolsky, & Theeuwes, 2012; Theeuwes, 2019; 2018). It was argued that through statistical learning (SL) observers extract regularities present in the environment which in turn biases attentional selection (Theeuwes, 2018). SL is considered to be a mechanism that enables the extraction of distributional properties from sensory input across time and space (Frost, Armstrong, Siegelman, & Christiansen, 2015). The learning of these statistical regularities makes it possible to interact more effectively with the visual world (Chun & Jiang, 1998). Furthermore, it is generally assumed that the adaptation to regularities that exist in world evolves largely unconscious,<sup>1</sup> and without the intention to learn nor awareness of learning (e.g., Fiser & Aslin, 2001; Turk-Browne, Jungé, & Scholl, 2005).

Given the view that SL is automatic and occurs without intention to learn, one question that has been heavily debated is whether people learn to extract these regularities from the environment even if the regularities are completely irrelevant for the task at hand. In other words, do people learn to extract regularities by simply looking at displays that have these regularities in them? In terms of attentional control, can learning take place even if extracting these regularities is not part of the task relevant top-down set?

Several studies have investigated the question of whether attention is needed for statistical learning to occur. Using auditory stimuli, some have shown that attention improves SL while others have shown that without attention there is no learning. For example, it was shown that if there is less attention available, auditory SL is negatively impacted (Toro, Sinnott, & Soto-Faraco, 2005). Also, instructions to attend to one auditory pattern (i.e., words) improved learning, at the expense of learning other features of the word stream (such as the grammar; see Finn, Lee, Kraus, & Hudson Kam, 2014). On the other hand, there are also studies showing that the role of attention in SL is limited. In a study by Saffran, Newport, Aslin, Tunick, and Barrueco (1997) children and adults had to listen to unsegmented artificial language while performing a cover task of creating computer illustrations. Participants did not know they were listening to an artificial language. Critically, even when they performed the additional cover

task, both adults and children learned this artificial language equally well. More recently, Batterink and Paller (2019) demonstrated that participants can learn statistical properties of language even if fully engaged in an unrelated and difficult visual n-back task.

While the emerging picture in auditory SL is that learning benefits but does not rely on attention to occur, the picture is far less clear-cut in the field of visual statistical learning (VSL). The seminal work of Turk-Browne et al. (2005) stood as the last word in the question of the role of attention in SL for nearly a decade. In their study, participants viewed sequentially presented streams of non-sense shapes. Participants had to perform a demanding n-back task while they viewed a centrally presented interleaved stream of shapes. Participants were instructed to attend one of the streams (i.e., shapes in the color green) while ignoring the other interleaved stream (i.e., shapes in the color red). The results indicated only learning of the triplet shapes of the attended stream and not of triplet shapes of the unattended stream. It was thus concluded that selective attention determines the input for statistical learning. However, a replication by Musz, Weber, and Thompson-Schill (2015) utilizing a larger participant group found that SL could indeed be observed in the unattended shape group, though the effect was greatly attenuated and possibly obscured by the original studies limited sample size. Further studies by Campbell, Zimmerman, Healey, Lee, and Hasher (2012) and Forest and Finn (2018) add to the growing tide of evidence that VSL may be more similar to auditory SL than was previously thought, where it greatly benefits from the presence of attention, but is not dependent on it to occur.

Several recent studies have provided evidence that VSL biases attentional selection. It has been shown that in visual search, participants learn that the target appears more often in one location than other locations, speeding up target detection (e.g., Ferrante et al., 2018; Geng & Behrmann, 2002, 2005; Jiang, Swallow, Rosenbaum, & Herzog, 2013). Clearly in these cases in which the regularities concern the location of the target, learning may be particularly effective because “looking for targets” is clearly part of the top-down set of the observers.

However recently in a series of experiments, it was demonstrated that participants can also learn regularities in the environment that are not directly part of the top-down task set. Using the additional singleton task, Wang & Theeuwes, (2018a, b, c; also Failing, Wang, & Theeuwes, 2019) showed that participants can learn to suppress a salient singleton that was systematically presented more often in one location than in all other locations. These findings show that participants can learn about regularities present in the displays even when looking for these regularities is not necessarily part of the task set (see also Ferrante et al., 2017; Goschy, Bakos, Müller, & Zehetleitner, 2014 for similar findings using a different paradigm). Indeed, the typical instruction in the additional singleton paradigm is to search as fast as possible for a shape singleton (i.e., a diamond between circles or a circle between diamonds). In order to find the target quickly, participants need to effectively suppress the influence of the

<sup>1</sup> Though any claim of unconscious learning should generally be approached cautiously as highlighted by Vadillo et al. (2019).

salient distractor on search. The Wang & Theeuwes, findings indicate that participants learn the distractor regularities such that they become better able to suppress distractor if these consistently appear more often in one location than in all other locations. Even though strictly speaking, ignoring these distractors is not part of the top-down instruction to observers, ignoring the distractor does in fact improve search and as such suppression of the distractor may at least be considered indirectly relevant for the task and the top-down goals of the participant.

In the current study we wanted to push the limits of whether participants can learn statistical regularities present in the environment even when these regularities are clearly not part of and not relevant to the explicit top-down set. The question we addressed is whether participants can learn regularities even if they are engaged in a task that has nothing to do with visual search.

## 2. Experiment 1

In Experiment 1 we presented displays that were similar to the Wang & Theeuwes, (2018a,b,c) displays in which a color singleton distractor was presented much more often in one location than in all other locations. Yet, instead of participants having to search for a target, they simply had to look at the display and decide whether the elements in the display form a global circle or a global diamond shape (see also Belopolsky & Theeuwes, 2010 who used a similar task). After being exposed to arrays containing the statistical regularity, participants next switched to a regular additional singleton task and searched for a diamond among circles or a circle among diamonds. During this search task, the statistical regularity was no longer present and, critically, the color distractor was equally likely to appear at any of the locations. The question addressed was whether participants learn the color distractor regularity during the learning phase in which they did not have to search for anything but only judge the overall arrangement of the elements. If so, we would expect to see some lingering bias during the test phase when participants searched for the shape singleton while ignore the color singleton.

### 2.1. Material and methods<sup>2</sup>

#### 2.1.1. Participants

25 naïve participants<sup>3</sup> (24 female; mean age 20.55) from Vrije Universiteit Amsterdam volunteered to participate in this

<sup>2</sup> We report how we determined our sample size, all data exclusions (if any), all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

<sup>3</sup> This sample size was chosen based on those used in previous experiments on SL using the additional singleton paradigm (Wang & Theeuwes, 2018a,b,c). A post-hoc sensitivity analysis focusing on the critical one-tailed t-test comparing high & low-probability distractor trials with an alpha of .05 and a power of .8 indicated a sensitivity to Cohen's *d* of .512 and above, indicating this design was best suited for detecting medium-to-large effects.

experiment through the universities online participant recruitment system in exchange for course credits. All participants had normal or corrected to normal vision and gave informed written consent before beginning the experiment. This study was approved by the Ethical Review Committee of the Faculty of Behavioral and Movement Sciences of Vrije Universiteit Amsterdam.

