AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: Human Perception and Performance

ISSN: 0096-1523

2022, Vol. 48, No. 3, 262–274
https://doi.org/10.1037/xhp0000987

Statistical Learning of Across-Trial Regularities During Serial Search

Ai-Su Li, Louisa Bogaerts, and Jan Theeuwes
Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam
Institute Brain and Behavior Amsterdam (iBBA), Amsterdam, the Netherlands

Previous studies have shown that attention becomes biased toward those locations that frequently contain a target and is biased away from locations that have a high probability to contain a distractor. A recent study showed that participants also learned regularities that exist across trials: Participants were faster to find the singleton when its location was predicted by the location of the target singleton on the previous trial. Note, however, that this across-trial statistical learning was only demonstrated for parallel search involving "pop-out" singleton targets. The current study investigated whether there is also learning of across-trial regularities when search is serial, using a T-among-Ls task. In Experiment 1, using search displays with a gray T-target among gray Ls, we found that participants did not learn the existing across-trial regularities. In Experiment 2 we used the same display and same regularities except that during the first half of the experiment the targets were colored red, allowing feature search. Critically, now participants did learn the across-trial regularities during pop-out feature search and the learned biases persisted when search was serial again. Participants were not aware of these regularities suggesting that learning was automatic and implicit. We propose that across-trial target-target associations learned during feature search shape a flexible priority map whereby the selection of the predicting location results in up-weighting of the predicted location on the next trial. This flexible priority map remained active even when search task changed dramatically from parallel to serial search.

Public Significance Statement

The present study investigates the boundary conditions of implicit learning across trials. We show that during slow and serial search participants are not able to learn across-trial statistical regularities, most likely because there is too much noise for learning to occur. When we created conditions that reduced noise and facilitated the learning of across-trial statistical regularities, we show that the learned target-association biases in the feature search could persist when there was much noise again during serial search.

Keywords: across-trial regularities, attentional bias, serial search, statistical learning

Supplemental materials: https://doi.org/10.1037/xhp0000987.supp

Our visual environment contains abundant information, both relevant and irrelevant. Finding relevant objects while ignoring irrelevant and distracting information is a fundamental skill of any organism that needs to survive in this cluttered environment. By selectively attending relevant information, while inhibiting information that is irrelevant, search can become efficient. Traditionally, attentional selection was considered to be the result of the interaction between the goals of the observer (*current selection goals*) and the physical properties of the visual environment (*salience of the objects*; Egeth & Yantis, 1997; Theeuwes, 2010). Recently, Awh et al. (2012) argued that, in many situations, selection is not merely the result of goals of the observer nor the result of stimulus-driven factors (i.e., bottom-up saliency). As a third factor, *selection history* was proposed. Selection history refers to previous experiences in searching a display eliciting enduring selection biases that are unrelated to the intentions of the observer or the salience of the objects in the visual field (Anderson et al., 2021; Theeuwes, 2018, 2019).

Many familiar phenomena belong to the family of selection history effects. For instance, intertrial repetitions of target and distractor features and/or their locations facilitates search on the next

Ai-Su Li https://orcid.org/0000-0002-2340-3422
Louisa Bogaerts https://orcid.org/0000-0001-6145-8662
Jan Theeuwes https://orcid.org/0000-0002-5849-7721

Jan Theeuwes and Louisa Bogaerts were supported by a European Research Council (ERC) advanced Grant 833029 –[LEARNATTEND]. Ai-Su Li was supported by a China Scholarship Council (CSC) scholarship [201806920032].

Data and analysis materials for all experiments are available at https://github.com/AisuLi/SL_across-trial_TT, and none of the experiments was preregistered.

Correspondence concerning this article should be addressed to Ai-Su Li, Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Van der Boechorststraat 7-9, 1081 BT Amsterdam, the Netherlands. Email: a2.li@vu.nl

trial, known as repetition priming (e.g., Allenmark et al., 2021; Lamy et al., 2008; Maljkovic & Nakayama, 1994). It was shown that priming can make a stimulus appear to be more salient even though it physically had the same salience as the stimulus that was not primed (e.g., Theeuwes & Van der Burg, 2013). In the same vein, stimulus features previously associated with a monetary reward are more salient than nonrewarded features, even if they are no longer rewarded, an effect known as value-driven attentional capture (e.g., Anderson, 2019; Anderson et al., 2011; Failing & Theeuwes, 2018; Qin et al., 2021).

It is also known that participants can learn statistical regularities present in the environment, which in turn affects attentional selection. Through a process known as *visual statistical learning* (VSL) participants extract the spatial and/or temporal regularities present in the visual input, affecting attention and perception (e.g., Anderson et al., 2021; Frost et al., 2019; Perruchet & Pacton, 2006). Following the seminal work on statistical learning by Saffran, Aslin, and Newport (1996), who demonstrated that infants can extract trisyllabic patterns from continuous speech, a large number of studies have shown that individuals of all ages possess remarkable abilities in successful tracking patterns of co-occurrence of temporally adjacent elements in input streams with different types of visual stimuli (e.g., Dennis et al., 2006; Fiser & Aslin, 2002; Howard et al., 2008; Olson & Chun, 2001; Thomas et al., 2018; Turk-Browne et al., 2005; Turk-Browne & Scholl, 2009). It has been claimed that this process of assimilating statistical patterns in the input usually occurs incidentally and automatically (e.g., Aslin et al., 1998; Fiser & Aslin, 2001; Turk-Browne et al., 2005). Providing a powerful learning mechanism, VSL plays a critical role in many cognitive domains (see Bogaerts, Frost, & Christiansen, 2020 for a discussion), such as language acquisition (e.g., Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996), object and scene perception (e.g., Fiser & Aslin, 2005; Turk-Browne et al., 2010), classification learning (e.g., Aron et al., 2006; Aron et al., 2004), and memory (e.g., Brady et al., 2009; Umemoto et al., 2010).

More directly related to the current study, several studies provided evidence that VSL can bias attentional selection in an implicit way. For instance, participants can learn that a certain location in a search display has a higher probability of containing targets (e.g., Addleman et al., 2018; Ferrante et al., 2018; Geng & Behrmann, 2002, 2005) or distractors (e.g., Goschy et al., 2014; Wang & Theeuwes, 2018; Won et al., 2019; Zhang et al., 2019), respectively, accelerating target detection or enhancing distractor suppression, which both result in more efficient search. It is generally assumed that VSL modulates the allocation of visual attention via dynamic weight adjustments within the spatial priority map (e.g., Anderson et al., 2021; Fecteau & Munoz, 2006; Itti & Koch, 2001; Zelinsky & Bisley, 2015), even though others have challenged this view (see Allenmark et al., 2019; Liesefeld & Müller, 2021; Sauter et al., 2018). Taking the study by Geng and Behrmann (2002) as an example, each spatial location is initially weighed equally due to the same physical saliency and features (i.e., similar letters). But as, over time, different targets are selected, the weights of target locations selected in the past episodes are accumulated so that the location selected most frequently obtained the highest weights within the priority map. As a result, search is facilitated when the target appears at this prioritized location. Likewise, the repeated suppression of a location containing distracting information results in a gradual de-prioritization. It is worth noting that the prioritization and de-prioritization of locations based on target and distractor regularities has been shown in a wide range of search paradigms, such as simple feature search (Ferrante et al., 2018), classical additional singleton paradigm (Wang & Theeuwes, 2018), additional singleton in different dimension with dense displays (Goschy et al., 2014). Noting that the effect of location probabilities on attentional selection shares many characteristics with the learning of (motor) habits, such as the incidental and gradual nature of the learning, and that it is not affected by working memory load (Gao & Theeuwes, 2020; Won & Jiang, 2015), led some to characterize it as "habit-like attention" (see Jiang, 2018; Jiang & Sisk, 2019, for a discussion).

