Spatial suppression due to statistical learning tracks the estimated spatial probability

Rongqi Lin[1] · Xinyu Li[1] · Benchi Wang[2,3,4,5] · Jan Theeuwes[6,7]

Abstract

People are sensitive to regularities in the environment. Recent studies employing the additional singleton paradigm showed that a singleton distractor that appeared more often in one specific location than in all other locations may lead to attentional suppression of high-probability distractor locations. This in turn effectively reduced the attentional capture effect by the salient distractor singleton. However, in basically all of these previous studies, the probability that the salient distractor was presented at this specific location was relatively high (i.e., 65%; or a ratio of 13:1 between high- and low-probability locations). The question we addressed here was whether participants still can learn the regularities in the display even when these regularities are quite subtle. We systematically manipulated the ratio of the distractor appearing at the high- and low-probability location from 2:1 to 8:1. We asked the question whether the suppression effect would depend on the probabilities of the distractor appearing in the high-probability location. The results showed that the suppression of the high-probability location was linearly related to the high-low-probability ratio. In other words, the more evidence that a distractor appears more often at a particular location, the stronger the suppression. This indicates that the distribution of attention is optimally adapted to the statistical regularities present in the display.

Introduction

Spatial attention is an important mechanism for visual selection. Previous research that focused on the target's probability learning showed that in visual search, spatial attention would drift to those locations in the search array where the target was displayed with higher probability relative to other locations (Jiang, Swallow, Rosenbaum, & Herzig, 2013; Jiang, Swallow, Won, Cistera, & Rosenbaum, 2014). In addition, locations that are likely to contain distractors are suppressed relative to other locations such that they compete less for attentional resources (Ferrante et al., 2018; Wang & Theeuwes, 2018a, b, c). It is generally assumed through statistical learning (SL) that people learn the statistical regularities regarding target and distractor probabilities biasing attentional selection. SL is defined as the ability to extract events that co-occur in our environment and to utilize this learned co-variance to deploy our attentional resources implicitly in an efficient manner (Schapiro & Turk-Browne, 2015). It is assumed that through SL the weights within the spatial priority map are adjusted

✉ Benchi Wang
  wangbenchi.swift@gmail.com

1 Department of Psychology, Zhejiang Normal University, Jinhua, China

2 Key Laboratory of Brain, Cognition and Education Sciences (South China Normal University), Ministry of Education, Guangzhou, China

3 Institute for Brain Research and Rehabilitation, South China Normal University, Guangzhou, China

4 Center for Studies of Psychological Application, South China Normal University, Guangzhou, China

5 Guangdong Key Laboratory of Mental Health and Cognitive Science, South China Normal University, Guangzhou, China

6 Department of Experimental and Applied Psychology, Vrije Universiteit, Amsterdam, the Netherlands

7 Institute for Brain and Behavior Amsterdam (iBBA), Amsterdam, the Netherlands

dynamically in order to optimize attentional selection (Bisley & Goldberg, 2010; Fecteau & Munoz, 2006; Itti & Koch, 2001; Theeuwes, 2018, 2019; Zelinsky & Bisley, 2015).

It is generally believed that spatial priority maps may encode the priority of individual visual locations by combining signals, including the individual's goal (Folk & Remington, 2008; Leber & Egeth, 2006), the visual object's saliency (Theeuwes, 1992, 2010), past selection history (Awh, Belopolsky, & Theeuwes, 2012; Failing & Theeuwes, 2018; Theeuwes, 2018), reward association, and other possible sources influencing the object's saliency (Bourgeois, Chelazzi, & Vuilleumier, 2016; Bucker, Silvis, Donk, & Theeuwes, 2015; Bucker & Theeuwes, 2017; Chelazzi, Perlato, Santandrea, & Libera, 2013; Hickey, Chelazzi, & Theeuwes, 2010; Della Libera, Perlato, & Chelazzi, 2011; Failing & Theeuwes, 2018; Nissens, Failing, & Theeuwes, 2016; Schmidt, Belopolsky, & Theeuwes, 2015). As a result, the weight of each location within the map will determine the priority of selection, assuming that the location with the highest weight is selected first, followed by the next highest weight, etc. This can explain why a target appearing at a high-probability location is selected more efficiently than targets appearing in other regions.

Recently, it was shown that the weights within the spatial priority map are not only adjusted according to the probability that a location contains a target, but also according to the probability that a location contains a distractor. For example, using the additional singleton paradigm (Theeuwes, 1991, 1992), Wang and Theeuwes (2018a) showed that when a salient distractor appeared more often in one specific location (high-probability location), it caused less attentional capture than when the same distractor appeared at any other locations. It was argued that due to SL relative to all other locations, the location that was likely to contain a distractor was suppressed proactively – that is, before display onset (Wang et al., 2019). This effect has now been replicated many times (Failing, Wang, & Theeuwes, 2019; Wang, van Driel, et al., 2019; Wang, Samara, & Theeuwes, 2019; Wang & Theeuwes, 2018a, b, c; Wang & Theeuwes, 2020; Theeuwes, 2019; see also Ferrante et al., 2018; Goschy, Bakos, Mueller, & Zehetleitner, 2014; Sauter, Liesefeld, Zehetleitner, & Müller, 2018).