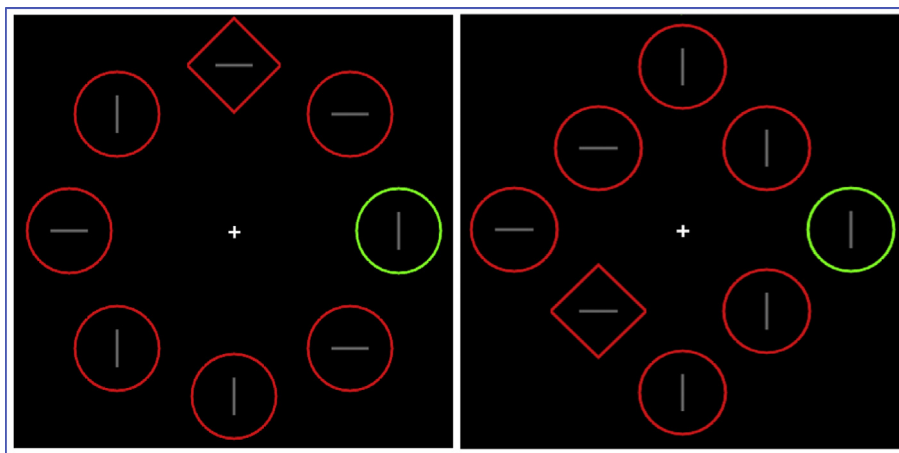
#### 2.1.2. Stimuli & apparatus

The experiment took place in a dimly lit lab; stimuli and responses were managed by HP Compaq 800 Elite computers with 21-inch liquid crystal display (LCD) colored monitors. The experiment was coded on a custom python script utilizing stimuli from the Psychopy library of psychophysical tools (Peirce, 2009). Participants were provided with a chinrest 71 cm away from the monitors and eye movements were monitored using an Eyelink 1,000 eye tracker to ensure fixation remained central at all times (sampling rate: 1,000hz; spatial resolution < .2°). Our stimuli were based on those used by Wang & Theeuwes, (2018a) which themselves were a version of the popular additional singleton paradigm (Theeuwes, 1992).

#### 2.1.3. Design & procedure

Our experiment had two parts which both utilized nearly identical stimuli. Every trial would begin with a 500 msec presentation of the fixation cross. Following this, an array of eight colored shapes were presented around the fixation cross for 3000 msec or until the participant gave a response. In the first three blocks of the experiment (the training blocks) the shapes could either be arranged in a circle or in a diamond, and participants had to indicate which it was by pressing the 'left' or 'up' directional keys respectively (see Fig. 1). In the final two blocks (the testing blocks) the stimuli were always presented in a circle, and participants had to do the classic additional search task (Theeuwes, 1992). On these trials, participants had to find the unique shape (either a circle among seven diamonds or vice versa) and indicate whether the line segment inside this shape was vertical or horizontal by pressing the 'up' or 'left' directional buttons as quickly as possible. Following the completion of each trial an inter-trial interval occurred which was a random duration between 500 msec and 750 msec. If participants moved their eyes, provided the wrong answer or failed to provide an answer within 3000 msec, an error message would be displayed and the erroneous trial would be appended to the end of the trial list to be presented again at the end of the block.

On each trial, the display consisted of either circles and one diamond or vice versa, and the shapes were equally likely to be red or green. On 67% of the trials a color singleton was present. This color singleton was one of the non-singleton shapes which would be a different color than all the other shapes (either red or green). During the training phase the color singleton was presented more often in one specific location than in all other locations. The color singleton was 13 times more likely to appear in this high probability location than the other seven locations. The color singleton was present in this location in 65% of all trials (52 trials) and evenly distributed among the remaining seven locations in the



**Fig. 1** – Each experiment started with three training blocks in which participants had to decide and respond whether the elements were arranged in a global circle (left) or a global diamond (right). In these displays and only during training, the color singleton was presented much more often in one location than in all other locations. Training blocks were followed by two testing blocks in which participants had to perform the classic additional singleton paradigm: searching for a shape singleton while ignoring the color singleton. During the testing blocks, the singleton was equally likely to appear at any of the location and the display was always arranged in a global circle. The target shape identity and color as well as target line orientation were random on every trial.

remaining trials (four presentations each). The high probability location could only be one of the four cardinal directions (top, bottom, left or right) and would never be one of the in-between locations (top-left, top-right, bottom-left, bottom-right).<sup>4</sup>

During the final two blocks, in which participants performed the singleton search task, all display elements were equally likely to appear at each location (i.e., there was no high probability location). The distractor appeared in each location the same number of times (10 times per location).

## 2.2. Results

Following the methods of Wang & Theeuwes, (2018a), all trials in which Reaction Times (RTs) were slower than 2,000 msec (2.39%) were excluded from the analysis. Additionally, all trials faster than 300 msec (.64%) were also excluded as this was deemed too fast of a RT to have properly performed the task. Furthermore, participants with error rates higher than 30% were removed (in this experiment, no participants were removed).

Due to the small number of trials in the test blocks in which the distractor was presented in the previously high-probability location during the search task (ten per block), and because we excluded trials which were too slow or too fast, it was possible that participants could end up having their average RT for high-probability trials calculated on a very small number of trials. For this reason, we chose to exclude any participants who had more than three trials in which the distractor was presented in the previously high-probability location due to being too slow or too fast (so

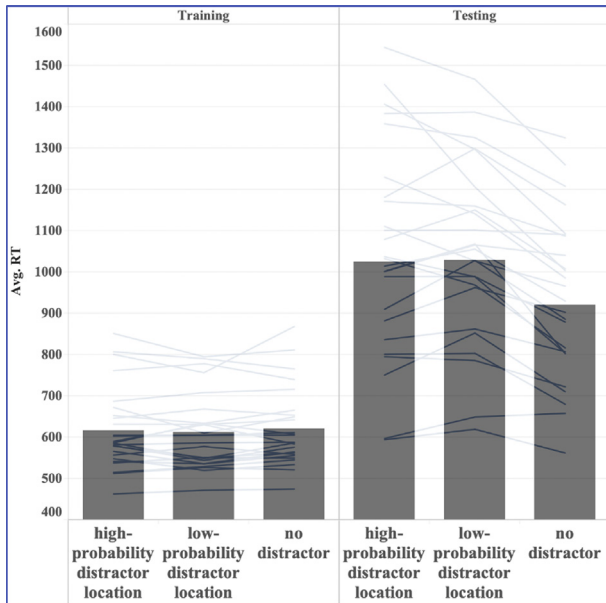
averages were calculated on seven or more trials). This resulted in one participant's result being excluded from block four and one from block five. Because these participants were often perfectly fine in all other respects of the experiment, we chose not to exclude their entire dataset from all our analyses, but rather to selectively exclude their data on the blocks in which their data was insufficient. For the combined analysis of the testing blocks (blocks 4 & 5) if a participant was excluded from either testing block for having too many high-probability trials excluded, then their entire dataset was excluded from the combined analysis of these blocks so as to not have an unbalanced representation of each block.<sup>5</sup>

Fig. 2 shows the differences in RTs when the colored singleton distractor was presented in the high probability location, in a low probability location, or was absent in the trial. One-way repeated measures analysis of variances (ANOVA's) using mean RTs across the distractor conditions (high-probability location, low-probability locations and no-distractor conditions) as factors showed a main effect of distractor condition [ $F(1.5,33) = 25.86, p < .001, \eta^2 = .052$  ( $X^2 = 8.513, p = .014, \epsilon = .708$ )].<sup>6</sup> However, planned comparisons between high-probability and low-probability distractor

<sup>5</sup> Because two participants were included in the individual block analyses that were not included in the combined block analyses, the results of the combined testing block analysis should not be taken to be a direct synthesis of the two individual block analyses.