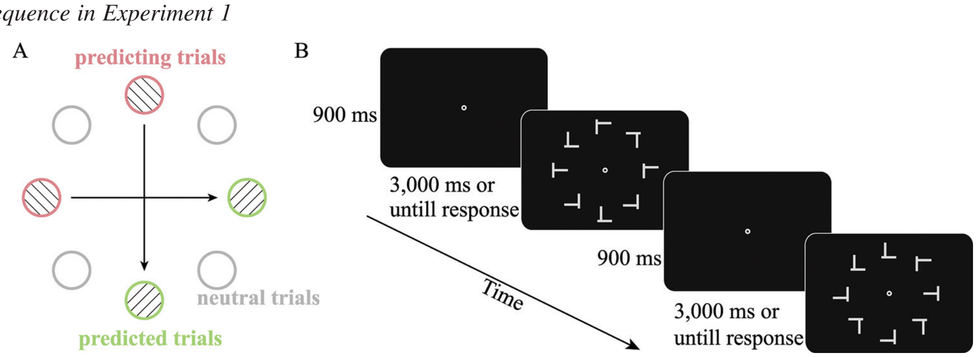
Such distributional regularities (pertaining the frequency of occurrence of a certain stimulus on a certain spatial location) are, however, clearly not "the only game in town." Another line of research that focuses on the VSL of the relationship among objects is known as contextual cueing (CC) which is considered to be an example of the spatial-binding relationship (see Goujon et al., 2015; Jiang et al., 2019; for a review). Compared with new search displays, displays that are repeated across trials (having the same spatial configuration of target and distractors) improve visual search performance, suggesting that the repeated configurations help to guide attention toward the associated target locations (e.g., Bergmann & Schubö, 2021; Chun & Jiang, 1998, 2003; Jiang & Wagner, 2004). In addition, pairs or triplets of visual stimuli that regularly co-occur in space can also be extracted and are then perceived as groups, increasing the perceptual capacity (e.g., Brady et al., 2009; Umemoto et al., 2010; Zhao et al., 2011; Zhao & Yu, 2016).

Certainly, more studies on VSL targeted the learning of sequential regularities (i.e., temporally co-occurring pairs, triplets, or longer sequences), owing to the functional significance for the visual system to use contextual information across time (e.g., Bogaerts, Richter, et al., 2020; Fiser & Aslin, 2002; Henin et al., 2021; Pacton et al., 2015; Pacton & Perruchet, 2008; Remillard, 2009; Schapiro et al., 2013; Turk-Browne et al., 2005; Turk-Browne & Scholl, 2009; Yu & Zhao, 2018; Zhao et al., 2013). For example, in the study by Fiser and Aslin (2002), participants viewed a movie with single shape horizontally moving back and forth until it was completely occluded by the vertical bar and switched to another shape. The shape sequence was not random but rather structured into triplets, so that after the occurrence of the first shape of a triplet the following shapes were predictable. It was found that observers were sensitive to this triplet structure containing temporal contextual information. This sensitivity to structure leads to implicit perceptual anticipation and faster object recognition (Turk-Browne et al., 2010). This raises the question if sequential, across-trial regularities can also modulate attentional selection in visual search. In other words, is it possible to use (implicit) knowledge of temporal structure to guide attention in space? This is by no means self-evident. Contrary to the effect of distributional regularities that has been commonly observed, the modulation of attentional allocation by across-trial regularities is not intertwined with repetition (location) priming nor can it result from changes in the weights of a single, static priority map. Instead, it requires that weights within the spatial priority map are dynamically adapted on a trial-to-trial basis.

In four different experiments, Li and Theeuwes (2020) demonstrated that attentional selection can indeed be biased by across-trial regularities concerning the target locations. Participants performed the additional singleton task (Theeuwes, 1991, 1992) searching for a shape singleton (i.e., a circle among diamonds or a diamond among circles) within a search array with eight items positioned in an imaginary circle. Although the target was equally likely to appear at one of eight locations, unbeknownst to participants, statistical regularities regarding target positions across trials were built in. The regularities were two temporal sequences with two target locations each (e.g., T1->T2), which meant the position of T1 on the current trial was 100% predictive of the position of T2 on the next trial. For example, a target singleton on the leftmost position (T1) was always followed by a target singleton on the rightmost position in the display (T2; see Figure 1A). The results showed that RTs for predictable T2 targets (which we will here label the *predicted* condition) were faster than those for T1 targets (labeled the *predicting* condition) or targets appearing on locations that were not included in the two regular temporal sequences (labeled the *neutral* condition), suggesting that observers had learned the across-trial regularities and used them to bias attentional selection. Moreover, this pattern of results remained intact when salient color distractors were present. More recently, Wang et al. (2021) investigated the effect of across-trial regularities regarding distractor locations using the same additional singleton task. Their study also indicated that participants learned these regularities and further showed that learning was very fast (with RT benefits present after just a few trials). In both studies, participants were basically unaware of the regularities present in the search displays. As such, we can conclude that participants can learn—in an automatic and implicit way—across-trial regularities, which in turn facilitates search.

Although Li and Theeuwes (2020) firmly established the across-trial VSL for target locations in all four experiments, each and every experiment involved parallel feature search in which the target was a singleton popping out from the display. When the target is a pop-out singleton, it is reasonable to assume that across-trial target-target location associations are easily formed because on each trial attention is immediately shifted to the target singleton. In this way, participants are able to associate a shift of attention to a location on trial T1 with a shift of attention to a particular other location on the following trial T2. Even though from a theoretical point of view this finding is crucial, it is obvious that in everyday life, there is hardly any pure pop-out search, and most of the time search is serial and inefficient. Hence, the aim of the present study was to gain further insight into the modulation of attention by target regularities that span multiple trials, by investigating whether across-trial facilitation also occurs when search is serial and inefficient. During this type of search, attention is deployed serially to each item in the display until the target is found (Treisman & Gelade, 1980; Wolfe, 1994). It can be hypothesized that such serial scanning might reduce or even prevent the formation of across-trial association between target locations, which would result in no across-trial biasing of attention. This would point to an important boundary condition for VSL. On the other hand, especially because search is so inefficient, making use of hidden regularities may improve search dramatically. If across-trial associations can be learned, their effect during serial search are expected to be much more substantial. The present study conducted two experiments involving search displays consisting of eight items presented at fixed locations on an imaginary circle. The built-in across-trial pair regularities were similar to those of Li and Theeuwes (2020) except that instead of searching for a pop-out target, participants searched for a rotated "T" target among seven rotated "L" distractors. The subtle difference between distractors and targets, both composed of vertical and horizontal lines but with a different spatial arrangement, forces search to be serial (Bergen & Julesz, 1983; Egeth & Dagenbach, 1991; Kwak et al., 1991). Indeed, the spatial-configuration search

Figure 1
Depiction of the Across-Trial Regularities Concerning Target Locations and the Stimulus Display Sequence in Experiment 1



Note. (A) Illustration of two regularity pairs: A predicting trial (denoted with red circles filled with left-oriented diagonal lines) predicts the location of the target on the next trial (labeled as predicted; denoted with green circles filled with right-oriented diagonal lines). The neutral condition (denoted with unfilled gray circles) consists of filler trials where the target appears randomly at other four locations. (B) Example of a stimulus display sequence with an across-trial regularity (a rotated "T" target on the far left is followed by a rotated "T" target on the far right). Note the target rotation was random from trial to trial. See the online article for the color version of this figure.

where the target is defined by the spatial arrangement of line segments, such as the T-among-Ls task, is generally considered to be the gold standard for true "serial" search (Wolfe, 1998).

Experiment 1

Experiment 1 was designed to determine whether participants could learn statistical regularities regarding the target positions across trials in serial search. If participants can extract across-trial regularities of target positions, RTs for trials that are predicted by the previous trial should be faster than comparable trials that are not predicted by a previous trial.

Method

Participants

Using G*Power (Faul et al., 2007), an a priori power analysis was conducted, with $\alpha = .05$, $1 - \beta = .9$, and the default value of $r = .5$ as the assumed correlation among repeated measures (note that this is conservative, see Brysbaert, 2019). Because no prior research regarding the effect of across-trial regularities on serial search was available,[1] we chose $f = .25$ (corresponding to $\eta_p^2 = .06$) as the smallest effect size of interest, reflecting a theoretically meaningful effect in psychological research (see also Brysbaert, 2019). Our effect of interest was the main effect of target regularity (specifically the difference between predicted and unpredicted conditions) in the two-way repeated-measures analysis of variance (RM-ANOVA) with block and target regularity as within-subject factors. The minimum sample size for this effect was calculated to be 44 participants. Considering the additional noise in online experiments, 59 participants (22 females and 37 males; M_{age} = 24.64 years, SD_{age} = 5.79) were recruited via Prolific (Palan & Schitter, 2018). Participants were rewarded £5.63 after finishing the whole experiment. Two participants who had overall accuracies lower than 75% were excluded. All participants reported normal or corrected-to-normal visual acuity and gave informed consent before the experiment. The study was approved by the Ethics Committee of Department of Experimental and Applied Psychology of Vrije Universiteit Amsterdam.

