Abbreviated title: Learned distractor inhibition

Neural mechanisms underlying distractor inhibition on the basis of feature and/or spatial expectations.

Dirk van Moorselaar[1,2], Nasim Daneshtalab[1] & Heleen A. Slagter[1,2]

1. Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, the Netherlands
2. Institute of Brain and Behaviour Amsterdam, the Netherlands

Number of pages: 35
Number of Figures: 7
Word count Abstract: 260
Word count Introduction: 1181
Word count Discussion: 2308

Conflict of Interest: The authors declare no competing financial interests

Acknowledgements

This research was supported by a European Research Council (ERC) starting grant (679399) to H.A.S. D.v.M. contributed to design, collected the data, performed the analyses, and contributed most of the writing. HS was closely involved in the design of the experiment and the analysis plan and made significant contributions to the writing. ND collected the data and provided feedback on the writing. We would like to thank Mingwei Zhou for her valuable help with data collection.

Please address correspondence to:
Dr. Dirk van Moorselaar
Department of Experimental and Applied Psychology
Vrije Universiteit Amsterdam
e-mail: dirkvanmoorselaar@gmail.com

Abstract

A rapidly growing body of research indicates that inhibition of distracting information may not be under flexible, top-down control, but instead heavily relies on expectations derived from past experience about the likelihood of events or statistical learning. Yet, how expectations about distracting information influence distractor inhibition at the neural level remains unclear. To determine how expectations about distractor features and/or location modulate distractor processing, we measured EEG while participants performed two variants of the additional singleton paradigm. Critically, in these different variants, target and distractor features either randomly swapped across trials, or were fixed, allowing for the development of distractor feature-based expectations. Moreover, participants performed these tasks both with and without a high probability distractor location. Our results show that both distractor feature- and location learning contributed to distractor inhibition, as indicated by corresponding reductions in distractor costs during visual search and an earlier distractor-evoked Pd ERP. Large individual differences further suggest a functional dissociation between early and late Pd components, with the former reflecting early sensory suppression related to intertrial priming and the latter reflecting suppression sensitive to statistical learning over a longer time scale. No increase in anticipatory alpha-band activity was observed over visual regions representing the expected distractor location, suggesting that statistical learning may rely on activity-silent mechanisms, such as synaptic plasticity, or that proactive spatial inhibition was not reliably induced with the current study design. Together, our findings provide further support to the idea that distractor suppression strongly relies on expectations derived from past experience and reveal the neural mechanisms underlying spatial and non-spatial distractor learning.

1. Introduction

Our visual environment is filled with distractions, ranging from billboards alongside the road to the blinking light on your smart phone demanding your attention. Whether such visual distractors automatically capture attention in a stimulus-driven fashion (Theeuwes, 2010), or alternatively, only capture attention in a goal-driven manner when they match the current attentional set (e.g.,)(Folk, Remington, & Johnston, 1992), has long been a topic of active debate. Yet, growing evidence highlights that this theoretical dichotomy cannot account for many situations in which strong attentional biases arise from previous selection episodes (Awh, Belopolsky, & Theeuwes, 2012). As a specific case, there are now many demonstrations that location-probability learning of not only the target but also of the distractor benefits visual search, such that salient distractors are more efficiently ignored at locations with a high distractor probability (Ferrante et al., 2018; Goschy, Bakos, Müller, & Zehetleitner, 2014; Leber, Gwinn, Hong, & O'Toole, 2016; Wang & Theeuwes, 2018b). Critically, such distractor suppression is not observed when the upcoming distractor location is cued on a trial-by-trial basis (Noonan et al., 2016; Wang & Theeuwes, 2018a). Together, these findings suggest that distractor inhibition may not be under voluntary, top-down control, but instead greatly relies on statistical learning or expectations derived from past experience about the certainty or likelihood of events (Noonan, Crittenden, Jensen, & Stokes, 2018; Theeuwes, 2019; van Moorselaar & Slagter, 2020).

Although there is now a general consensus that learning, either implicitly or explicitly, about the spatial distribution of salient, but task-irrelevant distractors in a search display minimizes their interference, the processing stage at which this learned suppression is realized appears flexible. Whereas in some visual search studies the distractor-location effect (i.e., reduced distractor interference at high- vs. low-probability distractor location) was accompanied by impaired target processing at high probability distractor locations, in other studies target processing was unaffected by the spatial distractor regularity across searches. A notable difference between these studies is that in many of the former studies, the defining target and distractor features swapped randomly

across trials (e.g.,)(Ferrante et al., 2018; Wang & Theeuwes, 2018b), such that learning could necessarily only occur on the spatial level, whereas in the latter studies, the color of the distractor was fixed across trials, hence also allowing for learning at the feature level (Allenmark, Zhang, Liesefeld, Shi, & Müller, 2019; Sauter, Liesefeld, Zehetleitner, & Müller, 2018; Zhang, Allenmark, Liesefeld, Shi, & Muller, 2019). This dissociation suggests that spatially specific suppression can shift from generic to feature-specific, when regularities across displays allow for learning at the both the spatial and the feature level (Failing, Feldmann-Wustefeld, Wang, Christian, & Theeuwes, 2019). Here, we set out to determine how spatial and feature expectations independently and/or in interaction influence distractor processing and interference at the neural level.

Previous event-related potential (ERP) studies support the notion that distractor learning at either the feature or the spatial level can prevent attentional capture by salient distractors. In search displays with randomly mixed distractor (vis-à-vis non-distractor) features (i.e., shape and color), salient distractors typically elicit an N2pc (Burra & Kerzel, 2013), an ERP index of attentional selection (Luck & Hillyard, 1994b), especially so after feature swaps (Hickey, Olivers, Meeter, & Theeuwes, 2011)(but see)(McDonald, Green, Jannati, & Di Lollo, 2013). By contrast, in fixed-feature variants of the same tasks, which allow for the development of feature expectations, the same distractors elicit a Pd, an ERP index linked to inhibition (Gaspelin & Luck, 2018b; Sawaki & Luck, 2010), indicating that salient distractors were effectively filtered out before they could capture attention (Burra & Kerzel, 2013; Jannati, Gaspar, & McDonald, 2013). Similarly, Wang, van Driel, Ort, and Theeuwes (2019) showed that at high probability distractor locations, singleton distractors elicited a Pd, whereas without this spatial bias (measured in an independent set of participants), the same distractors first elicited an N2pc. Thus, expectations at either the feature or spatial level are by itself sufficient to implement pre-attentive distractor suppression, raising the question how distractor suppression develops and is neurally implemented when feature and spatial expectations can be combined.

As outlined above, behavioral evidence suggests that generic spatial suppression only arises when the spatial distractor regularity cannot be combined with feature expectations. Consistent with this, in the Wang et al. (2019) study not only distractors, but also targets elicited a Pd at the high probability distractor location. What is more, both distractors and targets at high probability distractor locations also elicited an early Pd component, indicative of early suppression (Weaver, van Zoest, & Hickey, 2017). These findings suggest that learning-dependent spatial suppression was set up in advance in a feature blind manner. Consistent with this, alpha-band activity, a marker of top-down neural inhibition (Jensen & Mazaheri, 2010), tuned to the high probability distractor location already increased before search display onset. While such anticipatory and generic spatial suppression may be the relatively most optimal strategy under conditions in which only information about the likely distractor location is available, it is obviously more advantageous from an ecological perspective to selectively suppress only distractors based on their specific properties. In this respect it is notable that two recent EEG studies did not observe any changes in anticipatory alpha-band activity as a function of distractor location learning when target and distractor features did not swap across trial (Noonan et al., 2016; van Moorselaar & Slagter, 2019). Here, we therefore also examined anticipatory alpha-band modulations and target processing at high probability distractor locations in conditions where spatial expectations could or could not be combined with feature expectations, in addition to directly contrasting condition effects on distractor processing using ERPs.

Specifically, we measured EEG while observers performed, in separate sessions, fixed-feature and mixed-feature variants of the additional singleton paradigm (Theeuwes, 1992). In the fixed-feature version, distractor and target features were fixed throughout an experimental block and participants could therefore develop expectations at the feature level. By contrast, in the mixed-feature version, the shapes and colors of the target and distractors randomly swapped across trials so that participants could not form any predictions at the feature level (see Figure 1). Critically, in both variants, halfway through the experiment, and unbeknownst to the participant, we introduced a high probability distractor location. This allowed us to examine how feature and spatial

expectations may interact in modulating distractor processing and interference. Without any spatial regularities we expected distractors to elicit a Pd in the fixed-feature variant, but an N2pc, possibly followed by a subsequent Pd signaling reactive distractor inhibition (Liesefeld, Liesefeld, Töllner, & Müller, 2017; van Moorselaar & Slagter, 2019), in the mixed-feature variant. By contrast, with a spatial distractor bias we expected a distractor Pd irrespective of the feature variant, and no distractor-evoked N2pc. Yet, whether this Pd differed between feature variants was an open question. Moreover, following Wang et al. (2019), we expected a feature-blind Pd (i.e., a Pd elicited by both targets and distractors) at high probability distractor locations and increased anticipatory alpha power contralateral to the high probability distractor location, both indicative of generic spatial suppression, in the mixed-feature variant, but not in the fixed-feature variant, when expectations about distractor features could also be formed.

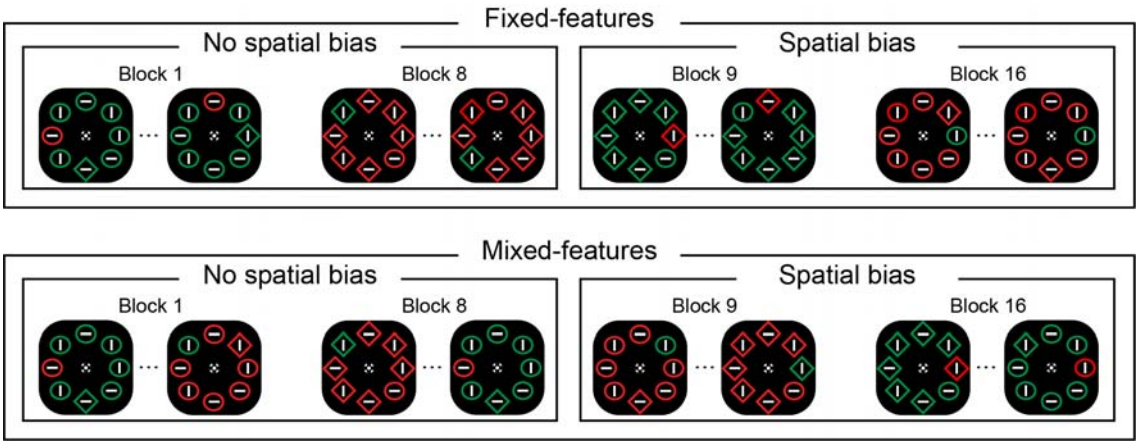

Figure 1 | Experimental paradigm. In each trial participants had to indicate the orientation (vertical or horizontal) of the line inside the unique shape (the target), a diamond among circles or vice versa. In fixed-feature conditions (top row) the display configuration was fixed within an experimental block, whereas in mixed-feature conditions (bottom row) target and non-target features randomly varied across trials. On a subset of trials, the search display also contained a singleton distractor color. Critically, in no spatial bias blocks (i.e., first half of experiment) this distractor appeared equally often at all search locations, whereas in spatial bias blocks it appeared with a higher probability on one specific location (i.e., high probability distractor location).