<sup>6</sup> Many of our ANOVA's failed to pass Mauchly's test for sphericity, necessitating corrections. The reason for this was likely due to the disparity in trial counts between high-probability distractor trials and low-probability/no-distractor trials. In all further analyses if sphericity was violated then Greenhouse-Geisser corrected  $F$  values used and the approximate chi-squared, significance value and epsilon are reported in brackets following the main analysis.

<sup>4</sup> This was done as the cardinal locations actually appeared in the exact same locations on the computer screen for both circle and diamond arrays.



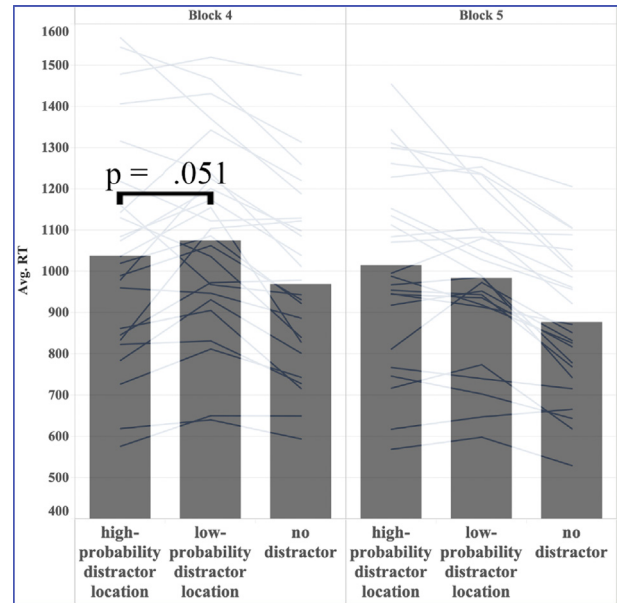
**Fig. 2** – RTs in training and testing blocks for the three distractor conditions. Individual participant scores are displayed as lines in the background.

locations showed no significant difference in RTs if the distractor was present in the previously highly probable location [one-tailed  $t(22) = 1.031, p = .157$ ]. No main effect was found when replicating the analysis for the training blocks [ $F(2,44) = 1.162, p = .322$ ].

A further analysis was undertaken in which each individual test block was analyzed in isolation (Fig. 3). The reason for this extra step was to see whether SL was indeed present in the early parts of testing (block four) as it is known that these effects can quickly fade away as evidence accumulates against the existence of a high probability location. For example, Wang & Theeuwes (2020) showed that participants start learning to reorient quite rapidly (within a few trials) when the high probability location changed between blocks. Two new unplanned repeated measures ANOVA's were conducted for the RTs on blocks four and five independently. Significant main effects were found for the RT's in blocks four and five [ $F(1.415,32.55) = 9.812, p = .001, \eta^2 = .035 (X^2 = 11.73, p = .003, \epsilon = .740)$ ;  $F(1.505,34.6) = 27.695, p < .001, \eta^2 = .084 (X^2 = 8.789, p = .012, \epsilon = .752)$  respectively]. However, critically, one-tailed  $t$ -tests showed marginally significant differences in RTs for distractors presented in the previously high-probability locations in block four [one-tailed  $t(23) = 1.703, p = .051, d = .348$ ] but not in block five [one-tailed  $t(23) = 1.481, p = .924$ ] suggesting that indeed participants showed a tendency to suppress the distractor when presented in the previously trained high probability location in the block immediately following training which rapidly faded until it was absent one block later.

### 2.3. Discussion

Our pre-planned analyses showed no significant differences between the RTs of trials in which the distractor was presented in the previously high-probability location from those



**Fig. 3** – The RTs in the two testing blocks for the distractor conditions. Only the two testing blocks are shown here as isolated block analyses were not conducted on the three training blocks.

in which it was presented elsewhere. However, further unplanned analyses of the testing blocks separately seemed to suggest that there may have been some learning taking place, but this learning quickly went extinct when the regularities were no longer in place. It was recently shown that participants very quickly adjust their statistical contingencies in the additional singleton task when the high-probability location changes from one to another location (Wang & Theeuwes, 2020). However, Britton and Anderson (2020) very recently showed that learned statistical bias should still be detectable after 180 trials of the additional singleton task without any high probability location. We therefore designed Experiment 2 to answer the question of exactly how quickly we should expect a learned contingency to fade when the regularities are no longer in place using our specific task and setup. We first ensured that participants learned the contingency, establishing a clear suppression of the high probability location. After this we changed the contingencies such that the distractor was equally likely to appear at each location, and observed how quickly the suppression effect waned. We expected that the effect would at least be present in the block immediately following training, but were unsure how long the effect would linger when the contingencies were no longer in place.

## 3. Experiment 2

Experiment 2 was intended to determine the speed with which SL dissipates following a paradigm in which we knew SL should occur. In Experiment 2, participants encountered nine blocks. Unlike Experiment 1, all blocks in Experiment 2 required participants to perform the additional singleton search. However, only the first six blocks used a high

probability distractor location while the final three presented singletons with no regularity whatsoever. Because the learning block used the additional search task, it can be claimed that this design indirectly evoked top-down attention.

### 3.1. Material and methods

A new set of 25 naïve subjects (22 female, mean age 20.55) were recruited in the exact same way as in Experiment 1. Experiment 2 began with an exact replication of Wang & Theeuwes, (2018a) where six blocks were presented with a high probability distractor location (training blocks). An additional three blocks were appended to the end of the experiment in which participants continued to perform the singleton search task but distractors were presented in all locations an equally often (testing blocks). In this experiment we did not use an eye tracker anymore as eye movement errors were relatively rare (~5% of trials in Experiment 1).