In the original Wang and Theeuwes (2018a) study, participants searched for a shape singleton (circle among diamonds or vice versa) while ignoring a task-irrelevant but salient color distractor (Theeuwes, 1991, 1992). The salient distractor systematically appeared at one specific location more than in all other locations. There was the classic attentional capture effect, as reaction times (RTs) were slower in distractor-present compared to distractor-absent trials. More importantly, however, when the salient distractor appeared at the high-probability location, RTs in distractor-present trials were significantly faster compared to when it appeared at any other locations (i.e., the low-probability distractor locations). Moreover, they found

RTs in distractor-absent trials were slower when the target appeared at the high-probability location compared to when it appeared at any of the other locations. Notably, even though the effects on RTs were quite strong, basically none of the participants could explicitly indicate the location that was most likely to contain a distractor. It is generally believed that these SL effects are implicitly not accessible to conscious awareness (see also Ferrante et al., 2018).

Although the findings of Wang and Theeuwes (2018a) have been replicated many times (Failing & Theeuwes, 2020; Failing et al., 2019; Gao & Theeuwes, 2020; Wang et al., 2019; Wang & Theeuwes, 2018b, c), in all these studies the salient distractor was about 13 times more likely to appear in one location than in all other locations. The question is whether such an extreme ratio of 13:1 is necessary to obtain the suppression effect. Indeed, studies investigating the probability of where the target could appear showed that this had a large effect on attentional selection. For example, Ferrante et al. (2018) showed that the spatial probability of the target directly affected distractor filtering efficiency and, vice versa, the spatial probability of the distractor affected the target selection efficiency, even though there was some imbalance between these effects.

The current study was designed to determine the limits of the distractor suppression by systematically manipulating the ratio between the high- and low-probability locations (high-low ratio). The question was whether participants can still learn the regularities in the display even when these regularities are quite subtle. As noted in Wang and Theeuwes (2018a) and all the other above-mentioned replications, the probability of the high-probability location was 65%, which meant the high-low ratio was 13:1. In the current study we systematically manipulated the high-low ratio from 2:1 to 8:1. The question we asked was whether the suppression effect would depend on the probabilities of the distractor appearing in the high-probability location.

The experiment was basically identical to Wang and Theeuwes (2018a). We used the same experimental setup, which involved a variant of the classic additional singleton task (Theeuwes, 1991, 1992) in which participants searched for a shape singleton (a circle between diamonds or a diamond between circles) while ignoring an irrelevant color singleton (either green or red; see Fig. 1). However, unlike Wang and Theeuwes (2018a) where the high-low ratio was 13:1, we manipulated the ratio from 2:1 to 8:1.

Method

The study was approved by both the Ethics Review Committee of the Vrije Universiteit Amsterdam and the Ethics Review Committee of Zhejiang Normal University.
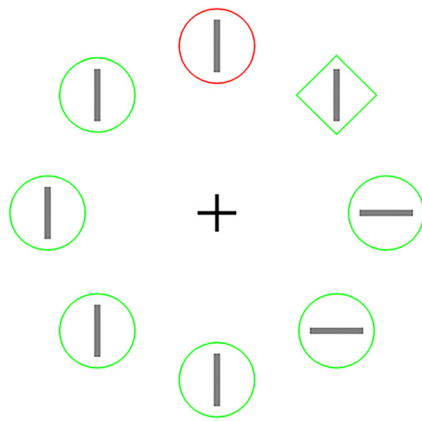
Fig. 1 The display setup and possible target and distractor locations

Participants

One hundred and twelve adults (102 females; mean age 19.5 years) from Zhejiang Normal University in China participated in the experiment for monetary compensation. They were equally divided into seven groups for different high probabilities and each group had 16 participants. Sample size was predetermined based on the difference between high- and low-probability location in Wang and Theeuwes (2018a), dz = 1.91. With 16 subjects and α = .001, power for the critical effect would be > 0.99. They provided signed informed consent, and all reported normal color vision and normal or corrected-to-normal visual acuity.

Apparatus and stimuli

Participants were tested in a dimly lit room with their chin on a chinrest located 70 cm from a 17-in. LCD monitor. The stimulus presentation and the response registration were controlled by custom scripts written in Python.

The search display consisted of eight discrete stimuli with different shapes (one circle vs. seven unfilled diamonds, or vice versa), each containing a vertical or horizontal gray line inside (0.3° × 1.4°, see Fig. 1). These stimuli were presented on an imaginary circle with a radius of 4°, centered at the fixation (a white cross that was visible throughout each trial), against a black background. The circle's radius was 1°, the unfilled diamond was subtended by 2° × 2°, and each had a red or green outline.

Procedure and design

Throughout the trial, a fixation cross appeared and remained visible, and participants had to keep fixation at this cross. After 500 ms, the search array was presented for 3,000 ms or until response. Participants had to search for one circle (target) among seven diamonds, or vice versa. They were required to indicate whether the line segment inside the target

was vertical or horizontal, by pressing the "up" or "left" key, respectively, and as fast as possible. The inter-trial interval (ITI) was randomly chosen from 500–750 ms. If the participant did not respond or if they pressed the wrong key, warning messages were shown.

The target was present on each trial, and it was equally likely to be a circle or a diamond. Across conditions, about two-thirds of the trials were distractor-present trials, in which a uniquely colored distractor (i.e., distractor singleton) was presented having the same shape as other distractors, but a different color (red or green with an equal probability). About one-third of the trials were distractor singleton-absent trials, in which no distractor singleton was presented. All conditions were randomized within each block.

The search element was presented at eight locations from the imaginary ring with 4° radius. The target appeared at each location with equal chance. The distractor singleton was presented more often in one location (high-probability location) relative to each of the other locations (i.e., low-probability location) in the distractor singleton-present condition. For each group, the high-probability location remained the same for each participant and was counterbalanced across participants. This implies that each location in the display was equally likely to be the high-probability location. The participants were not informed about the probability distribution.