1. *Methods*

Experiment and analyses were based on the OSF preregistration at Open Science Framework (https://osf.io/n35xa/). All data will be made publicly available there upon publication.

A planned number, based on the Wang et al. (2019) study (*N* = 18), of 24 participants (Mean age = 23 years, range 18 – 37; 4 men) participated in the experiment after replacement of 2 participants because EEG preprocessing (for details see below) resulted in exclusion of >30% of trials. Participants were seated in a sound-attenuated and dimly lit laboratory at a viewing distance of ~80 cm.

The paradigm was modeled after the Wang et al. (2019) study. Each trial started with a randomly jittered blank display (0-ms – 350-ms) followed by a 1200-ms fixation display. This display contained a shape that looks like a combination of bulls eye and cross hair, which has been shown to improve stable fixation (Thaler, Schütz, Goodale, & Gegenfurtner, 2013) and that remained visible throughout the trial. Subsequently, a search display appeared with eight shapes in a circular configuration around fixation (radius 4.3^{o}), which remained visible until response. Each display contained one circle (radius 1.05^{o}) and seven diamonds (2.7^{o} x 2.7^{o}), or vice versa, each with a red or green outline on a black background and a horizontal or vertical white line segment (1.6^{o}) in their center. Participants were instructed to keep their eyes at fixation and covertly search for the unique shape in the display and indicate the orientation of the line segment in this target shape via a button press (i.e., 'z' or '/' button) as quickly as possible while keeping the number of errors to a minimum. Whereas on 40% of trials all stimuli in the search display had the same color, in the other trials the outline of one of the distractor shapes had a different color than the other stimuli in the display. Participants were instructed to ignore this singleton distractor. Targets and singleton distractors were only presented along the vertical and horizontal axis. Critically, in the first half of the experiment, targets and distractors appeared equally often at these four locations (no spatial bias blocks). In the second half of the experiment[1], the singleton distractor appeared with 66.7% probability at one of the horizontal locations (counterbalanced across participants), while targets nevertheless continued to appear equally often on all locations (spatial bias blocks) collapsed over distractor-present and -absent trials.

[1] Due to a programming error, in five subjects the high probability location was introduced after 7 blocks and in another seven subjects after 9 blocks, instead of after eight blocks.

Crucially, in separate experimental sessions (separated by 5-7 days) participants performed the task in the mixed- or the fixed-feature condition (order counterbalanced across participants). The high probability distractor location was kept the same in both sessions. In the mixed-feature condition, targets were equally likely to be a circle among diamonds or a diamond among circles and the singleton distractor colors were also equally likely to be red or green. By contrast, in the fixed-feature condition, a new display configuration was introduced at the start of each block, but target and distractor features (i.e., shape and color) remained fixed throughout a block of trials (features counterbalanced across blocks).

Participants completed 16 blocks of 60 trials per session, each preceded by a sequence of training blocks (60 trials; no spatial bias) until they felt comfortable with the task. At the start of each block participants received written instructions about the dynamics of the upcoming set of trials (e.g., 'In the upcoming block the target and distractor will have random shapes and colors'). After each block participants received feedback on their performance (i.e., mean reaction time (RT) and accuracy) and were encouraged to take a break. At the end of the second session, a display with only circles, each with a unique number (i.e., 1-8) corresponding to one of the search locations was shown. Participants were asked to indicate (and if necessary, guess) which location they believed had contained the singleton distractor most frequently throughout the experiment.

Behavioral analysis. All data were preprocessed in a Python environment (Python Software Foundation, https://www.python.org/). As preregistered, behavioral analyses were limited to reaction time (RT) data of correct trials only. RTs were filtered in a two-step trimming procedure: trials with reaction times (RTs) shorter than 200 ms were excluded, after which data were trimmed on the basis of a cutoff value of 2.5 SD from the mean per participant. Remaining RTs were analyzed with repeated measures ANOVAs, where reported *p*-values are Greenhouse-Geiser corrected in case of sphericity violations, followed by planned comparisons with paired t-tests using JASP software *(JASP-TEAM, 2018).*

EEG recording and preprocessing. EEG data were recorded at a sampling rate of 512 Hz using a 64-electrode cap with electrodes placed according to the 10-10 system (Biosemi ActiveTwo system; biosemi.com) and from two earlobes (used as offline reference). Vertical and horizontal EOG (VEOG/HEOG) were recorded via external electrodes placed ~2 cm above and below the eye, and ~1 cm lateral to the external canthi, respectively. EEG from the separate sessions were preprocessed independently (as specified below) after which the two data sets were concatenated.

As preregistered, EEG data were re-referenced off-line to the average of the left and the right earlobe, and subsequently high-pass filtered using a zero-phase 'firwin' filter at 0.1 Hz to remove slow drifts. Continuous EEG was then epoched from -1500 – 1100-ms relative to search display onset (to avoid filter artefacts during preprocessing and time-frequency analysis), with trial rejection procedures being limited to smaller time windows (i.e., -1000 – 600-ms). The resulting epochs were baseline normalized using the whole epoch as a baseline. Prior to trial rejection, malfunctioning electrodes (M = 1, range = 0 - 6) were temporarily removed on the basis of visual inspection. EMG contaminated epochs were then identified using an adapted version of an automatic trial-rejection procedure as implemented in the Fieldtrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011). To specifically capture muscle activity we used a 110-140 Hz band-pass filter, and allowed for variable z-score cut-offs per subjects based on within-subject variance of z-scores (de Vries, van Driel, & Olivers, 2017) resulting in an average removal of 9.2% (range 2.3% - 22.4%) of trials thus identified as containing artifacts. Next, ICA as implemented in MNE (method = 'picard') was performed on the continuous EEG to identify and remove eye-blink components from the epoched data. Finally, malfunctioning electrodes were interpolated using spherical splines (Perrin, Pernier, Bertrand, & Echallier, 1989) before the data of the separate sessions was combined.

Throughout the EEG recording, eye movements were monitored using an Eyelink 1000 (SR Research), sampled at 500 Hz, and a head mount. Gaze data were analyzed online, so that every time the participant broke fixation, auditory feedback was provided, signaling to the participant to keep their gaze at fixation. In some cases, this auditory feedback was turned off, but only once

fixation was relatively stable, as it was considered distracting by a subset of participants when it did occur. The Eyelink data were also analyzed offline. As in our previous work (van Moorselaar & Slagter, 2019), to control for drifts in the eye-tracker, epochs without a saccade (Nyström & Holmqvist, 2010) in the 300 ms pre-display interval were shifted towards fixation. Subsequently, each data sample was converted into a single number specifying the visual deviation from fixation (in visual angle), after which epochs were summarized by a single value indicating the largest deviation in visual degrees from fixation measured in a continuous segment of data of at least 40 ms. Thus, each epoch was summarized by its maximal deviation in visual degrees. The results of the EEG analyses presented next are limited to trials with a fixation deviation <= 1° or in case of missing Eyelink data, trials with sudden sharp jumps in the HEOG as detected via a step algorithm with a window length of 200-ms, a step size of 10 ms, and with a threshold of $20\mu V$, which led to the exclusion of 7.7% of the cleaned data (range = 0.7% - 21.8%). All EEG preprocessing parameters per subject are summarized at OSF, alongside the raw data files.

Event-related potential (ERP) analysis. We examined effects of feature- and location-expectations on distractor and target processing using ERPs, which were computed for a prespecified time window from -100 to 600-ms separately per condition of interest. First, epochs were 30 Hz low-pass filtered and baseline corrected using a -100 to 0 ms pre-stimulus baseline period. To enable isolation of lateralized distractor- and target-specific ERP components, ERP analyses focused on trials in which the stimulus of interest (distractor or target) was presented to the left or right of fixation, while the other stimulus was presented on the vertical meridian. Waveforms evoked by the various search displays were collapsed across left and right visual hemifield and left and right electrodes to produce separate waveforms for contralateral and ipsilateral scalp regions. Lateralized difference waveforms were then computed by subtracting the ipsilateral waveform from the corresponding contralateral waveform. In spatial bias blocks, this analysis was limited to trials where the distractor appeared at high probability distractor locations.

Following Wang et al. (2019), we focused the analysis of distractor- and target-elicited Pd and/or N2pc ERP waveforms on electrodes PO7/8. Pd epochs of interest were centered (± 25-ms) around condition-specific positive peaks in distractor and target-tuned waveforms. In many studies, including in Wang et al. (2019), distractor learning is associated with a distractor-evoked early Pd followed by a second positive peak within the N2pc time range (late Pd). As visual inspection of our results suggested that distractor-specific positivities also contained both an early and a late component (Feldmann-Wüstefeld & Vogel, 2019; Feldmann-Wüstefeld, Brandhofer, & Schubö, 2016; Weaver et al., 2017), we searched for positive peaks in a 50 – 150-ms (early Pd) and a 200 – 400-ms (late Pd) in distractor tuned waveforms and only the 50 – 150-ms window in target tuned waveforms. N2pc epochs of interest were centered around condition-specific negative peaks in target-tuned waveforms within a 200 – 400-ms window.

Time-frequency analysis. We adopted the same analysis approach as reported in Wang et al. (2019). Using Morlet wavelet convolution, EEG time series were decomposed into their time-frequency representation for frequencies ranging from 1 to 40 Hz in 25 logarithmically spaced steps. To create complex Morlet wavelets, for each frequency a sine wave ($e^{i2\pi ft}$, where i is the complex operator, f is frequency, and t is time) was multiplied by a Gaussian ($e^{-t^2/2s^2}$, where s is the width of the Gaussian). To keep a good trade-off between temporal and frequency precision, the Gaussian width s varied (i.e., $s = \delta/(2\pi f)$, where δ represents the number of wavelet cycles, in 25 logarithmically spaced steps between 3 and 12). Frequency-domain convolution was then applied by multiplying the Fast Fourier Transform (FFT) of the EEG data and Morlet wavelets. The resulting signal was converted back to the time domain using the inverse FFT. Time-specific frequency power was defined as the squared magnitude of the complex signal resulting from the convolution.