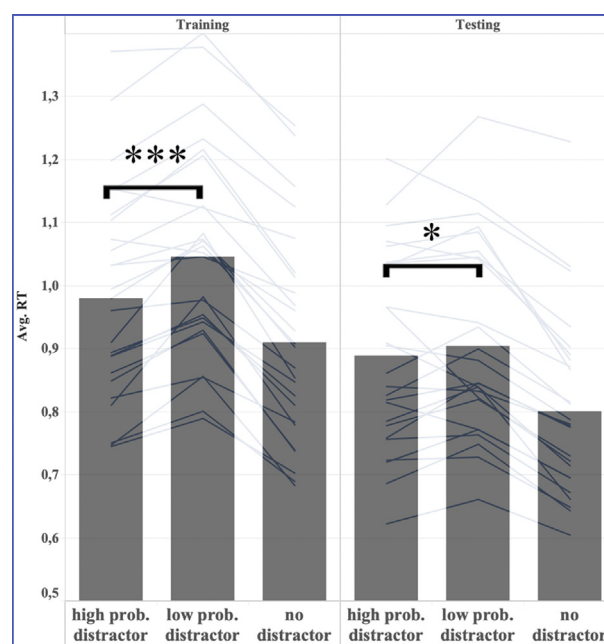
### 3.2. Results

The same exclusion criteria as Experiment 1 was used in this experiment. 3.74% of trials were excluded for being slower than 2,000 msec, .06% were excluded for being faster than 300 msec. Again, zero participants were excluded having error rates above 30%. One participant's data was excluded from block eight for having too few trials included in which the distractor appeared in the previously high probability location. One participant's data was excluded from block nine for the same reason. Again, these participant's data was excluded from the combined analysis of the three testing blocks.

A repeated measures ANOVA taking the RTs in the training blocks (blocks 1–6) with the three distractor conditions as factors showed a significant main effect [ $F(1.587, 38.1) = 124.87, p < .001, \eta^2 = .106 (X^2 = 6.923, p = .031, \epsilon = .794)$ ] as well as the comparison between trials in which the distractor was presented in the high vs. low probability locations [one-tailed  $t(24) = 6.609, p < .001, d = 1.322$ ] thereby perfectly replicating the results of Wang & Theeuwes, (2018a) and demonstrating without a doubt that participants had learned the regularities in the eight training blocks (Fig. 4, training column).

Repeating this analysis on the testing blocks revealed a main effect of distractor condition [ $F(1.423, 31.32) = 41.73, p < .001, \eta^2 = .075 (X^2 = 10.91, p = .004, \epsilon = .712)$ ], as well as a significant difference between high and low probability distractor trials [one-tailed  $t(22) = 1.921, p = .034, d = .401$ ] suggesting that the effect more strongly resisted extinction in Experiment 2 than in Experiment 1 (Fig. 4, testing column).

Furthermore, separating the testing blocks for individual analysis showed significant main effects of distractor condition in all three testing blocks [ $F(2, 48) = 27.82, p < .001, \eta^2 = .087; F(1.372, 31.57) = 17.06, p < .001, \eta^2 = .073 (X^2 = 13.45, p = .001, \epsilon = .686); F(1.148, 31.21) = 13.49, p < .001, \eta^2 = .1 (X^2 = 14.13, p < .001, \epsilon = .678)$  for blocks 7, 8 & 9 respectively]. However, planned comparisons between high and low probability distractor trials showed significant differences only in block seven, immediately following training [one-tailed  $t(24) = 1.792, p = .043, d = .358$ ] while the result was marginal in



**Fig. 4 – RTs in training and testing blocks for the three distractor conditions. (\* =  $p < .05$ , \*\* =  $p < .01$ ; \*\*\* =  $p < .001$ ).**

block eight and absent in block nine [one-tailed  $t(23) = 1.594, p = .062, d = .325; t < 1$  respectively, Fig. 5].

### 3.3. Discussion

The results of Experiment 2 showed that the learned attentional bias obsolesce when the regularities are no longer in place. While participants showed a strong learned bias in the first block of testing, this difference was no longer reliable after two blocks of testing. The fact that even after extensive training in which participants develop a strong bias to suppression the high probability location, the effects quickly dissipate when the contingencies are no longer in place, puts the results of Experiment 1 in another perspective. Indeed, the indication of suppression of the high probability location even when participants simply judged the spatial arrangement of the display elements during training suggests that some learning took place. In Experiment 4 we address this issue again in an experiment using a much larger sample size. Before we conducted this experiment, we first sought to replicate the results of Experiment 2 using an on-line procedure and with the reduced block count of Experiment 1.

## 4. Experiment 3

In Experiment 2 we employed six blocks of training while in Experiment 1 we only had three blocks. One may wonder how much learning would have taken place if only three blocks of training would have been applied. In Experiment 3 we ran three blocks of training involving visual search with the statistical regularity followed by two blocks of testing in which the regularities were no longer in place. Additionally, we

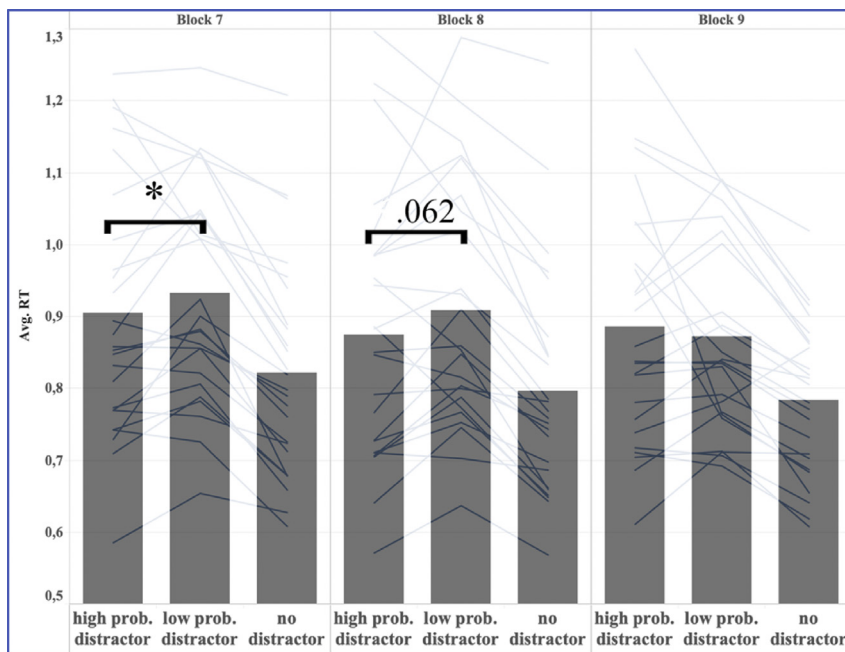


Fig. 5 – The RTs in the three testing blocks for the distractor conditions. (\* =  $p < .05$ , \*\* =  $p < .01$ ; \*\*\* =  $p < .001$ ).

choose to conduct this study on-line to determine whether we could replicate the basic Wang & Theeuwes, effect as an on-line study.

#### 4.1. Material and methods

30 anonymous naïve participants were recruited online through the website Prolific ([www.prolific.co](http://www.prolific.co)), a web-based platform for online experiments. In order to convert the original python scripted experiment to a form understandable by java-script-restricted online platforms, the experimental design tool OpenSesame was used (Mathôt, Schreij, & Theeuwes, 2012). The stimuli and procedure were reproduced as closely as possible in the new script. All participants indicated informed consent before beginning. The experiment lasted approximately 30 min and participants were compensated £3.15 for their time.

Participants were first instructed to find a quiet environment to perform the experiment in and were requested to perform the experiment on full screen. Following 12 practice trials participants performed the singleton search task for five blocks. The first three blocks participants encountered utilized a high probability location, while the last two presented the color singleton with no regularity. Additionally, unlike the previous two experiments, error trials were no longer recycled to the end of each block.

