Vision Research 188 (2021) 211–226

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Testing temporal integration of feature probability distributions using role-reversal effects in visual search☆

Ömer Dağlar Tanrıkulu [a,*], Andrey Chetverikov [b], Árni Kristjánsson [a,c]

[a] Faculty of Psychology, School of Health Sciences, University of Iceland, Reykjavik, Iceland
[b] Visual Computation Lab, Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behavior, Nijmegen, The Netherlands
[c] National Research University, Higher School of Economics, Moscow, Russian Federation

ARTICLE INFO

Keywords:
Visual ensembles
Visual search
Priming of pop-out
Summary statistics
Feature distributions
Optimal integration

ABSTRACT

The visual system is sensitive to statistical properties of complex scenes and can encode feature probability distributions in detail. But does the brain use these statistics to build probabilistic models of the ever-changing visual input? To investigate this, we examined how observers temporally integrate two different orientation distributions from sequentially presented visual search trials. If the encoded probabilistic information is used in a Bayesian optimal way, observers should weigh more reliable information more strongly, such as feature distributions with low variance. We therefore manipulated the variance of the two feature distributions. Participants performed sequential odd-one-out visual search for an oddly oriented line among distractors. During successive learning trials, the distractor orientations were sampled from two different Gaussian distributions on alternating trials. Then, observers performed a 'test trial' where the orientations of the target and distractors were switched, allowing us to assess observer's internal representation of distractor distributions based on changes in response times. In three experiments we observed that observer's search times on test trials depended mainly on the very last learning trial, indicating a strong recency effect. Since temporal integration has been previously observed with this method, we conclude that when the input is unreliable, the visual system relies more on the most recent stimulus. This indicates that the visual system prefers to utilize sensory history when the statistical properties of the environment are relatively stable.

1. Introduction

The visual system can be flexible and adaptable in the face of changes to the temporal context of visual input. Temporal effects on visual processing can be observed on various timescales, from developmental changes to recent sensory exposure. Studying such effects of sensory history provides important insights for understanding the nature of visual representations, the mechanisms that determine how they are formed, and the visual phenomenology resulting from them.

Many effects of recent sensory history on visual processing have been well studied, such as adaptation (Kohn, 2007; Webster, 2015), perceptual learning (Sasaki, Nanez, & Watanabe, 2010; Dosher & Lu, 2017), priming (Kristjánsson & Campana, 2010; Maljkovic & Nakayama, 1994), and serial dependence (Fischer & Whitney, 2014). While these temporal effects may not have completely distinct underlying mechanisms (e.g., Larsson & Smith, 2012; Walther, Schweinberger, Kaiser, & Kovács, 2013), they can be observed independently of one another.

Among these temporal effects, intertrial priming (Maljkovic & Nakayama, 1994), where observers' search times decrease as the same target feature is repeated in a series of odd-one-out search trials, provides an ideal paradigm to study how the visual system encodes complex probabilistic visual information. This is because such priming effects cannot be explained by either the low-level neural response dynamics of feature receptors, nor high-level influences from post-perceptual processes. For example, such priming effects are not retinotopic (Tower-Richardi, Leber, & Golomb, 2016), indicating that they cannot be explained by low-level adaptation mechanisms. They have also been observed independently of observers' expectations or perceptual learning (Maljkovic & Nakayama, 1996; Wang, Kristjansson, &

☆ The data from the experiments reported in this paper and the scripts for statistical analyses are available at https://osf.io/46hj9/?view_only=0cb56a37bdac4de389f460bcef34f08c.

* Corresponding author.
E-mail address: daglar@hi.is (Ö.D. Tanrıkulu).

https://doi.org/10.1016/j.visres.2021.07.012
Received 28 April 2020; Received in revised form 16 April 2021; Accepted 23 July 2021
Available online 7 August 2021
0042-6989/

Nakayama, 2005; Becker, 2008; Sigurdardottir, Kristjánsson, & Driver, 2008; Shurygina, Kristjánsson, Tudge, & Chetverikov, 2019; for review see Kristjánsson & Ásgeirsson, 2019) or priming of responses (Goolsby & Suzuki, 2001; Sigurdardottir et al., 2008). Kristjánsson and Driver (2008) demonstrated that similar priming effects in visual search also occur from repetition of distractor features, which means that the visual information is encoded with respect to the role it plays (e.g., target vs. distractors) in completing a perceptual task. Priming effects reflect accumulation of information about target and distractor features over successive search trials and therefore provide an effective experimental tool for investigating how the visual system accumulates and represents information in a dynamic environment (Chetverikov, Campana, & Kristjánsson, 2017d).

While target and distractor priming in visual search have been independently observed, they can also interact. For example, if an odd-one-out search target has similar features as the distractors on a preceding search trial, search times will increase (Kristjánsson & Driver, 2008; Lamy, Antebi, Aviani, & Carmel, 2008). Importantly, this decrease in search efficiency following a "role-reversal" of target and distractor features is modulated by the similarity between the target on the current search trial and distractors on preceding trials. Using a method later coined *feature distribution learning* (FDL), Chetverikov, Campana & Kristjánsson (2016, 2017b, 2017c) demonstrated that search times following role-reversals as a function of the orientation difference between the target after role-reversal and the mean orientation of preceding distractor lines could be used to probe observers' representations of distractor orientation distributions (Fig. 1, see Chetverikov, Hansmann-Roth, Tanrıkulu, and Kristjánsson (2019), for review). Their observers searched for an oddly oriented target in a series of learning trials where distractor priming was induced by sampling the distractor orientations from the same distribution. Role reversal effects were induced on test trials by swapping the target and distractor features. The strength of role reversals was manipulated by varying the orientation difference between the target on the test trial and the mean of the distractor distribution on the preceding learning trials.

A key finding in Chetverikov, Campana, and Kristjánsson (2016) was that search times on test trials as a function of the similarity between targets on test trials and distractors on preceding learning trials followed the shape of the probability distribution of distractor orientations on the learning trials. This occurred even for distractor distributions on learning trials with the same mean and range (or variance) but different shapes (Gaussian vs. uniform) or different skewness. Observers could encode surprisingly detailed information about visual feature distributions going beyond the summary statistics mean and variance. They can also encode complex orientation distributions, such as bimodal ones (Chetverikov et al., 2017b; Chetverikov et al., 2020), hue distributions of isoluminant colored items (Chetverikov, Campana, & Kristjánsson, 2017a; Hansmann-Roth, Chetverikov, & Kristjánsson, 2019), and when the search array appeared in the peripheral visual field (Tanrıkulu, Chetverikov, & Kristjánsson, 2020). Notably the learning required a certain minimum number of items on each search trial (Chetverikov, Campana, & Kristjánsson, 2017c).

Importantly, these results contrast with research on ensemble perception that suggests that the visual system extracts summary statistical information from ensembles of similar visual items (for reviews, see Alvarez, 2011; Haberman & Whitney, 2012; Whitney & Yamanashi Leib, 2018). According to this literature, observers cannot estimate higher-order properties of feature distributions such as skewness and kurtosis (Atchley & Andersen, 1995; Dakin & Watt, 1997; Dakin, 2015). Even though the FDL method does not show that observers can deliberately estimate such higher-order properties, it reveals that such properties are encoded by the visual system. This discrepancy is most likely due to the fact that the FDL method involves priming effects rather than observers' explicit judgments of ensemble properties. This implicit nature of FDL reveals how the visual system encodes visual feature distributions without requiring explicit access to this visual information (Hansmann-Roth, Kristjánsson, Whitney, & Chetverikov, 2021).

1.1. Current aims

The literature on feature ensemble encoding has mostly focused on spatial integration of features. However, in the real world, the visual input is continuous and dynamic. While the temporal integration of visual features *into* ensembles has been studied (Albrecht & Scholl, 2010; Chong & Treisman, 2003; Haberman, Harp, & Whitney, 2009; Hubert-Wallander & Boynton, 2015; Whiting & Oriet, 2011), little work is available on temporal integration of different ensembles. Oriet and Hozempa (2016) demonstrated that observers can accurately judge the statistical properties of sets of visual items presented temporally over an extended duration. Crawford, Corbin, and Landy (2019) showed that observers' judgments of ensemble properties of a display can be biased by previously seen displays. Here, we assessed such potential combination of spatial and temporal integration of visual feature distributions, without requiring explicit perceptual judgements.