Condition-averaged lateralized alpha-power (contra- vs. ipsilateral to high probability distractor location) in the pre-stimulus period (i.e., -1000 – 0-ms.) was evaluated via Z-transformed time-frequency power. Z-values were computed by a statistical transformation of real observed lateralized power for each contra- and ipsilateral channel pair with respect to the mean and

standard deviation of a "lateralized" power distribution under the null hypothesis of no difference between a contra- an ipsilateral channel pair. This distribution was obtained with a permutation procedure (N = 1000), where we swapped electrode pairs for a random subset of trials. This procedure yields a metric that shows interpretable dynamics over time and frequency and has the benefit of not having to select a pre-stimulus interval, the same interval that we are interested in, for baseline normalization.

Analysis software. Preprocessing and subsequent analyses were performed using custom-written analysis scripts, which were largely based on functionalities implemented within MNE (Gramfort et al., 2014). These scripts can be downloaded at https://github.com/dvanmoorselaar/DvM.

2. Results

Search times. Exclusion of incorrect responses (6.5%) and data trimming (3.1%) resulted in an overall loss of 9.6% of the behavioral data. We first analyzed distractor costs as a function of search condition (Figure 2A/B). A repeated measures ANOVA with the within subject factors Condition (mixed-features, fixed features) and Distractor (present, absent) using only data from no spatial bias blocks showed faster RTs in fixed- than in mixed-feature conditions (main effect Condition: F (1, 23) = 31.5, p < 0.001, $n_p^2 = 0.58$) and reliable distractor interference (main effect Distractor: F (1, 23) = 119.6, p < 0.001, $n_p^2 = 0.84$). Importantly, a Condition by Distractor interaction (F (1, 23) = 20.6, p < 0.001, $n_p^2 = 0.47$) showed that distractor interference was larger in the absence of feature foreknowledge, in the mixed-feature condition (862 vs. 937 ms; t (23) = 7.4, p < 0.001, d = 1.52), compared to when features were entirely predictable, in the fixed-feature condition (675 vs. 694 ms; t (23) = 5.6, p < 0.001, d = 1.14). Notably, while distractor interference was reduced, it was still far from complete in in the mixed-feature condition. Conceivably, this may have been the case because in each block observers were introduced to a new display configuration so that target and distracter features, while fixed within a given block, hence varied from block-to-block. Indeed, an exploratory

analysis (Figure 2B) in which trials per experimental block were grouped in bins of 10 trials, yielded a marginally significant three-way interaction (F (5, 115) = 2.0, p = 0.086, $n_p^2 = 0.079$), which was driven by reliable distractor costs across all bins in mixed-feature condition (all t's > 4.0, all p's < 0.001, all d's > 0.082), but only in the first two bins in fixed-feature condition (all t's > 2.7, all p's < 0.0013, all d's > 0.55) and not in bins 3 – 5 (all t's < 2.0, all p's > 0.057, all d's < 0.41). This finding is in line with the notion that distractor inhibition critically relies on statistical learning and that to resist distractor interference in in fixed-feature conditions, a specific target template needs to be accompanied by learning-dependent tuning to distractor features (Vatterott & Vecera, 2012).

We next determined how distractor interference was modulated when in addition distractor location learning was possible. To this end, we entered RTs into a repeated measures ANOVA with the within subjects' factors Block Type (no spatial bias, spatial bias), Condition (mixed-features, fixed-features) and Distractor Location (high probability distractor location, low probability distractor location). For this purpose, within no spatial bias blocks, distractor locations were artificially coded as 'high' and 'low' probable locations as the spatial bias was only introduced in the second half of the experiment. Overall RTs were faster in spatial bias blocks (main effect of Block Type: F (1, 23) = 62.7, p < 0.001, $n_p^2 = 0.73$) and in the fixed-feature condition (main effect of Condition: F (1, 23) = 38.6, p < 0.001, $n_p^2 = 0.63$). Crucially, as visualized in Figure 2C, and reflected in a Block Type by Distractor Location interaction (F (1, 23) = 14.2, p < 0.001, $n_p^2 = 0.38$), distractors were more efficiently ignored at high probability distractor locations in spatial bias blocks, both in mixed- (t (23) = 2.1, p = 0.047, d = 0.43) and fixed-feature conditions (t (23) = 2.3, p = 0.028, d = 0.48), yet no such benefit was observed in no spatial bias blocks (all t's < 1.0, all p's > 0.33, all d's < 0.20). Importantly, despite the spatial bias distractors continued to produce reliable interference (i.e., relative to no distractor trials) at both the low and high probability locations in mixed-feature blocks (all t's > 6.5, all p's < 0.001, all d's > 1.33) and in fixed-feature blocks, although while this effect was robust at low probability locations (t (1,23) = 3.3, p = 0.003, d = 0.68), it was only marginally significant at high probability distractor locations (t (1,23) = 2, p = 0.055, d = 0.41).

Together these findings suggest that both feature- and location-based distractor learning reduced distractor interference, and may interact in distractor inhibition.

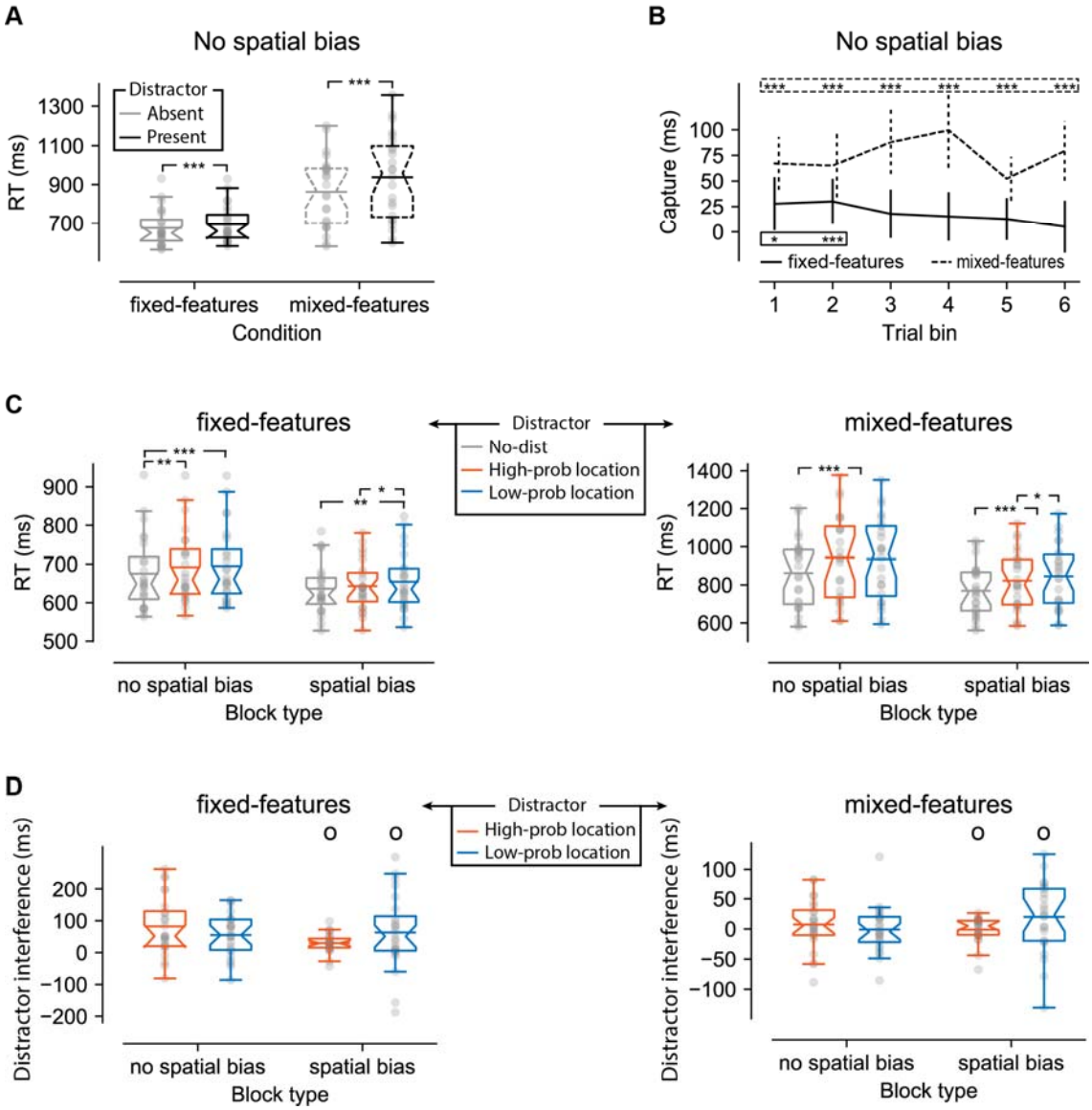

Figure 2 | Behavioral findings visualized by notched boxplots with solid horizontal lines corresponding to the mean. (**A**) Reaction times in no spatial bias blocks shown separately for distractor absent and distractor present trials in the fixed-features and mixed-features conditions. (**B**) Binned distractor interference (distractor present – distractor absent RT) per conditions. Each trial bin corresponds to 10 trials (bin 1: trial 1 – trial 10, bin 2: trial 11 – trial 20, etc). Error bars represent condition-specific, within-subject 95% confidence intervals (Morey, 2008) (**C**) Reaction times as a function of distractor location in no spatial bias and spatial bias blocks for fixed-features (left) and mixed feature (right) conditions. Note that in no spatial bias blocks, distractor locations were artificially coded as high and low probable distractor locations. (**D**) Distractor costs (i.e., difference with distractor absent trials) in trials where the distractor location was the same in the previous trial as a function of distractor location probability in no spatial bias and spatial bias blocks for fixed-features (left) and mixed feature (right) conditions. In these trials, significantly reduced distractor costs at high relative to low probability locations were observed only in the spatial bias blocks (indicated by ° in the figure), demonstrating that distractor repetition effects went beyond intertrial priming in spatial bias blocks.