The ratio between high- and low-probability location (high-low ratio) was 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, and 8:1 for different groups of participants. Each participant performed one practice block of 16 trials and six experimental blocks. However, the number of trials in each experimental block was slightly different between different groups, to ensure that the target could be equally often presented at each location (see Table 1 for the exact number of trials). Note that the time for finishing one block was about 5 min for each group, which were basically the same. Thus, we could still compare the learning effect over blocks between different groups.

Table 1 The exact number of trials per block for different experimental conditions

Group (ratio)	Distractor-present (trials)		Distractor-absent (trials)	Total (trials)
	High	Low		
1 (2:1)	16	56 (each 8)	32	104
2 (3:1)	24	56 (each 8)	40	120
3 (4:1)	32	56 (each 8)	40	128
4 (5:1)	35	49 (each 7)	40	124
5 (6:1)	36	42 (each 6)	40	118
6 (7:1)	42	42 (each 6)	40	124
7 (8:1)	40	35 (each 5)	40	115

After performing the search task, we tested participants' awareness regarding the high-probability location. They had to answer three questions: (1) They needed to indicate whether they were aware that the distractor was presented more often in one particular location. (2) If they answered "yes," they had to indicate which location was the high-probability location; if they answered "no," they had to guess the high-probability location. (3) They were asked to indicate the confidence in their answer on a 7-point scale (from not confident, i.e., 0% sure, to very confident, i.e., 100% sure).

Results

Incorrect trials and trials on which the response times (RTs) were larger or smaller than 2.5 standard deviations from the average RTs per participant were excluded from analyses across all experiments. Only small proportions of trials were excluded for different groups: 2.6%, 2.6%, 2.9%, 2.4%, 2.6%, 2.6%, and 2.8% trials were excluded for RT outliers, and 2.2%, 2.9%, 3.0%, 2.3%, 2.2%, 3.8%, and 2.8% trials were excluded for incorrect responses, for group one (with the high-low ratio of 2:1) to seven (with the high-low ratio of 8:1), respectively. The total number of excluded trials was not statistically different between different high-low ratios, $F(6, 105) = 2.06$, $p = .064$, $\eta_p^2 = .11$, $BF_{10} = 0.73$.

Distractor singleton present

The mean RTs for different high-low ratios in the distractor singleton-present condition are presented in Fig. 2a. With *distractor location* (high- vs. low-probability location) as a within-subjects factor and *high-low ratio* (2:1, 3:1, 4:1, 5:1, 6:1, 7:1, and 8:1) as a between-subjects factor, a mixed ANOVA revealed a significant effect for distractor location, $F(1, 105) = 153.96$, $p < .001$, $\eta_p^2 = 0.60$, but not for high-low ratio, $F(6, 105) = 1.71$, $p = .127$, $\eta_p^2 = 0.09$. However, a significant interaction was observed, $F(6, 105) = 5.42$, $p < .001$, $\eta_p^2 = 0.24$. Planned comparisons showed that only when the high-low ratio was 2:1, there was no difference between high- and low-probability locations, $t(15) = 1.18$, $p = .255$, Cohen's $d = 0.30$, $BF_{10} = 0.46$. For other ratios, the mean RTs were smaller for high-probability location compared to low-probability location, all ts > 3.12, ps < .007, indicating that participants learned to suppress the high-probability location when the high-low ratio was higher than 2:1. The analysis on error rates (see Online Supplementary Material (OSM) for details) did not show evidence for a speed-accuracy trade-off.

To clarify the significant interaction, we used the difference in mean RTs between high- and low-probability location (i.e., *the suppression effect*, here we used mean RTs for low-probability location minus that for high-probability location) to conduct further analysis, investigating whether the
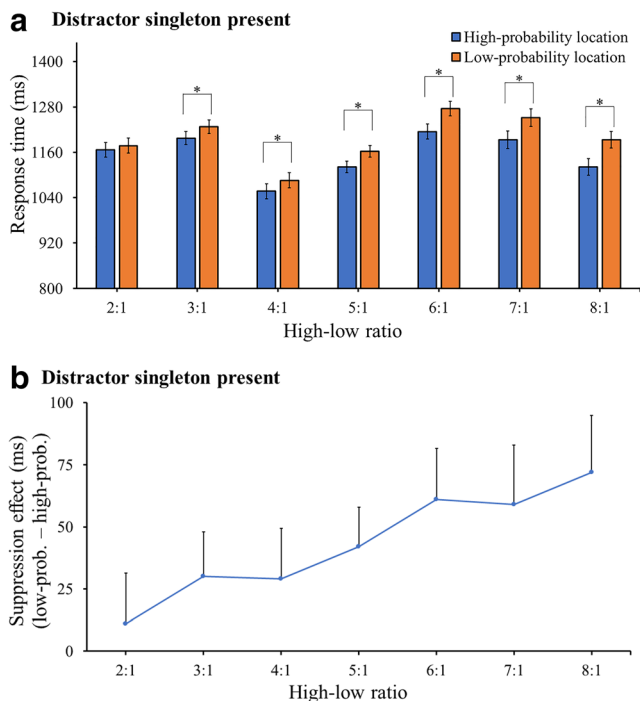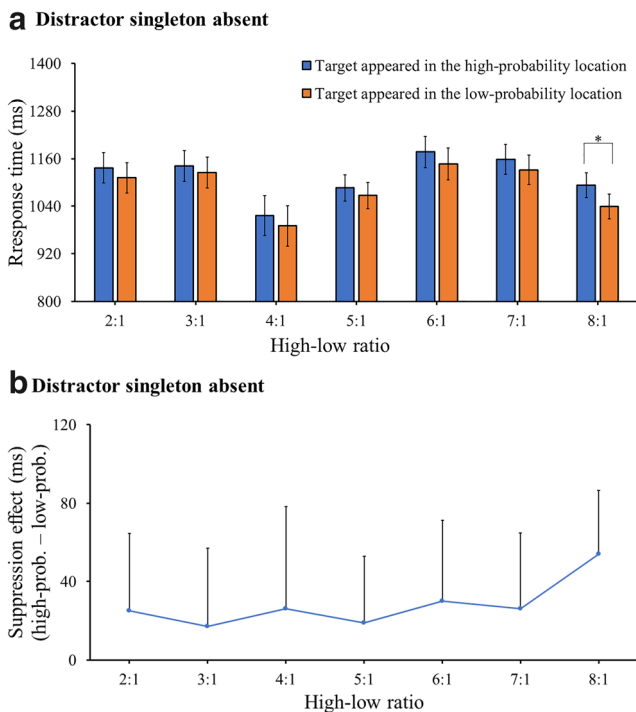