Apparatus and Stimuli

The experiment was programmed using OpenSesame (3.3.9b1) and run via OSWeb (Mathôt et al., 2012) and JATOS (Lange et al., 2015). Participants completed the task on their own computers or laptops. Considering the fact that resolution of monitors could vary, the resolution specified in the experiment was 1,024 × 768 pixels (px). All stimuli would be displayed on this "virtual monitor" in the center of the screen.

The search display contained one "T" target (rotated 90° to the left or right) and seven "L" distractors in light gray (RGB: 210/210/210) presented against a black (RGB: 30/30/30) background. Each stimulus subtended 64 × 64 px and was centered 210 px from a white (RGB: 255/255/255) fixation dot with a radius of 8 px. "L" distractors had a 12.5% offset in the line junction to increase search difficulty, forcing the search to be serial in nature. "L" distractors had 4 possible rotations (0°, 90°, 180°, or 270°).

Design and Procedure

The target "T" was tilted 90° to the left or right with equal probability and was present on each trial. The target was equally likely to appear at one of eight locations. On each trial, the rotation angle for different "L" distractors was randomly selected and assigned to the nontarget locations, with the constraint any given rotation angle could not be used more than twice in the same search array. All the trials were randomized within each block except the particular regularities regarding target locations across trials. Specifically, for half of the participants, if on a trial the target "T" was presented at the leftmost position of the display, it was always followed by a subsequent trial with the target presented at the rightmost position of the display. For the same group of participants, if the target "T" was presented at the top position in the display, it was always followed by the target at the bottom position on the following trial (see Figure 1A). For the other half of participants, regularity pairs with opposite directions (rightmost [R] → leftmost [L], bottom [B] → top [T]) were built in. The same regularity pair could not repeat back to back (i.e., the trial sequences RLRL and BTBT were not allowed). Notably, the regularities only concerned the spatial location of the target, the identity of the target (rotated to the left or right) varied randomly across trials.

Each trial began with the presentation of a fixation dot at the center of the screen. After 900 ms, the search array was presented, which remained on until the participant responded. Participants' task was to search for the rotated "T" (among the rotated "L" distractors) and report its rotation. They were instructed to maintain fixation on the central dot while doing the task and to press the appointed key ("Z" and "/" for respectively left and right rotation) as fast and as accurately as possible to report the target's orientation. If participants took too long to respond (> 3,000 ms), or responded incorrectly, the text display "Your response was wrong!" was presented for 800 ms together with a 800-Hz tone. At the end of each block, feedback regarding accuracy and mean RTs in the just completed block was given. Breaks between blocks were controlled by participants themselves (with a minimum of 30 seconds).

The whole task consisted of an experiment part (practice blocks plus eight experimental blocks) and a questionnaire part. The experimental blocks did not start until the performance in the practice block with all[2] practice trials in a random order (with no regularities present), had reached the criteria of accuracy > 85% as well as mean RT < 2,000 ms, which was to make sure participants got familiar with the task during the practice. Each experimental block contained 64 trials, yielding 16 predicting trials (predictive of the upcoming target location), 16 predicted trials (with a predictable target location), and 32 neutral trials (with a target on one of the locations that were not

[1] Note that previous work by Li and Theeuwes (2020) on the effect of across-trial regularities in parallel search reported a difference between predicted and unpredicted conditions with a large effect size of $f = 0.42$ (average across all experiments). Given our hypothesis that across-trial regularities might not drive search in the same way in serial search, we did not base the sample size calculation on this effect size. It is however worth noting that a power calculation with $n = 44$ and this effect size results in a power of 99.98%.

[2] Owing to a programming error, there were 20 trials in a practice block for the first 51 participants who performed on average 67 (SD = 49) practice trials. The trial count in a practice block was 40 for the final six participants who performed on average 93 (SD = 41) practice trials.

included in the two regular sequences). After finishing the experiment part, a questionnaire with three questions was filled in. First, participants were asked whether they had noticed a sequence of some target locations, such as one specific location that was always followed by another specific location. Subsequently, two eight-alternative forced-choice questions were asked. More specifically, participants were shown a search array with a target at the predicting location (among seven "L" distractors) on the left side of the screen, and there was an array of eight circles representing eight locations on the right side of the screen. They were asked to choose one location that they thought the target was most likely to appear on following the search array presented on the left side of the screen. The same question was asked for the second regularity pair. All three questions were followed by a confidence rating on a five-point scale (1 = *not certain at all*, 5 = *very certain*).

Results

In the current experiment there were three trial types: predicting, predicted, and neutral (see Figure 1A). Even though the predicting and neutral trials were both unpredicted by the previous trial, it is important to note that there is a difference between these trials. The predicting condition completely matched the predicted condition with respect to target spatial locations (i.e., leftmost/rightmost/top/bottom) whereas in the neutral condition the target randomly appeared at other four locations. Also, the neutral condition contained trials in which there may have been intertrial priming in which the location of the target was repeated across two consecutive trials. Therefore, we split the data into three conditions: predicted, predicting, and neutral (for the latter condition, all intertrial target location priming trials were removed).

Analysis

RTs analyses were limited to correct trials (89.26%) only. For the correct trials of each block of each participant, RTs were submitted to a nonrecursive trimming procedure (Vanselst & Jolicoeur, 1994) that uses cell size to determine a criterion number of *SD*s from the mean beyond which an observation is considered as an outlier (.40% of trials excluded). Then, trials with RTs < 200 ms (.01%) or with intertrial priming effect of the target location (6.38%) were also excluded from analysis (see Figure S1 in online supplemental materials for details of RTs as a function of across-trial distance of the target location). Finally, mean RTs and accuracies were each submitted into a two-way (block and target regularity) RM-ANOVA. Greenhouse-Geisser corrected *p*-values (p_c) were used in case of sphericity assumption violations. In addition, whenever a comparison using traditional null hypothesis testing was insignificant, we also quantified the Bayes factor (BF) using Bayesian hypothesis testing in JASP (Wagenmakers et al., 2018) to evaluate the strength of the evidence for the alternative hypothesis (H1) over the null hypothesis (H0). Data were visualized using Prism GraphPad and Adobe Illustrator.

Learning Effect

Overall RTs and mean RTs across blocks as a function of target regularity condition are illustrated in, respectively, panels B and C of Figure 2. A two-way RM-ANOVA with blocks (1–8) and regularity (predicted, predicting, neutral) on mean RTs only revealed the significant main effect of block, $F(7, 392) = 30.08$, $p_c < .001$, $\eta_p^2 = .35$. No significant effects were observed for the main effect of target regularity (predicted: 1529 ms, predicting: 1562 ms, neutral: 1549 ms), $F(2, 112) = .69$, $p_c = .50$, $\eta_p^2 = .01$, $BF_{01} = 6.19$ (moderate evidence for the absence of any condition differences) or the Block × Regularity interaction, $F(14, 784) = 1.30$, $p_c = .20$, $\eta_p^2 = .02$, $BF_{01} = 3720.89$.

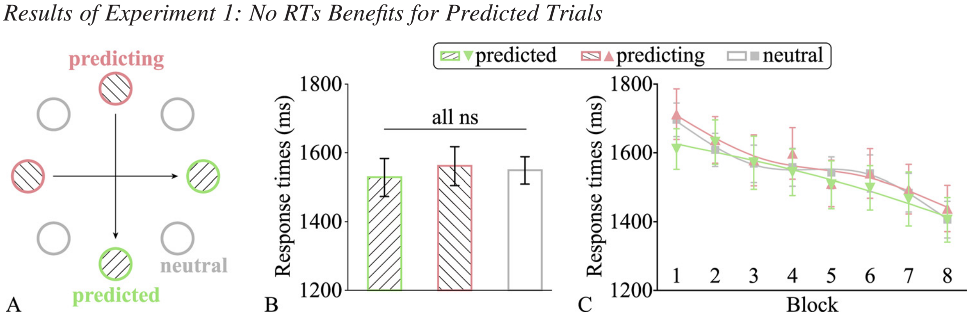
A RM-ANOVA on accuracies showed similar results. Only the main effect of block was significant, $F(7, 392) = 17.02$, $p_c < .001$, $\eta_p^2 = .23$. Neither the main effect of regularity (predicted: 89.12%, predicting: 87.98%, neutral: 89.47%), $F(2, 112) = 2.75$, $p = .07$, $\eta_p^2 = .05$, $BF_{01} = 1.80$, nor Block × Regularity interaction, $F(14, 784) = .60$, $p_c = .82$, $\eta_p^2 = .01$, $BF_{01} = 5293.94$, reached the significance.