Arguably, a significant proportion of the observed RT benefits in the preceding analysis are due to short-lasting, intertrial location priming (Maljkovic & Nakayama, 1994). However, previous studies have shown that distractor probability effects cannot solely be explained in terms of intertrial location priming (e.g.,)(Failing et al., 2019; Ferrante et al., 2018; Sauter et al., 2018). To examine whether the observed RT benefits remained reliable after controlling for intertrial priming, we repeated the same analysis but now excluding trials in which the distractor location repeated. Although numerically, the pattern remained the same, the Block Type by Distractor Location interaction was no longer significant (F (1, 23) = 1.8, p = 0.19, $n_p^2 = 0.074$). Also, in spatial bias blocks there no longer was a reliable difference between high and low probability locations, neither in fixed-features (t (23) = 1.0, p = 0.31, d = 0.21) nor in mixed-feature conditions (t (23) = 0.9, p = 0.39, d = 0.18). Next, we examined whether intertrial location distractor priming differed between blocks with and without a spatial bias, as one would not expect to be the case if no statistical learning occurred in the spatial bias blocks. For this purpose, we limited the analysis of distractor costs at high and low probability locations (i.e., difference with distractor absent trials) to trials where the distractor location repeated (see Figure 2D). This analysis yielded a Block Type by Distractor Location interaction (F (1, 23) = 5.8, p = 0.024, $n_p^2 = 0.20$) reflecting reduced distractor costs at high relative to low probability locations, but critically, only in the spatial bias blocks (main effect of Distractor location in spatial bias blocks: F (1, 23) = 5.0, p = 0.035, $n_p^2 = 0.18$; and in no spatial bias blocks: F (1, 23) = 2.2, p = 0.15, $n_p^2 = 0.15$). Together these findings demonstrate that while the observed spatial probability effects largely reflected intertrial priming, in contexts where the distractor location is more frequently repeated, regularities across multiple trials also contribute to performance.

To further examine the influence of distractor location regularities on spatial suppression, we repeated the preceding analyses, but now tuned to target locations and only including distractor

absent trials (Figure 3). Counter to previous findings (e.g.,)(Failing et al., 2019; Wang et al., 2019), there was no evidence that target processing was impaired at high probability distractor locations across conditions (Block Type by Distractor Location interaction: F (1, 23) = 0.048, p = 0.83, $n_p^2 = 0.002$), not even when the analysis was limited to the mixed-feature condition (t = 0.8, p = 0.44, d = 0.16), where we a priori did expect impaired target processing.

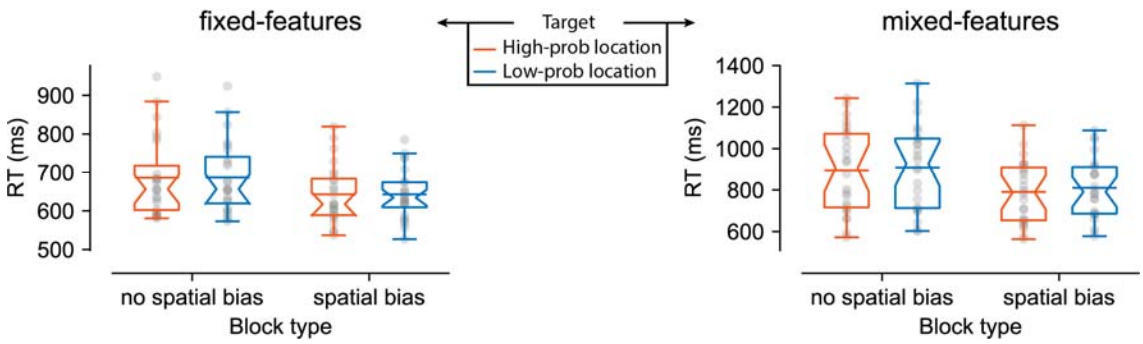

Figure 3 | Reaction times as a function of target location in no spatial bias and spatial bias blocks for fixed-features (left) and mixed feature (right) conditions. Data is based on distractor absent trials only.

Finally, we examined if participants could correctly identify the high probability distractor location. Eight out of 24 participants correctly identified the high probability distractor location (chance level = 25% or 6 participants; assuming that participants noticed that stimuli only appeared on cardinal axes) suggesting that the spatial bias may have been noted by at least a subset of participants. However, the average deviation from the high probability location was 2.1 and thus not above chance (chance level = 2). After excluding the eight participants that indicated the correct location, the results from the distractor and target location tuned ANOVA's did not change (i.e., same effects were reliable).

3.1 Interim discussion and exploratory analyses

Consistent with previous studies (e.g.,)(Bacon & Egeth, 1994; Theeuwes, 1992), in mixed-feature blocks the colored distractor singleton reliably interfered with attentional selection, while in fixed-feature blocks this interference was reduced, but nevertheless reliable. It should be noted though that in fixed-feature blocks, distractor interference was only evident at the start of each block when

observers were introduced to a new stimulus configuration, and it disappeared once they were accustomed to the new context, in line with the idea that statistical learning plays a critical role in distractor inhibition (van Moorselaar & Slagter, 2020). Also replicating previous findings, introducing a spatial bias greatly reduced distractor interference at high probability distractor locations. Yet, our findings unexpectedly differ from previous findings in two important aspects. First, distractor suppression was not accompanied by impaired target processing at high probability locations (Wang & Theeuwes, 2018b). Second, whereas previous studies have shown that spatial distractor suppression cannot solely be explained in terms of intertrial location priming (Failing et al., 2019; Ferrante et al., 2018), here distractor suppression at high probability locations were largely driven by lingering biases from the preceding trial. These two discrepancies are especially relevant as we aimed to investigate whether neural mechanisms underlying learning from spatial regularities differ under conditions that either allow for the development of feature expectations or not.

One notable difference with previous studies (Wang & Theeuwes, 2018b; Wang et al., 2019) is that in the currently study, the high probability location was not introduced immediately, but only halfway through the experiment. It has been shown that different strategies may be applied to the search task depending on the specifics of the individual paradigms (Zhang et al., 2019). Although speculative, in the current study, observers may generally have been less prone to learn more subtle probabilities once they already had developed a strategy specifically tuned to conditions without such probabilities. If distractor rejection is efficient in the absence of spatial regularities, there is also less need for spatial distractor learning, simply because distractors are already efficiently ignored. Indeed, an explorative correlational analysis between distractor interference in the first half of the experiment (i.e., distractor present vs. absent) and selection history effects (distractor at high vs. low probability location) yielded a robust positive correlation both in mixed- ($r_s = 0.43$, p = 0.038) and fixed-feature blocks ($r_s = 0.45$, p = 0.026), although the same correlation was not significant after controlling for intertrial priming (all r_s's > 0.18, all p's < 0.38). That is, participants who exhibited high distractor interference when its location could not be predicted in advance in the first half of

the experiment, generally were more sensitive to the high probability distractor manipulation in the second half of the experiment, consistent with the idea that the magnitude of selection history biases depends on the amount of distractor interference (Failing & Theeuwes, 2020; Sauter et al., 2018). This may suggest that whereas spatial statistical learning occurs more or less automatic when regularities are immediately introduced, as in previous studies (e.g.,)(Wang & Theeuwes, 2018b; Wang et al., 2019), it becomes increasingly sensitive to individual differences in initial task strategy when the spatial regularity is introduced at a later stage. To investigate these potential individual differences further we divided the participants in two groups ('learners' and 'non-learners') based on whether or not target processing was impaired at high probability distractor locations. As a priori, we expected this impaired target processing to be specific to mixed-feature conditions, we only used this condition to label individual participants as 'learners' and 'non-learners' on the basis of whether the target impairment score (i.e., target at high probability distractor location vs. target at low probability distractor location) was larger or equal to/smaller than zero, respectively. The logic here is that if this division is not simply taking two halves of a normal distribution around zero, but is truly diagnostic of individual differences, then spatial suppression should not only also go above and beyond intertrial priming in 'learners', but one would also expect neural markers of suppression to differ between groups (e.g., a Pd to targets at high probability distractor location in spatial bias blocks in the mixed-features condition in the learners, but not in the non-learners). Note that importantly, distinctions between 'learners' and 'non-learners' could not be explained by the order in which they completed both EEG sessions (i.e., 5 and 7 learners identified in session 1 and 2, respectively).

Intriguingly, the two groups, which were thus selected on the basis of a target impairment score in spatial bias blocks (i.e., second half of experiment), already differed in no spatial bias blocks (i.e., first half of experiment), with 'learners' exhibiting larger distractor costs than 'non-learners' (Distractor by group interaction: F (1, 22) = 5.1, p = 0.034, $n_p^2 = 0.19$), while being, although only marginally significant, slower overall (main effect Group: F (1, 22) = 4.0, p = 0.059, $n_p^2 = 0.15$). More

importantly, in spatial bias blocks, a Distractor Location by Group interaction (F (1, 22) = 5.7, p = 0.026, $n_p^2 = 0.21$) showed that 'learners' (high vs low probability distractor location fixed features: M RT= 654.5 ± SD = 69.6 vs. 672.0 ± SD = 81.0; high vs low probability distractor location mixed features: M RT = 872.8 ± SD = 147.1 vs. 917.5 ± SD = 146.2) were less distracted by distractors at high vs. low probability distractor locations than 'non-learners' (high vs low probability distractor location fixed features: M RT = 631.8 ± SD = 60.0 vs. 637.2 ± SD = 67.4; high vs low probability distractor location mixed features: M RT = 772.0 ± SD = 152.1 vs. 773.0 ± SD = 160.6). Moreover, the same interaction remained reliable after exclusion of trials in which the distractor location repeated (F (1, 22) = 5.6, p = 0.027, $n_p^2 = 0.20$), demonstrating that in 'learners' distractor suppression could not be explained by intertrial priming alone (F (1, 11) = 9.0, p = 0.012, $n_p^2 = 0.45$). Finally, although 'learners' were selected on the basis of the target impairment score in the mixed-feature condition ('learners' vs. 'non-learners: M = 33.6 ± SD = 11.7 vs. -50.7 ± SD = 44.8), they also exhibited, although only marginally significant (t (11) = 2.0, p = 0.07, d = 0.57), impaired target processing at high probability distractor location in the fixed-feature condition ('learners' vs. 'non-learners: M = 23.0 ± SD = 39.9 vs. -18.5 ± SD = 39.7). Together, these findings suggest that our group division was not an arbitrary split, but rather separated participants that were sensitive to the introduced statistical regularity over longer time scales from those where suppression did not go above and beyond intertrial priming. Therefore, in addition to the preregistered whole group ERP and time-frequency analyses, we also separate between 'learners' and 'non-learners' in these analyses.