While the visual system is clearly sensitive to details of feature distributions (Chetverikov et al., 2016; Chetverikov et al., 2017a; Chetverikov et al., 2017b; Chetverikov et al., 2017c), it is unclear how the visual system uses this information. Despite this observed sensitivity to detailed statistical information, observers might not be able to use it to build probabilistic representations that integrate information from the constant flow of visual input.[1] Even though previous work with the FDL method (Chetverikov, Campana, & Kristjánsson, 2017b; Chetverikov et al., 2020) suggests that observers can build probabilistic representation of feature distributions, it is still not known to what extent and in what ways this detailed probabilistic information is utilized by the visual system. We investigated this using FDL, but critically, we altered the dynamics of the learning trials. Our crucial manipulation was that distractor orientations on learning trials were sampled from two different Gaussian distributions with different means and standard deviations (*SD*) in *alternating* order. Search times on test trials (following a streak of such alternating distractor distributions) as a function of similarity between target and previous distractors will therefore reflect how observers integrate the two distractor distributions.

In an odd-one-out search task, observers are forced to process all search items to find out the most dissimilar one with respect to all the others. As observers perform a sequence of search trials during the learning streak, they gradually form a probabilistic template for distractors which helps in finding the target quicker. In other words, observers are building an estimate of the underlying distractor distribution throughout the learning streak, even though they are not explicitly instructed to do so (Chetverikov et al., 2019; Chetverikov et al., 2020). We therefore hypothesize that the *SD* of the distractor distributions would influence how the two distributions are integrated. The search trials with more variable distractors (higher *SD*s) would provide less reliable information about the mean of the underlying orientation distribution. In other words, if the visual system obtains samples from the search displays to estimate the mean of the distractors, this estimate will be less precise when the distractors are more variable. Therefore, if the visual system uses detailed probabilistic information about distractor distributions, this integration should accord with classic Bayesian integration principles (e.g., Knill & Richards, 1996; Körding & Wolpert, 2006; Knill, 2007), where observers weigh more reliable information more strongly, in our case the distribution with lower variance (Fig. 2A).

If the estimates of the distractor distributions are optimally updated between learning trials, more precise samples (i.e., trials with lower distractor variance) would receive higher weights. In such cases, we expect representations of the integrated feature distributions to fall

[1] For a more detailed discussion of the difference between the terms "sensitivity to" and "representation of" probabilistic information, see Tanrıkulu, Hansmann-Roth, Chetverikov, and Kristjánsson (2020).

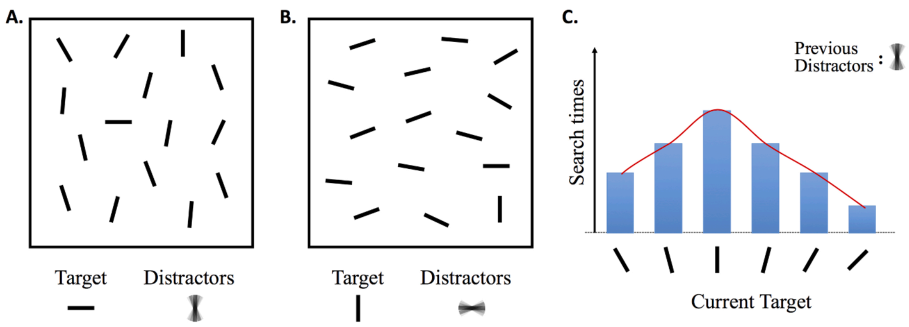

Fig. 1. The feature distribution learning (FDL) methodology used by Chetverikov et al. (2016). A. An example of a simple odd-one-out visual search display in orientation space. Whether a line is the target or not can only be determined if all the other lines can be grouped into a category of distractors, to which the target is unlikely to belong. Therefore, odd-one-out search requires perceptual grouping and encoding of the distractor orientations. B. The orientation values of the target and distractors in display A are swapped in display B to induce role reversal effects C. Hypothetical search times obtained from display B when it follows the one in A (i.e. when role-reversals occur). Search times following role-reversals depend on the similarity between the current target and the previous distractors. Manipulating this similarity and observing its effects on search times can reveal observers' representations of previous distractors (the red curve).

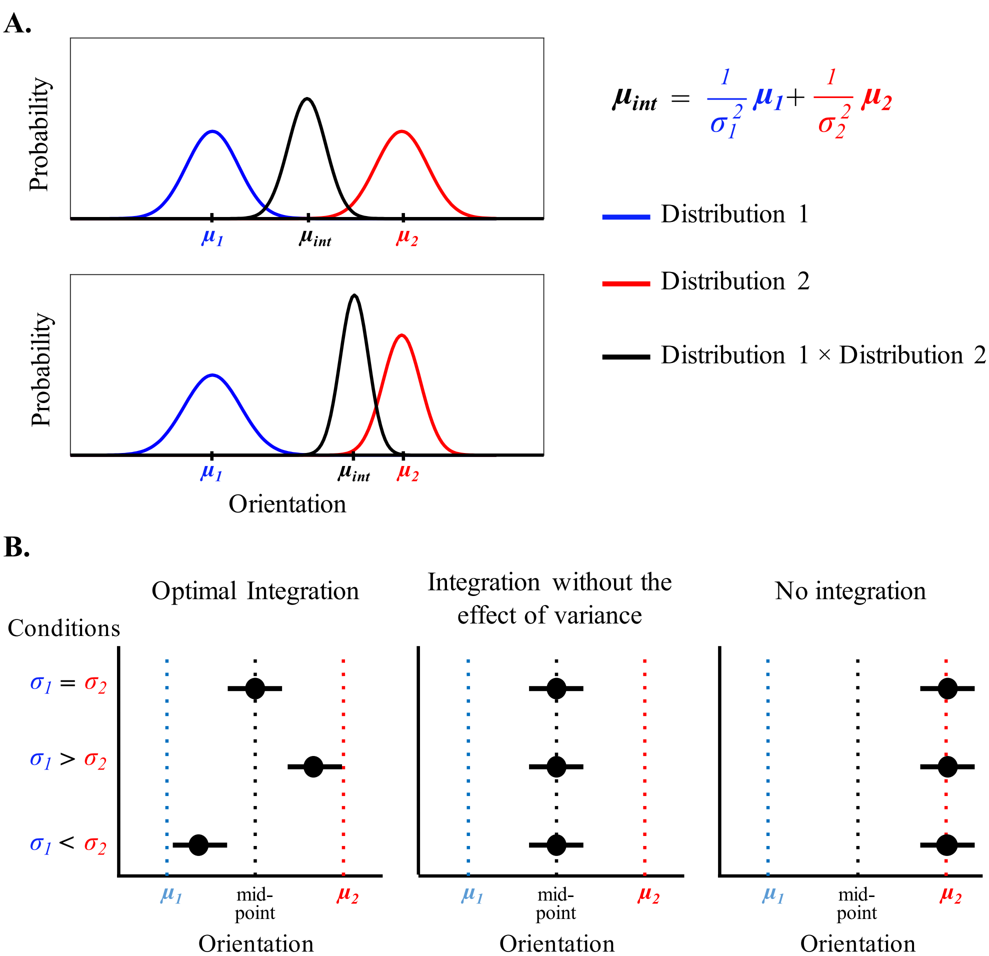

Fig. 2. A. In Bayesian optimal integration of probability distributions, the weights assigned to the mean of the integrated distributions are inversely proportional to their variance, as the equation at top right shows where μ and σ correspond to the mean and variance of the distributions, respectively. Top plot: Integration of two distributions with equal variances (red and blue curve) would yield a distribution (black curve) that is centered on the mid-point of the two distributions. Bottom plot: If one of the distributions (red curve) has a lower variance than the other, then their integration should yield a distribution (black curve) biased towards the distribution with the lower variance. We manipulated the variance of the two orientation distributions presented during learning trials and examined how it influences search times on test trials, which reflect how observers integrated the two orientation distributions. B. Three potential outcomes of this study. The x-axis is the orientation space where the means of the two distributions are marked by the dashed red and blue lines. The black dashed line indicates the mid-point of the two distribution means. μ_2 and μ_1 show the mean of the distractor distribution used on the last learning trial and on the trial before the last one, respectively. The y-axis shows the three different conditions in terms of the relation between the variances of the two distribution. The black data points correspond to the expected mean (μ_{int}) of the resultant distribution obtained by the integration of the two distributions. The plot on the left shows the expected results if the integration is optimal, where μ_{int} would be biased toward the distribution with the lower variance. The center plot shows the expected results if the integration is done without taking the variances of the two distributions into account, where μ_{int} would be on the mid-point of the two distribution means regardless of their variances. If no integration takes place, the expected result is shown by the plot on the right, given that the last orientation distribution presented to the observers was the second distribution (μ_2). Without any temporal integration, the visual system should only rely on the most recently observed orientation distribution.

between the means of the two distractor distributions but importantly be biased towards the one with lower variance. Conversely, if uncertainty is not taken into account, the integrated distribution should fall around the mid-point of the means of the two distractor distributions. Finally, if no temporal integration occurs, then search times on test trials should reflect the distractor distribution on the last learning trial. Predicted results from these three alternative outcomes are depicted in Fig. 2B. Of course, such alternatives are not fully mutually exclusive. For example, if the uncertainty in the representation increases with time due to additional noise during the maintenance, the last trial can have more weight in the integration process even if the stimulus distribution on that trial was noisier than on the trial before. Nevertheless, the three alternatives outlined here provide an overview of the potential outcomes of the study.