a Mean reaction times (RTs) for the high versus low distractor locations and different high-low ratios in the distractor singleton-present condition. **b** The suppression effect (mean RTs in low-probability location minus that in high-probability location). Error bars denote the within-subjects 95% confidence intervals

suppression effect varied systematically. Inspection of Fig. 2b seems to suggest that the suppression effect reached a plateau at the ratio of 6:1. Our results confirmed this pattern, showing that, compared to the ratio of 8:1, the suppression effect was smaller for the ratios from 2:1 to 5:1, all ts > 2.4, ps < .023; the suppression effect was statistically the same for the ratios of 6:1 and 7:1, both ts < 0.85, ps > .40, BF_{10}s < 0.43.

In addition, we compared the distractor singleton-present and -absent conditions. The results showed that the amount of attentional capture (i.e., mean RTs for distractor singleton-present condition minus that for distractor singleton-absent condition) did not vary between different ratios, $F(6, 105) = 2.16$, $p = .053$, $\eta_p^2 = 0.11$, $BF_{10} = 0.91$.

Distractor singleton absent

The mean RTs for different high-low ratios in the distractor singleton-absent condition are presented in Fig. 3a. With *target location* (high- vs. low-probability location) as a within-subjects factor and *high-low ratio* (2:1, 3:1, 4:1, 5:1, 6:1, 7:1, and 8:1) as a between-subjects factor, a mixed ANOVA revealed a significant main effect of target location, $F(1, 105) = 16.25$, $p < .001$, $\eta_p^2 = 0.13$, but not for high-low ratio, $F(6, 105) = 1.56$, $p = .165$, $\eta_p^2 = 0.08$, $BF_{10} = 0.48$. The interaction was also not reliable, $F(6, 105) = 0.43$, $p = .856$, $\eta_p^2 = 0.02$, $BF_{10} = 0.06$ (see Fig. 3a and b). Even though the effect of
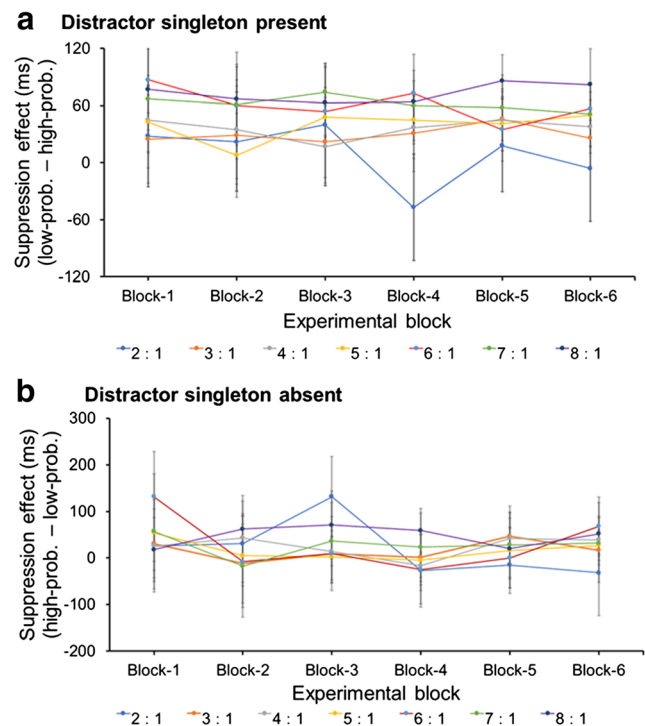
a Distractor singleton absent

b Distractor singleton absent

Fig. 3 a Mean reaction times (RTs) for the high versus low distractor locations and different high-low ratios in the distractor singleton-absent condition. **b** The suppression effect (mean RTs in high-probability location minus those in low-probability location). Error bars denote the within-subjects 95% confidence intervals

a Distractor singleton present

b Distractor singleton absent

Fig. 4 The suppression effect for different blocks and different high-low ratios in the distractor singleton-present (**a**) and distractor singleton-absent (**b**) condition. Error bars denote within-subjects 95% confidence intervals

high-low ratio was not significant, planned t-tests showed that only when the high-low ratio was 8:1, there was a significant difference between high- and low-probability location, $t(15)$ =3.64, $p = .002$, Cohen's $d = 0.91$. For other ratios, there was no statistical difference, all $ts < 1.60, ps > 0.131$, $BF_{10}s < 0.73$.

