ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Habituation to onsets is controlled by spatially selective distractor expectation

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:
Massimo Turatto, Matteo Valsecchi (2023). Habituation to onsets is controlled by spatially selective distractor expectation. JOURNAL OF EXPERIMENTAL PSYCHOLOGY-HUMAN PERCEPTION AND PERFORMANCE, 49(1), 145-158 [10.1037/xhp0001078].

Availability:
This version is available at: https://hdl.handle.net/11585/902876 since: 2024-04-29

Published:
DOI: http://doi.org/10.1037/xhp0001078

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/).
When citing, please refer to the published version.

(Article begins on next page)

08 September 2026

Habituation to onsets is controlled by spatially selective distractor expectation

Massimo Turatto[1] & Matteo Valsecchi[2]

[1] Center for Mind/Brain Sciences, University of Trento, Italy

[2] Department of Psychology, University of Bologna, Italy

Word count: 11719

Corresponding author:
Massimo Turatto
University of Trento
Corso Bettini, 31
38068 Rovereto, Italy
+39.0464.808673
massimo.turatto@unitn.it

Data and study materials are available at:

https://doi.org/10.5281/zenodo.7115824

This study has not been preregistered

Abstract

Habituation to onset distractors has been shown to be stronger the higher the distractor probability. However, since in previous studies distractor probability covaried with distractor numerosity, it was unclear whether habituation was controlled by a mechanism that relies on distractor expectation (Sokolov, 1963), or by a mechanism that is merely driven by the number of stimulations delivered to the nervous system (Groves & Thompson, 1970). To address this issue, we manipulated the probability of distractor occurrence at a fixed location, without varying the number of distractors being presented. The results of Experiment 1 clearly favored the Sokolov model of habituation, showing that habituation of capture is controlled by the level of distractor expectation for the same distractors number. Experiment 2 excluded that the pattern of habituation was determined by the difference in the temporal frequency of the distractor between higher and lower distractor rates. Furthermore, the results of Experiment 3 suggested that the amount of habituation of capture is mainly controlled by the local rather than by the global rate of the onset distractor occurrence, thus indicating that habituation of capture is largely spatially specific.

Public Significance Statement

The ability to habituate to distracting stimuli is a key feature of the human brain, as it prevents important mental resources to be wasted in attending to irrelevant albeit salient repetitive distractors that may be encountered in the environment. Although in recent years different cognitive mechanisms for distractor filtering have been proposed, our study shows that the mechanism that controls habituation provides a straightforward explanation for the reduced distractor interference that is observed with practice. Crucially, as predicted by Sokolov's model, habituation to salient onset stimuli is based on the general principle of prediction or expectation of the distractor occurrence. In other words, distraction is reduced when salient stimuli become expected or familiar, and seems to be specific for the visual field location where the distractor appears.

Introduction

For humans abrupt visual onsets are among the most powerful attentional attractors (Jonides & Yantis, 1988; Lamy & Egeth, 2003; Yantis & Jonides, 1990). This unique property is explained by the fact that since the earliest stages of visual analysis, like the retina, and up to the occipital cortex, the visual system presents neurons specifically dedicated to the detection of visual transients or luminance changes. This class of neurons form the *magnocellular system*, which starting from the retina projects to both cortical (e.g., the parietal cortex) and subcortical (e.g., the superior colliculus) structures directly involved in the reflexive or automatic covert and overt orienting of attention (Breitmeyer & Ganz, 1976). Another reason why onsets elicit a strong orienting response is that their occurrence may signal the appearance of new perceptual objects in the visual field, which deserve a rapid inspection to evaluate their motivational significance (Pavlov, 1927; Yantis & Jonides, 1996).

If on the one hand the attentional call triggered by a sudden visual onset is certainly a highly adaptive response, on the other hand it has been also shown that the resulting interference on a visual discrimination task is progressively reduced as exposure to the same irrelevant onset continues, a phenomenon that we have described as habituation of capture (Pascucci & Turatto, 2015; Turatto, Bonetti, & Pascucci, 2018; Turatto, Bonetti, Pascucci, et al., 2018; Turatto et al., 2019; Turatto & Pascucci, 2016). Furthermore, in line with the idea that habituation is controlled in a quite automatic fashion by the mere stimulus repetition, the attenuation of distraction takes place also when the onset is presented while participants are not engaged in any discriminative task, but are simply passively exposed to the irrelevant distractor (Turatto, Bonetti, Pascucci, et al., 2018). Habituation of capture, however, can also be spatially selective, as

it varies according to the rate of distractor occurrence at a given location (Valsecchi & Turatto, 2022). This result parallels those emerged from studies showing that the amount of capture elicited by a feature-singleton distractor decreases as its location probability increases, and has been suggested to be governed by a statistical learning (SL) process (Ferrante et al., 2018; Goschy et al., 2014; Sauter et al., 2018; Wang & Theeuwes, 2018b).

Habituation and SL are often presented as distinct explanations to account for the attenuation of capture observed when the distractor is repeatedly presented (Chelazzi et al., 2019; Geng et al., 2019). However, as we have recently argued (Valsecchi & Turatto, in press), a SL process concerning the occurrence of irrelevant sensory stimulation is a key component of the *stimulus-model comparator* theory proposed by (Sokolov, 1960, 1963), which was specifically devised to explain the habituation of the orienting reflex (OR). The Sokolov model combines neuronal and cognitive aspects, and postulates that the OR habituation emerges when the current sensory input matches the brain's expectation, with the latter being "extrapolated" from the statistics of the past stimulation. Conversely, a mismatch between the expected and the actual stimulation triggers an OR, which serves the purpose of getting more information from the environment by aligning attention with the surprising event (Sokolov et al., 2002). The idea that habituation is driven by expectation is also common to Wagner's model of habituation (1976, 1978, 1979), according to which the strength of the response evoked by a stimulus is attenuated when it is pre-represented (or *primed* in Wagner's terminology) in STM because of its recent occurrence.

Unlike Sokolov's model, which emphasizes the role of a cognitive factor like expectation in habituation, the *dual-process theory* proposed by (Groves & Thompson, 1970) is a different model according to which this habituation is not controlled by

expectation. Rather, what matters is the mere repetition of the same stimulation, which triggers excitatory and inhibitory neural responses in specific interneurons, whose effects accumulate across repetitions. More specifically, the model essentially postulates that each stimulation is processed by two different classes of "plastic" interneurons. The *Type H* neurons decrease their response as the stimulation progresses leading to habituation, whereas the *Type S* neurons increase their response, at least initially, as the stimulus is repeated thus favoring a short sensitization. On each stimulus presentation the two classes of interneurons interact with each other, and the final outcome of this interactions determines the net organism's response to the stimulation, which in general decreases with repetitions (Thompson & Spencer, 1966). Although these classes of interneurons and the relative synaptic mechanisms have been studied and identified in the cat's spinal cord, the authors explicitly assumed that analogous mechanisms are present in the brain (Groves & Thompson, 1970). Hence, in the model there is no room for cognitive processes like prediction, expectation, and SL in the control of habituation, which is instead governed by neuronal synaptic mechanisms that change the organism's response upon stimulus repetitions (Thompson, 2009).

Although the two models share some principles, for example both rely on some form of memory, they largely diverge on the role of cognition in habituation control. For this reason, it would seem important to address which one better accounts for the habituation of distracting effect caused by abrupt onsets. At first glance, the results of studies showing that habituation of capture is affected by the onset distractor probability, either across locations (Turatto & Pascucci, 2016) or at a specific location (Valsecchi & Turatto, 2022), might seem to give support to the more "cognitive" account of habituation provided by the Sokolov model. Indeed, the model predicts habituation of capture to vary with the degree of (distractor) expectation, so that capture should be

weaker when the distractor has been estimated to occur more frequently, because the OR toward a highly expected distractor (location) is particularly suppressed. Symmetrically, habituation of capture should be weaker when the distractor occurs less frequently, as this remains a relative unexpected or surprising event. The problem with this interpretation is that the different distractor probabilities, and consequently degrees of expectation, have been obtained by varying the number of distractor-present trials, for example at a given location, with respect to the number of distractor-absent trials (e.g., Valsecchi & Turatto, in press). In this way, the onset distractor probability (and expectation) is confounded with distractor numerosity, a problem that concerns not only onset distractors studies, but potentially also feature-singleton distractors experiments (e.g., Wang & Theeuwes, 2018). Previous studies have shown that by removing those trials in which the distractor was repeated at the same location the spatial modulation of attention is still evident, thus excluding a large role of a trial-by-trial location-based priming. Still, it remains true that participants were exposed to a larger number of distractors where these occurred more frequently than less frequently (e.g., Valsecchi & Turatto, in press), and this might have determined the degree of capture observed as a function of the distractor rate.

In other words, previous evidence showing that capture is modulated by the rate of the onset distractor is ambiguous as to whether this can be accounted for by a genuine distractor expectation (Sokolov, 1963), or alternatively by the number of onsets that impinged on the visual system (Groves & Thompson, 1970). Hence, to understand which model better accounts for the habituation of capture, in Experiment 1 we manipulated, across two distinct groups of participants, the onset local rate while presenting the same number of onsets in the two conditions. Under these circumstances, the (Groves & Thompson, 1970) model would predict an equal level of

capture habituation, as the two groups of participants are exposed to the same number of onsets, which would trigger an equal number of neural activations in the visual-attention system. By contrast, the Sokolov (1960, 1963) model would predict two distinct levels of habituation of capture, as the two conditions differ in terms of distractor expectation despite the number of onsets being presented is the same.

Experiment 1

The best model accounting for the attentional capture habituation was evaluated by testing two groups of participants, each presented with an onset distractor appearing at a single fixed location the same number of trials. The distractor probability was manipulated as a between-group condition by exposing participants in the two groups to the same number of distractor-present trials, but to a different number of distractor-absent trials.

By making the distractor appearing at a single fixed location we had also the opportunity to test whether habituation of attentional capture emerges also when there is no uncertainty with regard to the distractor location. This manipulation has never been tried before, given that in previous studies habituation to onsets always involved distractors appearing at different locations (Turatto & Pascucci, 2016)Valsecchi & Turatto, 2022).