Across-Trial Target Location Distance Analysis

Because our regularity pairs always contained a transition with a distance of four locations (e.g., a target at the top position was always followed by a target at the bottom position), a more even comparison would include only trials with a four-location distance as well between the current target position and the target position on the previous trial. When we restricted the across-trial target distance to 4 locations and performed a paired-samples *t* test comparing predicted and neutral trials, the result again indicated no

Figure 2

Results of Experiment 1: No RTs Benefits for Predicted Trials



Note. (A) Illustration of regularity pairs and the three conditions. Overall RTs (B) and RTs across blocks (C) as a function of target regularity (note that both predicting and neutral condition were unpredicted conditions). The error bars denote 95% confidence intervals. The lines in C represent the fitted data using the smoothing spline method in Prism GraphPad software. ns = not significant. See the online article for the color version of this figure.

significant difference between predicted (1,529 ms) and neutral (1,581 ms) trials, $t_{56} = -1.74$, $p = .09$, $\eta_p^2 = .05$, $BF_{01} = 1.67$, showing anecdotal evidence for the absence of the significance.

Awareness of the Regularities

Thirteen of 57 participants reported to have been aware of the across-trial association of the target location during the experiment, with a mean confidence score (CS) of 3.46 ± .88 ($M \pm SD$). Yet only two of them correctly chose both of the predicted locations (CS: 3.25 ± .35), whereas others indicated wrong locations (CS: 3.00 ± 1.00). The remaining 44 participants reported to be unaware of the across-trial association of the target location (CS: 3.43 ± 1.07). The mean CS regarding locations they chose was 1.93 ± .91.

Discussion

Even though this experiment had across-trial regularities regarding the target, it turns out that participants did not learn and use these regularities to improve search. Indeed, the time to find the target in predicted trials was not different than the time to find a target in trials in which the location was selected randomly. This finding is unlike what was previously reported for feature (pop-out) search, where robust RT benefits were found for predicted over unpredicted trials (Li & Theeuwes, 2020). There are in principle two possibilities why no benefit was found in the current task. First, it is possible that participants were not able to learn the regularities present in the display because of the serial nature of the current search task. Indeed, as searching for a target "T" among "L"s is considered to be a true serial search task, it is feasible that participants scan the display serially location by location, most likely in a more or less random order. Because search is random it may be difficult to form associations between the two locations across trials that make up the regularity. Second, it is possible that participants did implicitly learn the regularities but were unable to apply them because of the nature of the search task. Because search is serial observers may always use a random scanning order to find the target and hence even a learned regularity might not result in faster search times.

The aim of Experiment 2 was to test these possibilities. Therefore, we combined an initial learning phase involving feature search and a second phase involving serial search. During the first half of the experiment, we colored the target "T" red whereas the nontargets remained gray "L"s. This created conditions similar to the feature search task of Li and Theeuwes (2020), which did result in learning of the across-trial contingencies. During the second part of the experiment, we changed the task to a normal serial T-among-Ls task as all elements were gray again. We hypothesized that, during the first half of the experiment, participants may be able to learn the regularities. Note that the same regularities stayed in place during serial search because it is likely that the learned biases from across-trial regularities would fade quickly if the target locations would be randomly assigned (i.e., no regularities). In that case it would be impossible to determine whether the location-association biases learned during parallel search can persist during serial search or whether the learned associations would be rapidly unlearned during when the contingencies are no longer in place. Indeed, previous work involving distractor learning has shown that only a few trials are needed to unlearn particular contingencies (Wang & Theeuwes, 2020). Therefore, during serial search, we kept the same regularities as during parallel search and the critical question was whether the learned target-association biases during feature search could persist when search became serial.

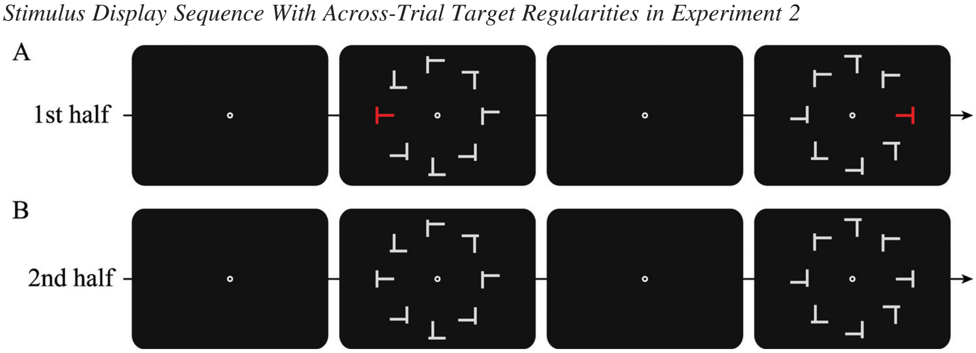
Experiment 2

Experiment 2 was identical to Experiment 1 except that during the first four blocks of the experiment, the target "T" was colored red, creating a pop-out which allowed parallel feature search to find the target. During the final 4 blocks of the experiment, both the target and distractors were colored gray, invoking the need for serial search through the display.

Method

The method was identical to that of Experiment 1, with the following changes: First, we still focused on the RTs difference between predicted and unpredicted conditions, the same power analysis as Experiment 1 showed a minimum of 44 participants was needed. Thus, a new set of fifty-four participants (25 females and 29 males, $M_{age} = 23.02$ years, $SD_{age} = 2.97$) were recruited via Prolific. Three participants were excluded because their overall accuracies in the serial search were lower than 75%. All

Figure 3
Stimulus Display Sequence With Across-Trial Target Regularities in Experiment 2



Note. The "T" target was red in the first half of Experiment 2 (A) but became gray in the second half (B). See the online article for the color version of this figure.

participants reported normal color vision as well as normal or corrected-to-normal visual acuity. Second, as illustrated in Figure 3A, the target was colored in red during the first 4 blocks of 64 trials each (a total of 256 trials), making the target salient, popping-out from the display. During the next 4 blocks of 64 trials each (a total of 256 trials), all elements including the target "T", were presented in gray (see Figure 3B). The statistical regularities remained the same throughout the two halves of the experiment and were identical to those of Experiment 1. Third, practice consisted of a 20-trial practice block of serial search with all stimuli in gray (repeated until they achieved the performance criteria as in Experiment 1; they performed on average 65 (*SD*: 43) practice trials), followed by a single 20-trial practice block of parallel search for a red target among gray distractors.

Results

Analysis

RTs were limited to trials with correct responses (92.22%). The same nonrecursive trimming procedure (Vanselst & Jolicoeur, 1994) as in Experiment 1 was used for outlier removal (1.17%). There were no trials with RTs faster than 200 ms. As above, trials with an immediate repetition of the target location were also excluded from analysis (6.32%).

Learning Effect

As illustrated in Figure 4A, a two-way RM-ANOVA on mean RTs with "half" (1st and 2nd half) and target regularity (predicted, predicting, neutral) as factors showed the significant main effect of half, $F(1, 50) = 2327.25$, $p < .001$, $\eta_p^2 = .98$, suggesting that there was a large difference in search times between parallel search (522 ms) and serial search (1,577 ms). The main effect of target regularity was significant as well, $F(2, 100) = 7.35$, $p = .001$, $\eta_p^2 = .13$. More importantly, the Half × Regularity interaction also reached significance, $F(2, 100) = 5.91$, $p = .004$, $\eta_p^2 = .11$. To further assess this interaction, separate one-way RM-ANOVA were conducted for parallel search (first half) and serial search (second half). There was no reliable effect of regularity (predicted: 519 ms, predicting: 522 ms, neutral: 524 ms) in the first half, $F(2, 100) = 2.10$, $p = .13$, $\eta_p^2 = .04$, $BF_{01} = 2.67$ (but see the analysis on trials whose target location was 4 locations away from the target location on the previous trial below). It can be noted that the lack of a significant regularity effect here is likely attributable to a ceiling effect, because the effect was significant in the first block (i.e., RTs in the predicted [531 ms] condition were faster than predicting [541 ms, $p = .023$, $\eta_p^2 = .10$] and neutral [541 ms, $p = .009$, $\eta_p^2 = .13$] condition), but as RTs further decreased in subsequent blocks, a difference between the conditions was no longer found (see Figure 4B). During the second half a significant main effect of target regularity was found, $F(2, 100) = 6.66$, $p = .002$, $\eta_p^2 = .12$. Post hoc tests showed faster RTs in predicted trials (1,525 ms) relative to predicting (1,588 ms, $p = .023$, $\eta_p^2 = .10$) or neutral (1,620 ms, $p = .002$, $\eta_p^2 = .18$) trials. There was no difference between the latter two ($p = .19$, $\eta_p^2 = .03$).