#### 4.2. Results

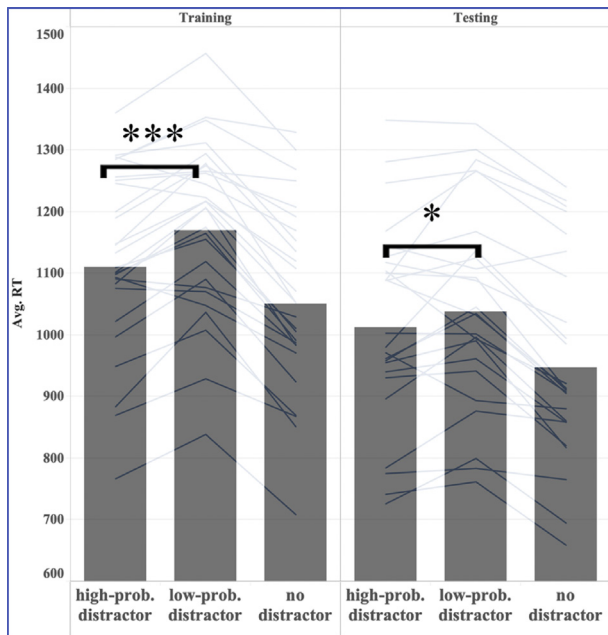
The exact same exclusion criteria as the previous two experiments was used in this experiment. 8.37% of trials were excluded for being slower than 2,000 msec, .4% were excluded for being faster than 300 msec and three participants were excluded for having an error rate above 30%.

Because error trials were not recycled in this version of the experiment, more participants had less than seven trials during the test blocks in which the distractor was presented in the previously high probability location. Two participant results were excluded from the analysis of block four and four from block five for this reason. Again, these participants were excluded from the combined testing block analysis.

Our analysis exactly mirrored that used in Experiment 2. A repeated measures ANOVA taking the RTs in the training blocks with the three distractor conditions as factors again showed a significant main effect [ $F(2,52) = 74.46, p < .001, \eta^2 = .11$ ] as well as the comparison between trials in which the distractor was presented in the high vs. low probability locations [one-tailed  $t(26) = 5.671, p < .001, d = 1.091$ ] thereby once again replicating the results of Wang & Theeuwes, (2018a) and demonstrating that such a result can be observed using an online study.

This same analysis on the testing blocks revealed a main effect of distractor condition [ $F(2,44) = 21.85, p < .001, \eta^2 = .055$ ], as well as a significant difference between high and low probability distractor trials [one-tailed  $t(22) = 2.259, p = .017, d = .471$ , Fig. 6].

Finally, separating the two testing blocks for analysis showed significant main effects of distractor condition in both blocks [ $F(2,48) = 20.92, p < .001, \eta^2 = .065$ ;  $F(1,385,30.48) = 10.75, p = .001, \eta^2 = .044$  ( $X^2 = 12.31, p = .002, \epsilon = .724$ ) for blocks 4 & 5 respectively]. But comparisons between high and low probability distractor conditions showed significant differences only in the block immediately following training (block four) which was no longer detected in the next block, suggesting that the effect went extinct after one block of no-regularity trials [one-tailed  $t(24) = 3.077, p = .003, d = .615$ ;  $t < 1$  respectively, Fig. 7].



**Fig. 6** – RTs in training and testing blocks for the three distractor conditions. (\* =  $p < .05$ , \*\* =  $p < .01$ ; \*\*\* =  $p < .001$ ).

#### 4.3. Discussion

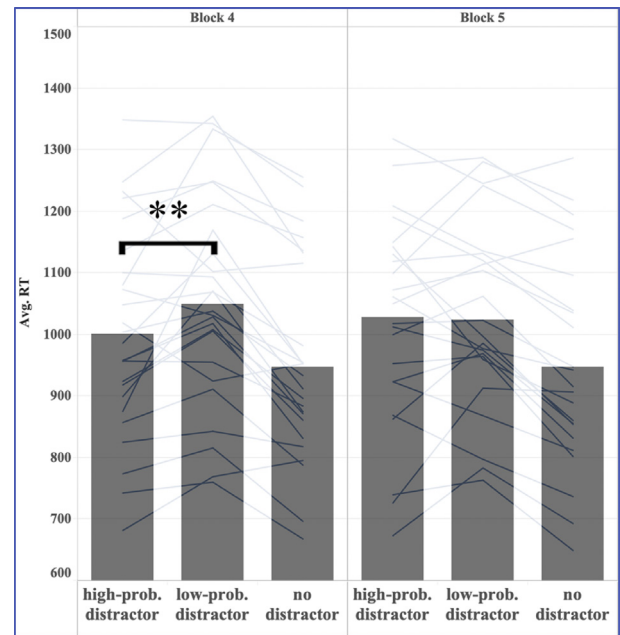
Experiment 3 shows that when only three blocks of training are given, suppression is only found in the block immediately following training but not any subsequent blocks, essentially replicating the findings of Experiment 2. Furthermore, Experiment 3 demonstrated that online experiments are an acceptable means of studying the SL effect using our paradigm, producing reasonable and interpretable results and reproducing the effects seen in lab settings.

## 5. Experiment 4

The results of Experiment 1 suggested that participants may have learned regularities of the colored distractor during the training blocks even when this was no part of their task. However, the results indicate only a trend. Therefore we chose to re-run the experiment but with a much larger sample size and on-line.

#### 5.1. Material and methods

120 anonymous naïve participants were once again recruited using the same web service as Experiment 3. All participants indicated informed consent before beginning. The experiment again lasted approximately 30 min and participants were compensated £3.15 or their time. A post-hoc sensitivity analysis for this sample size for our critical one-tailed t-test comparing high and low-probability distractor trial reaction times with an alpha of .05 and a power of .8 revealed this sample size was sensitive to Cohen's  $d$  effect sizes of .228 and above.



**Fig. 7** – The RTs in the two testing blocks for the distractor conditions. (\* =  $p < .05$ , \*\* =  $p < .01$ ; \*\*\* =  $p < .001$ ).