Methods

Participants

Participants (N = 88) were recruited online through the Prolific service (Prolific Academic Ltd, Oxford, UK). All of the data included in this study were collected in the first half of 2022. We required participants to be native speakers of English, to have

normal or corrected to normal vision and to run the experiment on a desktop computer. Since we did not have a priori hypotheses concerning the role of age or gender on the effects at hand, we did not impose further limitations for recruitment. The limitations concerning language skills were based on the fact that the instructions were delivered in English through the online server, but we would not assume that language or culture per se would affect the results in our paradigm, so we predict that our results would generalize to the whole adult population. The low-probability group and the high-probability group consisted of 44 participants each (low probability: mean age = 41.02 years, 31 females; high probability: mean age = 40.07 years, 14 females). The main conclusions from our study are drawn based on the main effect of Group, so we performed a sensitivity analysis targeting this particular effect. The sensitivity analysis performed using G*Power 3.1.9.7 (Faul et al., 2007) indicated that, given a mixed design with a 2-level between-participants factor and a 3-level within-participants factor, and assuming a .5 correlation between repeated measures, our total sample size of 88 participants allows to detect with 80% power between-participants effects that have η_p^2 = .0573 or larger.

All participants were informed about the general aim of the experiment, their task, and data handling procedures in the Prolific interface. They gave their consent by agreeing to be directed to the experiment url, and were paid either 3.75 or 2.5 GBP for their participation in the low- and high-probability groups, which amounted to approximately the same hourly rate, given that the experiment lasted approximately 30 minutes and 20 minutes respectively. All the experiments of the present study were carried out in accordance with the Declaration of Helsinki.

Stimuli and Procedure

The experiment was constructed using the PsychoPy3 version 2020.1.3 software (Peirce et al., 2019) and run online using the Pavlovia web hosting service (Open Science Tools Limited, Nottingham, UK).

In order to control the retinal size of the stimuli, at the beginning of the experiment we asked participants to position themselves at a distance that was a multiple of the length (84% of the screen width) of a reference segment presented on the screen. Both the stimuli and the reference segment were defined in screen coordinates, so that if a participant performed the experiment on a larger display, this would be compensated by the proportional increase in viewing distance. Stimulus sizes are reported in degrees of visual angle assuming the instructed viewing distance.

The stimuli and procedure were identical to those used in Valsecchi and Turatto (in press), and are depicted in Figure 1. The trial began with the onset of a gray central fixation spot (radius .38°), surrounded by eight gray circular placeholders. The placeholders had a radius of 3.4° and were placed at an eccentricity of 7.6° from the fixation point, and a line width of 2 pixels. In distractor-present trials, 850 ms after the onset of the initial array one of the placeholders briefly flashed for 150 ms by becoming a white circle (line width 10 pixels). In distractor-absent trials no onset was presented before the occurrence of the target, which consisted of the letter T rotated either clockwise or counterclockwise, and appearing inside one of the placeholders (1.5° wide, 2 pixels line width, same gray level as the placeholders) 1000 ms after the onset of the initial array.

The task required to press as fast as possible the “h” key if the T was oriented clockwise, and the “g” key if the T was oriented counterclockwise, but error minimization was also emphasized. The whole display disappeared after the participant’s response, or 2500 ms after target onset if no response was provided. In

case the response was incorrect or no response was provided within 2500 ms, the messages “Wrong!” or “Too Slow!” in red letters were presented for 300 ms during the 1-second inter-trial interval.

In order to avoid any confound between the locations of the target and those of the distractor, and to have an equal target distribution across locations independent from the distractor spatial probability (Valsecchi & Turatto, in press), the target and distractor appeared at non overlapping locations throughout the experiment. Out of the 8 positions defined by the placeholders, one was assigned to the distractor (chosen among the positions along the horizontal and vertical meridians), whereas the target could randomly appear at one of the 4 placeholders located in oblique directions.

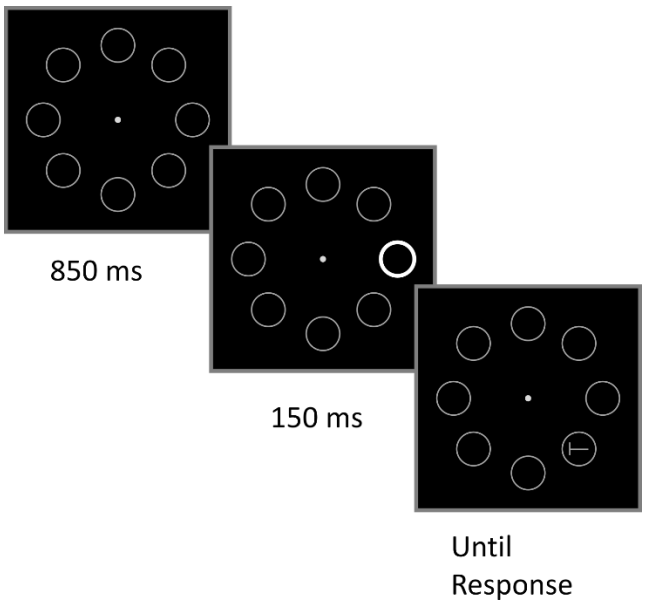

Figure 1. Stimuli and events of Experiment 1 depicting a distractor-present trial. Participants reported the orientation of the target letter T (left vs right) while trying to ignore the onset distractor. Notice that the locations of the target and distractor were completely segregated within and between trials. The distractor could appear only in a fixed placeholder (counterbalanced across participants) among the four located along the vertical or horizontal meridians, whereas the target could randomly appear in one of the four placeholders located along the oblique meridians.

The experimental design consisted of a practice phase of 28 trials, in which no distractors were presented, followed by three blocks of training trials. Specifically, during training all participants were administered 168 distractor-present trials, but whereas the low-probability group was also shown 336 distractor-absent trials, the high-probability group was exposed to only 84 distractor-absent trials. In other words, the low-probability (p = .333) group was exposed to 168 distractors out of 504 trials, whereas the high-probability (p = .666) group was exposed to 168 distractors out of 252 trials. The location where the distractor could appear, which remained fixed across the trials, was chosen for each participant between the four possible distractor locations (along the vertical and horizontal meridians) and was balanced across participants and conditions.

Transparency and openness

We report a sensitivity analysis based on our sample size, all data exclusions (if any), all manipulations, and all measures in the study. The data analysis was performed using Matlab R2021A and JASP 0.16.0.0. The data, data analysis scripts and experimental scripts for the experiments reported in this study are available for reviewers at the following link: https://www.dropbox.com/s/0ju69zln1w1sy4o/ZENODODATA2ndSubmission.zip?dl=0

The final version of the data will be made available through the Zenodo repository. This study's design and its analysis were not pre-registered.

Results

Pre-processing of RTs involved removing incorrect responses and applying an outlier-removal procedure based on Median Absolute Deviation (Leys et al., 2013) for each participant and cell design, with a threshold of 5 MADs. We further removed RTs faster than 200 ms from the analyses. Overall, we rejected 1.2% of the trials in the low-probability group, and 1.5% of the trials in the high-probability group. Statistical analyses were performed directly on capture effects, namely on the differences between RTs in distractor-present and distractor-absent trials. For each ANOVA, we report the Greenhouse-Geisser corrected results whenever a significant Mauchy's test of sphericity signaled a violation of the assumptions.

Analysis including all trials

The first analysis that we performed, both on RTs and error rates, was based on the whole dataset of trials, namely it included, by design, twice as many trials in the low-probability condition compared to the high-probability condition, but the same number of distractors.

Figure 2a depicts the onset distraction effect on RTs measured in the two conditions across blocks[1]. The first thing to note is that the overall level of distraction was clearly weaker in the high-probability condition compared to the low-probability

[1] We choose to directly report the analysis of the capture effects (i.e., the RT difference between distractor-absent and distractor-present trials) because distractor-absent RTs were largely unaffected by our experimental manipulations. On average distractor-absent trials RTs decreased from 578 to 574 ms from Block 1 to Block 3, and were 578 and 572 ms in the low- and high-probability conditions. With distractor-absent RTs neither the main effect of Block or Condition, or their interaction was significant (all *p*s > .532). Additionally, RTs in distractor-absent trials were not significantly affected by whether the target appeared in a location close to the distractor or far from it (as could be expected if the spatial suppression of the distractor location extended to the nearby space). RTs were on average faster by 1.4 ms and slower by 0.3 ms when the target was in a placeholder near the distractor position in the high- and low-probability conditions respectively. Neither the effect of Target Position, the effect of Condition, nor their interaction were significant (all *p*s > .758).

condition. In addition, distraction decreased across blocks in both conditions, but the decrement was more marked in low-probability condition. To substantiate these observations, mean correct RTs were entered in a mixed ANOVA with Distractor probability (low vs. high) as between-participants factor, and Block (1-3) as within-participants factor. The effect of Distractor Probability, $F(1,86) = 20.57$, $p < .001$, $\eta_p^2 = .193$, of Block, $F(1.852,159.247) = 12.19$, $p < .001$, $\eta_p^2 = .124$, and their interaction, $F(1.852,159.247) = 4.429$, $p = .016$, $\eta_p^2 = .049$ were all significant. Further analyses were conducted in order to qualify the significant interaction. First we performed two within-participants one-way ANOVA separately for the two probability conditions, with Block as factor, which was significant in the low-probability condition, $F(2,86) = 17.172$, $p < .001$, $\eta_p^2 = .285$, but not in the high-probability condition $F(2,86) = 1.071$, $p = .347$, $\eta_p^2 = .024$. Secondly, this difference among distractor-probability conditions were confirmed by pairwise comparisons (*t* tests, Bonferroni corrected) showing that capture was weaker in the high-probability condition than in the low-probability condition in the first block, $t(86) = 4.297$, $p < .001$, $d = 0.916$, in the second block, $t(86) = 4.4926$, $p < .001$, $d = 0.957$, and in the third block, $t(86) = 2.496$, $p = .043$, $d = 0.532$.

The results for error rates (computed as the difference between % errors in distractor-present and distractor-absent trials) are presented in Figure 2b. The Distractor probability x Block analysis failed to provide any significant result (all *p*s > .208), thus excluding the possibility that the RT patter was due to a speed-accuracy trade-off. In general, accuracy was very high, with an error rate of 2.72 % in the low-probability condition and 3.95% in the high-probability condition.