This significant influence of the learned regularities during the second half raises the question the regularities learned in the first half are simply immediately used in serial search, or they are gradually learned within serial search blocks. To examine the time course of the effect we divided the RTs data during the second half (serial search) into separate blocks (see Figure 4C). A two-way RM-ANOVA with blocks of serial search (1–4) and regularity (predicted, predicting, neutral) as factors did not reveal a significant Block × Regularity interaction, $F(6, 300) = .70$, $p = .65$, $\eta_p^2 = .01$, $BF_{01} = 184.49$ (strong evidence for the absence of an interaction).

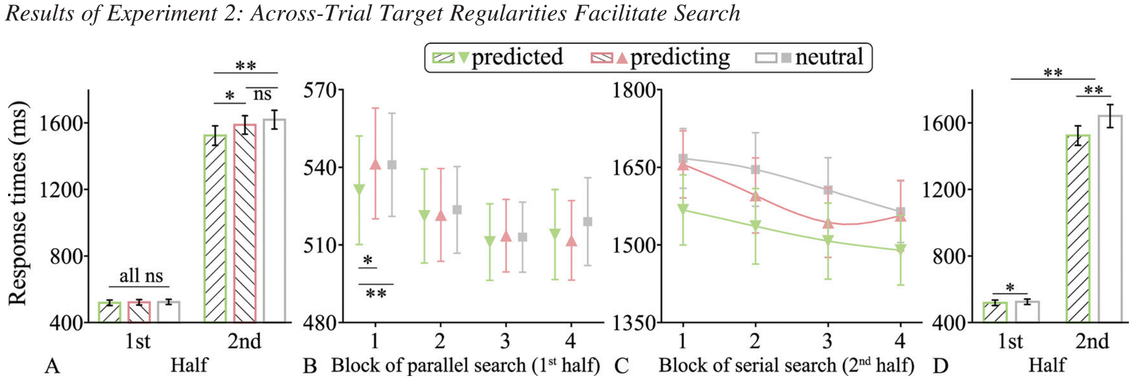
The RM-ANOVA on accuracies revealed the significant main effect of half, $F(1, 50) = 116.74$, $p < .001$, $\eta_p^2 = .70$, with higher accuracy in the first half (95.97%) than the second half (87.85%). The main effect of regularity did not reach the significance, $F(2, 100) = 3.05$, $p = .052$, $\eta_p^2 = .06$, $BF_{01} = 10.70$ (strong evidence for the absence of any difference). Accuracies in the predicted, predicting, and neutral condition were respectively 91.36%, 91.76% and 92.60%. The two-way interaction was insignificant as well, $F(2, 100) = .70$, $p = .50$, $\eta_p^2 = .01$, $BF_{01} = 10.03$.

Across-Trial Target Location Distance Analysis

With a restriction of target-to-target distance of 4 locations, a 2 (half) × 2 (regularity) RM-ANOVA was conducted. The result

Figure 4

Results of Experiment 2: Across-Trial Target Regularities Facilitate Search



Note. Overall RTs as a function of target regularity, in the first and second half of Experiment 2 (A). Mean RTs under separate blocks in three conditions in the first half (B) and second half (C). Mean RTs as a function of regularity when the across-trial target location distance was 4 locations. The error bars denote 95% confidence intervals. The lines in C reflect the fitted data using smoothing spline method in Prism GraphPad software. ns = not significant. * $p < .05$. ** $p < .01$. See the online article for the color version of this figure.

revealed the significant main effect of target regularity, $F(1, 50) = 12.55$, $p = .001$, $\eta_p^2 = .20$, with faster RTs in predicted trials (1,022 ms) relative to neutral trials (1,084 ms). The Half × Regularity interaction also reached significance, $F(1, 50) = 9.80$, $p = .003$, $\eta_p^2 = .16$. As illustrated in Figure 4D, the additional simple-effect analysis showed that the difference was significant not only in the first half, $F(1, 50) = 4.62$, $p = .037$, $\eta_p^2 = .09$, but also highly significant in the second half, $F(1, 50) = 11.20$, $p = .002$, $\eta_p^2 = .18$. These results support the interpretation that participants learned the across-trial regularities regarding target locations during the parallel search and this bias persisted during the serial search.

Awareness of the Regularities

Twelve of 51 participants indicated that they were aware of the across-trial association of the target location during the experiment (CS: 4.08 ± .90). Of those, only one participant correctly chose both predicted locations (CS: 4.5), whereas the other 11 participants indicated incorrect locations (CS: 2.73 ± 1.08). The remaining 39 participants indicated that they were unaware of the trial-to-trial target location association (CS: 2.11 ± .93). Note that when we removed the one "aware" participant, the observed pattern of results remained qualitatively identical.

Discussion

Experiment 2 was designed to disentangle the learning and use of statistical regularities. To that end we first had observers engage in parallel search with the embedded across-trial regularities, which was then followed by serial search with those same regularities. We showed that during the part of the experiment that involved parallel search, during the first block, participants learned and used the regularities as they were faster for predicted trials than for unpredicted (both predicting and neutral) trials. When controlling for the across-trial distance of the target locations (i.e., four locations), there was a reliable difference between predicted and neutral trials (see *across-trial target location distance analysis* and Figure 4D). Critically, after being exposed to feature search during the first half of the experiment, the learned across-trial target-target associations generalized to serial search. It is important to note that the display during the second half of this experiment was exactly the same as the display in Experiment 1; the only difference was that in the current experiment participants had the opportunity to first learn the regularities during feature search. The finding that the facilitation of search for predictable target locations generalized to the subsequent phase of serial search leads to the conclusion that the learned across-trial biases in the pop-out feature search could persist during serial search.

General Discussion

The present study shows that when searching serially through a display, participants were not able to learn relatively rare across-trial target-target statistical regularities (two pairs, four of eight locations). Indeed, when the location of the target on a given trial predicted the location of the target on the next trial, participants were not faster compared with when the target locations were randomly assigned across trials. However, by creating conditions which facilitated learning of the statistical regularities across trials (employing pop-out feature search), the target-to-target associations were learned and they persisted during serial search, facilitating target detection when presented at predicted locations.

Even though it is well known that intertrial repetition priming is a prime example of how selection history affects visual search (e.g., Anderson et al., 2021; Failing & Theeuwes, 2018), it is important to note that the current target-to-target RT benefits are not related to what is known as intertrial priming. Several forms of intertrial priming are recognized; for example, there are intertrial RT benefits when a target appears at the same location on consecutive trials (Maljkovic & Nakayama, 1996), or there are benefits when the target has the same defining features across trials (Maljkovic & Nakayama, 1994). In our study, a particular target location predicted the target location on the following trial, but this was never the same location. Instead, it was always the location on the opposite side of the display. In addition, across trials the features of the target (its orientation, and therefore the response) were randomly assigned which prevented any intertrial feature priming. Clearly, the current findings cannot be explained in terms of some form of intertrial priming. Note that previous work investigating the distributional regularities of the target/distractor location by presenting the target or distractor much more often in one location than in all other locations do suffer from this potential shortcoming of repetition intertrial priming as the target or distractor has to be presented at the same location repeatedly (see also Goschy et al., 2014; Huang et al., 2021; Kabata & Matsumoto, 2012). The current findings can only be explained by assuming that observers have learned the embedded across-trial regularities without the need to assume any low-level repetition priming effects.

We used a T-among-Ls search task, a common operationalization for what has been labeled "spatial-configuration search" which is considered the gold standard for true serial search (Wolfe, 1998). Indeed, it has been claimed that using this type of search, participants attend one item at a time until the target is found (Wolfe, 2003). If we assume that in our experiments search is truly done in a one-by-one manner, this would imply that participants would find the target on average after attending 4 items (as our display size is 8). On those trials in which the location of the target on the current trial predicts the location of target on the upcoming (predicted) trial (T1->T2 regularities), after attending T1 on the current trial they will not directly attend T2 on the next trial but rather attend on average 4 items before encountering T2. Forming T1->T2 associations might hence be prevented as other nontarget items are attended between attending T1 and T2.