Learning to suppress over time

To determine the speed with which learning took place and whether this depended on the high-low ratio, we conducted mixed ANOVAs with *experimental block* (block-1, block-2, block-3, block-4, block-5, block-6) as a within-subjects factor and *high-low ratio* (2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1) as a between-subjects factor for distractor-present and -absent conditions separately. In the distractor present condition, the results revealed a significant effect for the high-low ratio, $F(6, 105) = 5.53$, $p < .001$, $\eta_p^2 = 0.24$, but not for experimental block, $F(5, 525) = 0.35$, $p = .88$, $\eta_p^2 < 0.01$, $BF_{10} < 0.01$, nor for the interaction, $F(30, 525) = 1.0$, $p = .475$, $\eta_p^2 = 0.05$, $BF_{10} = 0.01$ (Fig. 4a). In the distractor absent condition, there were no significant effects for the high-low ratio, $F(6, 105) = 0.28$, $p = .945$, $\eta_p^2 = 0.02$, $BF_{10} < 0.01$, experimental block, $F(5, 525) = 1.51$, $p = .185$, $\eta_p^2 = 0.01$, $BF_{10} = 0.03$, nor the interaction, $F(30, 525) = 0.79$, $p = .780$, $\eta_p^2 = 0.04$, $BF_{10} = 0.04$ (Fig. 4b). These results indicate that people quickly learn to suppress the high probability in the first small block (about 5

min), and this effect remained stable over the remaining blocks.

Inter-trial effect

Since the high-low ratio was greater than 1, there were more repetition trials for the high-probability condition. For example, in order to have an 8:1 ratio, the distractor needed to be repeated much more often in the high-probability location than when the ratio is only 2:1. Therefore, it is possible that the increase in suppression depending on the ratio is simply the result of more distractor-location repetition trials when the ratio is high. To address this issue, we determined whether this distractor-location repetition (inter-trial priming) played a role in obtaining the above-described effects. A mixed ANOVA with *repeat condition* (repeated vs. non-repeated), *high-low ratio* (2:1, 3:1, 4:1, 5:1, 6:1, 7:1, and 8:1) and *distractor location* (high- vs. low-probability location) as factors, showed a significant main effect for the distractor location, $F(1, 105) = 143.31$, $p < .001$, $\eta_p^2 = 0.58$, but not for the ratio, $F(6, 105) = 1.72$, $p = .124$, $\eta_p^2 = 0.09$, $BF_{10} = 0.45$, nor for the repeat condition (repeated vs. non-repeated), $F(1, 105) = 3.62$, $p = .06$, $\eta_p^2 = 0.03$, $BF_{10} = 0.11$. The mean RTs for the repeated trials was 1,175 ms, and that for the non-repeated trials was 1,176 ms. The interaction between the distractor location and the high-low ratio was reliable, $F(6, 105) = 4.78$, $p < .001$, $\eta_p^2 = 0.22$. None of the other interactions

involving the repeat condition were significant, all $ps > .132$, $BF_{10}s < 0.05$ (Fig. 5).

We also compared the suppression effect between the high- and low-probability location for the repeated and non-repeated condition, with *repeat condition* (repeated vs. non-repeated) and *high-low ratio* (2:1, 3:1, 4:1, 5:1, 6:1, 7:1, and 8:1) as factors, a mixed ANOVA revealed a significant main effect for the high-low ratio, $F(6, 105) = 4.77$, $p < .001$, $\eta_p^2 = .21$; but not for the repeat condition, $F(1, 105) = 1.30$, $p = .256$, $\eta_p^2 = .01$, $BF_{10} = 0.26$, nor for the interaction, $F(6, 105) = 1.13$, $p = .352$, $\eta_p^2 = .06$, $BF_{10} = 0.14$ (Fig. 6).

Awareness analysis

From group 1 (with the ratio of 2:1) to group 7 (with the ratio of 8:1), respectively, three, five, seven, ten, six, seven, and ten participants claimed that the distractor was present more often at one location than at other locations. However, only none, two, two, three, three, four, and five participants correctly identified the high-probability location with low average confidence scores of 0, 2.5, 2.5, 4.3, 3.6, 3.0, and 4.2 in group 1 to group 7, respectively. It suggests little, if any, awareness of the regularities present in the display even when the ratio is relatively high. Note that the number of participants correctly identifying the high-probability location increased significantly with the high-low ratio, $r = 0.98$, $p < .001$, while the average confidence of correctly identifying the high-probability location did not increase with the high-low ratio, $r = 0.58$, $p = .228$.

We considered those participants who correctly identified the high-probability location to be the *aware* group, and the remaining participants the *unaware* group. When we included awareness (aware vs. unaware) as a factor in the analysis on the suppression effect in the distractor singleton-present condition, there was no effect of awareness. With the high-low ratio (3:1, 4:1, 5:1, 6:1, 7:1, and 8:1) and awareness (aware vs.
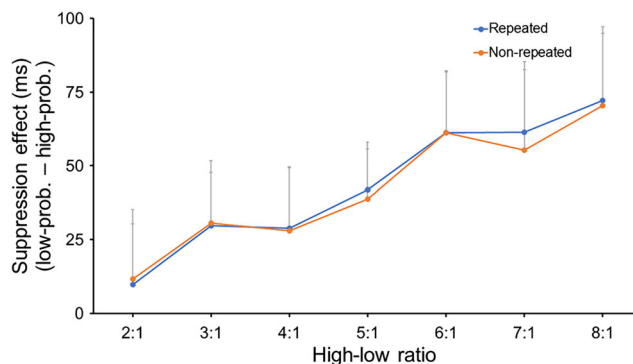