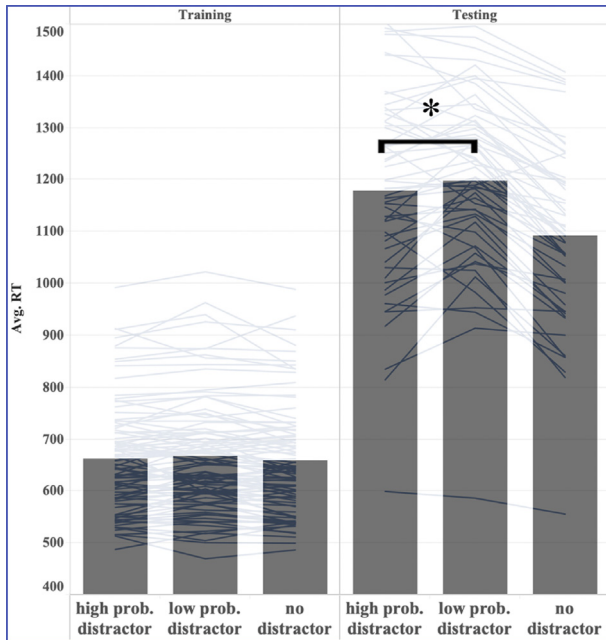
#### 5.2. Results

While 120 participants completed our experiment online, due to a bug in our code some participants finished the experiment before completing all 600 trials. As a result of this, 16 participants were excluded from the analysis of block 5 for having completed less than 75% of block 5's trials. These participants data were not excluded from the overall analysis as often they were complete in all other respects. Otherwise, the same exclusion criteria as in Experiments 3 were applied. 2.46% of trials were excluded for being slower than 2,000 msec, .11% were excluded for being faster than 300 msec and 12 participants were excluded for having an error rate above 30%. A large number of participants were excluded from the individual block analyses due to having too few correct trials in which the distractor was in the previously high-probability location. 28 were excluded from block four and a different 28 from block five. Presumably these large numbers reflect fatigue from running a psychophysical experiment online.<sup>7</sup> Again, these participants were excluded from the testing block's combined analysis.

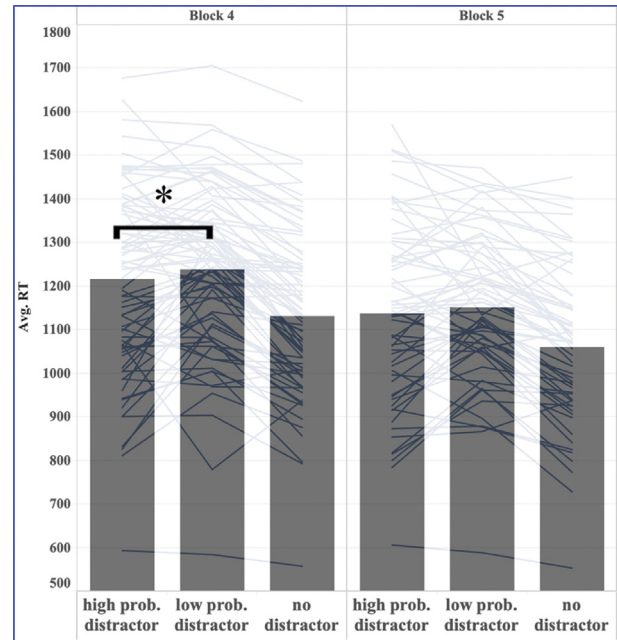
A repeated measures ANOVA of the RTs in the testing blocks taking the three distractor conditions as factors again showed a significant main effect of distractor condition [ $F(1.658, 89.552) = 69.267, p < .001, \eta^2 = .068 (X^2 = 12.253, p = .002, e = .829)$ ]. Additionally, this analysis did find a significant

<sup>7</sup> Interestingly, a much smaller percent of participants struggled to in the latter blocks of Experiment 3, which was also online and presumably employed a more difficult task. Our interpretation of this is that the effort of suddenly switching from a very easy task to a much harder task late in an experiment caused more participants to struggle than simply performing one hard but consistent task for the entire experiment.





**Fig. 8 – RTs in training and testing blocks for the three distractor conditions. (\* =  $p < .05$ , \*\* =  $p < .01$ ; \*\*\* =  $p < .001$ ).**



**Fig. 9 – The RTs in the two testing blocks for the distractor conditions. (\* =  $p < .05$ , \*\* =  $p < .01$ ; \*\*\* =  $p < .001$ ).**

difference between high and low distractor probability trials [one-tailed  $t(54) = 2.239$ ,  $p = .015$ ,  $d = .302$ , see Fig. 8].

Separating the two testing blocks into independent analyses showed significant main effects for distractor condition in both [ $F(1.625, 128.39) = 40.77$ ,  $p < .001$ ,  $\eta^2 = .054$  ( $X^2 = 20.45$ ,  $p = .001$ ,  $\epsilon = .813$ );  $F(1.663, 103.09) = 28.67$ ,  $p < .001$ ,  $\eta^2 = .046$  ( $X^2 = 13.83$ ,  $p < .001$ ,  $\epsilon = .831$ ) for blocks 4 & 5 respectively]. Additional comparisons between high and low probability distractor conditions showed significant differences in the block immediately following training (block four) but not the subsequent block [one-tailed  $t(79) = 2.016$ ,  $p = .024$ ,  $d = .224$ ; one-tailed  $t(62) = 1.093$ ,  $p = .139$ , see Fig. 9].

Additionally, an unplanned repeated measures ANOVA using the three distractor conditions as factors was conducted on the training blocks and surprisingly revealed an extremely weak but highly significant effect of distractor conditions in these blocks as well [ $F(1.825, 171.57) = 6.926$ ,  $p = .002$ ,  $\eta^2 = .001$  ( $X^2 = 9.363$ ,  $p = .009$ ,  $\epsilon = .913$ )] where the same pattern of results as in the testing blocks began to emerge with a large enough sample despite a total lack of top-down attention directed to the regularities of the coloured singleton.

## 6. General discussion

The question we addressed in this study was whether participants learn particular statistical regularities present in the environment even when these regularities are not part of the top-down task-set. In other words, do participants learn regularities simply because they are exposed to them, even when attending these regularities is unrelated to the current task

and goals. The results of Exp. 1 and 4 indicate that even under these circumstances participants do learn these contingencies. We show that after being exposed to these contingencies during an unrelated task in which there are no targets nor distractors, participants showed a suppression effect from the previously learned contingencies when switching to a task in which they search for a target and suppress a distractor. During search, when participants search for a shape singleton while ignoring a color singleton (the additional singleton task), we show that participants are better able to ignore the color singleton when it appears in the location it was more likely to appear during the previous unrelated training block. This was despite the fact that during the search the distractor was actually equally likely to appear at any of the eight locations, and thus the previously learned attentional bias served no behavioral benefit. The current study also shows that the effects are small and wane quickly when these regularities are no longer in place. Critically, our Experiments 2 and 3 show that this quick waning also occurs when there is clear evidence that the regularities are well learned during training when they are relevant for the task. Indeed, even when participants show a strong evidence for suppressing the high probability location during training sessions, the effect quickly wanes and only stays around for about 120 trials when contingencies are no longer in place, thereby replicating the approximate timescale found by Britton and Anderson (2020).

The current study's finding is consistent with a number of recent studies which assert that VSL can occur even when attention is not focused at the task at hand. Both the works of Musz et al. (2015) as well as Campbell et al. (2012) found evidence for attention free VSL when using a serial shape presentation task where participants were instructed to attend only to a subset of shapes (e.g. attend to only the red shapes). Both studies found significant familiarity scores as well as

speeded RTs in relation to trained shape contingencies, but not their foil counterparts, regardless of their membership in the attended shape set or not. However, in both studies the strength of the effect for attended shapes were stronger than the unattended shapes, highlighting top-down attentions beneficial relationship to VSL. [Forest and Finn \(2018\)](#) found a similar beneficial role for top-down attention using stimuli that were rapidly changing shapes embedded within changing colored squares; finding that SL scaled greatly depending on instruction set. Participants learned attended-to regularities better than unattended regularities and regularities passively viewed with no explicit instructions. Participants also learned the passively viewed regularities better than the unattended regularities, showing that VSL scaled nicely to the availability and orientation of attention. The above studies all used a form of serial shape presentation tasks, thus making our study the first in this body of work to demonstrate implicit VSL with an attention-based search task.