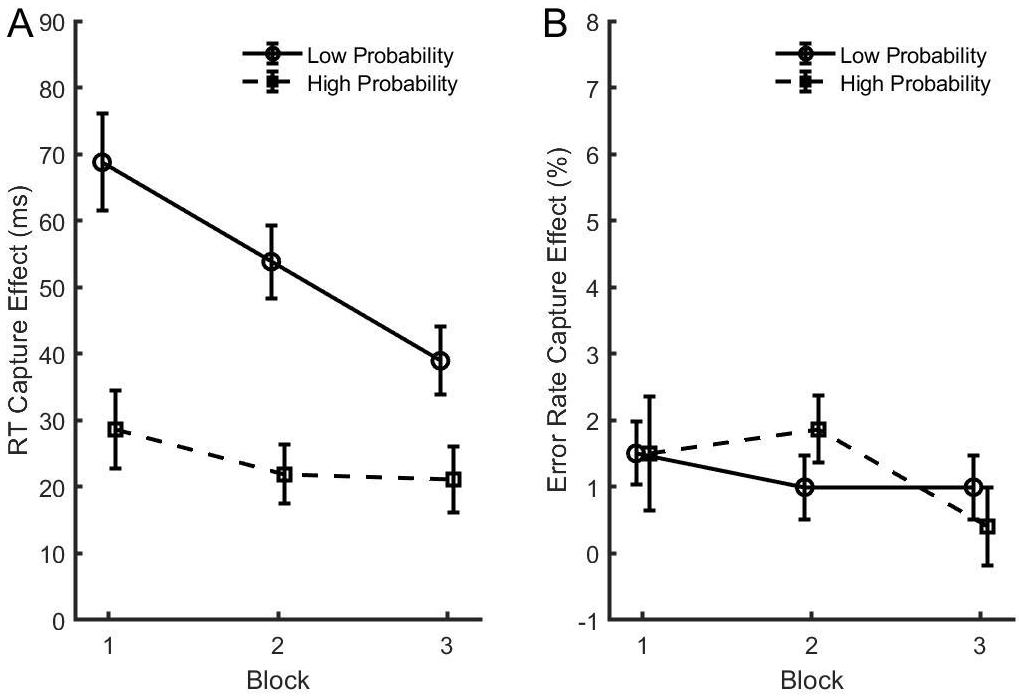

Figure 2. Capture effect, measured as distractor-absent trials minus distractor-present trials differences in either RTs (left panel) or error rates (right panel), as a function of distractor probability and block. The RT data show the habituation of capture functions for the high- and low-probability conditions, with habituation being stronger and closer to the asymptotic level in the former than in the latter condition. The error rates data show minimal systematic differences across conditions, thus excluding any speed-accuracy trade off. Error bars represent between-participants Standard Errors of the Mean (SEM).

The RT pattern emerged was in agreement with the possibility that the onset distracting effect is subject to habituation (Turatto & Pascucci, 2016). Two observations support this conclusion: first, distraction was weaker the higher the rate or probability of onset occurrence; and second, distraction overall decreased across blocks as onsets presentation progressed. The steepness of the capture decrement was larger in the low-probability condition than in the high-probability condition, which suggests that at the higher rate of onset presentation habituation of capture already almost reached the asymptotic level in the first block. By contrast, the habituation function was still in its descending part when the onset appeared at a lower rate, as in the low-probability

condition (also see the General discussion section). In other words, for the higher rate of distractor occurrence habituation was very fast, which also indicated that the estimation of the distractor statistics required relatively few trials, in agreement with our previous findings (Valsecchi & Turatto, 2021). Crucially, however, since participants were exposed to the same number of onset distractors in the two probability conditions, the two different amounts of habituation and relative functions support the (Sokolov, 1960, 1963) model rather than the (Groves & Thompson, 1970) model, indicating that distraction attenuation was driven by onset expectation rather than by the number of onset stimuli being delivered.

Analysis excluding distractor-repetition trials

Although the previous analysis indicated that habituation was controlled by distractor expectation, an alternative possibility is that capture was weaker in the high-probability condition compared to the low-probability condition because of inter-trial priming. One may argue, indeed, that participants were more likely to encounter two consecutive distractor-present trials in the same (and only) location in the high-probability condition than in the low-probability condition, given that in the former case the same number of distractors were presented interleaved with a smaller number of distractor-absent trials. To rule out the possibility that inter-trial priming accounted for the different patterns of capture emerged as a function of distractor probability, we re-analyzed the data by including only trials in which the distractor was not repeated (for a similar approach see, for example, Ferrante et al., 2018; Wang & Theeuwes, 2018b)[2]. The RTs results are depicted in Figure 3. The corresponding ANOVA yielded a

[2] In principle, the influence of exposure to distractors in the trial history could extend beyond trial N-1. In order to investigate this possibility, limitedly to the data from the low-probability condition in Experiment 1, we isolated RTs corresponding to distractor-present trials that were preceded by a

significant main effect of Distractor probability, $F(1,86) = 16.363$, $p < .001$, $\eta_p^2 = .16$, a significant main effect of Block, $F(2,172) = 16.335$, $p < .001$, $\eta_p^2 = .16$, and a significant two-way interaction, $F(2,172) = 3.23$, $p = .042$, $\eta_p^2 = .036$. We further performed two within-participants one-way ANOVA separately for the two probability conditions with Block as factor, which was significant in the low-probability condition, $F(2,86) = 20.38$, $p < .001$, $\eta_p^2 = .322$, but not in the high-probability condition $F(2,86) = 2.165$, $p = .121$, $\eta_p^2 = .048$. Pairwise comparisons (*t* tests, Bonferroni corrected) showed that capture was weaker in the high-probability than in the low-probability condition in the first block, $t(86) = 4.002$, $p < .001$, $d = 0.853$, in the second block, $t(86) = 3.475$, $p = .002$, $d = 0.74$, but only close to significance in the third block, $t(86) = 2.405$, $p = .054$, $d = 0.512$.

This pattern of results closely mimicked that of the previous analysis including all trials, and confirmed that above any possible role of inter-trial priming, habituation of capture was mainly driven by onset expectation. Again, analyzing the error rates as a

sequence of 0 to 4 distractor-absent trials. Combining all blocks, this adds on average to 49 trials directly preceded by a distractor, with individual differences due to the randomness of trials ordering. The number of available trials decreases as the sequence of absent trials becomes longer, with an average value of 12.1 trials for sequences of 4 distractor-absent trials. The average RTs in the 0-4 sequences were 639, 647, 625, 642 and 632 ms, respectively, which shows that there was no indication for a general increase or decrease in RTs as a function of the sequence length. The corresponding data were submitted to a linear-mixed model regression analysis with sequence length as fixed-effect predictor and participant as random factor using the lme4 library in R (Bates et al., 2015). This resulted in a slope of -1.91 ms/trial for sequence length, and the corresponding 95% confidence interval [-4.783 0.968] included 0, thus confirming that the sequence length had no reliable effects on RTs.

function of Block and Probability did not produce any significant result (all *p*s > .151)

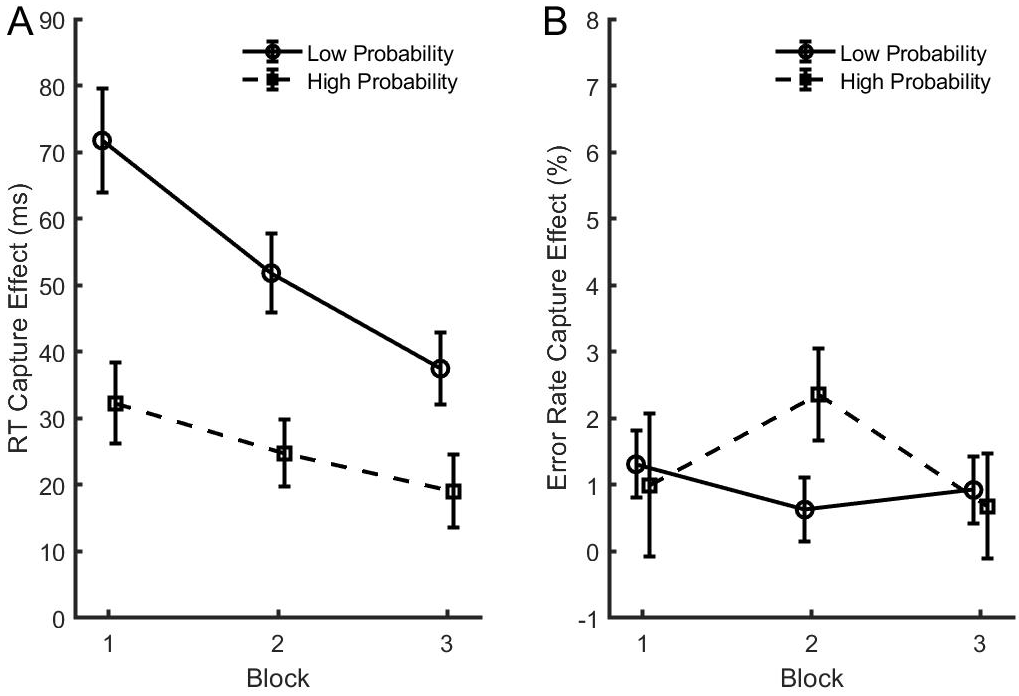

Figure 3. Capture effect, measured as distractor-absent trials minus distractor-present trials differences in either RTs (left panel) or error rates (right panel), as a function of distractor probability and block. The RT data show the habituation of capture functions for the high- and low-probability conditions, with habituation being stronger and closer to the asymptotic level in the former than in the latter condition. The error rates data show minimal systematic differences across conditions, thus excluding any speed-accuracy trade off. Error bars represent between-participants SEM.

Experiment 2

The results of Experiment 1 indicated that the larger habituation of capture observed when the distractor appeared more frequently was controlled by expectation, as the number of distractors was the same in the high- and low-probability conditions. This interpretation was substantiated also by the fact that the habituation patterns did not change when we removed any possible inter-trial priming effect that may have arisen from distractor-repetition trials. Yet, although this analysis removed the effects of distractor repetition trials from RTs, still in the high-probability condition

participants were exposed to an irrelevant onset occurring at a higher temporal frequency as compared to the low-probability condition. Given that habituation is largely affected by the temporal frequency of stimulation (Thompson, 2009), in the present experiment we wanted to rule out the possibility that the results of Experiment 1 were contaminated by the different distractor temporal frequency associated with the different probability conditions.