2.2 ERPs.

We next examined effects of distractor feature and location expectations on the lateralized Pd and N2pc ERP components. Visual inspection of Figure 4 shows that in the no spatial bias blocks, a positive difference, corresponding to the Pd, emerged to lateral distractors with midline targets. Notably, while this positivity consisted of an early and a late component in the fixed-feature condition, in the mixed-feature condition, in which participants could not predict what the distractor

would look like in advance, only a late Pd can be observed. By contrast, in the spatial bias blocks, at the high probability distractor location, the distractor Pd was virtually identical in mixed- and fixed-feature conditions.

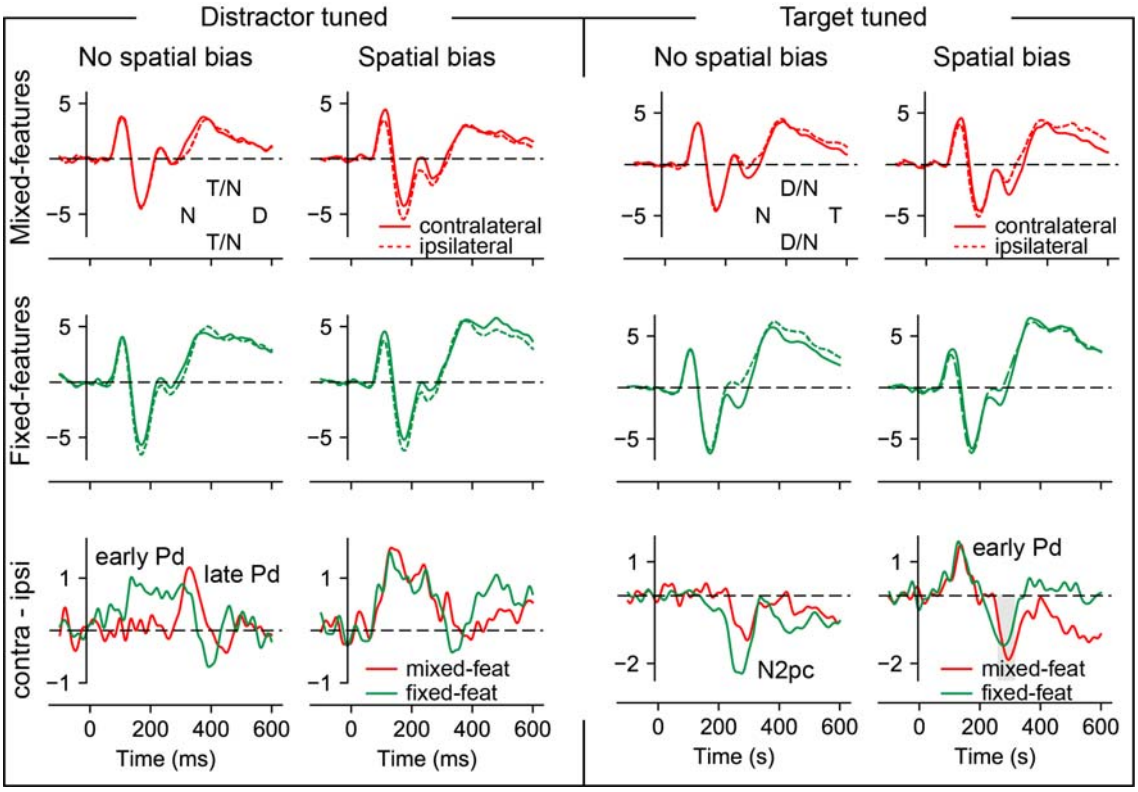

Figure 4 | Electrophysiological results from search trials with lateral distractors (distractor tuned ERPs; columns 1 and 2) and with lateral targets (target tuned ERPS; columns 3 and 4), separately for no spatial bias (columns 1 and 3) and spatial bias blocks (columns 2 and 4). Rows 1 and 2 show the contra- and ipsilateral waveforms for the mixed- and fixed-features conditions, respectively. Row 3 shows the difference between contra- and ipsilateral waveforms for the mixed- and fixed-features conditions. Microvolts are plotted on the y-axes. Inserts shows the schematic target (T), distractor (D) and non-target N) positions.

Differences in contra- vs. ipsilateral voltage elicited by lateralized distractors at "high" probability distractor locations at electrodes PO7/8 were therefore analyzed separately in the early and late Pd windows. The early Pd peaked on average in no spatial bias blocks at 100 ms (mixed-feature) and at 137 ms (fixed-feature), and in spatial bias at 133 ms (mixed-feature) and at 129 ms (fixed-feature). A repeated measures ANOVA with peak amplitude in a 50 ms window surrounding the peak as the dependent variable and Condition (mixed features, fixed features), Block Type (no spatial bias, spatial bias) and Hemifield (contralateral, ipsilateral) revealed that the main effect of Hemifield (F (1, 23) = 4.5, p = 0.045, $n_p^2 = 0.16$), was accompanied by a significant three-way

interaction (F (1, 23) = 8.1, p = 0.009, $n_p^2 = 0.26$). This confirmed that the early Pd was reliably larger in fixed- than mixed-feature conditions, but only in no spatial bias blocks (t (23) = 3.7, p = 0.001, d = 0.75), and not at high probability distractor locations in spatial bias blocks (t (23) = 0.1, p = 0.93, d = 0.02). The late Pd peaked on average slightly later in no spatial bias blocks at 328 ms (mixed-feature) and at 305 ms (fixed-feature), than in spatial bias at 240 ms (mixed-feature) and at 246 ms (fixed-feature), but did not differ in amplitude across conditions. This was statistically reflected in the fact that in the late Pd analysis, the main effect of Hemifield (F (1, 23) = 4.4, p = 0.047, $n_p^2 = 0.16$) was no longer accompanied by any interactions (all F's < 0.7, all p's > 0.40). Moreover, an exploratory peak latency analysis[2] within the late Pd window showed significant main effects of Condition and Block Type (all F's > 4.7, all p's < 0.041 all n_p^2's > 0.17), reflecting an earlier late Pd as a function of feature and spatial expectations, respectively. In addition, this analysis revealed a Condition by Block Type interaction (F (1, 23) = 5.6, p = 0.027, $n_p^2 = 0.20$), which reflected that fact that whereas the late Pd peaked later in mixed-features (M = 314-ms) than in fixed-features (M = 277-ms) conditions (t (23) = 2.7, p = 0.012, d = 0.56) in no spatial bias blocks, this difference between mixed-features (M = 269-ms) and fixed-features (M = 265-ms) conditions disappeared in spatial bias blocks (t (23) = 0.4, p = 0.67, d = 0.088). Together, these results suggest that an early suppressive mechanism, as reflected in the early Pd, can only develop in mixed-feature conditions as a function of spatial learning, and that a later stage suppressive mechanism, as captured by the late Pd, onsets earlier in time when expectations can be formed at either the feature or the spatial level.

The presence of a Pd in fixed-feature conditions irrespective of the spatial bias suggests that the observed reduction in distractor interference in the fixed-feature condition cannot solely be explained by a more precise target template, but may also reflect feature-based experience-dependent distractor inhibition. In line with this, behaviorally, in fixed feature conditions, distractor interference gradually decreased over the course of a block as observers increasingly gained

[2] Here we chose peak-latency rather than area latency to minimize the risk that the analysis was confounded by overlapping activity from the preceding early Pd component (Liesefeld, 2018).

experience with the distracting information (Figure 2B). One would hence expect that in no spatial bias blocks, the Pd also emerged gradually. To test this prediction, following the behavioral analysis where distractor interference gradually decreased within each block, in an explorative analysis, we analyzed the signed area under the curve (i.e., the positive area; baseline is zero) in a 100 – 400-ms window using only non-spatial bias blocks with a repeated measures ANOVA with the within subjects' factors Condition (mixed-features, fixed-features) and Trial Bin (1 – 20, 21 – 60). Here, we focused on area under the curve rather than mean amplitude as this method is less sensitive to potential latency differences between individuals as a function of learning. A reliable interaction (F (1, 23) = 5.5, p = 0.028, $n_p^2 = 0.19$) demonstrated that in fixed-feature blocks the Pd area was indeed larger in the final 40 trials of the block then in the first 20 trials (t (23) = 2.4, p = 0.027, d = 0.48), whereas no such difference was observed in mixed-feature conditions (t (23)= 1.3, p = 0.21, d = 0.27) (Figure 5). Intriguingly, while the same feature-based learning effect was absent in spatial bias blocks (F (1,23) = 1.2, p = 0.29, $n_p^2 = 0.049$), in fixed-feature conditions the Pd area was larger in spatial bias blocks than in non-spatial blocks, not only for the final 40 (t (23) = 2.5, p = 0.019, d = 0.51), but already for the first 20 trials in a block (t (23) = 3.4, p = 0.003, d = 0.69). Together these findings suggest that the inhibitory mechanism reflected by the Pd was dependent on both feature-based and location-based information in the fixed feature conditions, whereas in mixed-feature conditions, distractor suppression was purely spatially informed.

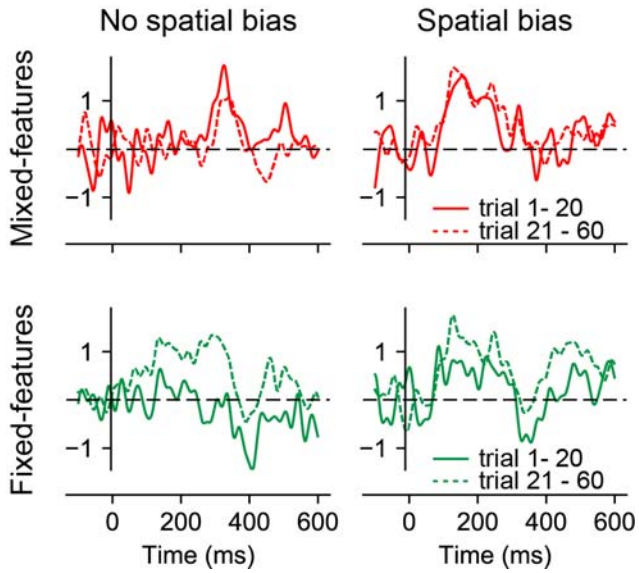

Figure 5 | In fixed features conditions in no spatial bias blocks the Pd only emerged after some experience was gained with the distracting information, suggestive of feature-based learning. Results from exploratory analysis: difference between contra- and ipsilateral waveforms across conditions separately for first 20 (solid lines) and last 40 trials (dashed lines) within a block. Microvolts are plotted on the y-axes.