A crucial factor that determines how the visual system integrates information from two unreliable sources is the degree of conflict between them. For example, studies demonstrating optimal Bayesian integration of perceptual cues in human observers have carefully limited the discrepancy between the two conflicting cues (e.g., Ernst & Banks, 2002; Hillis, Watt, Landy, & Banks, 2004), because if a cue conflict becomes large enough the visual system can completely discount one of the cues (Banks & Backus, 1998; Blake, Bülthoff, & Sheinberg, 1993). A similar pattern occurs for ensemble representations. Utochkin and Yurevich (2016) examined the effect of segmentability of distractor orientation distributions in odd-one-out visual search while keeping the heterogeneity of the distractors constant. They found that when the distractor orientation distribution was segmentable into separate groups (i.e., when the orientation distance between the distractors was large) search was inefficient, indicating that observers did not treat the distractors as a single orientation ensemble, but instead encoded them as separate distractor groups. Search performance improved, however, when the orientation difference between the distractors was low so that the distractor distribution became unsegmentable, suggesting that observers represented the distractors as a single orientation distribution. In other words, independent of heterogeneity, the orientation distance between groups of lines determines whether the visual system integrates them into a single feature ensemble, or treats them as separate feature ensembles. Given that distance between the distributions and their and standard deviations can have a significant influence on the integration, we ran three experiments in which we varied the overlap between the two distractor distributions across different experiments. This ensured that the results we obtained were not due to the distance between the two distractor distributions chosen for that particular experiment.

2. General method

2.1. Overview

We modified the FDL methodology described in Chetverikov et al. (2019) by using two different distractor distributions during a single learning streak, where distractor orientations on each trial of the learning streak were sampled from two different probability distributions in alternating order. All three experiments had essentially the same design and procedure, except that the *SD* and the orientation distance between the means of the two distractor distributions differed across experiments. In Experiment 1 and 3, the *SD*'s of the two distractor distributions were 5° and 15°, whereas in Experiment 2 they were 8° and 15°. The orientation differences between the two distractor distributions were 30°, 20° and 12° in Experiments 1, 2 and 3, respectively. These differences across experiments allowed us to test our method with different degrees of overlap between the two distractor distributions.

2.2. Participants

All experiments were performed in accordance with the requirements of the Declaration of Helsinki and of the local ethics committee. Sample size and number of trials for each experiment were chosen considering the results of previous FDL studies (Chetverikov et al., 2016; Chetverikov et al., 2017a; Chetverikov et al., 2017b; Chetverikov et al., 2017c; Tanrıkulu et al., 2020) as an informed minimum. Instead of using traditional null-hypothesis frequentist testing, we used Bayes Factor analysis to estimate the amount of evidence in favour of our hypotheses regarding the integration of feature distributions.

2.3. Stimuli and design

The display contained a search array of 36 white lines arranged in a 6 by 6 invisible grid on a grey background (Fig. 3). The grid was centered on the screen and extended to 13°×13°. The length of each line was 1° and random jitter (±0.5°) was added to the horizontal and vertical coordinates of each line.

We used blocks of search trials where each block had a streak of learning trials (varied from two to eight trials) and a test trial. During a learning streak, the orientations of the distractor lines were sampled from two separate Gaussian distributions in alternating order (Fig. 4). The *SD*s of the two distractor orientation distributions were manipulated within an experiment, and each distribution was truncated 2 × *SD* away from its mean to prevent influences from accidental outliers. To keep the range of different distribution types constant, we set the orientation of two distractor lines to the minimal and maximal values of the range of the distributions. The mid-point of the mean of the two distractor distributions was chosen randomly from 0° to 180° across learning streaks, but kept fixed during a single streak and their means were equidistant from this mid-point (Fig. 4A). The distance between the mean of the distributions and the mid-point differed in each experiment. Target orientation on learning trials was determined randomly with the restriction that its distance in orientation space was at least 75° from the mid-point of the two distractor distributions (to keep the search task moderately easy for observers).

A single test trial followed the learning streak, where the orientation difference between the target on the test trial and the mid-point of the two distractor distribution means from the preceding learning streak was manipulated. This CT-PD (Current Target – Previous Distractors) distance determines the similarity between the current target and previous distractors. The CT-PD distance was manipulated throughout each experiment yielding a uniform distribution of CT-PD values across orientation space. For this, we divided orientation space into 12 bins with each bin having a range of 15°. On each test trial, a CT-PD distance was randomly determined from the bins with the restriction that at the end of each experiment there were equal numbers of CT-PD distances

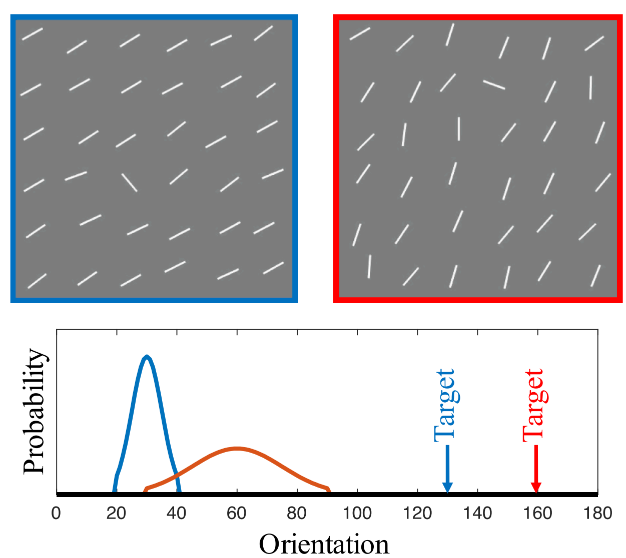

Fig. 3. Two example search displays and the underlying distractor probability distributions for distractor sets (shown below). The colored frames around the search arrays are added to indicate the corresponding target and distractor features, and were not included in the experimental stimulus. On the left, the orientations of the distractor lines are sampled from a Gaussian distribution centered on 30° (blue curve) with SD = 5°. The target orientation is 130°. On the right, the orientations of the distractors are sampled from a Gaussian distribution centered on 60° (red curve) with SD = 15°. The target orientation is 160°.

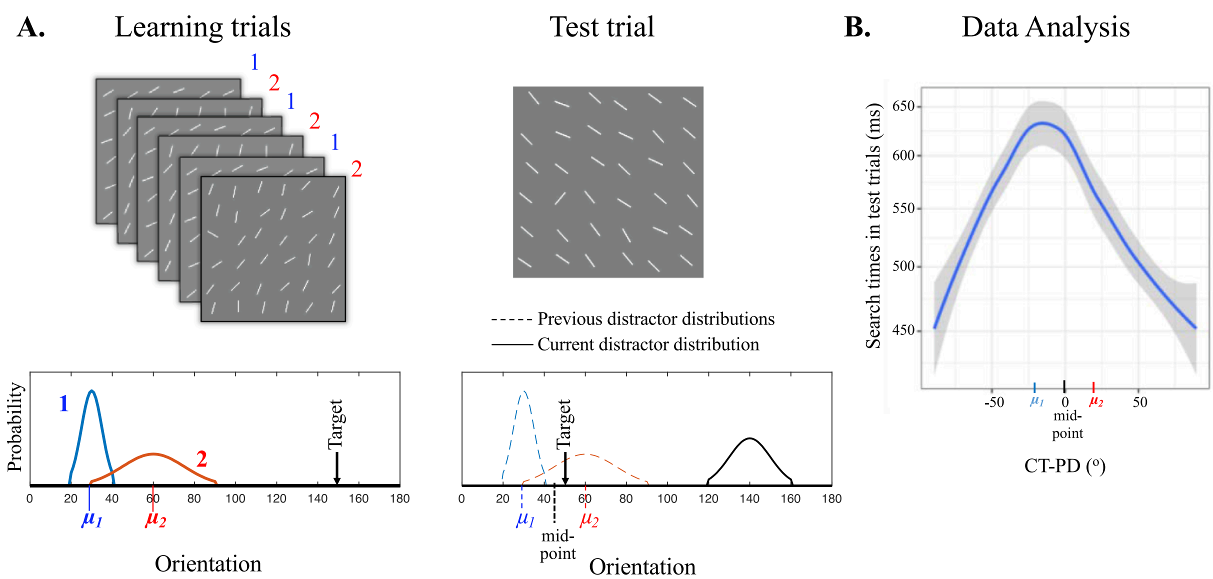

Fig. 4. A. An example block from Experiment 1. On learning trials, distractor orientations were sampled from two different Gaussian distributions in alternating order. On a test trial role reversal effects are induced at different CT-PD (Current Target – Previous Distractors) distances. In this example, the mid-point of the means of the previous distractor distributions of the learning streak is 45°, and the target orientation on the test trial is 50°, which yields a CT-PD value of 5°. Within different blocks, the target on the test trial is chosen at different places in orientation space to yield a uniform coverage of CT-PD values. B. An example curve obtained by plotting search times on test trials as a function of CT-PD distances. The mean of this curve corresponds to the most expected distractor orientation given the two distractor distributions on the learning trials. In three experiments, we manipulated the SD of the two distractor distributions on the learning trials to investigate their effect on the most expected orientation obtained from such CT-PD curves.