To this aim, the distractor was shown only at the high probability (at a fixed location), but with a temporal frequency equivalent to the low probability. If habituation of capture is mainly controlled by distractor expectation, then we envisage to replicate the same RT pattern of the high-probability condition of Experiment 1; conversely, if habituation is determined by the distractor temporal frequency, we expect to replicate the results of the low-probability condition.

Methods

Participants

Participants (N = 44, mean age = 36.09, 20 females) were recruited following the same procedure used for Experiment 1. Compensation was 3.75 GBP for an approximate experiment duration of under 30 min. Notice that since the comparisons performed on the data involve a mixed 2 (between) x 3 (within) design, the same sensitivity analysis applies as described for Experiment 1. None of the datasets had to be discarded for not reaching the accuracy level of 80%.

Stimuli and Procedure

Stimuli and procedures were as in Experiment 1. In particular, the distractor probability was set to .666, but the Inter Trial Interval (ITI), i.e., the interval between the response in a given trial and the appearance of the fixation point and placeholders

in the next trial, was lengthened from 1000 to 3600 ms, in order to have the distractor appearing, on average, with the same temporal frequency of the low-probability condition of Experiment 1. The current ITI was calculated as follows: in Experiment 1 the pre-target onset time was 1000 ms, the RT (including wrong trials) in the low-probability condition was on average approximately 600 ms, and the ITI was 1000 ms, yielding a total trial duration of 2600 ms. Hence, in the present experiment the ITI was increased by 2600 ms, which ensured that the distractor appeared at the high probability but approximately with the same temporal frequency of the low-probability condition in Experiment 1.

The experimental design was identical to the one of the high-probability condition of Experiment 1 (three blocks of 84 trials, for a total of 168 distractor-present trials).

Results

Analysis including all trials

Figure 4 depicts the average onset distracting effect in Experiment 2 (high probability with low temporal frequency), along with the results from the low- and high-probability conditions of Experiment 1 for comparisons. At a first inspection it seems evident that the pattern of results of Experiment 2 closely resembled the high-probability condition of Experiment 1, but was clearly different from the low-probability condition of the same experiment. To substantiate these observations, we compared in two separate ANOVAs the results of Experiment 2 with each distractor-probability condition of Experiment 1. In each ANOVA the within-participants factor was Block and between-participants factor was Experiment.

The ANOVA comparing the RTs in Experiment 2 and in the high-probability condition of Experiment 1 yielded a significant main effect of Block, $F(2,172) = 4.828$, $p = .009$, $\eta_p^2 = .053$, but neither the main effect of Experiment, $F(1,86) = 0.367$, $p = .546$, $\eta_p^2 = .004$, nor the two-way interaction, $F(2,172) = 1.04$, $p = .356$, $\eta_p^2 = .012$, were significant. A Bayes Factor analysis yielded evidence against the inclusion of the factor Experiment in the data model ($BF_{01} = 3.719$). The corresponding ANOVA performed on the error rates did not yield any significant results (all *p*s > .103).

By contrast, the ANOVA comparing the RTs in Experiment 2 and in the low-probability condition of Experiment 1 yielded a significant main effect of Block, $F(2,172) = 13.923$, $p < .001$, $\eta_p^2 = .193$, and of Experiment, $F(1,86) = 21.704$, $p < .001$, $\eta_p^2 = .202$, but no significant two-way interaction, $F(2,172) = 2.754$, $p = .067$, $\eta_p^2 = .031$. The corresponding analysis performed on the error rates did not yield any significant results (all *p*s > .076). The overall error rate in Experiment 2 was 2.72%.

The results clearly indicated that the main factor contributing to the patterns of habituation of capture observed in Experiment 1 was distractor expectation, not the distractor temporal frequency which covaried with distractor probability. Indeed, in Experiment 2, where the distractor appeared at the high-probability but with the same temporal frequency of the low-probability condition, the amount of capture was equal to that found in the high-probability condition of Experiment 1.

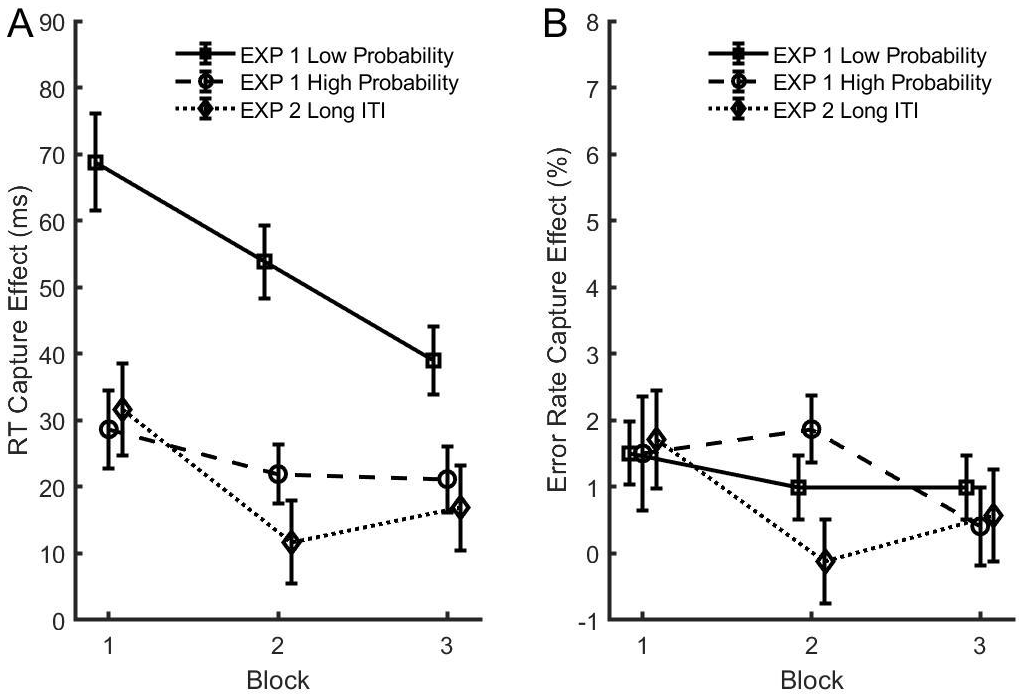

Figure 4. Capture effect measured in Experiment 1 (single distractor location: low- and high-probability) and in Experiment 2 (high-probability but long ITI) as distractor-absent trials minus distractor-present trials differences in either RTs (left panel) or error rates (right panel), as a function of distractor probability and block. The RT data (left) show that the capture effect in Experiment 2 was similar to that in the high-probability condition of Experiment 1, which was matched in terms of distractor probability, but smaller than the one in the low-probability condition of Experiment 1, which was matched in terms of distractor temporal frequency, suggesting that distractor expectation controlled the amount of capture. The error rate data (right) show minimal systematic differences across conditions, thus excluding any speed-accuracy trade off. Error bars represent between-participants SEM.

Analysis excluding distractor-repetition trials

Once again, we performed the analyses conducted above by including only distractor-present trials preceded by a distractor-absent trial in order to control for a possible effect of distractor repetition. The corresponding results are presented in Figure 5.

The ANOVA comparing the RTs in Experiment 2 and in the low-probability condition of Experiment 1 yielded a significant main effect of Experiment, $F(1,86) = 12.513$, $p < .001$, $\eta_p^2 = .127$, a significant main effect of Block, $F(2,172) = 15.535$, $p < .001$, $\eta_p^2 = .153$, but no significant two-way interaction, $F(2,172) = 0.426$, $p = .654$, $\eta_p^2 = .005$. The corresponding analysis performed on the error rates did not yield any significant results (all *p*s > .358).

Similar to the analysis including distractor repetition trials, the ANOVA comparing the RTs in Experiment 2 and in the high-probability condition of Experiment 1 yielded a significant main effect of Block, $F(2,172) = 6.025$, $p = .003$, $\eta_p^2 = .065$, but neither a significant main effect of Experiment, $F(1,86) = 0.016$, $p = .901$, $\eta_p^2 < .001$, nor significant two-way interaction, $F(2,172) = 0.745$, $p = .476$, $\eta_p^2 = .009$. A Bayes Factor analysis provided evidence against the inclusion of the factor Experiment in the data model ($BF_{01} = 4.929$). The corresponding ANOVA performed on error rates did not yield any significant results (all *p*s > .276).

In sum, the results closely paralleled those obtained including all trials, confirming that habituation was mostly controlled by distractor expectation regardless of the temporal frequency of distractor presentation.

