Supplementary Information

" Statistical regularities modulate attentional capture" by Benchi Wang and Jan Theeuwes.

Error rates analysis

Attentional capture effect

The results on error rates mimicked those on RT. A one-way ANOVA on mean error rates with *distractor condition* (high-probability location, low-probability location, and no distractor) as a factor showed a main effect, $F(2, 46) = 28.75$, $p < .001$, partial $\eta^2 = .56$. Planned comparisons showed that participants made more errors when a distractor was present at the high-probability location then when no distractor was present, $t(23) = 4.74$, $p < .001$. The same was found for a distractor presented at the low-probability location, $t(23) = 5.7$, $p < .001$. Crucially, also significantly less errors were made when a distractor was presented at the high- versus low-probability location, $t(23) = 4.92$, $p < .001$, suggesting that the suppression of the distractor resulted in more accurate behavior.

The spatial distribution of the suppression effect

The results on error rates mimicked those for RTs. A one-way ANOVA on mean RTs showed a significant main effect for *distance* (dist-0, dist-1, dist-2, dist-3, and dist-4), $F(4, 92) = 6.68$, $p < .001$, partial $\eta^2 = .23$. Moreover, we fitted the data with a linear function and used its slope to determine whether error rate changed with distance. The slope (1.52% per display element) was significantly larger than zero, $t(23) = 4.83$, $p < .001$, suggesting a spatial gradient of the suppression effect.

Practicing of overcoming distraction?

One might question that the suppression effect observed in the present study was due to the practicing of overcoming distraction, rather than the underlying statistical learning. If this is the case, the suppression effect would change over time. To test this possibility, the variable *block* was involved in the analysis. A repeated measures ANOVA on mean RTs with *block* (1-6) and *distractor condition* (high-probability location, low-probability location, and no distractor) as factors showed main effects for *block*, $F(5, 115) = 12.27$, $p < .001$, partial $\eta^2 = .35$, and *distractor condition*, $F(2, 46) = 133.61$, $p < .001$, partial $\eta^2 = .85$. However, importantly, the interaction was not significant, $F < 1$, suggesting the suppression effect did not change over time.

The results on error rates mimicked those for RT. A repeated measures ANOVA on mean error rates with *block* (1-6) and *distractor condition* (high-probability location, low-probability location, and no distractor) as factors showed main effects for *block*, $F(5, 115) = 8.15$, $p < .001$, partial $\eta^2 = .26$, and *distractor condition*, $F(2, 46) = 28.89$, $p < .001$, partial $\eta^2 = .56$. However, importantly, the interaction was not significant, $F < 1.04$, suggesting the suppression effect did not change over time. Taken together, these results suggest that the suppression effect observed here was not due to the practicing of overcoming the distraction.

Furthermore, we determined whether the spatial distribution of the suppression effect also emerged early in time. To test this, the variable *block* was included in the analysis. A repeated measures ANOVA on mean RTs with *block* (1-6) and *distance* (dist-0, dist-1, dist-2, dist-3, and dist-4) as factors showed main effects for *block*, $F(5, 115) = 10.38$, $p < .001$, partial $\eta^2 = .31$, and *distance*, $F(5, 115) = 13.28$, $p < .001$, partial $\eta^2 = .37$. However, importantly, the interaction was not reliable, $F < 1$, suggesting the spatial distribution of the suppression effect on RTs was present early in the experiment and did not change over time.

Short-term location-based priming?

We did additional analyses to determine whether the effect was driven by short-term location-based priming. It is well known that the effect of a distractor is reduced if it is repeated over consecutive trials (e.g., Maljkovic & Nakayama, 1996). To that end, we analyzed the data for trials in which the distractor singleton was repeated at the high-probability location versus when it was not repeated at that location. The mean RTs for repeated trials (from high-likely distractor location to high-likely distractor location) was 818 ms and the mean RTs for non-repeated trials (from low-likely distractor location to high-likely distractor location) was 832 ms. Even though numerically there was a difference, the effect was statistically not reliable $t(23) = 1.89$, $p = .071$. Moreover, there was also a spatial gradient of the suppression effect when analyzing only trials in which the location was not repeated, $F(4, 92) = 12.93$, $p < .001$, partial $\eta^2 = .36$ with a linear slope of 16.7 ms per display element which was different from zero, $t(23) =$

5.66, $p < .001$. On the basis of these results, we conclude that short-term location-based priming of the distractor location may have contributed to the effect, but it is unlikely to be the main factor driving the effect.

Implicit statistical learning or top-town strategy?

We analyzed the data separately for those participants that were able to identify correctly (15 people) or incorrectly (17 people) the high probability distractor location. A repeated measures ANOVA on mean RTs (see Figure S1) with *distractor condition* as a within-subjects and *group* (correctly identified vs. incorrectly identified) as a between-subjects factor showed main effect for *distractor condition*, $F(2, 60) = 85.27$, $p < .001$, partial $\eta^2 = .74$, but not of *group*, $F < 1$. Crucially, the factor *group* did not interact with *distractor condition*, $F < 1$, indicating that whether or not participants were able to correctly identify the high-probability distractor location had no effect on the pattern of results (see Figure S1). The same pattern was found for the efficiency of selecting the target in the no distractor condition. There was an effect of *target location*, $F(1, 30) = 22.03$, $p < .001$, partial $\eta^2 = .42$, but not of *group*, $F < 1$, and no interaction, $F < 1$. As is clear from Figure S2, the selecting of the target was less efficient when it appeared at the high-probability distractor location relative to the low-probability distractor location regardless of whether or not participants correctly identified the high-probability distractor location. Overall, the additional experiment basically replicates the results of the main experiment.