We next examined effects of feature and spatial expectations on target processing. Lateralized ERPs evoked by lateral targets with midline (or absent) distractors are shown in Figure 4. As can be seen, targets evoked a clear N2pc in all conditions. In no spatial bias blocks, this N2pc was larger in amplitude in fixed-feature than in mixed-feature conditions. As expected based on Wang et al. (2019), after introduction of the spatial bias, targets at high probability distractor locations first appeared to elicit an early Pd, both in the fixed-features and mixed-features condition. Yet, counter to Wang et al. (2019), this early Pd was immediately followed by an N2pc (rather than a late Pd in their study).

Voltages elicited by lateralized targets within the N2pc interval (295 ± 25-ms for no spatial bias/mixed features; 275 ± 25-ms for no spatial bias/fixed features; 295 ± 25-ms for spatial bias/ mixed features; 279 ± 25-ms for spatial bias/fixed features) were analyzed as above. While the N2pc was reliable across all conditions (main effect Hemifield: F (1, 23) = 9.5, p = 0.005, $n_p^2 = 0.29$), there was no evidence that it differed between conditions or was reduced in amplitude at high probability distractor locations, as neither Condition nor Block Type interacted with Hemifield (all F's < 1.9, all p's > 0.18). At the same time, despite being clearly visible in the group-average lateralized

waveforms (Figure 4), an analysis of the early pd interval did not yield a Block Type by Hemifield interaction (F (1,23) = 1.5, p = 0.24, $n_p^2 = 0.061$), nor did planned-pairwise comparisons identify hemispheric differences in any of the conditions (all t's < 1.5, all p's > 0.15). This suggests that at a group level the early Pd at high probability distractor locations was not reliably evoked by targets across subjects (but see below).

One potential explanation for the finding that the target-evoked Pd at high probability distractor locations was not reliable, is that it is selectively elicited by a subset of participants. Note that our exploratory behavioral analysis showed that some participants demonstrated spatial statistical learning and impaired target processing, while other participants only exhibited intertrial priming. Therefore, in an exploratory analysis using the condition specific time windows as in the preceding analyses, we included Group ('learners', 'non-learners') as a between subjects' factor to explore whether the early target-evoked Pd was selectively elicited by the learners. Surprisingly, however, a Group by Block Type interaction (F (1, 22) = 5.2, p = 0.032, $n_p^2 = 0.019$) and a main effect of Group (F (1, 22) = 4.8, p = 0.039, $n_p^2 = 0.018$) showed that 'non-learners' rather than 'learners' were driving the target Pd effect (see Figure 6), strongly suggesting that the early Pd at the high probability distractor location discussed above thus largely reflects intertrial priming, as some studies have also suggested (Gokce, Geyer, Finke, Müller, & Töllner, 2014), instead of spatial statistical learning. Consistent with this, in spatial bias blocks the early Pd in response to distractors was virtually absent in 'learners', but especially pronounced in 'non-learners' (main effect of Group: F (1, 22) = 4.2, p = 0.053, $n_p^2 = 0.016$). Intriguingly, the opposite pattern was observed for the late Pd, which appeared largely driven by learners, although this effect was not statistically significant (F (1,22) = 0.18, p = 0.67, $n_p^2 = 0.008$). This apparent dissociation may suggest that early and late distractor Pd's observed at the group level in spatial bias blocks did not reflect two sequential suppressive processes, but instead individual differences in early and late suppression mechanisms. Moreover, the current results may suggest that the early Pd is especially sensitive to intertrial priming.

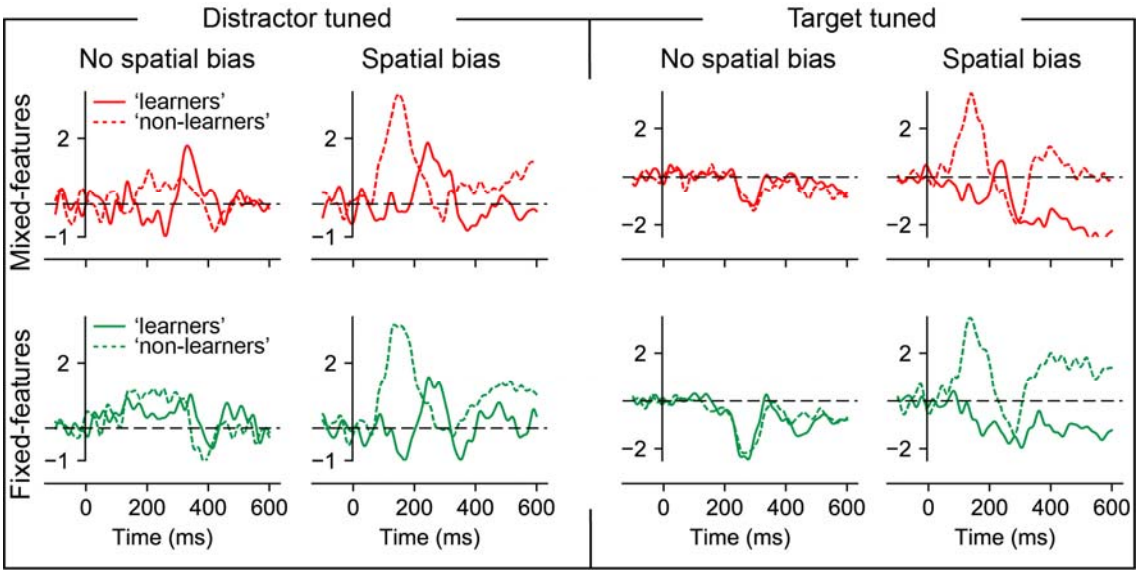

Figure 6 | Individual differences in spatial statistical learning were associated with individual differences in the early and late Pd. Results from exploratory analysis: difference between contra- and ipsilateral waveforms across conditions separately for 'learners' and 'non-learners'. Microvolts are plotted on the y-axes.

2.3 *Time-frequency analysis.*

To investigate whether the introduction of the high probability distractor location lead to anticipatory spatial inhibition, we focused our analysis on pre-stimulus alpha-band oscillatory activity (7.5 – 12 Hz; Figure 7A/B). Counter to Wang et al. (2019), after introduction of the spatial bias, no reliable increase in pre-stimulus alpha-band power was observed at contralateral compared to ipsilateral electrodes relative to the high probability distractor location in the mixed-feature condition, nor in the fixed-feature condition. That is, cluster-based permutation tests across time and alpha frequencies between spatial and non-spatial blocks separately for the fixed-features and mixed-features conditions did not yield any reliable differences, neither did the same tests across time on averaged alpha power. Also, a Condition (mixed-features, fixed-features) by Block Type (no spatial bias, spatial bias) repeated measures ANOVA on averaged alpha power within the anticipatory window (-1000 – 0ms) did not show reliable effects (all F's < 2.0, all p's > 0.17).

It is possible that the development of generic suppression by alpha-band activity is dependent on spatial statistical learning. In an exploratory analysis we therefore next investigated whether lateralized alpha power modulations were specific to participants who exhibited slower

search times when the target was presented at the high probability distractor location (i.e., the learners). As visualized in Figure 7B, however, here again there was no evidence that lateralized alpha power was increased in spatial bias blocks relative to no spatial bias blocks, neither in 'learners' nor in 'non-learners' (all Group interactions: all F's < 1.0, all p's > 0.33). Together, and in line with findings from two other EEG studies (Noonan et al., 2016; van Moorselaar & Slagter, 2019), our findings thus argue against the idea that activity of brain regions representing the high probability distractor location can be suppressed in advance, at least as indexed by modulations of alpha-band activity.

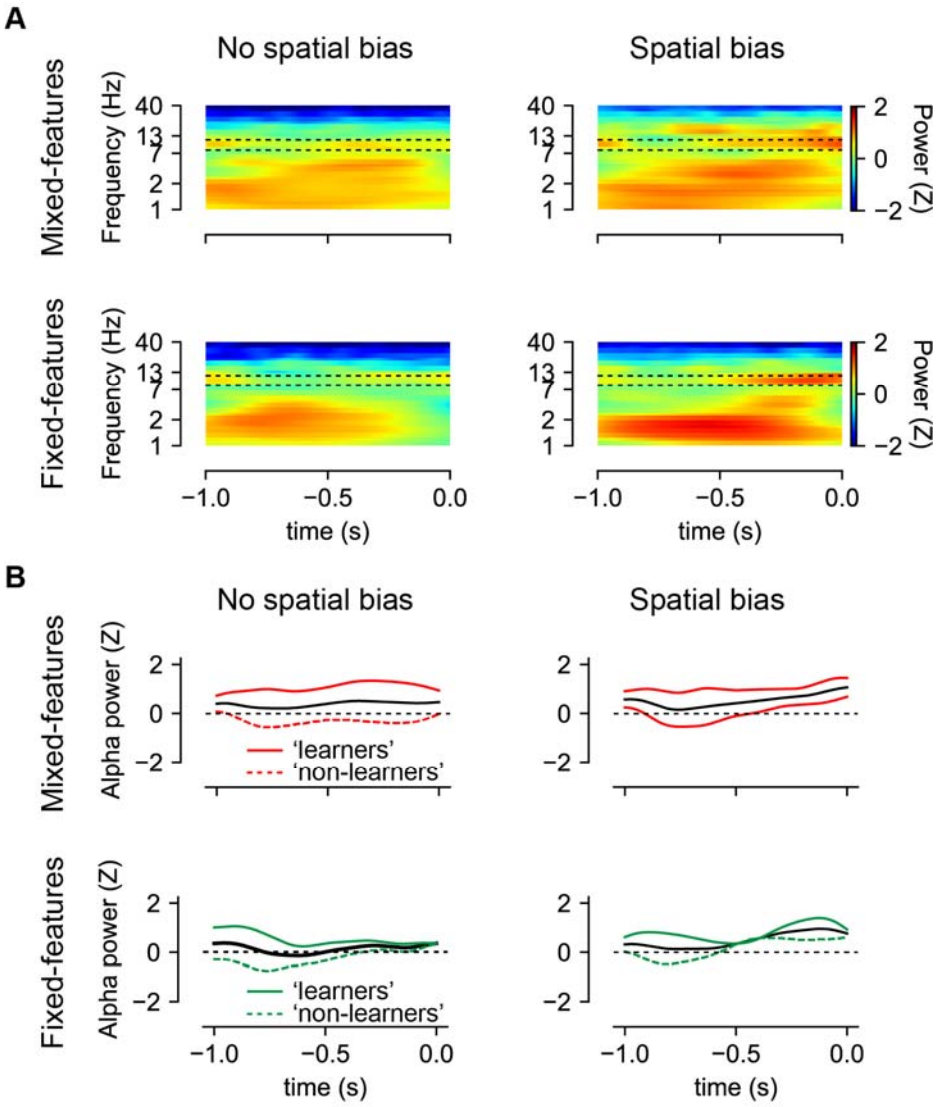

Figure 7| Distractor location learning was not associated with any changes in pre-stimulus alpha-band activity contralateral to the likely distractor location. Shown is posterior lateralized Z scored power. (**A**) Time-frequency representation of contralateral minus ipsilateral (relative to the high probability distractor location) power averaged over electrodes O1/2, PO3/4 and PO7/8. Dotted black outlines show the frequency bands of interest (7.5 – 12 Hz) (**B**) Time series of contralateral minus ipsilateral averaged alpha power (7.5 – 12 Hz) averaged over electrodes O1/2, PO3/4 and PO7/8. Solid black lines show the group mean, whereas colored lines show the effect separately for participants labeled as 'learners' and 'non-learners'.