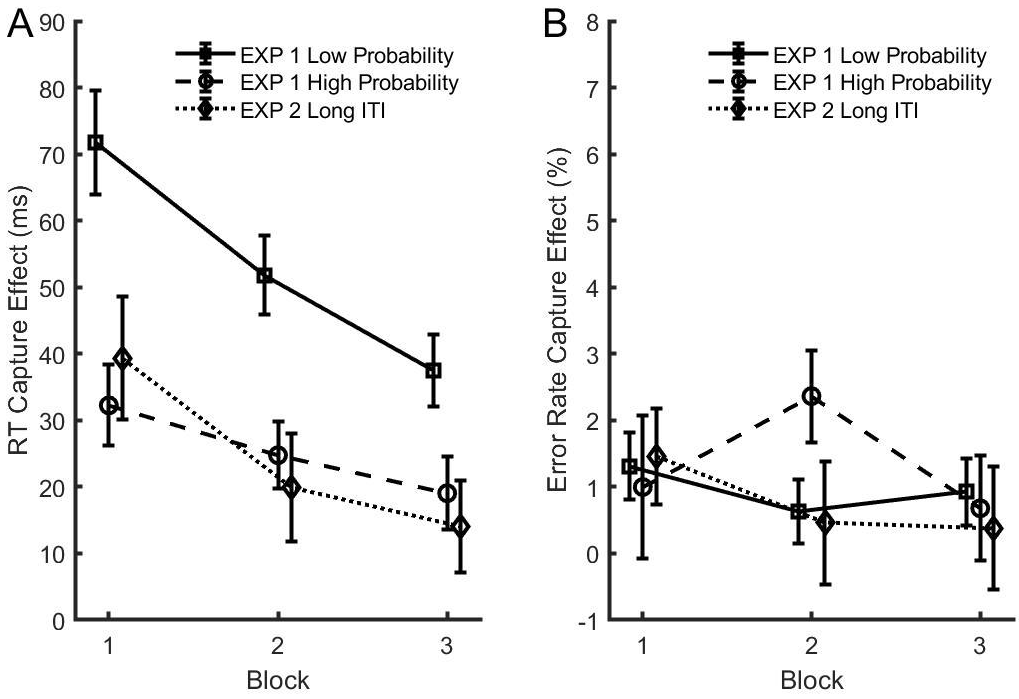

Figure 5. Capture effect in no distractor-repetition trials in Experiments 1 (low- and high-probability) and 2 (high-probability and long ITI). Values are distractor-absent trials minus distractor-present trials differences in either RTs (left panel) or error rates (right panel), as a function of distractor probability and block. The RT data (left) confirm the finding from the whole dataset of trials, whereby the capture effect in Experiment 2 was similar to the high-probability condition of Experiment 1. The error rate data (right) show minimal systematic differences across conditions, thus excluding any speed-accuracy trade off. Error bars represent between-participants SEM.

Experiment 3

Experiment 1 attested that habituation to a peripheral onset is mainly governed by distractor expectation rather than by the number of distractors presented, thus supporting the Sokolov model (1960, 1963) rather than the Groves and Thompson (1970) model. This conclusion was additionally supported by the results of Experiment 2, showing that the effect of distractor probability on habituation of capture was not accounted for by the different distractor temporal frequency in the high- and low-probability conditions. In addition, because in Experiment 1 the distractor appeared with a fixed probability and location, this also suggests that habituation to onsets could

be spatially selective, as recently shown by Valsecchi and Turatto (2022). However, in Experiment 1 the local distractor probability coincided also with the overall distractor probability, and therefore it was not clear whether habituation was mainly driven by the local or global distractor probability (also see, Valsecchi & Turatto, 2021).

To evaluate the contribution of the local and global onset expectation, in the present experiment the distractor could appear at two possible distinct locations with the same local low-probability (p = .333). However, since there were two distractor locations, the global distractor probability (p = .666) was identical to the high-probability condition of Experiment 1. Hence, if habituation to onsets was mainly driven by spatially specific distractor expectation, in Experiment 3 we predicted to observe the same habituation pattern of the low-probability condition of Experiment 1, where the distractor appeared at a single location. By contrast, if habituation to onsets in was due to the overall level of distractor expectation, in Experiment 3 we predicted the same habituation pattern of the high-probability condition of Experiment 1.

Methods

Participants

Participants (N = 44, mean age = 46.92, 28 females) were recruited following the same procedure used for Experiment 1. Compensation was 3.75 GBP for an approximate experiment duration of under 30 min. Notice that since the comparisons performed on the data involve a mixed 2 (between) x 3 (within) design, the same sensitivity analysis applies as described for Experiment 1. None of the datasets had to be discarded for not reaching the accuracy level of 80%.

Stimuli and Procedure

Stimuli and procedures were as in Experiment 1 (see Figure 1), except for the fact that across trials the distractor could appear with equal probability (p = .333) at one of two possible locations, opposite to each other. The couples of locations (right-left or up-down) were balanced across participants. As in Experiment 1, targets appeared with equal probability in each of the oblique locations.

The experimental design consisted of a practice phase of 28 trials, in which no distractors were presented, followed by three blocks of 168 trials each, for a total of 504 trials. Excluding practice, participants were overall submitted to 168 distractor-absent trials and 336 distractor-present trials (global distractor probability = .666), with distractors appearing 168 times at each of the two designated locations (local distractor probability = .333). This means that when the results of Experiment 1 and 2 are compared, the number of distractors presented at a single location in each block of trials is equivalent.

Results

Analysis including all trials

Figure 6 depicts the average onset distracting effect from the two low-probability locations of Experiment 3, along with the results from the low- and high-probability conditions with a single location of Experiment 1 for comparisons. At a first inspection it seems evident that the pattern of results of Experiment 3 closely resembled the low-probability condition of Experiment 1, but was markedly different from the high-probability condition of the same experiment. To substantiate these observations, we compared in two separate ANOVAs the results of Experiment 3 with each distractor-probability condition of Experiment 1. In each ANOVA the within-participants factor was Block and between-participants factor was Experiment.

The ANOVA comparing the RTs in Experiment 3 and in the low-probability condition of Experiment 1 yielded a significant main effect of Block, $F(2,172) = 20.359$, $p = .01$, $\eta_p^2 = .191$, but neither the main effect of Experiment, $F(1,86) = 1.672$, $p = .2$, $\eta_p^2 = .019$, nor the two-way interaction, $F(2,172) = 2.094$, $p = .126$, $\eta_p^2 = .024$, were significant. A Bayes Factor analysis indicated that the present data do not provide clear evidence against (or in favor of) the inclusion of the factor Experiment in the data model ($BF_{01} = 1.781$). Again, the corresponding ANOVA performed on the error rates did not yield any significant results (all ps > .094).

By contrast, the ANOVA comparing the RTs in Experiment 3 and in the high-probability condition of Experiment 1 yielded a significant main effect of Experiment, $F(1,86) = 11.211$, $p < .001$, $\eta_p^2 = .115$, a significant main effect of Block, $F(1.858,159.817) = 4.864$, $p = .01$, $\eta_p^2 = .054$, but no significant two-way interaction, $F(1.858,159.817) = 0.596$, $p = .54$, $\eta_p^2 = .007$. The corresponding analysis performed on the error rates did not yield any significant results (all ps > .067). The overall error rate in Experiment 3 was 3.62%.

Overall, the results indicated that the local level of distractor probability, not the global one, was the major factor contributing to the RT pattern observed in Experiment 3, namely to habituation to onsets, when participants were exposed to distractors randomly appearing at two locations.

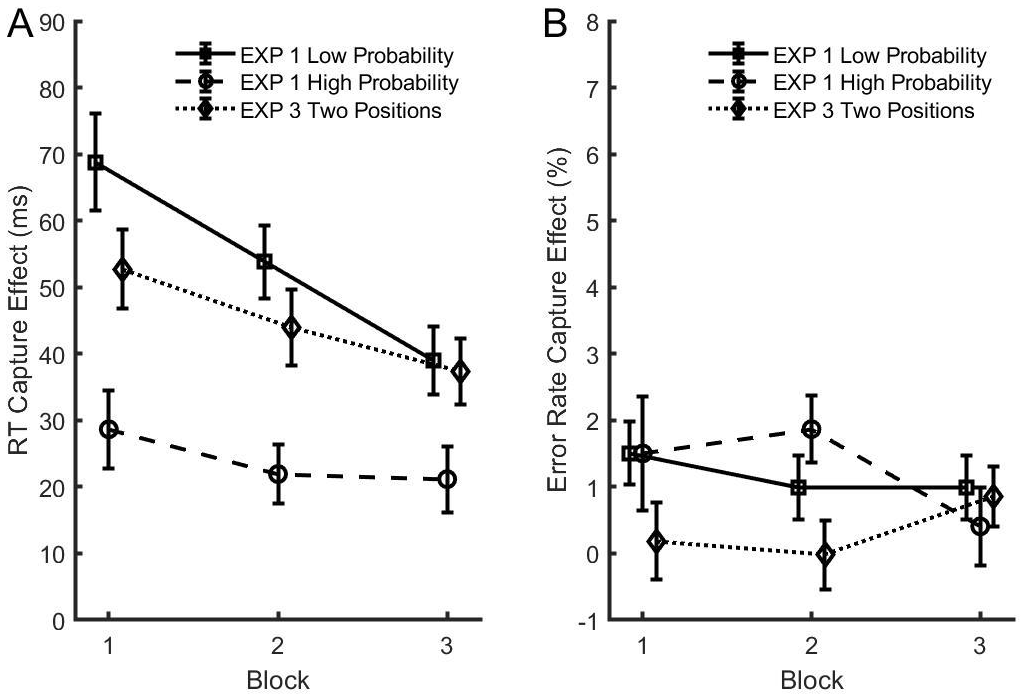

Figure 6. Capture effect measured in Experiment 1 (single distractor location: low- and high-probability) and in Experiment 3 (two distractor locations, each low-probability) as distractor-absent trials minus distractor-present trials differences in either RTs (left panel) or error rates (right panel), as a function of distractor probability and block. The RT data (left) show that the capture effect in Experiment 3 was similar to that in the low-probability condition of Experiment 1, which is matched in terms of local distractor probability, thus suggesting that habituation is mostly determined by local distractor expectation. The error rate data (right) show minimal systematic differences across conditions, thus excluding any speed-accuracy trade off. Error bars represent between-participants SEM.

Analysis excluding distractor-repetition trials

In order to exclude a possible role of inter-trial priming, we performed the same analyses conducted above by including only distractor-present trials preceded by a distractor-absent trial. The corresponding results are presented in Figure 7.

The ANOVA comparing the RTs in Experiment 3 and in the high-probability condition of Experiment 1 yielded a significant main effect of Experiment, $F(1,86) = 10.69$, $p = .002$, $\eta_p^2 = .111$, a significant main effect of Block, $F(2,172) = 8.656$, $p < .001$, $\eta_p^2 = .091$, but no significant two-way interaction, $F(2,172) = 0.771$, $p = .464$, $\eta_p^2 = .009$.

The corresponding analysis performed on the error rates did not yield any significant results (all *p*s > .213).

Similar to the analysis including switch trials, the ANOVA comparing the RTs in Experiment 3 and in the low-probability condition of Experiment 1 yielded a significant main effect of Block, $F(2,172) = 24.74$, $p < .001$, $\eta_p^2 = .223$, but neither a significant main effect of Experiment, $F(1,86) = 0.716$, $p = .4$, $\eta_p^2 = .008$, nor a significant two-way interaction, $F(2,172) = 0.897$, $p = .41$, $\eta_p^2 = .01$. A Bayes Factor analysis provided evidence against the inclusion of the factor Experiment in the data model ($BF_{01} = 2.933$). The corresponding ANOVA performed on error rates did not yield any significant results (all *p*s > .296).