from each bin. The target orientation on each test trial was determined by the chosen CT-PD distance. Distractor orientations on a test trial were sampled from a Gaussian distribution ($SD = 10°$, truncated similarly to learning trials, and the orientations of two distractors were set to the minimal and maximal values of its range) whose mean was randomly determined with the restriction that the target to distractor-mean distance was at least 60°.

2.4. Procedure and materials

Participants sat 57 cm away from a 24-inch LCD monitor with a resolution of 1920 × 1080 pixels, connected to a Windows 7 PC. All experiments were presented with MATLAB Psychtoolbox (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007). Participants searched for the oddly oriented line among the 36 oriented lines and indicated with key presses whether the oddly-oriented line (i.e., target) was in the upper or lower three rows of the search array. The position of the target within the search array was randomized. Participants were allowed to make eye-movements, and instructed to respond as quickly and as accurately as possible. If they made a mistake, the word "ERROR!" appeared for 1 s at center in red, otherwise, the next search trial immediately followed. A score was calculated for each response based on response time and accuracy to motivate participants (for correct responses: $score = 10 + (1 - RT)*10$; for errors:$score = -|10 + (1 - RT)*10| - 10$; where RT is response time in sec; the score rewards fast correct responses). The score from each trial was presented in the top left corner of the screen in green for correct and red for incorrect responses. Total accumulated scores were shown to participants at each break and at the end of the sessions. All experiments were run in accordance with the Declaration of Helsinki and the requirements of the local ethics committee. All observers signed an informed consent form before participating.

Trials with incorrect responses and with exceptionally high (>3 s) and low (<200 ms) search times were excluded from all analyses on search times. For the analyses done on test trial search times (i.e., role reversal effects), test trials following a learning streak with at least one incorrect response on the last two learning trials were also excluded.[2]

3. Experiment 1

3.1. Participants

Ten participants (five females, age $M = 32.7$) who had normal or corrected-to-normal visual acuity took part in the study. Two participants were the authors ODT and AK. All signed an informed consent form before participating and were paid for participation (except the two authors).

3.2. Method

The experiment started with blocks of trials with a learning streak followed by a test trial (see General Method for details). To test the effects of learning streak length, the streaks were either short (2 trials) or long (6 or 8 trials). On learning trials, the *SD* of the distractor distribution alternated either between 5° and 15° or stayed constant, which yielded four conditions: "15 – 15", "15 – 5", "5 – 15" and "5 – 5", where the first number indicates the *SD* of the distractor distribution used on the first trial of the learning streak (distribution 1 in Fig. 4), and the second the *SD* of the distribution used on the last trial of the learning streak (distribution 2 in Fig. 4). One of the distributions was shifted clockwise and the other counter-clockwise from the mid-point. All these factors were randomized between blocks, and counterbalanced for each condition and participant. Each participant completed 8448 search trials, or 1536 blocks: 4 (*SD* manipulation: "15 – 15", "15 – 5", "5 – 15", "5 – 5") × 2 (which distribution is clockwise of the mid-point: distribution 1

[2] For all three experiments, similar results were obtained when such trials were not excluded. However, excluding such trials is important to observe a clearer effect of role-reversals.

or distribution 2) × 12 (CT-PD bins) × 2 (learning streak length: short or long) × 8 (repetition).

The experiment was completed in four approximately 45 min sessions. Four breaks split the session into five parts, where observers could rest as needed. Two naïve observers participated in a full session for practice before the first experimental session because of exceptionally high response times (>3 s) initially. Other participants completed 100 practice trials at the start of the first session, and 50 practice trials for the following sessions.

3.3. Results

Trials with search times higher than 3 s and lower than 200 ms were excluded (around 0.1% of all trials). Search times were log-transformed for all analyses, and trials where participants made an error were excluded from search time analyses.

3.3.1. Average search performance

Table 1 shows the average search times and accuracy for distractor distributions with different *SD*s. A one-way repeated measures ANOVA yielded a significant effect of distribution *SD* on search times ($F(2,18) = 82.19$, $p < 0.001$, $\eta^2 = 0.2$) and accuracy ($F(2,18) = 27.29$, $p < 0.001$, $\eta^2 = 0.48$). As expected, distractor distributions with larger *SD*s yielded longer search times and lower accuracy, presumably because of larger distractor orientation variability in the display.

3.3.2. Repetition effects

Search times as a function of trial number during the learning streaks are shown in Fig. 5 for each distractor *SD* condition. In Fig. 5A, conditions "15 – 5" and "5 – 15" were combined to compare repetition effects in learning streaks with alternating distractor *SD*s to learning streaks with a single *SD*. Search times from learning streaks with alternating distractor *SD*s are shown in more detail in Fig. 5B. A linear mixed effects regression with Helmert contrasts was fit to search times with trial number as a fixed effect and participants as a random effect. This allowed us to compare the search time of each trial number with the average of the following trials. Search times decreased significantly only after the first learning trial before reaching a plateau for the "15 – 15" condition ($B = 0.06$, $t = 6.66$, $p < 0.001$), but this plateau occurred after the third trial for the "5 – 5" condition ($B = 0.01$, $t = 2.82$, $p < 0.01$). For the learning streaks where *SD* alternated ("5 – 15" and "15 – 5"), search times decreased after the second trial before reaching a plateau ($B = 0.01$, $t = 3.23$, $p < 0.01$). Accuracy significantly increased on the first two trials of the "15 – 15" condition ($B = 0.21$, $Z = 2.57$, $p = 0.01$), whereas it increased only after the first trial for the learning streaks with alternating *SD*s ($B = 0.36$, $Z = 5.05$, $p < 0.001$). There were no repetition benefits for accuracy in the "5 – 5" condition, most likely reflecting ceiling effects at the first trials of the learning streak. For learning streaks where the *SD* of the distractor distribution alternated, the *SD* of the distractor distribution determined search times and accuracy (Fig. 5B), as the alternating search times and accuracy values show. Repetition effects were only observed after the first trial regardless of the first distractor distribution *SD* in a streak, as the difference in search times between the first and third trials of the conditions "15 – 5" and "5 – 15" shows (Fig. 5B, upper row).

Table 1
Search times and accuracy in Experiment 1 as a function of distractor distribution SD. The trials in which the distractor SD was 10° correspond to test trials, whereas the other ones (SD of 5° and 15°) correspond to learning trials.

Distractor Distribution SD	Accuracy M	SD	Search times of correct responses (ms) M	SD
5°	0.97	0.01	558	67
10°	0.95	0.02	629	90
15°	0.91	0.04	667	112

3.3.3. Integration of distributions

Besides excluding error trials and trials with exceptionally high (>3s) and low (<200 ms) search times, test trials following a learning streak with at least one incorrect response on the last two trials were also excluded (11% of the remaining test trials).

We flipped the CT-PD values around the mid-point (by multiplying them with −1) only for test trials which follow a learning streak that ends with a distractor distribution whose mean was counter-clockwise of the mid-point. This enabled us to analyze the search times on test trials when all learning streaks in each condition end with a distractor distribution whose mean is clock-wise of the mid-point (distribution 2 in Fig. 6 whose mean is indicated with the red dotted line). This was done to avoid having two types of learning streaks where the mean of the distractor distribution on the last trial falls on opposite sides of the mid-point. If observers do not integrate the distributions but instead rely on the distractor distribution on the last learning trial, then having these two types of learning streaks could lead to average circular means falling around the mid-point, not because the distributions got integrated, but simply because the data from both sides of the mid-point was aggregated.