Additionally, these findings compliment nicely the work in audio statistical learning where it has been found that SL occurs without directed attention (e.g. [Saffran et al., 1997](#); [Batterink & Paller, 2019](#)) though it has also been suggested that SL improves when the stimuli is represented by the top-down attentional set ([Toro et al., 2005](#)).

It is important to emphasize that these results do not suggest that participants picked up the statistical regularities of their environment in the absence of *attention* but rather they did so in the absence of goal directed *top-down* attention. The task utilized in this experiment inherently required participants to attend to the contingent stimuli, assuming participants divided their attention across the visual field employing what has been called a large “attentional window” ([Belopolsky, Zwann, Theeuwes, & Kramer, 2007](#); [Treisman, 2006](#)). Even though not part of the top-down set, it is feasible that during training when judging the overall arrangement of the elements in the visual field, the color singleton did capture attention. Indeed, it known that in the additional singleton paradigm, attention is captured exogenously in a bottom-up way ([Theeuwes, 1991, 1992](#)), indicating during training attention may have been captured by the color singleton giving rise to observed SL (as supported by the findings in Experiment 4 that a small but significant speedup was observed in the training phase when a large enough sample was available for analysis). Furthermore, the results of [Zhao, Al-Aidroos, and Turk-Browne \(2013\)](#) showed that attention was implicitly oriented towards a location containing a structured stream of visual items relative to a location containing a random stream of visual items even though the regularity itself was irrelevant to the task at hand, suggesting that attention can indeed be automatically attracted towards environmental regularities. This capture of attention in turn may lead to implicit learning that the color singleton is presented more often in one location than in all other locations, which can then be teased out in the test phase in which the color singleton becomes the distractor and thus implicitly is connected to the top-down attentional set. In other words, while our results clearly show that SL may occur in the absence of top-down attention, we cannot rule out that bottom-up capture may still play a role.

The current results are related to studies investigating the effect of habituation on attentional capture. Dating back to [Sokolov \(1963\)](#), it has been suggested that due to repetitive exposure, the orienting of attention towards salient stimuli may habituate. For example, several recent studies have shown habituation can lead to attenuated capture effects, whether from irrelevant sudden onsets ([Chelazzi, Marini, Pascucci, & Turatto, 2019](#); [Turatto, Bonetti, Pascucci, & Chelazzi, 2018](#); [Turatto & Pascucci, 2016](#)), or irrelevant color distractors during serial search ([De Tommaso & Turatto, 2019](#); [Won & Geng, 2020](#)). Specifically, [De Tommaso and Turatto \(2019\)](#) demonstrated the potential in reinterpreting the results of attentional capture research as cases of habituation - showing distractor suppression could be learned only when a feature search mode was engaged, not when participants employed the singleton detection mode. In contrast to [De Tommaso and Turatto \(2019\)](#), the current study did find a reduction in attention capture even though the singleton detection mode was employed. Note however that the reduction in capture was observed for color distractors presented at the high probability location suggesting that during singleton search, habituation may only be possible when it concerns spatial regularities.

Generally, it is believed that any repeatedly presented stimuli should become habituated leading to a reduced attentional orienting response which should, in principle, lead to faster reaction times in any concurrent task ([Sokolov, 1963](#)). Thus a habituation interpretation of the current study would have predicted RT differences in the training phase as well as the testing phase, as was found in the unplanned analysis of Experiment 4's training blocks. Furthermore, while research has shown habituated responses lasting over long time periods (e.g. [Turatto, Bonetti, and Pascucci \(2018\)](#) showed a complete habituation of attentional capture that lasted for at least two weeks after training), alternative results have shown that a rapid extinction of habituated responses is possible given enough intervening trials in which the habituated element is excluded, suggesting that habituated responses can rapidly extinguish under certain criteria ([Turatto & Pascucci, 2016](#), Experiment 3b).

In the current study we pushed the limits of statistical learning and tested whether participants can learn statistical regularities by simply being passively exposed to these regularities. When performing a passive task with neither targets nor distractors, participants still showed a suppression effect from the previously learned contingencies when they engaged in searching for a shape singleton while ignoring the color singleton. We conclude that participants learned the statistical regularities of the singleton presentations despite such learning having no explicit or explicit connection to a top-down attentional set, thereby demonstrating that visual statistical learning can occur in the absence of top-down attention.

---

### Author notes

JT was supported by a European Research Council advanced grant 833029 – [LEARNATTEND]. Correspondence should be

addressed to Docky Duncan, Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands. Email: [D.H.Duncan@vu.nl](mailto:D.H.Duncan@vu.nl).

## Open practices

The study in this article earned an Open Data badge for transparent practices. Data for the study are available at <https://osf.io/a6ge4/>.