Fig. 6 The suppression effect (mean reaction times in the high-probability location minus those in the low-probability location) of different high-low ratio groups between the repeated and non-repeated conditions. Error bars denote the within-subjects 95% confidence intervals

unaware) as between-subjects factors, a mixed ANOVA only showed a main effect for the high-low ratio, $F(5, 84) = 3.58$, $p = .005$, $\eta_p^2 = .18$, but not for awareness, $F(1, 84) = 2.30$, $p = .133$, $\eta_p^2 = .03$, $BF_{10} = 0.14$, nor for the interaction, $F(5, 84) = 1.23$, $p = .304$, $\eta_p^2 = .07$, $BF_{10} = 0.35$, suggesting that being aware of the high-probability location is not necessary for obtaining the learning effects, and the learning can be implicit (findings identical to Wang & Theeuwes, 2018a). Furthermore, when conducting this analysis for the unaware group only the results remained the same (see OSM for details).

Discussion

In the current study we systematically manipulated the ratio of the distractor appearing at the high- vs. low-probability locations (high-low ratio) from 2:1 to 8:1. We showed that participants could learn the distractor's probability distribution even when the regularities were quite subtle. Indeed, for any high-
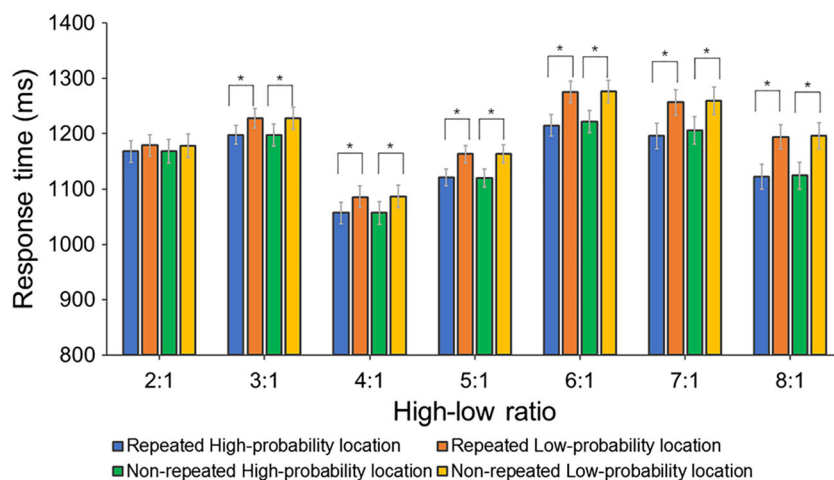
Fig. 5 Mean response times of different high-low ratio groups in the distractor singleton-present condition for repeated and non-repeated conditions. Error bars denote the within-subjects 95% confidence intervals

low ratio that was larger than 2:1, there was suppression of the high-probability location. Elegantly, the suppression effect became stronger with more evidence for the high-probability location (going from 2:1 to 8:1). Even though in all previous replications of the original Wang and Theeuwes (2018a) study the evidence for the high-probability location was much larger (using a ratio of 13:1; see Failing et al., 2019; Theeuwes, 2019; Wang, Samara, et al., 2019; Wang & Theeuwes, 2018a, b, c, 2020), the current findings indicate that the effect is actually quite robust and much more subtle than previously assumed. The current study also shows suppression of the target when it happens to be presented at the high-probability location as originally reported by Wang and Theeuwes (2018a). Critically, this effect was present for all ratios as there was only a main effect of target location that did not interact with ratio.

It should be noted that in distractor-absent trials, when the target singleton happened to be presented at the high-probability location there was only significant suppression for the highest ratio (8:1). For all other ratios numerically, there was a suppression effect but it did not reach significance (see Fig. 3b). In our previous studies we used a ratio of 13:1 and we always found that when the target was presented at the high-probability location it was suppressed (see Wang & Theeuwes, 2018a, b, c). However, some studies failed to find such a target suppression effect (Allenmark, Zhang, Liesefeld, Shi, & Müller, 2019; van Moorselaar, Daneshtalab, & Slagter, 2020; Zhang, Allenmark, Liesefeld, Shi, & Müller, 2019). It is not immediately clear why the suppression of the target singleton is found in some studies and not in others; but the current findings suggest that if probabilities are small such that there is no strong incentive to suppress the high-probability location, this effect may simply be too weak to become reliable. However, it is also possible that the effect on target suppression is weak because in the current paradigm the speed of disengagement from the high-probability location may play a role (see Wang, Samara, et al., 2019).

It is important to note that the current findings cannot be explained by inter-trial priming (Maljkovic & Nakayama, 1994). When we removed all repeat trials, the results remained basically the same, suggesting that inter-trial priming had little, if any, effect. It is important to note that this finding is unlike that reported by Goschy et al. (2014), who showed that inter-trial priming (i.e., distractor position repetitions) did play a large role in distractor suppression. For example, in their Experiment 2, when the location of the distractor was repeated, the suppression was so strong that inference by the distractor was completely eliminated and RTs were comparable to the no-distractor condition. Even though there is some resemblance between Goschy et al. and the current study, it should be noted that the tasks employed were quite different. Goschy et al. used a search task consisting of 36 vertical lines

divided across the visual field. With so many lines presented within a 6° radius, there may be interactions between these line segments related to crowding, and it is known that repetition priming affects crowding (Kristjánsson, Heimisson, Róbertsson, & Whitney, 2013). This may explain why in Goschy et al. (2014) inter-trial priming played such a large role.