The findings previously reported by Ono et al. (2005), who examined the boundary of across-trial temporal CC, are in line with such an interpretation. Ono et al. (2005) had participants search for a rotated "T" among eleven rotated "L"s presented at locations randomly chosen within a 8 × 6 grid matrix. Unbeknownst to the participants, various across-trial regularities of the spatial layout were built in separate experiments. When a specific target location on trial $N - 1$ predicted another specific target location on the following trial N (Experiment 2), no learning effect was observed. Similarly, when distractor locations on the previous trial predicted the specific target location on the current trial (Experiment 3), there was no learning as well. VSL only occurred under circumstances where a repeated target-distractor configuration was predictive of the specific target location on the following trial (Experiment 1). Ono et al. (2005) attributed the lack of learning in Experiments 2 and 3 to the disruptive effects produced by the random variation. A consistent target–target association, for example, was insufficient for VSL to occur because the random variation caused by the attended nontargets in the layout hindered observers from learning. Therefore, they concluded that the visual system is capable of detecting consistent associations (*ubiquitous statistical learning*), but the detection efficiency depends on the co-occurring noise (*noise-sensitive statistical learning*). In our Experiment 1, searching for a "T" among "L"s obviously also generated a lot of noise, preventing the detection of consistent associations.

Notably, the noise in our Experiment 1 was different from that induced by the additional singleton paradigm (Theeuwes, 1991, 1992) in Experiments 3 and 4 of Li and Theeuwes (2020). In their experiments, attention was allocated to the target location immediately or, when attention was captured by a colored singleton distractor, as soon as attention was disengaged from the distractor location. In other words, at maximum one distractor location was attended between two successive target locations, which still made nonadjacent target-target association possible to occur. This is in line with observations in the domain of procedural habit-like learning, where the learning of a motor sequence is found to be less efficient when random elements are inserted within the predictable sequence (e.g., Howard & Howard, 1997; Nemeth et al., 2013). It is also in line with previous work on the learning of regular patterns that occur in nonadjacent elements. It was demonstrated that dependencies between the first and the third elements (A*x*B, where *x* is a random element) could be detected by infants (e.g., Gómez, 2002) and nonhuman primates (e.g., Newport et al., 2004), whereas the ones with several more intervening random elements (A*xxx*B) could not be learned (e.g., Gómez, 2002; Grama et al., 2013). In addition, higher perceptual similarity between the predicting and predicted stimuli (that in turn were perceptually different from the intervening elements) was shown to help the learning of nonadjacent regularities (e.g., Gómez, 2002; Newport & Aslin, 2004; Wilson et al., 2020). This might explain why the colored singleton distractor in the experiments by Li and Theeuwes (2020), which was perceptually different from the consistent gray targets, did not disrupt the learning of across-trial target-target associations. By contrast, in the T-among-Ls serial search task, the other elements of the search display (which can be considered intervening elements due to the serial scanning) were perceptually very similar to the targets that made up the target-target pair regularities, hence this might have kept observers from detecting these regularities.

Typically, in feature search, the target stands out from the nontarget items, which means that the nontarget items are basically not attended. This implies that as soon as the display is presented, spatial attention moves to the location of the feature (the colored T) that stands out. This makes it possible to learn the across-trial target-target location associations. As speculated by Li and Theeuwes (2020), once the associations between the locations are formed the weights within the spatial priority map of selection are adjusted such that following the selection of a target positioned at a predicting location, the weights of the location on the upcoming trial that is predicted by the previous trial are up-regulated. This in turn will result in faster response times when targets are presented at the predicted location. We speculate here that this mechanism of across-trial up-regulating the weights of predicted location remains operative even when the task switches from feature search to serial search.

The results of Experiment 2 suggest that there is transfer of the priority map that incorporated the predictive relation between the target locations, to a different version of the task that involved a very different type of search. Note the term "transfer" here might not be completely correct because the statistical regularities regarding target locations across trials remained the same throughout the whole experiment. However, the exposure to the across-trial regularities during serial search was not sufficient for VSL to occur. A more credible possibility, also supported by our time-resolved analysis, is that observers first learned the regularities during parallel search which in turn modified the spatial priority map. This map remained in place when search became serial. Our results are, in this sense, in agreement with previous findings that participants learn nonadjacent statistical regularities (A*x*B, where *x* is a random element) better if they are first exposed to the adjacent version of the regularity (AB; e.g., Lai & Poletiek, 2011; Lany & Gómez, 2008).

Similarly consistent with our findings, some previous results indeed found that regularities that were implicitly learned remained in place even when the task changed or when the regularities were no longer present (e.g., Duncan & Theeuwes, 2020; Jiang et al., 2013, 2015; Jiang & Won, 2015; Sauter et al., 2019). For example, Jiang et al. (2015) showed that regularities regarding the location of the target learned during a training session generalized to a similar search tasks with different stimuli (i.e., from the T-among-Ls task to the 2-among-5s task) and with different difficulty (i.e., different offsets between two segments of nontarget "L" in different phases). Duncan and Theeuwes (2020) showed that when participants performed two different tasks involving the same display, regularities that were present during one task stayed in place while performing another task. The learned spatial biases could persist more than one day even if when the spatial distribution was even (Jiang et al., 2013; Sauter et al., 2019). There is also evidence that implicitly learned suppression of high-probability distractor locations generalizes across different contexts (Britton & Anderson, 2020; de Waard et al., 2021). Compared with these previous studies, our Experiment 2 went a step further by changing the nature of the search task entirely (even though the search displays were almost identical). Our results suggest that the spatial priority map "shaped" during parallel search can transfer to serial search.

We also measured awareness by asking participants, upon completion of the experiment, whether they were aware of any sequence regarding target locations such as that one target location was always followed by another target location. We also had them answer eight-alternative forced-choice questions to test their explicit knowledge of the regularities. As in most previous studies investigating VSL (e.g., Chun & Jiang, 1998; Ferrante et al., 2018; Geng & Behrmann, 2005; Jiang & Swallow, 2013; Turk-Browne et al., 2009; Wang & Theeuwes, 2018), only a few participants reported to be aware of a sequential relationship between target locations across trials. Most participants were not able to identify what would be the predicted target location on the subsequent trial when the predicting target location was given. As with most VSL experiments, our effects seem to be driven by implicit rather than explicit knowledge of across-trial regularities. An additional indication that the learning we observed was implicit, and the facilitation of search was not the result of top-down expectation comes from the search times that were observed. It is likely that if participants would have explicit and aware knowledge of where the target location would be on the upcoming trial, the RTs on predicted trials should be much lower as there would be no need to search serially through the display. One would expect that the search times for predicted trials should be close to what we have found for feature search. This is clearly not the case. Rather, there seems to be a subtle bias which drives the focus of serial search slightly faster to the predicted location than to the random locations.

In sum, we conclude that during serial search participants are not able to learn across-trial statistical regularities regarding the target locations. However, by introducing feature search we created conditions that facilitated the learning of these across-trial statistical regularities, which remained in place when the search switched to serial. We propose that the target–target associations learned during feature search led to the formation of a flexible priority map: once the predicting location was selected, the weights of the predicted location were up-weighted thereby biasing attentional priority for the upcoming trial. This flexible priority map remained active even when the search task changed from parallel to serial search.