## REFERENCES

- Awh, E., Belopolsky, A., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, 16(8), 437–443. <https://doi.org/10.1016/j.tics.2012.06.010>
- Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, 55, 485–496. <https://doi.org/10.3758/BF03205306>
- Batterink, L. J., & Paller, K. A. (2019). Statistical learning of speech regularities can occur outside the focus of attention. *Cortex*, 115, 56–71. <https://doi.org/10.1016/j.cortex.2019.01.013>
- Belopolsky, A. V., & Theeuwes, J. (2010). No capture outside the attentional window. *Vision Research*, 10(23), 2543–2550. <https://doi.org/10.1016/j.visres.2010.08.023>
- Belopolsky, A. V., Zwann, L., Theeuwes, J., & Kramer, A. F. (2007). The size of an attentional window modulates attentional capture by color singletons. *Psychonomic Bulletin & Review*, 14(5), 934–938. <https://doi.org/10.3758/BF03194124>
- Britton, M. K., & Anderson, B. A. (2020). Specificity and persistence of statistical learning in distractor suppression. *Journal of Experimental Psychology: Human Perception and Performance*, 46(3), 324–334. <https://doi.org/10.1037/xhp0000718>
- Campbell, K. L., Zimmerman, S., Healey, M. K., Lee, M. M., & Hasher, L. (2012). Age differences in visual statistical learning. *Psychology and Aging*, 27, 650–656. <https://doi.org/10.1037/a0026780>
- Chelazzi, L., Marini, F., Pascucci, D., & Turatto, M. (2019). Getting rid of visual distractors: The why, when, how and where. *Current Opinion in Psychology*. <https://doi.org/10.1016/j.copsyc.2019.02.004>
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, 36(1), 28–71. <https://doi.org/10.1006/cogp.1998.0681>
- De Tommaso, M., & Turatto, M. (2019). Learning to ignore salient distractors: Attentional set and habituation. *Visual Cognition*, 27(3–4), 214–226. <https://doi.org/10.1080/13506285.2019.1583298>
- Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology*, 48, 269–297. <https://doi.org/10.1146/annurev.psych.48.1.269>
- Failing, M., Wang, B., & Theeuwes, J. (2019). Spatial suppression due to statistical regularities is driven by distractor suppression not by target activation. *Attention, Perception, & Psychophysics*. <https://doi.org/10.3758/s13414-019-01704-9>
- Ferrante, O., Patacca, A., Di Caro, V., Della Libera, C., Santandrea, E., & Chelazzi, L. (2018). Alternating spatial priority maps via statistical learning of target selection and distractor filtering. *Cortex*, 102, 67–95. <https://doi.org/10.1016/j.cortex.2017.09.027>
- Ferrante, O., Patacca, A., Di Caro, V., Santandrea, E., Della Libera, C., & Chelazzi, L. (2017). Statistical learning of distractor suppression. *Journal of Vision*, 17(10). <https://doi.org/10.1016/j.cortex.2017.09.027>, 674–674.
- Finn, A. S., Lee, T., Kraus, A., & Hudson Kam, C. L. (2014). When it hurts (and helps) to try: The role of effort in language learning. *Plos One*, 9(7). <https://doi.org/10.1371/journal.pone.0101806>
- Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, 12, 499–504. <https://doi.org/10.1111/1467-9280.00392>
- Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology-Human Perception and Performance*, 18(4), 1030–1044. <https://doi.org/10.1037/0096-1523.18.4.1030>
- Forest, T. A., & Finn, A. S. (2018). Attention selectively boosts learning of statistical structure. In *Proceedings of the 40th annual conference of the cognitive science society* (pp. 1674–1680). Madison, WI: Cognitive Science Society.
- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, 19(3), 117–125. <https://doi.org/10.1016/j.tics.2014.12.010>
- Geng, J. J., & Behrmann, M. (2002). Probability cuing of target location facilitates visual search implicitly in normal participants and patients with hemispatial neglect. *Psychological Science*, 13(6), 520–525. <https://doi.org/10.1111/1467-9280.00491>
- Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception and Psychophysics*, 67, 1252–1268. <https://doi.org/10.3758/BF03193557>
- Goschy, H., Bakos, S., Müller, H., & Zehetleitner, M. (2014). Probability cueing of distractor locations: Both intertrial facilitation and statistical learning mediate interference reduction. *Frontiers in Psychology*, 5, 1195. <https://doi.org/10.3389/fpsyg.2014.01195>
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2(3), 194–203. <https://doi.org/10.1038/35058500>
- Jiang, Y. V., Swallow, K. M., Rosenbaum, G. M., & Herzig, C. (2013). Rapid acquisition but slow extinction of an attentional bias in space. *Journal of Experimental Psychology: Human Perception & Performance*, 39, 87–99. <https://doi.org/10.1037/a0027611>
- Leber, A. B., & Egeth, H. E. (2006). It's under control: Top-down search strategies can override attentional capture. *Psychonomic Bulletin & Review*, 13(1), 132–138. <https://doi.org/10.3758/BF03193824>
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314–324. <https://doi.org/10.3758/s13428-011-0168-7>
- Musz, E., Weber, M. J., & Thompson-Schill, S. L. (2015). Visual statistical learning is not reliably modulated by selective attention to isolated events. *Attention, Perception, & Psychophysics*, 77(1), 78–96. <https://doi.org/10.3758/s13414-014-0757-5>
- Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2(10), 1–8. <https://doi.org/10.3389/neuro.11.010.2008>
- Posner, M. I. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology*, 32(1), 3–25. <https://doi.org/10.1080/00335558008248231>
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrucco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, 8, 101–105. <https://doi.org/10.1111/j.1467-9280.1997.tb00690.x>

- Sokolov, E. N. (1963). *Perception and the conditioned reflex*. Oxford: Pergamon Press.
- Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics*, 50(2), 184–193. <https://doi.org/10.3758/BF03212219>
- Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, 51(6), 599–606. <https://doi.org/10.3758/BF03211656>
- Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, 135(2), 77–99. <https://doi.org/10.1016/j.actpsy.2010.02.006>
- Theeuwes, J. (2018). Visual selection: Usually fast and automatic; seldom slow and volitional. *Journal of Cognition*, 1, 29. <https://doi.org/10.5334/joc.13>
- Theeuwes, J. (2019). Goal-driven, stimulus-driven and history-driven selection. *Current Opinion in Psychology*, 29, 97–101. <https://doi.org/10.1016/j.copsyc.2018.12.024>
- Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, 97, B25–B34. <https://doi.org/10.1016/j.cognition.2005.01.006>
- Treisman, A. (2006). How the deployment of attention determines what we see. *Visual Cognition*, 14, 411–443. <https://doi.org/10.1080/13506280500195250>
- Turatto, M., Bonetti, F., & Pascucci, D. (2018). Filtering visual onsets via habituation: A context-specific long-term memory of irrelevant stimuli. *Psychonomic Bulletin & Review*, 25, 1028–1034. <https://doi.org/10.3758/s13423-017-1320-x>
- Turatto, M., Bonetti, F., Pascucci, D., & Chelazzi, L. (2018). Desensitizing the attention system to distraction while idling: A new latent learning phenomenon in the visual attention domain. *Journal of Experimental Psychology: General*, 147(12), 1827–1850. <https://doi.org/10.1037/xge0000503>
- Turatto, M., & Pascucci, D. (2016). Short-term and long-term plasticity in the visual-attention system: Evidence from habituation of attentional capture. *Neurobiology of Learning and Memory*, 130, 159–169. <https://doi.org/10.1016/j.nlm.2016.02.010>
- Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, 134, 552–564. <https://doi.org/10.1037/0096-3445.134.4.552>
- Vadillo, M. A., Linssen, D., Orgaz, C., Parsons, S., & Shanks, D. R. (2019). Unconscious or underpowered? Probabilistic cuing of visual attention. *Journal of Experimental Psychology: General*, 149(1), 160–181. <https://doi.org/10.1037/xge0000632>
- Wang, B., & Theeuwes, J. (2020). Implicit attentional biases in a changing environment. *Acta Psychologica* (in press).
- Wang, B., & Theeuwes, J. (2018a). Statistical regularities modulate attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, 44(1), 13–17. <https://doi.org/10.1037/xhp0000472>
- Wang, B., & Theeuwes, J. (2018b). How to inhibit a distractor location? Statistical learning versus active, top-down suppression. *Attention, Perception, & Psychophysics*. <https://doi.org/10.3758/s13414-018-1493-z>
- Wang, B., & Theeuwes, J. (2018c). Statistical regularities modulate attentional capture independent of search strategy. *Attention, Perception, & Psychophysics*. <https://doi.org/10.3758/s13414-018-1562-352>
- Won, B., & Geng, J. J. (2020). Passive exposure attenuates distraction during visual search. *Journal of Experimental Psychology: General*. Preprint, 80. <https://doi.org/10.31234/osf.io/py4sx>
- Zhao, J., Al-Aidroos, N., & Turk-Browne, N. B. (2013). Attention is spontaneously biased toward regularities. *Psychological Science*, 24, 667–677. <https://doi.org/10.1177/0956797612460407>