Learning streak length did not significantly affect search times from test trials ($F(1,9) = 0.74$, $p = 0.4$, $\eta^2 < 0.001$) or the circular means of CT-PD curves ($F(1,9) = 1.14$, $p = 0.3$, $\eta^2 = 0.09$), and we therefore combined the data from the two learning streak lengths in the following analyses. Fig. 6A shows representative search times on test trials as a function of similarity between the target and distractors on the preceding learning trials (i.e., CT-PD curves) for one participant. As expected, search times on test trials were highest when the target orientation came from the distractor distributions used on the preceding learning trials (i.e., when CT-PD is close to the mid-point, 0°). These CT-PD curves (Fig. 6A) reflect how participants integrated information from the distractor distributions from the learning streak. In order to assess whether the *SD* of the distractor distribution biased this integration, we calculated the circular mean of the CT-PD curves for each observer and *SD* condition (Fig. 6B). The calculated circular means mostly fell around the mean of the distractor distribution that was used on the last trial of learning streaks ($\mu = 15°$, red dotted line in Fig. 6B).

We conducted Bayesian analysis using the "BayesFactor" package in R (Morey & Rouder, 2018) with its default parameter values. We first conducted a two-way Bayesian hypothesis testing (for the details of the computation, see Rouder, Morey, Speckman, & Province, 2012) in which the *SD* (5° or 15°) of the distractor distribution on the last and of the one before the last[3] learning trial were the two main factors, whereas participants were added as a random factor. The effects of the two main factors were assessed by comparing a model with both main factors to a model with only one of the factors. The interaction between the two factors was assessed by comparing the full model to a model with only the two main factors. For all the three factors, the evidence assessed by the Bayes Factor (*BF*) was in favour of the null model (*SD* of the last trial, $BF_{excl} = 3$; *SD* of the trial before the last, $BF_{excl} = 2$; their interaction, $BF_{excl} = 2.4$). We also tested whether the circular CT-PD means for these two factors were different than the mean of the distractor distribution used on the last learning trial ($\mu = 15°$, red dotted line in Fig. 6C). A one-sample test compared an alternative model ($\mu \neq 15°$) to a null model ($\mu = 15°$). There was positive evidence in favour of the alternative model when the *SD* of the last trial was 15° ($BF_{10} = 3.4$), and when the *SD* of the trial one before the last was 15° ($BF_{10} = 5.2$). However, the evidence in

[3] Note that because the learning streaks had even lengths, the first learning trial has the same underlying distractor distribution as the trial before the last one. Therefore, examining the effect of the *SD* of the trial before the last learning trial is also equivalent to examining the effect of the *SD* of the first learning trial. Similarly, examining the effect of with which *SD* (5° or 15°) a learning streak ends (i.e. *SD* of the last trial) is equivalent to examining the effect of the *SD* of the second learning trial in a streak.

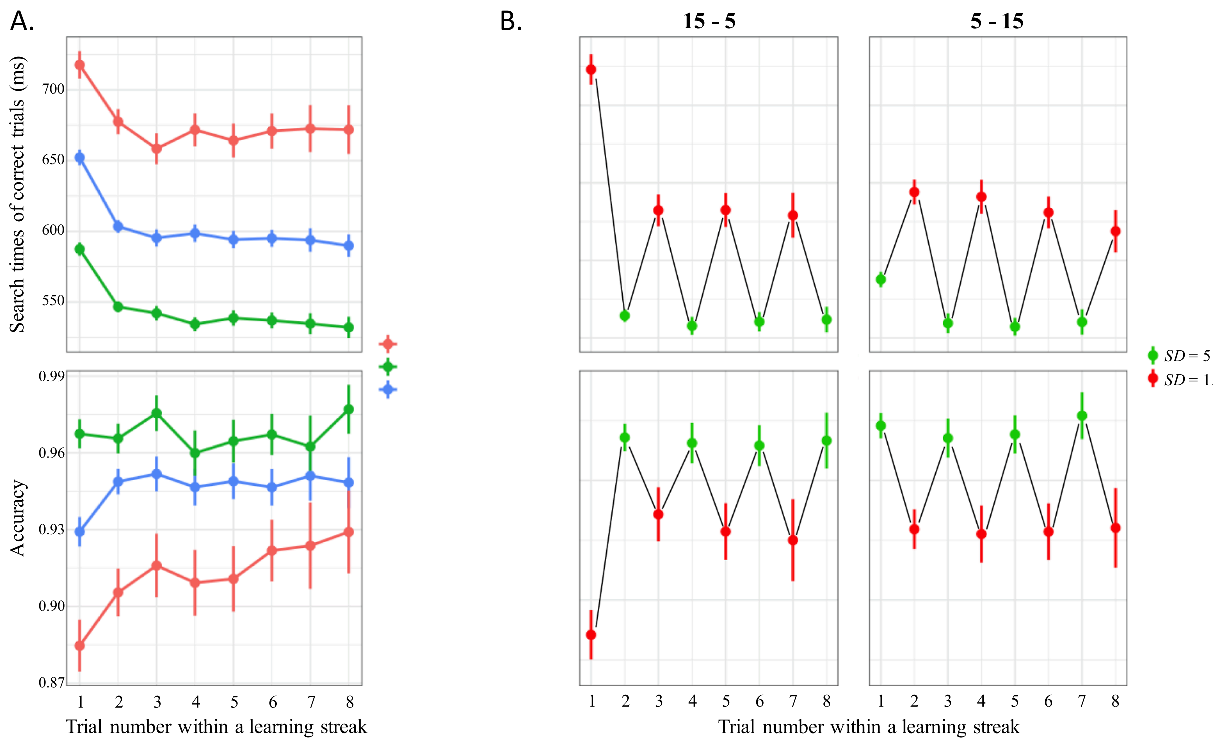

Fig. 5. A. Search times and accuracy in Experiment 1 as a function of trial numbers in the learning streak for different SD pairings. The conditions where the SD of the distractors alternate throughout the streak are combined for this plot. The error bars denote 95% confidence intervals. B. Search times and accuracy for the conditions where the SD of the distractor distributions alternates during the learning streak. The first column on the left shows the search times (top) and accuracy (bottom) for the learning streaks starting with a trial that includes a distractor distribution with SD = 15°. The column on the right shows the search times and accuracy for learning streaks starting with a trial that includes a distractor distribution with SD = 5°. The error bars denote 95% confidence intervals.

favour of the alternative model was low when the *SD* of the last trial was 5° ($BF_{10} = 1.7$), and when the *SD* of the trial one before the last was 5° ($BF_{10} = 1.2$).

We also calculated the circular *SD* of CT-PD curves obtained from each condition and participant (Fig. 6D). We examined the effect of the *SD* of the last and of the one before the last learning trial on the calculated circular *SD's* of the CT-PD curves. A similar two-way Bayesian analysis revealed positive evidence ($BF_{incl} = 10.7$) in favour of the main effect of the *SD* of the last trial. As Fig. 6D shows, the calculated circular *SDs* were notably higher when the last trial of the learning streak included a distractor distribution with an *SD* of 15° (i.e., conditions "15 – 15" and "5 – 15"), than an *SD* of 5° (i.e., conditions "15 – 5" and "5 – 5"). However, the main effect of the *SD* of the trial before the last one and the interaction of the two factors yielded negligible evidence in favour of the null model ($BF_{excl} = 2.5$ and 1.7, respectively).

3.4. Discussion

The largest role-reversal effects (i.e., highest search times) on test trials were observed when the target orientation was similar to the mean orientation of the distractor distribution on the last trial of the preceding learning streak. This indicates that observers were able to spatially integrate the distractor orientations on the last learning trial and encode their orientation distribution. At the same time, this also reveals a dominating influence of the last learning trial before the test trial. The evidence from the one-sample tests for the slight bias towards the distractor used in the trial before the last suggests a recency effect as opposed to a complete lack of integration in which observers would be only taking the last distractor distribution into account. However, the *SD* of the distractor distribution on the last trial had no effect on this bias.

Notably, the circular *SDs* of the CT-PD curves were significantly influenced by the *SD* of the distractor distribution on the last trial of a learning streak. This is in line with the recency effect observed on the mean of the CT-PD curves. When the *SD* of distractors on the last learning trials was 15°, the CT-PD curves obtained from the participants had higher *SDs* than those when the *SD* of the distractors on the last learning trials was 5°. This result also matches the findings of Chetverikov et al. (2016), who observed that search time slopes as a function of CT-PD distance on test trials become shallower as the *SD* of the distractor distribution on the learning trials increases. This is expected because larger *SDs* correspond to wider probability distributions, which, all things being equal, results in flatter CT-PD curves. However, in all conditions, the calculated *SD* of the CT-PD curves were much higher than the actual *SD*'s of the distractor distributions.