3. General discussion

The aim of the present study was to gain more insight into the neural mechanisms underlying distractor inhibition as a function of whether it's specific features and/or likely location could be predicted in advance or not. To this end, observers searched for a unique shape among

homogeneous non-target shapes, either in conditions where target and non-target shapes randomly swapped or were fixed (and hence predictable) across a sequence of trials. They performed these search tasks both when the distractor location was not, and when it was predictable in advance. We found that the ability to suppress distracting information was dependent on both distractor feature and location learning, although large individual differences were observed in the extent to which individuals exhibited spatial statistical learning. In no-spatial bias blocks, distractor interference was greatly reduced and associated with an early Pd and an earlier late Pd when distractor and target features were fixed and could thus be predicted compared to when they varied across trials. These findings suggest that distractors could be inhibited faster and more efficiently when distractor features were predictable in advance. Exploratory analyses showed that these effects were dependent on feature-based learning, as distractor costs were present at the start of each block and then gradually decreased, an effect that was accompanied by the emergence of the distractor-evoked Pd.

Distractor inhibition was in addition affected by spatial biases, as in both fixed-features and mixed-features conditions, distractor interference was reduced at high-probability distractor locations. Moreover, the early and late distractor-evoked Pd no longer differed in amplitude or latency between conditions. However, we also found that at the group level, intertrial priming rather than spatial statistical learning could account for the observed decrease in distractor interference in spatial bias blocks. Exploratory analyses showed that while some participants, those in which target processing was not slowed down at the high probability distractor location, indeed did not exhibit spatial statistical learning ('non-learners'), other participants, in which target processing was impaired at the high probability distractor location, did exhibit spatial statistical learning ('learners'). That is, in 'learners' the observed reduction in distractor interference in spatial bias blocks could not solely be explained by intertrial priming. Interestingly, post-hoc analyses of the ERP data showed that 'non-learners' only exhibited the early Pd, while, although not reaching statistical significance, 'learners' only exhibited the typical late Pd. Albeit speculative, these findings may suggest that the

early Pd selectively reflects changes in sensory processing related to intertrial priming, while the late Pd may more selectively reflect distractor inhibition and be more sensitive to distractor-specific spatial statistical learning. Finally, no location learning-related changes in anticipatory alpha-band activity were observed, suggesting that distractor location learning did not result in generic proactive spatial suppression, in line with two previous studies (Noonan et al., 2016; van Moorselaar & Slagter, 2019), but in contrast to another recent study (Wang et al., 2019). Together, the current findings corroborate the idea that distractor suppression, be it feature- or space-based, strongly depends on previous experiences with distracting information (Noonan et al., 2018; van Moorselaar & Slagter, 2020) and reveal the neural mechanisms underlying learned inhibition. In the below, we discuss these findings in more detail in light of the existing literature.

Although distractor interference was reduced in the fixed-feature condition, as in Vatterott and Vecera (2012), it was not eliminated until sufficient experience rejecting a specific distractor accumulated (Stilwell, Bahle, & Vecera, 2019). Similarly, we found that the Pd gradually emerged over the course of a fixed-features block, and was not immediately evoked after a few distractor repetitions. These findings suggest that the observed effects of feature predictability do not simply reflect intertrial priming and conceivably reflect learning over a longer time scale. Many spatial statistical learning studies have also shown that the benefit of distractor predictability cannot solely be accounted for in terms of intertrial priming (e.g.,) (Failing et al., 2019; Ferrante et al., 2018; Sauter et al., 2018), although it should be noted that the standard control (exclude D_{n-1} - D_n) still allows for influences from more distant trials in the past (Maljkovic & Nakayama, 1994). By contrast, here the benefit of distractor location predictability was no longer robust after elimination of all trials in which the distractor location repeated, even though we also observed that intertrial distractor location priming effects were larger in spatial bias than no-spatial bias blocks. Moreover, counter to many studies (e.g.,)(Ferrante et al., 2018; Wang & Theeuwes, 2018b), we also did not observe impaired target processing at the high probability distractor location. One notable difference between the current study and other studies that did observe these two markers of

spatial statistical learning is that here the spatial regularity was not introduced immediately, but only halfway into the experiment. Previous studies have already shown that statistical learning effects persist after the spatial bias is removed (Ferrante et al., 2018; Sauter, Liesefeld, & Müller, 2019). By contrast, the current findings may suggest that presenting spatially neutral blocks first prevents statistical learning in subsequent spatial bias blocks, possibly because in these neutral blocks at least a subset of observers develop a strategy that makes them less sensitive to the subsequently introduced distractor location regularity. Indeed, we found that the same set of participants that were distracted most by the singleton distractor in the first half of the experiment, were also the most sensitive to subsequent spatial distractor regularities. In these participants, in spatial bias blocks irrespective of feature foreknowledge, distractor suppression did not only go above and beyond intertrial priming, but was also accompanied by impaired target processing. Future research is necessary to establish whether initial exposure neutral task settings indeed interferes with later learning and to what extent this effect is modulated by individual differences.

Mirroring our behavioral findings that distractor suppression was dependent on feature-based and space-based learning, the distractor Pd only emerged once feature and/or spatial expectations were formed. Extending previous studies (Burra & Kerzel, 2013; Gaspelin & Luck, 2018a), in the fixed-feature condition, both an early and late Pd were evoked by distractors irrespective of spatial regularities. Yet, these neural inhibitory mechanisms were feature learning dependent, as in no spatial bias blocks both Pd components were absent in the beginning of the block when distractors still interfered with target selection, and only emerged once distractor interference gradually decreased. Moreover, within fixed features blocks, the Pd area was increased at both the start and end of spatial bias blocks, demonstrating that spatial and feature expectations can interactively shape distractor suppression. There is some disagreement on the presence of distractor Pd in mixed-feature conditions, with some studies reporting an N2pc confirming attentional capture by the singleton distractors (Burra & Kerzel, 2013; Hickey, McDonald, & Theeuwes, 2006), whereas other studies observed a Pd albeit selectively on fast RT trials (McDonald

et al., 2013). Here, at least on a group level, we observed a positivity that peaked relatively late, suggestive of a delayed late Pd. As in the Wang et al. (2019) study, however, an early and late Pd emerged at the high probability distractor location, demonstrating that suppressive mechanisms onset earlier in time when expectations can be formed at the spatial level.

A notable aspect of the current findings is that spatial and feature expectations were associated with a larger early Pd and/or an earlier-onset late Pd. This pattern appears to be at odds with two influential mechanistic explanations of distractor suppression, predictive coding (Friston, 2009; Noonan et al., 2018) and habituation (Thompson, 2009; Turatto & Pascucci, 2016), which both posit that previous experiences with distracting information attenuate sensory processing. That is, both accounts postulate that current sensory input is compared to stored representations that are grounded in the expected frequency and contextual properties of distractors, and that the greater the match between the two, the more the sensory response is attenuated or explained away. In this scenario, one would thus predicted repeated distractors to evoke smaller responses and elicit less interference in the system, reducing the need for distractor inhibition. Consistent with this, several previous studies have reported that the amplitude of the Pd decreases as a function of expectations (Heuer & Schubö, 2019; van Moorselaar & Slagter, 2019). To date, it is unclear why in certain contexts, as in the present study, the Pd increases, whereas in others, it decreases. Future work is necessary to establish whether an increased Pd in response to a repeated, predictable distractor is indeed paradoxical.

In many studies reporting a distractor Pd, the Pd can be separated into an early and a late component (e.g.,)(Wang et al., 2019; Weaver et al., 2017). Similarly, the distractor Pd observed here clearly contained two separate peaks, suggestive of two sequential inhibitory processes. Whereas initial reports linked the early Pd (also referred to as PPC or lateralized P1) to lateral asymmetries in stimulus processing (Luck & Hillyard, 1994a) or saliency driven selection (Jannati et al., 2013), later studies suggest that the early Pd contributes to active suppression (Gokce et al., 2014), possibly via intertrial priming (Gokce et al., 2014). This has led to the idea that the early Pd reflects negative

weights on a saliency map (Feldmann-Wüstefeld & Vogel, 2019). Our finding that the early Pd was specific to conditions that allowed for distractor learning, either at the feature or the spatial level, are directly in line with such an interpretation. Critically, however, our exploratory findings also demonstrate that while the early Pd may signal some pre-attentive suppression, by itself the early Pd is insufficient to filter out new sensory input. This is illustrated by the fact that an early target Pd was only observed in subjects that did not show impaired target processing at high probability distractor locations ('non-learners'). By contrast, in the Wang et al. (2019) study, in which target processing was impaired at high probability distractor locations, the early target Pd was followed by a late target Pd. This discrepancy strongly suggests that the late Pd reflects an additional inhibitory process that ultimately determines whether automatic attentional selection by salient distractors can be prevented (Gaspelin & Luck, 2018b). It is notable in this respect that 'non-learners' who did not exhibit impaired target processing and whose performance improvements in spatial blocks could be accounted for by inter-trial priming, only showed an early distractor Pd, while 'learners', who did exhibit impaired target processing and spatial statistical learning above and beyond intertrial priming, only exhibited hints of the late Pd. This further suggests a functional dissociation between the early and late Pd. Albeit speculative, given the post-hoc nature of these latter findings and the small samples sizes of the two groups, the early Pd may hence reflect sensory suppression related to intertrial priming in retinotopically organized visual regions (given that the eyes were at fixation at all times), while the late Pd may reflect suppression related to spatial statistical learning at a higher level in the processing hierarchy, where information is integrated over a longer time scale, at the spatiotopic level. Future research is necessary to more conclusively determine the functional significance of the early and late Pd, and to more systematically examine individual differences in statistical learning.