References

Addleman, D. A., Tao, J., Remington, R. W., & Jiang, Y. V. (2018). Explicit goal-driven attention, unlike implicitly learned attention, spreads to secondary tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(3), 356–366. https://doi.org/10.1037/xhp0000457

Allenmark, F., Gokce, A., Geyer, T., Zinchenko, A., Müller, H. J., & Shi, Z. (2021). Inter-trial effects in priming of pop-out: Comparison of computational updating models. *PLoS Computational Biology*, *17*(9), e1009332. https://doi.org/10.1371/journal.pcbi.1009332

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

Anderson, B. A. (2019). Neurobiology of value-driven attention. *Current Opinion in Psychology*, *29*, 27–33. https://doi.org/10.1016/j.copsyc.2018.11.004

Anderson, B. A., Kim, H., Kim, A. J., Liao, M. R., Mrkonja, L., Clement, A., & Grégoire, L. (2021). The past, present, and future of selection history. *Neuroscience and Biobehavioral Reviews*, *130*, 326–350. https://doi.org/10.1016/j.neubiorev.2021.09.004

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(25), 10367–10371. https://doi.org/10.1073/pnas.1104047108

Aron, A. R., Gluck, M. A., & Poldrack, R. A. (2006). Long-term test-retest reliability of functional MRI in a classification learning task. *NeuroImage*, *29*(3), 1000–1006. https://doi.org/10.1016/j.neuroimage.2005.08.010

Aron, A. R., Shohamy, D., Clark, J., Myers, C., Gluck, M. A., & Poldrack, R. A. (2004). Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *Journal of Neurophysiology*, *92*(2), 1144–1152. https://doi.org/10.1152/jn.01209.2003

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*(4), 321–324. https://doi.org/10.1111/1467-9280.00063

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443. https://doi.org/10.1016/j.tics.2012.06.010

Bergen, J. R., & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature*, *303*(5919), 696–698. https://doi.org/10.1038/303696a0

Bergmann, N., & Schubö, A. (2021). Local and global context repetitions in contextual cueing. *Journal of Vision*, *21*(10), 9. https://doi.org/10.1167/jov.21.10.9

Bogaerts, L., Frost, R., & Christiansen, M. H. (2020). Integrating statistical learning into cognitive science. *Journal of Memory and Language*, *115*, 104167. https://doi.org/10.1016/j.jml.2020.104167

Bogaerts, L., Richter, C. G., Landau, A. N., & Frost, R. (2020). Beta-band activity is a signature of statistical learning. *The Journal of Neuroscience*, *40*(39), 7523–7530. https://doi.org/10.1523/JNEUROSCI.0771-20.2020

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, *138*(4), 487–502. https://doi.org/10.1037/a0016797

Britton, M. K., & Anderson, B. A. (2020). Specificity and persistence of statistical learning in distractor suppression. *Journal of Experimental Psychology: Human Perception and Performance*, *46*(3), 324–334. https://doi.org/10.1037/xhp0000718

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, *2*(1), 16. https://doi.org/10.5334/joc.72

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*(1), 28–71. https://doi.org/10.1006/cogp.1998.0681

Chun, M. M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(2), 224–234. https://doi.org/10.1037/0278-7393.29.2.224

de Waard, J., Bogaerts, L., van Moorselaar, D., & Theeuwes, J. (2021). Surprisingly inflexible: Statistically learned suppression of distractors generalizes across contexts. *Attention, Perception & Psychophysics*. Advance online publication. https://doi.org/10.3758/s13414-021-02387-x

Dennis, N. A., Howard, J. H., Jr., & Howard, D. V. (2006). Implicit sequence learning without motor sequencing in young and old adults. *Experimental Brain Research*, *175*(1), 153–164. https://doi.org/10.1007/s00221-006-0534-3

Duncan, D., & Theeuwes, J. (2020). Statistical learning in the absence of explicit top-down attention. *Cortex*, *131*, 54–65. https://doi.org/10.1016/j.cortex.2020.07.006

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology*, *48*, 269–297. https://doi.org/10.1146/annurev.psych.48.1.269

Egeth, H., & Dagenbach, D. (1991). Parallel versus serial processing in visual search: Further evidence from subadditive effects of visual quality. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 551–560. https://doi.org/10.1037/0096-1523.17.2.551

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

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

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

Ferrante, O., Patacca, A., Di Caro, V., Della Libera, C., Santandrea, E., & Chelazzi, L. (2018). Altering spatial priority maps via statistical learning of target selection and distractor filtering. *Cortex*, *102*, 67–95. https://doi.org/10.1016/j.cortex.2017.09.027

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, *12*(6), 499–504. https://doi.org/10.1111/1467-9280.00392

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 458–467. https://doi.org/10.1037/0278-7393.28.3.458

Fiser, J., & Aslin, R. N. (2005). Encoding multielement scenes: Statistical learning of visual feature hierarchies. *Journal of Experimental Psychology: General*, *134*(4), 521–537. https://doi.org/10.1037/0096-3445.134.4.521

Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, *145*(12), 1128–1153. https://doi.org/10.1037/bul0000210

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

Geng, J. J., & Behrmann, M. (2002). Probability cuing of target location facilitates visual search implicitly in normal participants and patients with hemispatial neglect. *Psychological Science*, *13*(6), 520–525. https://doi.org/10.1111/1467-9280.00491

Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics*, *67*(7), 1252–1268. https://doi.org/10.3758/BF03193557

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431–436. https://doi.org/10.1111/1467-9280.00476

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. https://doi.org/10.3389/fpsyg.2014.01195

Goujon, A., Didierjean, A., & Thorpe, S. (2015). Investigating implicit statistical learning mechanisms through contextual cueing. *Trends in Cognitive Sciences*, *19*(9), 524–533. https://doi.org/10.1016/j.tics.2015.07.009

Grama, I., Wijnen, F., & Kerkhoff, A. (2013). *Constraints on non-adjacent dependency-learning: Distance matters an artificial grammar learning study with adults*. https://www.bu.edu/bucld/files/2013/06/Grama.pdf

Henin, S., Turk-Browne, N. B., Friedman, D., Liu, A., Dugan, P., Flinker, A., Doyle, W., Devinsky, O., & Melloni, L. (2021). Learning hierarchical sequence representations across human cortex and hippocampus. *Science Advances*. Advance online publication. https://doi.org/10.1126/sciadv.abc4530

Howard, J. H., Howard, D. V., Dennis, N. A., & Kelly, A. J. (2008). Implicit learning of predictive relationships in three-element visual sequences by young and old adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(5), 1139–1157. https://doi.org/10.1037/a0012797

Howard, J. H., Jr., & Howard, D. V. (1997). Age differences in implicit learning of higher order dependencies in serial patterns. *Psychology and Aging*, *12*(4), 634–656. https://doi.org/10.1037/0882-7974.12.4.634

Huang, C., Donk, M., & Theeuwes, J. (2021). *Proactive enhancement and suppression elicited by statistical regularities in visual search*. Manuscript submitted for publication.

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

Jiang, Y. V. (2018). Habitual versus goal-driven attention. *Cortex*, *102*, 107–120. https://doi.org/10.1016/j.cortex.2017.06.018

Jiang, Y. V., & Sisk, C. A. (2019). Habit-like attention. *Current Opinion in Psychology*, *29*, 65–70. https://doi.org/10.1016/j.copsyc.2018.11.014

Jiang, Y. V., & Swallow, K. M. (2013). Spatial reference frame of incidentally learned attention. *Cognition*, *126*(3), 378–390. https://doi.org/10.1016/j.cognition.2012.10.011

Jiang, Y. V., & Won, B. Y. (2015). Spatial scale, rather than nature of task or locomotion, modulates the spatial reference frame of attention. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(3), 866–878. https://doi.org/10.1037/xhp0000056

Jiang, Y. V., Sisk, C. A., & Toh, Y. N. (2019). Implicit guidance of attention in contextual cueing: Neuropsychological and developmental evidence. *Neuroscience and Biobehavioral Reviews*, *105*, 115–125. https://doi.org/10.1016/j.neubiorev.2019.07.002

Jiang, Y. V., Swallow, K. M., Rosenbaum, G. M., & Herzig, C. (2013). Rapid acquisition but slow extinction of an attentional bias in space. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 87–99. https://doi.org/10.1037/a0027611

Jiang, Y. V., Swallow, K. M., Won, B. Y., Cistera, J. D., & Rosenbaum, G. M. (2015). Task specificity of attention training: The case of probability cuing. *Attention, Perception & Psychophysics*, *77*(1), 50–66. https://doi.org/10.3758/s13414-014-0747-7

Jiang, Y., & Wagner, L. C. (2004). What is learned in spatial contextual cuing—Configuration or individual locations? *Perception & Psychophysics*, *66*(3), 454–463. https://doi.org/10.3758/BF03194893

Kabata, T., & Matsumoto, E. (2012). Cueing effects of target location probability and repetition. *Vision Research*, *73*, 23–29. https://doi.org/10.1016/j.visres.2012.09.014

Kwak, H. W., Dagenbach, D., & Egeth, H. (1991). Further evidence for a time-independent shift of the focus of attention. *Perception & Psychophysics*, *49*(5), 473–480. https://doi.org/10.3758/BF03212181