4. Experiment 2

Having observed the recency effect in the first experiment, we decided to decrease the orientation distance and also increase the overlap between the two distractor distributions in Experiment 2 to test whether it would facilitate integration. We decreased the orientation distance between the two distributions to 20° (as opposed to 30° in Experiment 1). Another potential factor that might influence integration could be the *SD* used for the low variance distribution. The distractor distribution with *SD* = 5° created a seemingly homogenous search display. This low level of uncertainty in one set of trials can potentially discourage the visual system from integrating it with information from other trials. Therefore, in Experiment 2, we increased the *SD* of the low variance distribution to 8°. This change not only introduced more uncertainty, but also increased the overlap between the two distributions, which could potentially facilitate integration.

4.1. Participants

Ten participants (5 females, age $M = 31.5$) with normal or corrected-to-normal visual acuity participated. Nine had participated in

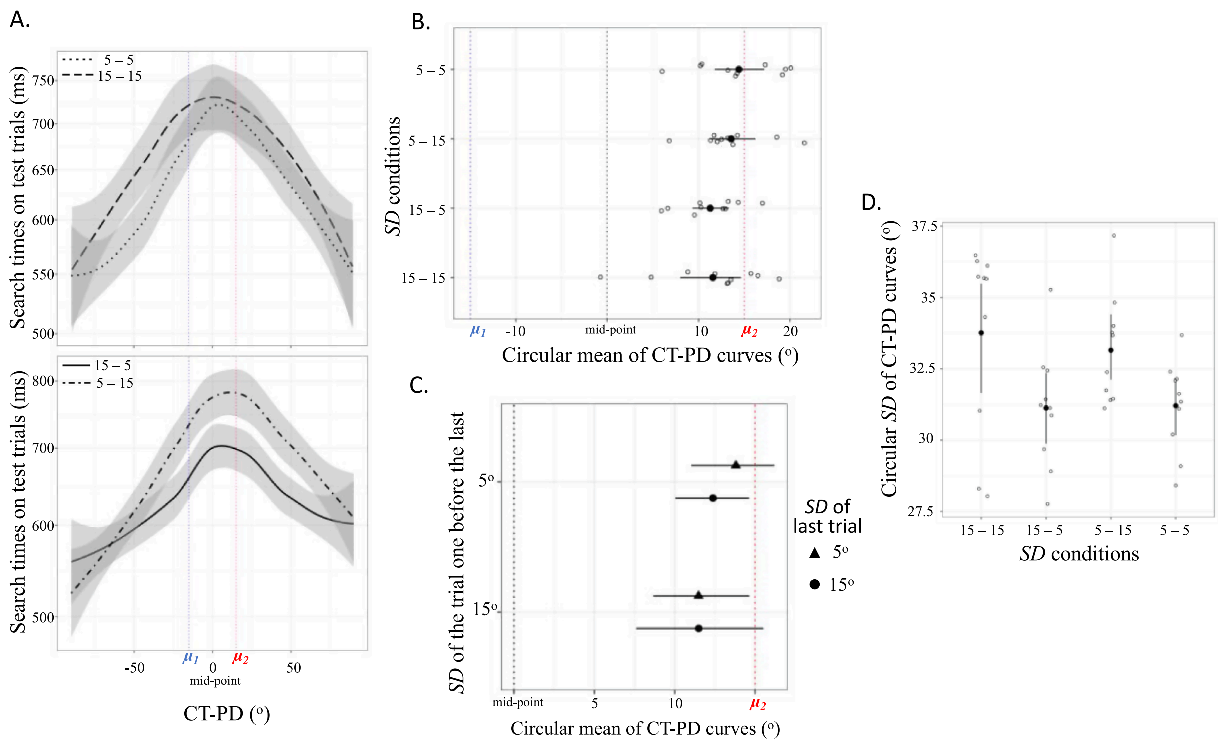

Fig. 6. A. Example of a CT-PD curve (i.e., search times on test trials as a function of CT-PD distances) obtained from one participant in Experiment 1. The upper plot includes the CT-PD curves from the conditions where the SD of the two distractor distributions were equal, where the one on the bottom shows the conditions where SD of the distractor distributions alternated. The red (μ_2) and the blue (μ_1) dashed line indicate the means of the distractor distribution on the last and on the one before the last learning trial, respectively. The grey region around the curve indicates the 95% confidence interval of the LOESS (locally estimated scatterplot smoothing) regression fitted to the data. B. Circular means of the CT-PD curves for each condition of Experiment 1. The black filled circles indicate the average circular means with error bars indicating 95% confidence intervals, whereas the smaller open circles indicate the circular means obtained from individuals. The red (μ_2) and the blue (μ_1) dashed line indicate the mean of the distractor distribution on the last and on the one before the last learning trial, respectively. C. Circular means of the CT-PD curves depending on the SD of the distractor distribution on the last or the one before the last learning trial. Error bars indicate 95% confidence intervals. The red dashed line indicates the mean of the distractor distribution in the last learning trial (μ_2), whereas the black one indicates the mid-point of the two distractor distributions. D. Circular SDs of the CT-PD curves obtained for each condition of Experiment 1. The black filled circles indicate the average circular SDs with error bars indicating 95% confidence intervals, whereas the smaller open circles indicate the circular SDs calculated from individuals. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Experiment 1 and two were the authors ODT and AK. All participants signed a consent form before participating and were paid for their participation (except the two authors).

4.2. Method

There were only two differences in the design from Experiment 1. Firstly, the distractor distribution with the lower variance had an *SD* of 8°, yielding four conditions “15 – 15”, “15 – 8”, “8 – 15” and “8 – 8”. Secondly, the orientation difference between the means of the two distractor distributions was decreased to 20°. One participant who was not familiar with the search task was given a full practice session before the experimental sessions.

4.3. Results

The same trial exclusion conditions as in Experiment 1 were applied.

4.3.1. Average search performance

There was a significant effect of distractor distribution *SD* on both search times ($F(2,18) = 124.63, p < 0.001, \eta^2 = 0.14$) and on accuracy ($F(2,18) = 19.09$, $p < 0.001$, $\eta^2 = 0.25$). As in Experiment 1, distractor distributions with larger *SD*s yielded longer search times and lower accuracy (Table 2).

Table 2
Search times and accuracy in Experiment 2 as a function of distractor distribution SD. The trials in which the distractor SD was 10° correspond to test trials, whereas the other ones (SD of 8° or 15°) to learning trials.

Distractor Distribution SD	Accuracy M	SD	Search time of correct responses (ms) M	SD
8°	0.97	0.01	590	66
10°	0.96	0.03	656	80
15°	0.94	0.03	659	90

4.3.2. Repetition effects

Fig. 7 shows average search times within learning streaks as a function of trial number. The conditions where the *SD* of the distractor distribution alternates within a learning streak (i.e., “15 – 8” and “8 – 15”) were combined in Fig. 7A (see Fig. 7B for more details for these conditions). Linear mixed effects regressions with Helmert contrasts, comparing each trial with the average of the following trials, revealed that for all three conditions search times significantly decreased after the first learning trial (For 15 – 15: $B = 0.08$, $t = 13.91$, $p < 0.001$; for 8 – 8: $B = 0.1$, $t = 15.25$, $p < 0.001$; for alternating SD streaks: $B = 0.08$, $t = 16.00$, $p < 0.001$). However, no significant decrease was observed after the first trial. Similar analyses on accuracy revealed significant increases only after the first trial for conditions “15 – 15” ($B = 0.17$, $Z = 4.50$, $p < 0.001$), and for the condition where the SD of the distractor distribution alternated within a streak ($B = 0.24$, $Z = 2.70$, $p < 0.01$). There was no