In sum, the results closely paralleled those obtained including all trials, confirming that habituation seems mostly controlled by the local distractor expectancy, even when possible additional inter-trial priming effects are discounted.

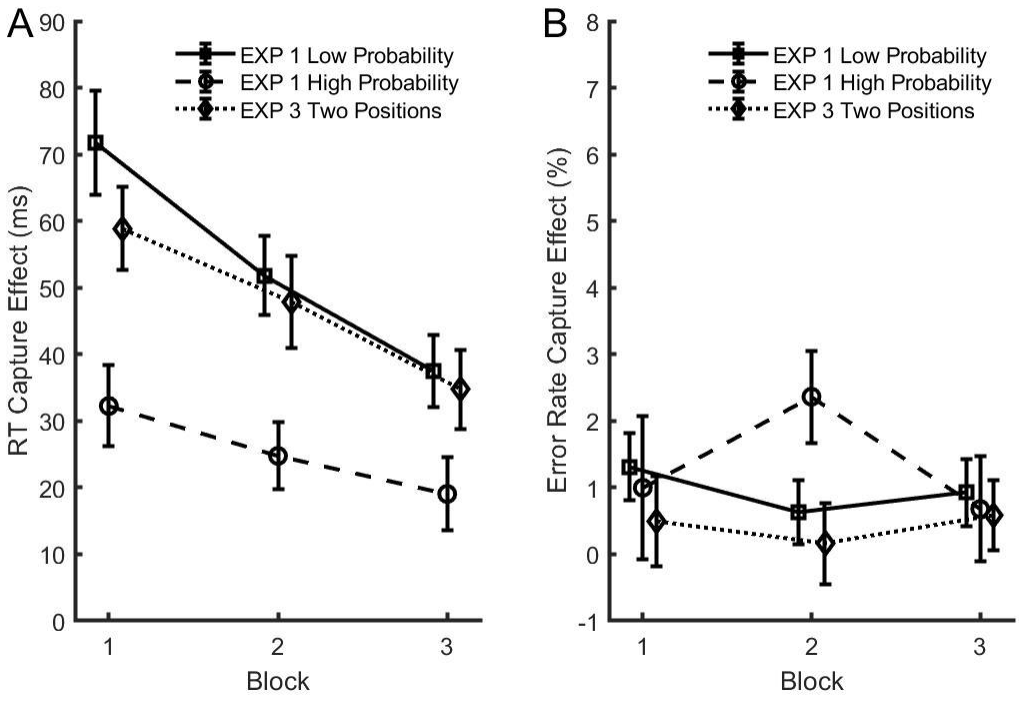

Figure 7. Capture effect in switch trials in Experiments 1 (low- and high-probability) and 2 (two positions, each low-probability). Values are distractor-absent trials minus distractor-present trials

differences in either RTs (left panel) or error rates (right panel), as a function of distractor probability and block. The RT data (left) confirm the finding from the whole dataset of trials, as the capture effect in Experiment 3 was similar to that in the low-probability condition of Experiment 1. The error rate data (right) show minimal systematic differences across conditions, thus excluding any speed-accuracy trade off. Error bars represent between-participants SEM.

General discussion

Habituation describes the progressive waning of a response to a repeated (usually irrelevant) stimulation, a decrement that is not accounted for by sensory or motor fatigue (Thompson, 2009). In recent years, we have shown that the distraction elicited by an irrelevant peripheral onset decreases as exposure to the salient event progresses, and that such decrement exhibits some key features of habituation (Turatto & Pascucci, 2016), including the fact that it is specific for the context in which it takes place (Turatto, Bonetti, & Pascucci, 2018; Turatto et al., 2019). Habituation, however, is just a behavioral phenomenon that requires an explanation.

In support of Sokolov's model: the role of prediction and information entropy in attentional orienting and its habituation

Three main models or theories of habituation have been historically proposed, although more recent computational models also exist (Ramaswami, 2014; Vogel et al., 2019). Sokolov's (1963) and Wagner's (1976) models invoke cognitive factors to explain the response attenuation, whereas in the Groves and Thompson (1970) model habituation results from the synaptic interaction between two different classes of interneurons (*Type H* and *Type S*). In particular, Sokolov's theory was elaborated to account for the habituation of the OR, of which the orienting of attention is one key component. Essentially, the theory assumes that the OR, and hence the attentional

orienting, occurs whenever the organism is presented with a *surprising* or *unexpected* stimulus. The latter is surprising to the extent that it does not match the *expected one,* which is extrapolated from the statistics of past events. By contrast, the more the input matches the expectation the more the response normally evoked by the stimulus is suppressed on any subsequent reiteration, leading to habituation (Valsecchi & Turatto, 2022). In other words, as stressed by Bridgeman (2013), Sokolov's model anticipated the general notion of *prediction* as a key principle in neuroscience, and more specifically the idea of *prediction error* (the difference between input and expectation) as the driving force of learning. This notion then became central in conditioning studies (Rescorla & Wagner, 1972), and more recently in different domains of cognitive neurosciences (Clark, 2013; Den Ouden et al., 2012). According to this view, the OR reflex is seen as a behavioral and cognitive strategy to acquire information from the environment, in line with the idea that unexpected events have information content, or *entropy,* as postulated by the *Information theory* devised by Shannon (1948). Conversely, habituation arises when through repetition the input confirms the brain's prediction, namely when the amount of *entropy* carried by the sensory input progressively becomes negligible (Shannon, 1948; Sokolov, 1960; Sokolov et al., 2002). In agreement with this interpretation, Allenmark et al. (2022) have recently shown that the amount of interference caused by a feature-singleton distractor appearing at a given location increases linearly with the amount of information entropy (determined by the distractor probability) associated with the distractor.

By contrast, the dual-process theory proposed by Groves and Thompson (1970) simply assumes that the response is attenuated when an identical stimulation is repeatedly delivered to the nervous system. This is because each stimulation activates the same neural pathways and synaptic transmissions, whose activity physiologically

declines with repetition (the *Type H* interneurons). So, habituation is explained by the observation that in some interneurons repetition leads to a decline in the physiological response. For this model the only relevant variable is the number of identical or similar stimuli that are delivered, in a given time interval, to the nervous system, and there is no room for any cognitive process like expectation, SL or alike.

In our previous studies showing an effect of distractor probability on the habituation of capture the number of distractors presented in the high vs low-probability conditions differed (Valsecchi & Turatto, 2022), similarly to what is commonly found in SL of feature-singleton distractor location studies (e.g., Ferrante et al., 2018; Wang & Theeuwes, 2018b). Hence, it was unclear which model better accounted for the stronger habituation of capture observed with the higher onset rate as compared to the lower onset rate, either across locations (Turatto & Pascucci, 2016), or at specific locations (Valsecchi & Turatto, 2022). The results of the present experiment were largely in favor of the Sokolov model, as we found a stronger habituation to onsets in the high- compared to the low-probability condition when they were paired in terms of number of distractors presented at the same location. The analysis in which we removed the effect of onset repetition across two consecutive trials (e.g., Ferrante et al., 2018; Wang & Theeuwes, 2018b), confirmed that the different patterns of habituation observed in the different probability conditions were not due to an inter-trial priming effect generated by distractor repetition trials, which are more frequent when a larger number of distractor are presented as in the high-probability condition. So, even when any facilitatory effect of consecutive distractor repetitions on RTs was discounted, a stronger habituation of capture was found where the distractor was more expected to occur, as predicted by the Sokolov model.

We have so far stressed the role of expectation in governing the habituation of the OR, arguing that the more a distractor is expected the less information it conveys and the more it is ignored (Sokolov et al., 2002). However, Allenmark et al. (2022) have argued that habituation of capture is simply governed by the discrepancy between the learned statistics and the current input, without invoking any role for expectation. The difference between our view and that of Allenmark and colleagues is an old question already discussed by Sokolov, who explicitly assumed that the neural model of the irrelevant sensory input should not be conceived as a "*passive stable engram",* but as a mechanism that can extrapolate or predict the incoming input (Sokolov, 1963). Hence, in line with Sokolov's proposal we have argued that expectation plays a key role in habituation, a conclusion that is also supported by evidence showing that habituation of capture is context specific, which is explained by the fact that context acts as a predictor of the distractor occurrence (Turatto, Bonetti, & Pascucci, 2018; Turatto et al., 2019). However, further studies are needed to clarify whether expectation does indeed play a key role in habituation of capture, or whether a more "passive" mechanism like that invoked by Allenmark et al. (2022) is sufficient to control habituation.

Some specifications and an alternative account to habituation

Our results favor a cognitive model of habituation of capture (and the OR in general) like that proposed by Sokolov (1960; 1963). However, this is not to say that such model always provides the best explanation for any response decrement obtained with an iterative stimulation. Indeed, it is very likely that the mechanism underlying the observed habituation may depend also on the type of response under investigation and the animal species considered. For example, in the *C. elegans*, a small worm with about 300 neurons, which is an animal model of paramount importance in neuroscience for

the study of habituation and other types of learning (Rankin, 2004), habituation of the tap-withdrawal response (and its recovery) as a function of the inter-stimulus-interval (ISI) seems better accounted for by the Groves and Thompson (1970) model. This model emphasizes the importance of the ISI directly affecting the memory decay of the stimulus representation, and therefore the response suppression. By contrast, the results are less in agreement with a Sokolovian-like model based on the expected pattern of stimulation (Broster & Rankin, 1994; Rankin & Broster, 1992). Furthermore, by claiming that habituation of capture was controlled by distractor expectation rather than by the delivered number of distractors we do not mean that the latter is not, in absolute terms, an important factor contributing to habituation in general. Indeed, first it is quite obvious that there is no habituation unless the stimulus is repeated, and second it is also evident that the longer the series of stimulation the higher the probability that learning is complete and that habituation reaches its maximum, namely an asymptotic level. However, the distinction between distractor numbers vs distractor expectation arises when the covert attentional capture response, and its variation, are measured. This is because the distractor effect is measured (only indirectly) by subtracting the average RTs in distractor-absent trials from distractor-present trials. The impact of the distractor rate on habituation of capture can then be evaluated by manipulating the number of distractor-present trials over a fixed number of trials (Valsecchi & Turatto, 2022), which however introduces the confound of the number of distractors being presented, or by varying the distractor rate while keeping the same number of distractors in the different probability conditions as done here. It remains true, however, that although distractor expectation is the main factor controlling the degree of habituation of capture, the absolute number of distractors delivered can still be a factor affecting habituation. Indeed, as it is especially evident in the low-probability

condition the distractor interference continued to slowly decline as exposure to the onset progressed.