Lai, J., & Poletiek, F. H. (2011). The impact of adjacent-dependencies and staged-input on the learnability of center-embedded hierarchical structures. *Cognition*, *118*(2), 265–273. https://doi.org/10.1016/j.cognition.2010.11.011

Lamy, D., Antebi, C., Aviani, N., & Carmel, T. (2008). Priming of Pop-out provides reliable measures of target activation and distractor inhibition in selective attention. *Vision Research*, *48*(1), 30–41. https://doi.org/10.1016/j.visres.2007.10.009

Lange, K., Kühn, S., & Filevich, E. (2015). "Just Another Tool for Online Studies" (JATOS): An easy solution for setup and management of web servers supporting online studies. *PLoS ONE*, *10*(6). e0130834. https://doi.org/10.1371/journal.pone.0130834

Lany, J., & Gómez, R. L. (2008). Twelve-month-old infants benefit from prior experience in statistical learning. *Psychological Science*, *19*(12), 1247–1252. https://doi.org/10.1111/j.1467-9280.2008.02233.x

Li, A. S., & Theeuwes, J. (2020). Statistical regularities across trials bias attentional selection. *Journal of Experimental Psychology: Human Perception and Performance*, *46*(8), 860–870. https://doi.org/10.1037/xhp0000753

Liesefeld, H. R., & Müller, H. J. (2021). Modulations of saliency signals at two hierarchical levels of priority computation revealed by spatial statistical distractor learning. *Journal of Experimental Psychology: General*, *150*(4), 710–728. https://doi.org/10.1037/xge0000970

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

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

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

Nemeth, D., Janacsek, K., & Fiser, J. (2013). Age-dependent and coordinated shift in performance between implicit and explicit skill learning.

Frontiers in Computational Neuroscience, *7*, 147. https://doi.org/10.3389/fncom.2013.00147

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127–162. https://doi.org/10.1016/S0010-0285(03)00128-2

Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004). Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology*, *49*(2), 85–117. https://doi.org/10.1016/j.cogpsych.2003.12.002

Olson, I. R., & Chun, M. M. (2001). Temporal contextual cuing of visual attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(5), 1299–1313. https://doi.org/10.1037/0278-7393.27.5.1299

Ono, F., Jiang, Y., & Kawahara, J. I. (2005). Intertrial temporal contextual cuing: Association across successive visual search trials guides spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(4), 703–712. https://doi.org/10.1037/0096-1523.31.4.703

Pacton, S., & Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 80–96. https://doi.org/10.1037/0278-7393.34.1.80

Pacton, S., Sobaco, A., & Perruchet, P. (2015). Is an attention-based associative account of adjacent and nonadjacent dependency learning valid? *Acta Psychologica*, *157*, 195–199. https://doi.org/10.1016/j.actpsy.2015.03.002

Palan, S., & Schitter, C. (2018). Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238. https://doi.org/10.1016/j.tics.2006.03.006

Qin, N., Gu, R., Xue, J., Chen, C., & Zhang, M. (2021). Reward-driven attention alters perceived salience. *Journal of Vision*, *21*(1), 7. https://doi.org/10.1167/jov.21.1.7

Remillard, G. (2009). Pure perceptual-based sequence learning: A role for visuospatial attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 528–541. https://doi.org/10.1037/a0014646

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*(4), 606–621. https://doi.org/10.1006/jmla.1996.0032

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*, *45*(11), 2080–2097. https://doi.org/10.1037/xlm0000691

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

Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. *Nature Neuroscience*, *16*(4), 486–492. https://doi.org/10.1038/nn.3331

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics*, *50*(2), 184–193. https://doi.org/10.3758/BF03212219

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*(6), 599–606. https://doi.org/10.3758/BF03211656

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. https://doi.org/10.1016/j.actpsy.2010.02.006

Theeuwes, J. (2018). Visual selection: Usually fast and automatic; seldom slow and volitional. *Journal of Cognition*, *1*(1), 29. https://doi.org/10.5334/joc.13

Theeuwes, J. (2019). Goal-driven, stimulus-driven, and history-driven selection. *Current Opinion in Psychology*, *29*, 97–101. https://doi.org/10.1016/j.copsyc.2018.12.024

Theeuwes, J., & Van der Burg, E. (2013). Priming makes a stimulus more salient. *Journal of Vision*, *13*(3), 21. https://doi.org/10.1167/13.3.21

Thomas, C., Didierjean, A., Maquestiaux, F., & Goujon, A. (2018). On the limits of statistical learning: Intertrial contextual cueing is confined to temporally close contingencies. *Attention, Perception & Psychophysics*, *80*(6), 1420–1435. https://doi.org/10.3758/s13414-018-1519-6

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Turk-Browne, N. B., & Scholl, B. J. (2009). Flexible visual statistical learning: Transfer across space and time. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 195–202. https://doi.org/10.1037/0096-1523.35.1.195

Turk-Browne, N. B., Jungé, J., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552–564. https://doi.org/10.1037/0096-3445.134.4.552

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*, *21*(10), 1934–1945. https://doi.org/10.1162/jocn.2009.21131

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *The Journal of Neuroscience*, *30*(33), 11177–11187. https://doi.org/10.1523/JNEUROSCI.0858-10.2010

Umemoto, A., Scolari, M., Vogel, E. K., & Awh, E. (2010). Statistical learning induces discrete shifts in the allocation of working memory resources. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1419–1429. https://doi.org/10.1037/a0019324

Vanselst, M., & Jolicoeur, P. (1994). A solution to the effect of sample-size on outlier elimination. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *47A*(3), 631–650. https://doi.org/10.1080/14640749408401131

Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Smira, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3

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

Wang, B., & Theeuwes, J. (2020). Implicit attentional biases in a changing environment. *Acta Psychologica*, *206*, 103064. https://doi.org/10.1016/j.actpsy.2020.103064

Wang, L., Wang, B., & Theeuwes, J. (2021). Across-trial spatial suppression in visual search. *Attention, Perception & Psychophysics*, *83*(7), 2744–2752. https://doi.org/10.3758/s13414-021-02341-x

Wilson, B., Spierings, M., Ravignani, A., Mueller, J. L., Mintz, T. H., Wijnen, F., van der Kant, A., Smith, K., & Rey, A. (2020). Non-adjacent dependency learning in humans and other animals. *Topics in Cognitive Science*, *12*(3), 843–858. https://doi.org/10.1111/tops.12381

Wolfe, J. M. (1994). Guided Search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238. https://doi.org/10.3758/BF03200774

Wolfe, J. M. (1998). What can 1 Million trials tell us about visual search? *Psychological Science*, *9*(1), 33–39. https://doi.org/10.1111/1467-9280.00006

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences*, *7*(2), 70–76. https://doi.org/10.1016/S1364-6613(02)00024-4

Won, B. Y., & Jiang, Y. V. (2015). Spatial working memory interferes with explicit, but not probabilistic cuing of spatial attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 787–806. https://doi.org/10.1037/xlm0000040

Won, B. Y., Kosoyan, M., & Geng, J. J. (2019). Evidence for second-order singleton suppression based on probabilistic expectations. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(1), 125–138. https://doi.org/10.1037/xhp0000594

Yu, R. Q., & Zhao, J. (2018). Implicit updating of object representation via temporal associations. *Cognition*, *181*, 127–134. https://doi.org/10.1016/j.cognition.2018.08.015

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

Zhang, B., Allenmark, F., Liesefeld, H. R., Shi, Z., & Müller, H. J. (2019). Probability cueing of singleton-distractor locations in visual search: Priority-map- versus dimension-based inhibition? *Journal of Experimental Psychology: Human Perception and Performance*, *45*(9), 1146–1163. https://doi.org/10.1037/xhp0000652

Zhao, J., & Yu, R. Q. (2016). Statistical regularities reduce perceived numerosity. *Cognition*, *146*, 217–222. https://doi.org/10.1016/j.cognition.2015.09.018

Zhao, J., Al-Aidroos, N., & Turk-Browne, N. B. (2013). Attention is spontaneously biased toward regularities. *Psychological Science*, *24*(5), 667–677. https://doi.org/10.1177/0956797612460407

Zhao, J., Ngo, N., McKendrick, R., & Turk-Browne, N. B. (2011). Mutual interference between statistical summary perception and statistical learning. *Psychological Science*, *22*(9), 1212–1219. https://doi.org/10.1177/0956797611419304

Received August 16, 2021
Revision received October 26, 2021
Accepted November 28, 2021 ■