Supplementary figures

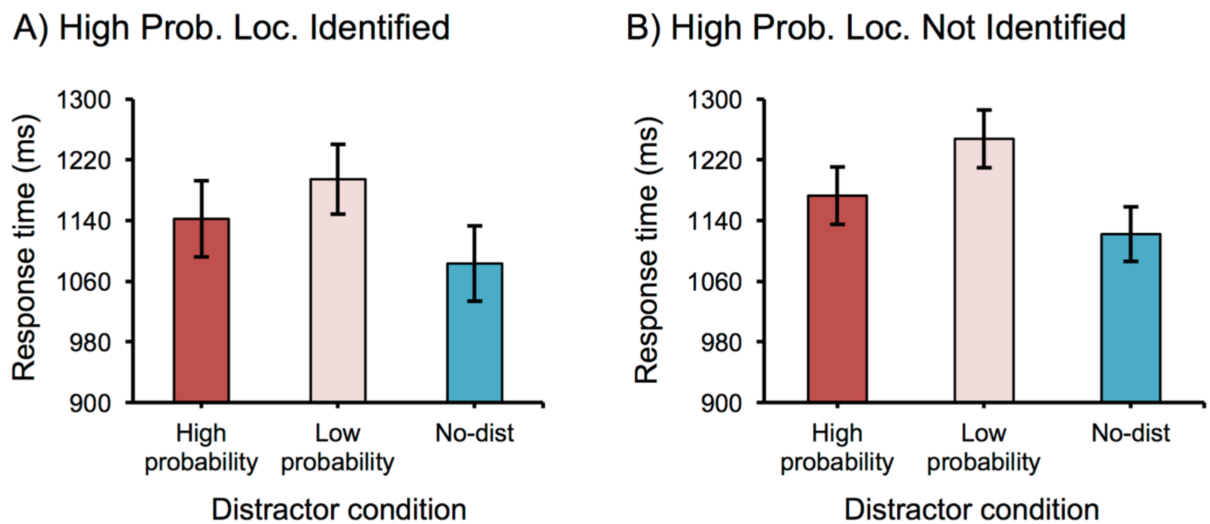

Figure S1. The mean RTs between different distractor conditions in correctly identified group (left panel) and incorrectly identified group (right panel). Error bars denote ± 1 SEM.

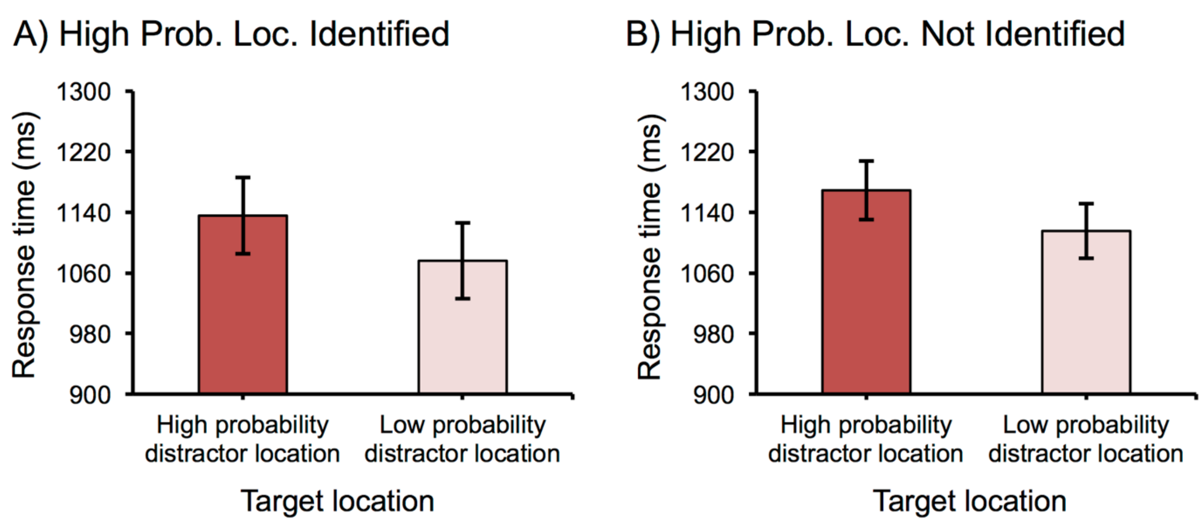

Figure S2. The mean RTs in the no distractor condition. Left panel showed the performance in correctly identified group, and right panel showed the performance in incorrectly identified group. Error bars denote ± 1 SEM.

Supplementary references

Maljkovic, V., & Nakayama, K. (1996). Priming of pop-out: II. The role of position. *Perception & Psychophysics*, *58*(7), 977–991. doi:10.3758/BF03206826