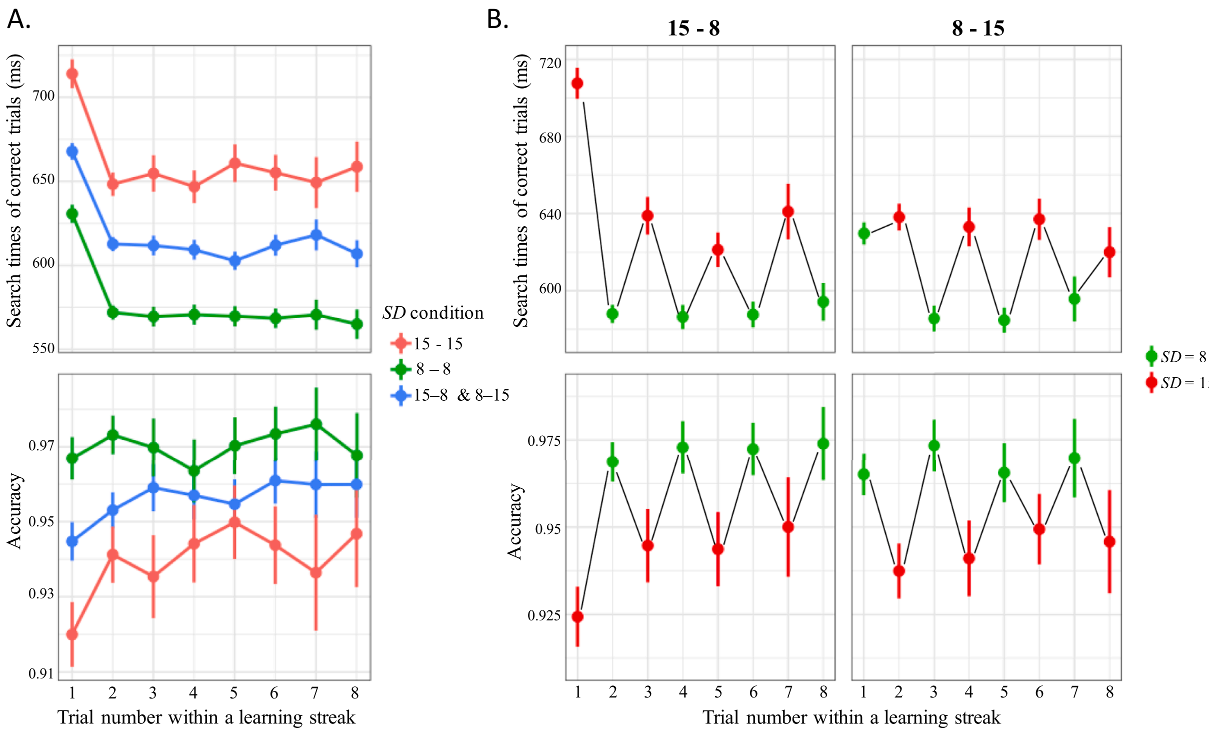

Fig. 7. A. Search times and accuracy as a function of learning trial number in Experiment 2 for different SD pairings. The conditions where the SD of the distractors alternate through the streak are combined. B. Search times and accuracy for the conditions where the SD of the distractor distribution alternates during the learning streak. The first column on the left shows the search times (top) and accuracy (bottom) for the condition "15–8", whereas the column on the right side denote the "8–15" condition. Error bars denote 95% confidence intervals.

repetition effect for the "8 – 8" condition ($B = 0.17$, $Z = 1.25$, $p = 0.21$), most likely due to a ceiling effect. Search times and accuracy for the conditions "8 – 15" and "15 – 8" can be seen in more detail in Fig. 7B.

4.3.3. Integration of distributions

In addition to excluding trials with incorrect responses and exceptionally high (>3s) and low (<200 ms) search times, 7% of the remaining test trials were also excluded (test trials following a learning streak with at least one incorrect response in the last two trials).

As for Experiment 1, we flipped the CT-PD values around the mid-point on test trials following a learning streak ending with a distractor distribution centered counter-clockwise of the mid-point. The data in Fig. 8 are therefore plotted so that all learning streaks end with a distractor distribution centered clock-wise (distribution 2 in Fig. 8 whose mean is indicated with the red dotted line). Again, there was no effect of learning streak length on search times on test trials ($F(1,9) = 0.14$, $p = 0.7$, $\eta^2 < 0.001$) nor on the mean of CT-PD curves ($F(1,9) = 0.01$, $p = 0.9$, $\eta^2 < 0.001$), so the data from different learning streaks lengths were combined.

Fig. 8A shows a representative CT-PD curve (i.e., search times on test trials as a function of similarity between target and distractors on preceding learning trials) for one participant. Fig. 8B shows the circular mean of the CT-PD curves obtained from the participants for each condition. As in Experiment 1, the means of the CT-PD curves were mostly centered around the mean of the distractor distribution from the last trial of a learning streak. We conducted the same Bayesian analysis as we did in Experiment 1 where we examined the effect of the *SD* of the distractor distribution on the last learning trials, as well as *SD* of the one before the last trial. For the effect of the *SD* of the last trial, the BF_{excl} in favor of the null model was 3, whereas for the effect of the *SD* of the one before the last the evidence was in favor of the alternative hypothesis ($BF_{incl} = 2.6$). There was no evidence in favor of either hypotheses for the interaction of the two factors ($BF_{incl} = 1$). We also tested whether the circular CT-PD means for the two factors were different than the mean of the distractor distribution used on the last learning trial ($\mu = 10°$, red dotted line in Fig. 8B). We compared a point null model ($\mu = 10°$) to the alternative model of $\mu \neq 10°$. There was positive evidence in favour of the null model when the *SD* of the last trial was 15° ($BF_{01} = 3.7$), and when it was 8° ($BF_{01} = 4.1$). The evidence in favor of the null was low when the *SD* of the trial before the last one was 15° ($BF_{01} = 1.7$), or when it was 8° ($BF_{01} = 1.2$).

Fig. 8D shows the calculated circular *SD*s of the CT-PD curves obtained from each condition. By using the same Bayesian analysis, we examined the effect of the *SD* of the distractor distributions (last vs. one before last) on the calculated circular *SD* of the CT-PD curves. There was positive evidence ($BF_{incl} = 57$) for the effect of the *SD* of the last learning trial. However, the evidence was in favor of the null model ($BF_{excl} = 3.3$) for the effect of the *SD* of the trial before the last learning trial. For the interaction of the two factors, there was almost no evidence for either hypothesis ($BF_{incl} = 1.1$ in favor of the interaction). Overall, the *SD* of the CT-PD curves was notably higher when the *SD* of the distractor distribution on the last learning trial was 15° (conditions 8 – 15 and 15 – 15) than when it was 8° (conditions 8 – 15 and 15 – 15).

4.4. Discussion

In Experiment 2, we reduced both the difference between the *SD*'s of the two distractor distributions on the learning trials (8° & 15° vs. 5° & 15°), as well as the distance between the means of the two distributions, compared to Experiment 1, causing larger overlap between the two distractor distributions. The results were nevertheless similar to Experiment 1. The role reversal effects depended on the last trial of the learning streak, since the mean of the CT-PD curves did not differ from the mean of the distractor distribution on that trial. The *SD*s calculated from the CT-PD curves, which were higher when the distractor distribution on the last learning trial had a higher *SD* (=15°), also show this.