While it is typically assumed that pre-attentive suppression as indexed by the (early) Pd relies upon proactive inhibition, to date very few studies have also examined anticipatory markers of suppression. Of the few studies that did, despite the prevailing view that alpha oscillations

implement active top-down inhibition (Jensen & Mazaheri, 2010), only a single study observed an increase in anticipatory alpha-band activity contralateral to the likely distractor location that was followed by a subsequent Pd to any stimulus presented at that location (Wang et al., 2019). By contrast, two other studies observed expectation-based distractor suppression as indexed via (early) Pd modulations in the absence of any distractor location-related anticipatory alpha lateralization (Noonan et al., 2016; van Moorselaar & Slagter, 2019). A notable difference that might explain this discrepancy in findings is that in these latter two studies, distractor (and target) features were fixed, whereas in the former study they varied from trial to trial, rendering it possible that generic space-based proactive suppression only develops when distractors cannot be predicted at the feature level. While the current study was designed to address this possibility, the present results do not allow for drawing any resolute conclusions as counter to Wang et al. (2019), in the spatial bias blocks the observed distractor benefits did not go above and beyond intertrial priming, nor were they accompanied by impaired target processing. These findings are not indicative of generic space-based proactive suppression, leaving it unclear whether anticipatory alpha-band modulations implement generic space-based proactive suppression or not. Yet, even the participants that were most sensitive to spatial statistical learning did not show any modulation of anticipatory alpha activity. Although one should be careful in interpreting null findings and there may be alternative explanations why alpha modulations were absent in 'learners', the current findings add to a rapidly growing literature suggesting that alpha-band oscillations do not play a causal role in proactive distractor suppression (Antonov, Chakravarthi, & Andersen, 2020; Foster & Awh, 2018).

To conclude, here we find that the ability to inhibit distracting information is strongly dependent on both feature-based and space-based distractor learning, as indicated by corresponding reductions in distractor costs during visual search and in the amplitude and latency of the distractor-evoked Pd ERP component. Furthermore, exploratory analyses revealed individual differences in sensitivity to distractor location regularities that may suggest a functional dissociation between the early and late Pd: while the early Pd may reflect early sensory suppression related to

intertrial priming, the late Pd may reflect later spatiotopic suppression related to statistical learning over a longer time scale. Finally, our findings suggest that current distractor learning may be influenced by the preceding task context, as distractor location repetition effects in spatial bias blocks preceded by no-spatial bias blocks were largely driven by intertrial priming.

References

Allenmark, F., Zhang, B., Liesefeld, H. R., Shi, Z., & Müller, H. J. (2019). Probability cueing of singleton-distractor regions in visual search: the locus of spatial distractor suppression is determined by colour swapping. *Visual Cognition*, 1-19.

Antonov, P. A., Chakravarthi, R., & Andersen, S. K. (2020). Too little, too late, and in the wrong place: Alpha band activity does not reflect an active mechanism of selective attention. *Neuroimage*, 117006.

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in cognitive sciences, 16*(8), 437-443.

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & psychophysics, 55*(5), 485-496.

Burra, N., & Kerzel, D. (2013). Attentional capture during visual search is attenuated by target predictability: evidence from the N2pc, Pd, and topographic segmentation. *Psychophysiology, 50*(5), 422-430.

de Vries, I. E., van Driel, J., & Olivers, C. N. (2017). Posterior alpha EEG dynamics dissociate current from future goals in working memory guided visual search. *Journal of Neuroscience*, 2945-2916.

Failing, M., Feldmann-Wustefeld, T., Wang, B., Christian, O., & Theeuwes, J. (2019). Statistical regularities induce spatial as well as feature-specific suppression. *Journal of Experimental Psychology: Human Perception and Performance*.

Failing, M., & Theeuwes, J. (2020). More capture, more suppression: Distractor suppression due to statistical regularities is determined by the magnitude of attentional capture. *Psychonomic Bulletin & Review, 27*(1), 86-95.

Feldmann-Wüstefeld, T., & Vogel, E. K. (2019). Neural evidence for the contribution of active suppression during working memory filtering. *Cerebral Cortex, 29*(2), 529-543.

Feldmann-Wüstefeld, T., Brandhofer, R., & Schubö, A. (2016). Rewarded visual items capture attention only in heterogeneous contexts. *Psychophysiology, 53*(7), 1063-1073.

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.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 1030.

Foster, J. J., & Awh, E. (2018). The role of alpha oscillations in spatial attention: Limited evidence for a suppression account. *Current opinion in psychology*.

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in cognitive sciences, 13*(7), 293-301.

Gaspelin, N., & Luck, S. J. (2018a). Combined Electrophysiological and Behavioral Evidence for the Suppression of Salient Distractors. *Journal of cognitive neuroscience*, 1-16.

Gaspelin, N., & Luck, S. J. (2018b). The role of inhibition in avoiding distraction by salient stimuli. *Trends in cognitive sciences*, *22*(1), 79-92.

Gokce, A., Geyer, T., Finke, K., Müller, H. J., & Töllner, T. (2014). What pops out in positional priming of pop-out: insights from event-related EEG lateralizations. *Frontiers in psychology*, *5*, 688.

Goschy, H., Bakos, S., Müller, H. J., & Zehetleitner, M. (2014). Probability cueing of distractor locations: both intertrial facilitation and statistical learning mediate interference reduction. *Frontiers in psychology*, *5*, 1195.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., . . . Hämäläinen, M. S. (2014). MNE software for processing MEG and EEG data. *Neuroimage*, *86*, 446-460.

Heuer, A., & Schubö, A. (2019). Cueing distraction: electrophysiological evidence for anticipatory active suppression of distractor location. *Psychological research*, 1-11.

Hickey, C., McDonald, J. J., & Theeuwes, J. (2006). Electrophysiological evidence of the capture of visual attention. *Journal of cognitive neuroscience*, *18*(4), 604-613.

Hickey, C., Olivers, C., Meeter, M., & Theeuwes, J. (2011). Feature priming and the capture of visual attention: Linking two ambiguity resolution hypotheses. *Brain research*, *1370*, 175-184.

Jannati, A., Gaspar, J. M., & McDonald, J. J. (2013). Tracking target and distractor processing in fixed-feature visual search: evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(6), 1713.

JASP-TEAM. (2018). JASP (Version 0.8.2) [Computer software].

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in human neuroscience*, *4*, 186.

Leber, A. B., Gwinn, R. E., Hong, Y., & O'Toole, R. J. (2016). Implicitly learned suppression of irrelevant spatial locations. *Psychonomic Bulletin & Review*, *23*(6), 1873-1881.

Liesefeld, H. R. (2018). Estimating the timing of cognitive operations with MEG/EEG latency measures: a primer, a brief tutorial, and an implementation of various methods. *Frontiers in neuroscience*, *12*, 765.

Liesefeld, H. R., Liesefeld, A. M., Töllner, T., & Müller, H. J. (2017). Attentional capture in visual search: Capture and post-capture dynamics revealed by EEG. *Neuroimage*, *156*, 166-173.

Luck, S. J., & Hillyard, S. A. (1994a). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*(3), 291-308.

Luck, S. J., & Hillyard, S. A. (1994b). Spatial filtering during visual search: evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 1000.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & cognition*, *22*(6), 657-672.

McDonald, J. J., Green, J. J., Jannati, A., & Di Lollo, V. (2013). On the electrophysiological evidence for the capture of visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 849.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *reason*, *4*(2), 61-64.

Noonan, M. P., Adamian, N., Pike, A., Printzlau, F., Crittenden, B. M., & Stokes, M. G. (2016). Distinct mechanisms for distractor suppression and target facilitation. *Journal of Neuroscience, 36*(6), 1797-1807.

Noonan, M. P., Crittenden, B. M., Jensen, O., & Stokes, M. G. (2018). Selective inhibition of distracting input. *Behavioural brain research, 355*, 36-47.

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior research methods, 42*(1), 188-204.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience, 2011*, 1.

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and clinical neurophysiology, 72*(2), 184-187.

Sauter, M., Liesefeld, H. R., & Müller, H. J. (2019). Learning to suppress salient distractors in the target dimension: Region-based inhibition is persistent and transfers to distractors in a nontarget dimension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

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

Sawaki, R., & Luck, S. J. (2010). Capture versus suppression of attention by salient singletons: Electrophysiological evidence for an automatic attend-to-me signal. *Attention, Perception, & Psychophysics, 72*(6), 1455-1470.

Stilwell, B. T., Bahle, B., & Vecera, S. P. (2019). Feature-based statistical regularities of distractors modulate attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 45*(3), 419.

Thaler, L., Schütz, A. C., Goodale, M. A., & Gegenfurtner, K. R. (2013). What is the best fixation target? The effect of target shape on stability of fixational eye movements. *Vision research, 76*, 31-42.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & psychophysics, 51*(6), 599-606.

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica, 135*(2), 77-99.

Theeuwes, J. (2019). Goal-driven, stimulus-driven, and history-driven selection. *Current opinion in psychology, 29*, 97-101.

Thompson, R. F. (2009). Habituation: a history. *Neurobiology of Learning and Memory, 92*(2), 127.

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.

van Moorselaar, D., & Slagter, H. A. (2019). Learning What Is Irrelevant or Relevant: Expectations Facilitate Distractor Inhibition and Target Facilitation through Distinct Neural Mechanisms. *Journal of Neuroscience, 39*(35), 6953-6967. doi:https://doi.org/10.1523/JNEUROSCI.0593-19.2019

van Moorselaar, D., & Slagter, H. A. (2020). Inhibition in selective attention. *Annals of the New York Academy of Sciences*.

Vatterott, D. B., & Vecera, S. P. (2012). Experience-dependent attentional tuning of distractor rejection. *Psychonomic Bulletin & Review, 19*(5), 871-878.

Wang, B., & Theeuwes, J. (2018a). How to inhibit a distractor location? Statistical learning versus active, top-down suppression. *Attention, Perception, & Psychophysics*, 1-11.
Wang, B., & Theeuwes, J. (2018b). Statistical regularities modulate attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 44*(1), 13.
Wang, B., van Driel, J., Ort, E., & Theeuwes, J. (2019). Anticipatory distractor suppression elicited by statistical regularities in visual search. *Journal of cognitive neuroscience*, 1-14.
Weaver, M. D., van Zoest, W., & Hickey, C. (2017). A temporal dependency account of attentional inhibition in oculomotor control. *Neuroimage, 147*, 880-894.
Zhang, B., Allenmark, F., Liesefeld, H. R., Shi, Z., & Muller, H. (2019). Probability cueing of singleton-distractor locations in visual search: priority-map-or dimension-based inhibition? *Journal of Experimental Psychology: Human Perception and Performance*.