Our results regarding the speed of learning are also important as it indicates that learning was extremely fast (all learning took place in the first block). In other words, learning to suppress did not develop faster when there was more evidence in favor of the high-probability location (i.e., at the high ratio). The suppression was stronger for the high ratio but this level of suppression was reached very quickly and stayed the same during the course of the experiment. This later finding is consistent with a recent study by Wang and Theeuwes (2020) that showed that participants quickly adapt to the spatial suppression when the location containing the distractor changes more often from one location to the next during the course of a block of trials.

Previous studies have examined how the allocation of attention depended on the target probability. The classic study by Shaw and Shaw (1977) showed that the allocation of attention depended very much on the target location probability. Participants viewed briefly presented displays and discriminated a single target letter that could appear in any of eight locations arranged in a virtual circle around the fixation point. In two locations, the target location probability was 25%, in four locations it was 10%, and in the remaining two locations it was 5%. The results showed that participants allocated attention optimally according to the probability distributions of the target locations. Geng and Behrmann (2002) also investigated the effect of the target location probability. In their task, participants searched for a target letter among five distractors. Eighty percent of the targets were presented anywhere on side of the display, while 20% appeared on the other half side. Not surprisingly, participants allocated attention depending on these probabilities to improve their search. Various other studies have shown that the probabilities of the target appearing at particular locations have a large effect on performance (e.g., Fecteau, Korjoukov, & Roelfsema, 2009; Geng & Behrmann, 2005; Hoffmann & Kunde, 1999; Miller, 1988). The present studies showed that the probability of the distractor appearing at a particular location also has a robust effect on attentional selection (see also, Ferrante et al., 2018; Goschy et al., 2014). Consistent with previous studies investigating target probability, the current study investigates distractor probabilities showing that participants optimally allocate attention resources according to the probabilities present in the display. We assume that this optimization of attentional selection to the probabilities of targets and distractors present in the visual field is not strategic in nature nor the result of repetition priming but instead due to statistical learning, i.e., the effects of

lingering biases of previous selection episodes (Theeuwes, 2018, 2019).

The present study also showed that awareness did not play a large role in obtaining the current results. Indeed, consistent with previous studies (Wang & Theeuwes, 2018a), participants who showed some awareness of the regularities behaved exactly the same as those participants who were unaware of them. This is consistent with the idea that these suppression effects are basically implicit, operating outside conscious control (see Theeuwes, 2018, for a discussion). In addition, when participants are asked to suppress a location in a true top-down manner, results show that participants are not able to do so (Wang & Theeuwes, 2018b).

In summary, we found more suppression when the evidence for a distractor appearing at a particular location is stronger. The results indicate that the distribution of attention is optimally adapted to the statistical regularities present in the display.

References

Allenmark, F., Zhang, B., Liesefeld, H. R., Shi, Z., & Müller, H. J. (2019). Probability cueing of singleton-distractor regions in visual search: The locus of spatial distractor suppression is determined by colour swapping. *Visual Cognition*, 27(5–8), 576–594. https://doi.org/10.1080/13506285.2019.1666953

Awh, E., Belopolsky, A. V., & 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

Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience*, 33(1), 1–21. https://doi.org/10.1146/annurev-neuro-060909-152823

Bourgeois, A., Chelazzi, L., & Vuilleumier, P. (2016). How motivation and reward learning modulate selective attention. *Progress in Brain Research, 229*, 325. https://doi.org/10.1016/bs.pbr.2016.06.004

Bucker, B., Silvis, J. D., Donk, M., & Theeuwes, J. (2015). Reward modulates oculomotor competition between differently valued stimuli. *Vision Research*, 108, 103–112. https://doi.org/10.1016/j.visres.2015.01.020

Bucker, B., & Theeuwes, J. (2017). Pavlovian reward learning underlies value driven attentional capture. *Attention, Perception, & Psychophysics*, 79(2), 415–428. https://doi.org/10.3758/s13414-016-1241-1

Chelazzi, L., Perlato, A., Santandrea, E., & Libera, C. Della. (2013). Rewards teach visual selective attention. *Vision Research*, 85, 58–72. https://doi.org/10.1016/j.visres.2012.12.005

Della Libera, C., Perlato, A., & Chelazzi, L. (2011). Dissociable effects of reward on attentional learning: From passive associations to active

monitoring. *PLoS ONE*, 6(4), 2–7. https://doi.org/10.1371/journal.pone.0019460

Failing, M., & Theeuwes, J. (2018). Selection history: How reward modulates selectivity of visual attention. *Psychonomic Bulletin and Review*, 25(2), 514–538. https://doi.org/10.3758/s13423-017-1380-y

Failing, M., & Theeuwes, J. (2020). More capture, more suppression: Distractor suppression due to statistical regularities is determined by the magnitude of attentional capture. *Psychonomic Bulletin and Review*, 27(1), 86–95. https://doi.org/10.3758/s13423-019-01672-z

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*, 81(5), 1405–1414. https://doi.org/10.3758/s13414-019-01704-9

Fecteau, J. H., Korjoukov, I., & Roelfsema, P. R. (2009). Location and color biases have different influences on selective attention. *Vision Research*, 49(9), 996–1005. https://doi.org/10.1016/j.visres.2009.03.013

Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: A priority map for target selection. *Trends in Cognitive Sciences*, 10(8), 382–390. https://doi.org/10.1016/j.tics.2006.06.011