Experiment 2 also showed that the stronger habituation of capture found in the high-probability condition could not be accounted for by the fact that in this condition the onset appeared at a higher temporal frequency compared to the low-probability condition. In principle the distractor temporal frequency might have been a factor given that one of the main robust characteristics of habituation is its dependance on the temporal rate of stimulation (e.g., Broster & Rankin, 1994; Geer, 1966; Thompson, 2009). However, when the distractor was presented at the higher probability but with a temporal frequency equivalent to the lower probability, the pattern of habituation resembled that of the higher probability condition. This is not to claim, however, that in general the temporal frequency of distractor presentation is irrelevant in habituation of capture, but only that within the limits of the ISI manipulation used in the present paradigm the distractor temporal frequency played no major role in the observed patterns of habituation of capture.

Although we interpreted the distractor interference decrement as an instance of habituation governed by a Sokolovian-like model relying on distractor expectation, an alternative account could be considered. The different amount of capture observed in the two probability conditions could be explained by the fact that they differed in terms of the number of distractor-absent trials interspersed among distractor-present trials. According to Müller et al. (2009), during the attentive task participants with practice develop a distractor-handling "routine" (i.e. a suppressive mechanism) to reduce the distractor interference. Specifically, this routine would be learned and automatized in distractor-present trials, but prone to forgetting in distractor-absent trials, which by definition are more frequent the lower the distractor rate. Hence, the more frequent the

distractor-absent trials the less “available” or “automatized” this routine becomes, which leads to a greater distractor interference, as for instance in our Experiment 1. In line with the proposed mechanism, Müller et al. (2009) found that a feature-singleton distractor interfered less on a given trial N if it was present also on trial N-1, which offered an occasion to consolidate or practice the distractor-handling routine. Furthermore, the larger capture observed when the distractor followed a distractor-absent trial was even larger in blocks in which the distractor appeared with a lower rate compared to a higher rate, as the latter condition overall allowed fewer occasions (i.e., distractor-absent trials) to forget the routine. However, it seems to us that such trial-based distractor-handling routine explanation would not be in principle incompatible with the Sokolov habituation account. The reason is that while this account predicts that the degree of expectation should be determined by the history of stimulation, it does not define the number of previous trials or events that should be integrated to generate a given distractor expectation. If this number is low, individual trials closer in time to the current event are bound to gain proportionally high relevance in determining the corresponding prediction. In any case, even in the low-distractor probability condition it is not difficult to imagine that a distractor could be slightly more expected when preceded by distractor-present trials than by distractor-absent trials. There reason could be that the system does not know whether the estimated probability will remain constant or will change, and therefore it must continuously monitor the sequence of events to adjust its predictions. At any rate, we conducted a further analysis of the inter-trial interference effect in the low-probability condition of Experiment 1, which failed to provide evidence for distraction costs associated with a recent history of trials without distractor presentation (see Footnote 2).

Distractor probability and habituation rate

One might note that in addition to find a stronger habituation for the high-probability condition we should have also found a faster habituation compared to the low-probability condition, thus confirming our recent findings where the high- and low-probability conditions were manipulated in a within-participants design by using two distinct distractor locations (Valsecchi & Turatto, 2022). The reason is that habituation is usually faster, in addition to be more pronounced, the higher frequency of stimulation (Broster & Rankin, 1994; Geer, 1966; Rankin & Broster, 1992; Thompson, 2009). Therefore, one would have predicted a steeper slope of the RT x block function in the high-probability condition than in the low-probability condition, whereas if anything we found the opposite (see for example, Figure 2 panel A). However, this apparent paradoxical outcome can be explained as follows. In classic habituation studies the response evoked by the irrelevant stimulus is measured directly and a short series of stimuli is typically sufficient to generate a clear habituation function. For example, in humans a clear habituation function of the skin conductive response associated with the OR triggered by an auditory stimulus is obtained with few presentations (Steiner & Barry, 2014). By contrast, the measurement of the capture response in the attentional paradigms typically requires to average several tens, if not hundreds, of RTs in distractor present vs absent trials, due to the high RT intrinsic variability. This means that if the habituation function reaches the asymptotic level quickly because the distractor is frequent, this steeper part of the curve is not detected because it is averaged out with the remaining flatter and larger portion of the curve by the mean block RTs. By contrast, for low distractor rates the habituation function declines more slowly, possibly without even reaching the asymptotic level if the number of trials is not enough, so that the block RT averaging does not change much of the curve profile. One

could certainly plot and analyze the data by binning the RTs in shorter bins of fewer trials, but this comes at the expense of largely increasing the noise in the RT results.

Changing the distractor salience

The aim of Experiment 1 was to distinguish between the Sokolov and the Groves and Thompson models of habituation of onset capture. However, on the assumption that onsets saliency are represented in the same priority or dimension-specific maps where feature-singleton distractors are encoded (Ferrante et al., 2018; Itti & Koch, 2001; Liesefeld & Müller, 2019; Zelinsky & Bisley, 2015), the results could also shed light on how distraction suppression might take place in this architecture. Specifically, the results seem to suggest that the attentional capture modulation by distractor probability does not arise from the fact that at each distractor presentation the corresponding peak of activation in the priority map is simply downweighed in a top-down fashion as a consequence of the erroneous attentional selection, with the attenuation being stronger the more distractors are presented. Indeed, if this were the mechanism at play, we would not have found a differential capture in two probability conditions, as they were equivalent in terms of number of distractors, and hence of occasions for suppression (but see the possibility discussed above about the forgetting of the distractor-handling routine). Rather, the different capture observed with an equal number of distractors in the two probability conditions suggests that the system generates an expectation of the distractor on the basis of its rate of occurrence at a given location. However, the distractor expectation does not need to be conscious, but on the contrary, and in agreement with what Sokolov already suggested (Sokolov et al., 2002), we believe that distractor expectation is estimated implicitly, and that such computation might take place in early-visual areas (Won et al., 2020; Zhang et al., 2022).

Then, in keeping with Sokolov's hypothesis, a suppression is automatically exerted on the neural activity generated by each presentation of the irrelevant input or distractor. The more the distractor matches the expectation the more it will be suppressed, and the repeated suppressive activity at the distractor location progressively changes the priority map in a plastic fashion (Ferrante et al., 2018; Turatto & Valsecchi, 2021; Wang & Theeuwes, 2018b). A key difference between our proposed view, based on Sokolov's model, and that advocated by Allenmark et al. (2022), is that we assume that the distractor suppression mechanism leading to habituation is triggered by the mere repetitive presentation of the irrelevant sensory input. In other words, we do not necessarily assume that the habituation mechanism is engaged because the distractor interferes with target selection, as shown by habituation in passive viewing (Turatto, Bonetti, Pascucci, et al., 2018; Won & Geng, 2020), whereas this assumption is central in Allenmark et al.'s view.

Are the suppressive resources optimally and strategically distributed?

Another peculiarity of our study is that we have found a different degree of habituation to an onset distractor appearing with different rates of occurrence at a fixed location in separate groups of participants. By contrast, the modulation of attentional capture by a distractor appearing with different rates at two or more locations has been previously tested in the same group of participants (Ferrante et al., 2018; Goschy et al., 2014; Sauter et al., 2018; Wang & Theeuwes, 2018b). The results invariably showed a larger attenuation of capture where the distractor appeared more frequently. A possible explanation for this finding is to reasonably assume that humans possess a finite quantity of suppressive resources, which are strategically and optimally allocated as a function of the spatial rate of distractor occurrence, with more suppression exerted

where the distractor is more likely to occur, and less where it is less likely to interfere with target selection and discrimination. This view echoes that proposed to explain the results of studies investigating the probability cueing of target location, and showing that target processing is more efficient where it appears more frequently (Geng & Behrmann, 2002; Shaw & Shaw, 1977). Based on the original studies of (Koopman, 1957) on the optimal distribution of searching effort for an object appearing with different probabilities among various locations, humans are assumed to be capable to estimate the target locations probabilities and to strategically and optimally distribute the finite attentional resources accordingly. However, the idea that also the suppressive resources are *optimally* distributed to match the distractor probability at different locations seems to be challenged by our present findings. Indeed, if the function of the suppressive mechanism is that of attenuating as much as possible the distractor interference, it is not clear why in our study each group of participants did not apply all the possible suppressive resources available at the fixed distractor location. Indeed, this strategy that would have ensured the maximum distractor filtering, and in turn the largest capture attenuation, regardless of the distractor rate. Note that this strategy would have been even more reasonable and optimal if one considers that the location of the distractor was not only fixed, but also completely segregated from that of the target. By contrast, the fact that capture attenuation was different in the two groups of participants suggests a different scenario. The mechanism of distractor filtering appears largely determined by the specific spatial distractor probability, and does not operate by optimally dividing all the possible suppressive resources available across locations, or by concentrating all of them in the same distractor location when this is fixed. This is in agreement with a habituation mechanism that runs automatically (Sokolov et al., 2002), in the sense of being non-intentional, by suppressing the OR toward the

irrelevant salient stimulus only as a function of how much this is expected to occur at a given location (Sokolov, 1963). In addition, this also fits with the idea that distractor rejection is not under voluntary top-down control (Wang & Theeuwes, 2018a). That the mechanism responsible for reducing distraction on the basis of the local distractor probability does not appear to be under strategic or flexible control is also coherent with the findings by Bogaerts, van Moorselaar and Theeuwes (2022). The authors found that even if participants could predict the appearance of a feature-singleton distractor within a specific trial, because distractor-present and distractor-absent trials alternated in sequence, or because a cue signaled the likelihood of distractor, the amount of capture was not affected.

Local vs global habituation of capture

The present findings, strongly indicating that habituation to onsets was mainly driven by the local rate of distractor occurrence (for similar findings with feature-singleton distractor see Allenmark et al., 2022), may seem in partial disagreement with those of Valsecchi and Turatto (2021). In our previous study we reported that capture attenuation to a repeatedly presented feature-singleton distractor was modulated by both its local and global probability. There are, however, several methodological differences between the two studies that may make a comparison less than straightforward. To begin with, here we used an onset distractor whereas Valsecchi and Turatto (2021) presented a feature-singleton distractor (also see Allenmark et al., 2022), and the rejecting mechanisms for the two types of distractors may not be the same; second, in the present study the global change in distractor probability was manipulated in two separate experiments (confront the .333 local probability condition of Experiment 1 and Experiment 3, with the latter presenting two local probabilities of

.333, but a global probability of .666), whereas in the Valsecchi and Turatto (2021) study the change in the global distractor probability was more drastic and achieved within the same experiment, as it occurred across different blocks of trials; third, we cannot exclude that our previous evidence in favor of a role of the global distractor probability could have been due to the fact that the location of the target and that of the distractor were not segregated as in the present study. Under those conditions any change in the distractor probability distribution across locations also changed the target probability, and this may have affected the observed findings. So, further studies are needed to clarify whether habituation to a salient distractor is controlled only by its local rate of occurrence (Allenmark et al., 2022), or alternatively whether also the global distractor rate may be a significant, albeit perhaps minor factor.

References

Allenmark, F., Zhang, B., Shi, Z., & Muller, H. J. (2022). Learning to suppress likely distractor locations in visual search is driven by the local distractor frequency. *Journal of Experimental Psychology: Human Perception and Performance*, *in press*.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*, *83*(1), 1–36. https://doi.org/10.1037/0033-295X.83.1.1

Bridgeman, B. (2013). Applications of predictive control in neuroscience. *Behavioral and Brain Sciences*, *36*(3), 208. https://doi.org/10.1017/S0140525X12002282

Broster, B. S., & Rankin, C. H. (1994). Effects of changing interstimulus interval during habituation in Caenorhabditis elegans. *Behavioral Neuroscience*, *108*(6), 1019–1029. https://doi.org/10.1037/0735-7044.108.6.1019

Chelazzi, L., Marini, F., Pascucci, D., & Turatto, M. (2019). Getting rid of visual distractors: The why, when, how, and where. *Current Opinion in Psychology*, *29*, 135–147. https://doi.org/10.1016/j.copsyc.2019.02.004

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Den Ouden, H. E. M., Kok, P., & de Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology*, *3*, 1–12. https://doi.org/10.3389/fpsyg.2012.00548

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and social, behavioral, and biomedical sciences. *Journal of Materials and Environmental Science*, *39*(2), 175–191.

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

Geer, J. H. (1966). Effect of interstimulus intervals and rest-period length upon habituation of the orienting response. *Journal of Experimental Psychology*, *72*(4), 617–619. https://doi.org/10.1037/h0023760

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*, *28*(6), 600–606. https://doi.org/10.1111/1467-9280.00491

Geng, J. J., Won, B. Y., & Carlisle, N. B. (2019). Distractor Ignoring: Strategies, Learning, and Passive Filtering. *Current Directions in Psychological Science*, *28*, 600–606. https://doi.org/10.1177/0963721419867099

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

Groves, P. M., & Thompson, R. F. (1970). Habituation: A dual-process theory. *Psychological Review*, *77*(5), 419–450. https://doi.org/10.1037/h0029810

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*(3), 194–203. https://doi.org/10.1038/35058500

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, *43*(4), 346–354. https://doi.org/10.3758/BF03208805

Koopman, B. O. (1957). The theory of search; Pt. Ill: The optimum distribution of searching effort. *Operation Research*, *5*, 613–626.

Lamy, D., & Egeth, H. E. (2003). Attentional Capture in Singleton-Detection and Feature-Search Modes. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(5), 1003–1020. https://doi.org/10.1037/0096-1523.29.5.1003

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Liesefeld, H. R., & Müller, H. J. (2019). Distractor handling via dimension weighting. *Current Opinion in Psychology*, *29*, 160–167. https://doi.org/10.1016/j.copsyc.2019.03.003

Müller, H. J., Geyer, T., Zehetleitner, M., & Krummenacher, J. (2009). Attentional Capture by Salient Color Singleton Distractors Is Modulated by Top-Down Dimensional Set. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 1–16. https://doi.org/10.1037/0096-1523.35.1.1

Pascucci, D., & Turatto, M. (2015). The distracting impact of repeated visible and invisible onsets on focused attention. *Journal of Experimental Psychology: Human*

Perception and Performance, *41*(3), 879–892. https://doi.org/10.1037/xhp0000025

Pavlov, I. (1927). *Conditioned reflexes*. Oxford University Press.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Ramaswami, M. (2014). Network plasticity in adaptive filtering and behavioral habituation. *Neuron*, *82*(6), 1216–1229. https://doi.org/10.1016/j.neuron.2014.04.035

Rankin, C. H. (2004). Invertebrate learning: What can't a worm learn? *Current Biology*, *14*, R617–R618. https://doi.org/10.1016/j.cub.2004.07.044

Rankin, C. H., & Broster, B. S. (1992). Factors Affecting Habituation and Recovery From Habituation in the Nematode Caenorhabditis elegans. *Behavioral Neuroscience*, *106*(2), 239–249. https://doi.org/10.1037/0735-7044.106.2.239

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In B. AH & P. WF (Eds.), *Classical Conditioning II: Current Research and Theory* (pp. 64–99). New York, NY: Appleton Century Crofts.

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, and Psychophysics*, *80*, 622–642. https://doi.org/10.3758/s13414-017-1477-4

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*, 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

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

Sokolov, E. N. (1960). Neural model and the orienting influence. In B. MA (Ed.), *The central nervous system and behavior*. Macy Foundation.

Sokolov, E. N. (1963). Higher Nervous Functions: The Orienting Reflex. *Annual Review of Physiology*, *25*(1), 545–580. https://doi.org/10.1146/annurev.ph.25.030163.002553

Sokolov, E. N., Spinks, J., Naatanen, R., & Lyytinen, H. (2002). *The Orienting Response in Information Processing*. Lawrence Erlbaum Associates. https://doi.org/10.4324/9781410601490

Steiner, G. Z., & Barry, R. J. (2014). The mechanism of dishabituation. *Frontiers in Integrative Neurosciences*, *8*, 1–8. https://doi.org/10.3389/fnint.2014.00014

Thompson, R. F. (2009). Habituation: A history. *Neurobiology of Learning and Memory*, *92*(2), 127–134. https://doi.org/10.1016/j.nlm.2008.07.011

Thompson, R. F., & Spencer, W. A. (1966). Habituation: A model phenomenon for the study of neuronal substrates of behavior. *Psychological Review*, *73*(1), 16–43. https://doi.org/10.1037/h0022681

Turatto, M., Bonetti, F., Chiandetti, C., & Pascucci, D. (2019). Context-specific distractors rejection: Contextual cues control long-term habituation of attentional capture by abrupt onsets. *Visual Cognition*, *27*, 291–304. https://doi.org/10.1080/13506285.2019.1580233

Turatto, M., Bonetti, F., & Pascucci, D. (2018). Filtering visual onsets via habituation: A context-specific long-term memory of irrelevant stimuli. *Psychonomic Bulletin and Review*, *25*(3), 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

Turatto, M., & Valsecchi, M. (2021). Impaired selection of a previously ignored singleton: Evidence for salience map plastic changes. *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1177/17470218211047859

Valsecchi, M., & Turatto, M. (2021). Distractor filtering is affected by local and global distractor probability, emerges very rapidly but is resistant to extinction. *Attention, Perception, & Psychophysics*, *83*, 2458–2472. https://doi.org/10.3758/s13414-021-02303-3

Valsecchi, M., & Turatto, M. (2022). Habituation to abrupt onset distractors with different spatial occurrence probability. *Attention, Perception, & Psychophysics*, in press.

Vogel, E. H., Ponce, F. P., & Wagner, A. R. (2019). The development and present status of the SOP model of associative learning. *Quarterly Journal of Experimental Psychology*, *72*(2), 346–374. https://doi.org/10.1177/1747021818777074

Wagner, A. R. (1976). Priming in STM: An information-processing mechanism for self-generated or retrieval-generated depression in performance. In T. J. Tighe & R. N. Leaton (Eds.), *Habituation: Perspectives from Child Development, Animal Behavior, and Neurophysiology* (pp. 95–128). Erlabaum.

Wang, B., & Theeuwes, J. (2018a). How to inhibit a distractor location? Statistical learning versus active, top-down suppression. *Attention, Perception, and Psychophysics*, *80*, 860–870. https://doi.org/10.3758/s13414-018-1493-z

Wang, B., & Theeuwes, J. (2018b). Statistical regularities modulate attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(1), 13–17. https://doi.org/10.1037/xhp0000472

Won, B.-Y., Forloines, M., Zhou, Z., & Geng, J. J. (2020). Changes in visual cortical processing attenuate singleton distraction during visual search. *Cortex*, *132*, 309–321. https://doi.org/10.1016/j.cortex.2020.08.025

Won, B.-Y., & Geng, J. J. (2020). Passive Exposure Attenuates Distraction During Visual Search. *Journal of Experimental Psychology: General*, In press. https://doi.org/10.1037/xge0000760

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 121–134. https://doi.org/10.1037/0096-1523.16.1.121

Yantis, S., & Jonides, J. (1996). Attentional Capture by Abrupt Onsets: New Perceptual Objects or Visual Masking? *Journal of Experimental Psychology: Human Perception and Performance*, *22*(6), 1505–1513. https://doi.org/10.1037/0096-1523.22.6.1505

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., Weidner, R., Allenmark, F., Bertleff, S., Fink, G. R., Shi, Z., & Müller, H. J. (2022). Statistical Learning of Frequent Distractor Locations in Visual Search Involves

Regional Signal Suppression in Early Visual Cortex. *Cerebral Cortex*, *32*(13), 2729–2744. https://doi.org/10.1093/cercor/bhab377