Ferrante, O., Patacca, A., Di Caro, V., Della Libera, C., Santandrea, E., & Chelazzi, L. (2018). Altering 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

Folk, C. L., & Remington, R. W. (2008). Bottom-up priming of top-down attentional control settings. *Visual Cognition*, 16(2–3), 215–231. https://doi.org/10.1080/13506280701458804

Gao, Y., & Theeuwes, J. (2020). Learning to suppress a distractor is not affected by working memory load. *Psychonomic Bulletin and Review*, 27(1), 96–104. https://doi.org/10.3758/s13423-019-01679-6

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 & Psychophysics* 67, 1252–1268. https://doi.org/10.3758/BF03193557

Goschy, H., Bakos, S., Mueller, H. J., & Zehetleitner, M. (2014). Probability cueing of distractor locations: Both intertrial facilitation and statistical learning mediate interference reduction. *Frontiers in Psychology*, 5, 1–11. https://doi.org/10.3389/fpsyg.2014.01195

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward guides vision when it's your thing: Trait reward-seeking in reward-mediated visual priming. *PLoS ONE*, 5(11), 1–5. https://doi.org/10.1371/journal.pone.0014087

Hoffmann, J., & Kunde, W. (1999). Location-specific target expectancies in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 25(4), 1127–1141. https://doi.org/10.1037/0096-1523.25.4.1127

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2(3), 194-203.

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 and Performance*, 39(1), 87–99. https://doi.org/10.1037/a0027611

Jiang, Y. V., Swallow, K. M., Won, B. Y., Cistera, J. D., & Rosenbaum, G. M. (2014). Task specificity of attention training: The case of probability cuing. *Attention, Perception, & Psychophysics*, 77(1), 50–66. https://doi.org/10.3758/s13414-014-0747-7

Kristjánsson, Á., Heimisson, P. R., Róbertsson, G. F., & Whitney, D. (2013). Attentional priming releases crowding. *Attention, Perception, & Psychophysics*, 75(7), 1323–1329. https://doi.org/10.3758/s13414-013-0558-2

Leber, A.B., Egeth, H.E. (2006). It's under control: Top-down search strategies can override attentional capture. *Psychonomic Bulletin & Review*, 13, 132–138. https://doi.org/10.3758/BF03193824

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition, 22*(6), 657–672. https://doi.org/10.3758/BF03209251

Miller, J. (1988). Components of the location probability effect in visual search tasks. *Journal of Experimental Psychology: Human Perception and Performance, 14*(3), 453–471. https://doi.org/10.1037/0096-1523.14.3.453

Nissens, T., Failing, M., & Theeuwes, J. (2016). People look at the object they fear: Oculomotor capture by stimuli that signal threat. *Cognition and Emotion,* 1–8. https://doi.org/10.1080/02699931.2016.1248905

Sauter, M., Liesefeld, H. R., Zehetleitner, M., & Müller, H. J. (2018). Region-based shielding of visual search from salient distractors: Target detection is impaired with same- but not different-dimension distractors. *Attention, Perception, & Psychophysics, 80*(3), 622–642. https://doi.org/10.3758/s13414-017-1477-4

Schapiro, A., & Turk-Browne, N. (2015). Statistical learning. In *Mapping: An encyclopedic reference* (Vol. 3, pp. 501–506). New York, Academic Press. https://doi.org/10.1016/B978-0-12-397025-1.00276-1

Schmidt, L. J., Belopolsky, A. V., & Theeuwes, J. (2015). Attentional capture by signals of threat. *Cognition and Emotion, 29*(4), 687–694. https://doi.org/10.1080/02699931.2014.924484

Shaw, M. L., & Shaw, P. (1977). Optimal allocation of cognitive resources to spatial locations. *Journal of Experimental Psychology: Human Perception and Performance, 3*(2), 201–211. https://doi.org/10.1037/0096-1523.3.2.201

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics, 49*(1), 83–90. https://doi.org/10.3758/BF03211619

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*(1), 1–15. 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

van Moorselaar, D., Daneshtalab, N., & Slagter, H. (2020). *Neural mechanisms underlying distractor inhibition on the basis of feature and/or spatial expectations*. bioRxiv.

Wang, B., Samara, I., & Theeuwes, J. (2019). Statistical regularities bias overt attention. *Attention, Perception, & Psychophysics, 81*(6), 1813–1821. https://doi.org/10.3758/s13414-019-01708-5

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, 80*(4), 860–870. 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, 80*(7), 1763–1774. https://doi.org/10.3758/s13414-018-1562-3

Wang, B., & Theeuwes, J. (2020). Implicit attentional biases in a changing environment. *Acta Psychologica* (in press).

Wang, B., van Driel, J., Ort, E., & Theeuwes, J. (2019). Anticipatory distractor suppression elicited by statistical regularities in visual search. *Journal of Cognitive Neuroscience, 31*(10), 1535–1548. https://doi.org/10.1162/jocn_a_01433

Zelinsky, G. J., & Bisley, J. W. (2015). The what, where, and why of priority maps and their interactions with visual working memory. *Annals of the New York Academy of Sciences, 1339*(1), 154–164. https://doi.org/10.1111/nyas.12606

Zhang, B., Allenmark, F., Liesefeld, H. R., Shi, Z., & Müller, H. J. (2019). Probability cueing of singleton-distractor locations in visual search: Priority-map-versus dimension-based inhibition? *Journal of Experimental Psychology: Human Perception and Performance, 45*(9), 1146.