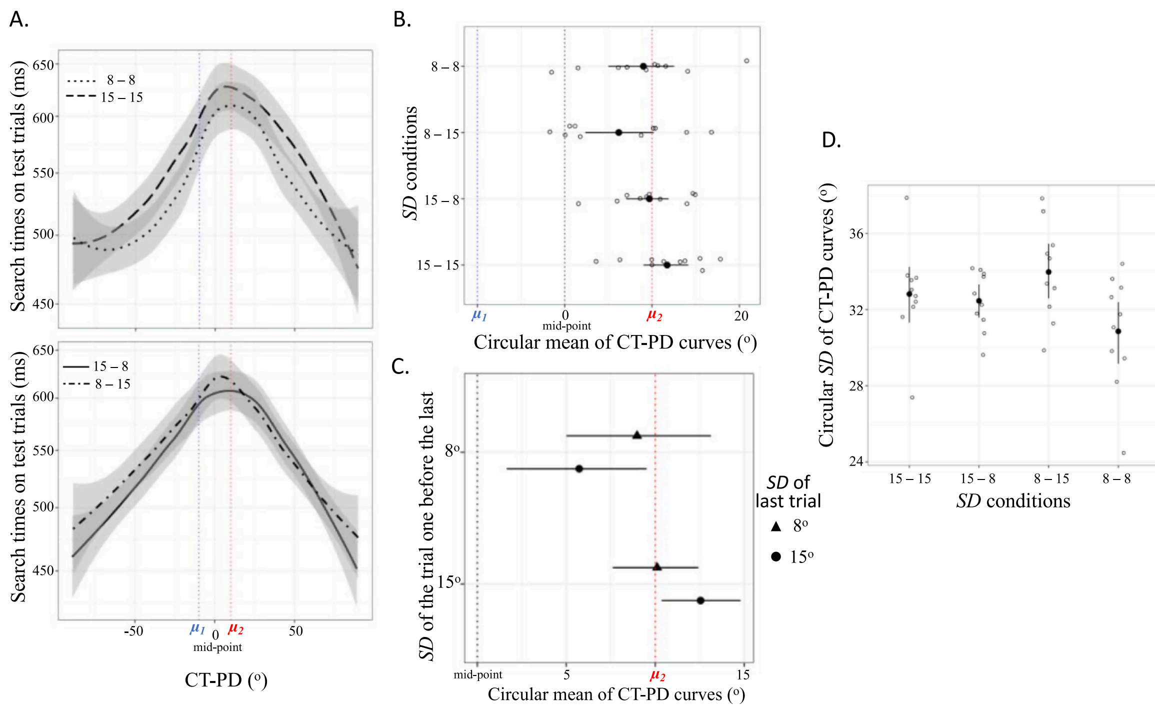

Fig. 8. A. Example of a CT-PD curve (i.e. search times on test trials as a function of CT-PD distances) obtained from one participant in Experiment 2. The upper plot includes the CT-PD curves from the conditions where the SD of the two distractor distributions were equal, while the one at the bottom shows the conditions where SD of the distractor distributions alternated. The red (μ_2) and the blue (μ_1) dashed line indicate the mean of the distractor distribution in the last and in the one before the last learning trial, respectively. The grey fill around the curve indicates the 95% confidence interval of the LOESS (locally estimated scatterplot smoothing) regression fitted to the data. B. Circular means of the CT-PD curves for each condition. The black filled circles indicate the average circular means, whereas the smaller open circles indicate the circular means obtained from individuals. The red (μ_2) and the blue (μ_1) dashed line indicate the means of the distractor distribution on the last trial and on the one before the last learning trial, respectively. C. Circular means of the CT-PD curves depending on the SD of the distractor distribution on the last or the one before the last learning trial. Error bars indicate 95% confidence intervals. The red dashed line indicates the mean of the distractor distribution in the last learning trial (μ_2), whereas the black one indicates the mid-point of the two distractor distributions. D. Circular SDs of the CT-PD curves obtained for each condition of Experiment 2. The black filled circles indicate the average circular SDs, whereas the smaller open circles indicate the circular SDs calculated from individuals. Error bars denote 95% confidence intervals. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Experiment 3

In Experiment 3, we further decreased the orientation distance between the two distractor distributions in order to examine whether it would influence integration. The orientation distance between the distributions was set to 12° in Experiment 3. Since increasing the *SD* of the low variance distribution in Experiment 2 did not facilitate integration, in this experiment we went back to using *SD* of 5° for the low variance distractor distribution. In addition to this, we did not manipulate the learning streak length in Experiment 3 since no influence of streak length was observed in the first two experiments.

5.1. Participants

Ten participants (7 females, age $M = 29.8$) with normal or corrected-to-normal visual acuity took part in the study. Seven had participated in Experiment 1 and 2. One participant was the author ODT. All were volunteers and signed a consent form before participating.

5.2. Method

The design and the procedure were similar to Experiment 1 and 2, with the following exceptions: The orientation distance between the two distractor distributions was set to 12°. In addition to this, the learning streak length was not manipulated in Experiment 3, but we included two different learning streak lengths (four or six trials) to decrease potential anticipation of upcoming test trials (Shurygina et al., 2019). *SD*'s of the distractor distributions on the learning trials were either 5° or 15°. Each participant completed 6912 search trials, a total of 1152 blocks: 4 (*SD* manipulation: "15 – 15", "15 – 5", "5 – 15", "5 – 5") × 2 (learning streak starts with: distribution 1 or distribution 2) × 12 (CT-PD bins) × 2 (learning streak length: 4 or 6) × 6 (repetition).

The experiment was completed in two ~55 min sessions with 5 breaks equally splitting the sessions. All participants completed 100 practice trials at the beginning of the first session, and 50 at the beginning of the second.

5.3. Results

The same trial exclusion conditions as in Experiment 1 were applied.

5.3.1. Average search performance

As before, there were significant effects of distractor distribution *SD's* on search times ($F(2,18) = 148.53$, $p < 0.001$, $\eta^2 = 0.22$), and accuracy ($F(2,18) = 29.2$, $p < 0.001$, $\eta^2 = 0.39$). Distractor distributions

Table 3
Search times and accuracy in Experiment 3 as a function of distractor distribution SD. The trials in which the distractor SD was 10° correspond to test trials, whereas the other ones (SD of 5° or 15°) to learning trials.

Distractor Distribution SD	Accuracy M	SD	Search time of correct responses (ms) M	SD
5°	0.97	0.01	531	58
10°	0.96	0.02	618	80
15°	0.94	0.03	612	84

with larger *SDs* yielded lower accuracy (Table 3). Search times for distractor distributions with $SD = 5°$ were significantly lower than on other trial types. Search times were highest for trials with distractor distribution $SD = 10°$, reflecting role-reversal effects since such trials were only used on test trials.

5.3.2. Repetition effects

The conditions with alternating distractor distribution *SDs* during a learning streak (5 – 15 and 15 – 5) were combined in Fig. 9A, but Fig. 9B shows search times from these two conditions in more detail. A linear mixed effects regression with Helmert contrasts showed that search times significantly decreased only after the first trial of the learning streaks for conditions "5 – 5" ($B = 0.09$, $t = 10.95$, $p < 0.001$) and "15 – 15" ($B = 0.1$, $t = 10.57$, $p < 0.001$). For the combination of the conditions "5 – 15" and "15 – 5" (the blue line in the upper plot of Fig. 9A), the significant decrease in search times also occurred after the second learning trial ($B = 0.01$, $t = 3.49$, $p < 0.001$), before reaching a plateau. The same analysis on accuracy yielded a significant increase after the first trial for the combination of the "5 – 15" and "15 – 5" conditions ($B = 0.19$, $Z = 2.04$, $p = 0.04$). There was also a significant increase after the second trial for the "15 – 15" condition ($B = 0.29$, $Z = 2.61$, $p = 0.01$). Similar to previous experiments, no repetition effect was observed for accuracy in the "5 – 5" condition.

5.3.3. Integration of distributions

Besides incorrect test trials and trials with exceptionally high and low search times, 8% of the remaining test trials were excluded due to at least one incorrect response in the last two trials of their preceding learning streak. As in previous experiments, the CT-PD values were flipped so that all learning streaks end with a trial where the distractor distribution is centered clockwise of the mid-point.

Fig. 10A shows a representative CT-PD curve for one participant, while Fig. 10B shows the circular mean of CT-PD curves for each different condition of the experiment. The same two-way Bayesian analysis as in previous experiments were performed to examine the effect of the *SD* of distractors of the last trial and of the one before the last trial. The evidence was in favor of the null model both for the effect of the *SD* on last trial ($BF_{excl} = 3.2$) and for the effect of the *SD* on the trial before the last ($BF_{excl} = 2.7$). The BF_{incl} in favor of the interaction of the two factors was 3.4. When the *SD* of both the last trial and the one before the last were equal ("15 – 15" and "5 – 5"), the circular means of the CT-PD curves fell around the mean of the distractor distribution of the last trial. A slight bias towards the mid-point can be seen when they were not equal ("5 – 15" and "15 – 5"). When we tested whether the circular means of CT-PD curves differ ($\mu \neq 6°$) from the mean of the distractor distribution in the last trial of the learning streak (the red dotted line in Fig. 10C), we observed positive evidence in favour of the null model ($\mu = 6°$) when the *SD* of the last trial was 15° ($BF_{01} = 3.5$), and when the *SD* of the trial before the last one was 15° ($BF_{01} = 3.8$). The evidence in favor of the null model was low when the *SD* of the last trial was 5° ($BF_{01} = 1.8$), and when the *SD* of the trial before the last one was 5° ($BF_{01} = 2.2$).

Fig. 10D shows the circular *SDs* of the CT-PD curves for each condition. A two-way Bayesian analysis revealed that there was strong evidence ($BF_{incl} = 319$) for the effect of the *SD* of the last trial. This effect was due to the higher *SDs* of the CT-PD curves from learning streaks ending with a distractor $SD = 15°$ (i.e. 5 – 15 and 15 – 15). The evidence was in favour of the null model for the effect of the *SD* of the trial before the last ($BF_{excl} = 2.7$), as well as for the interaction of the two factors ($BF_{excl} = 2.1$).

5.4. Discussion

In Experiment 3, we tried to facilitate integration of the two distractor distributions by reducing the distance between the means of the two distractor distributions. The results were, however, in essence similar to the previous two experiments. The distractor distribution used in the last learning trial determined the structure of the CT-PD curves, with little signs of integration of information from the previous trials. The only influence of the *SD* of the distractor distribution on role-

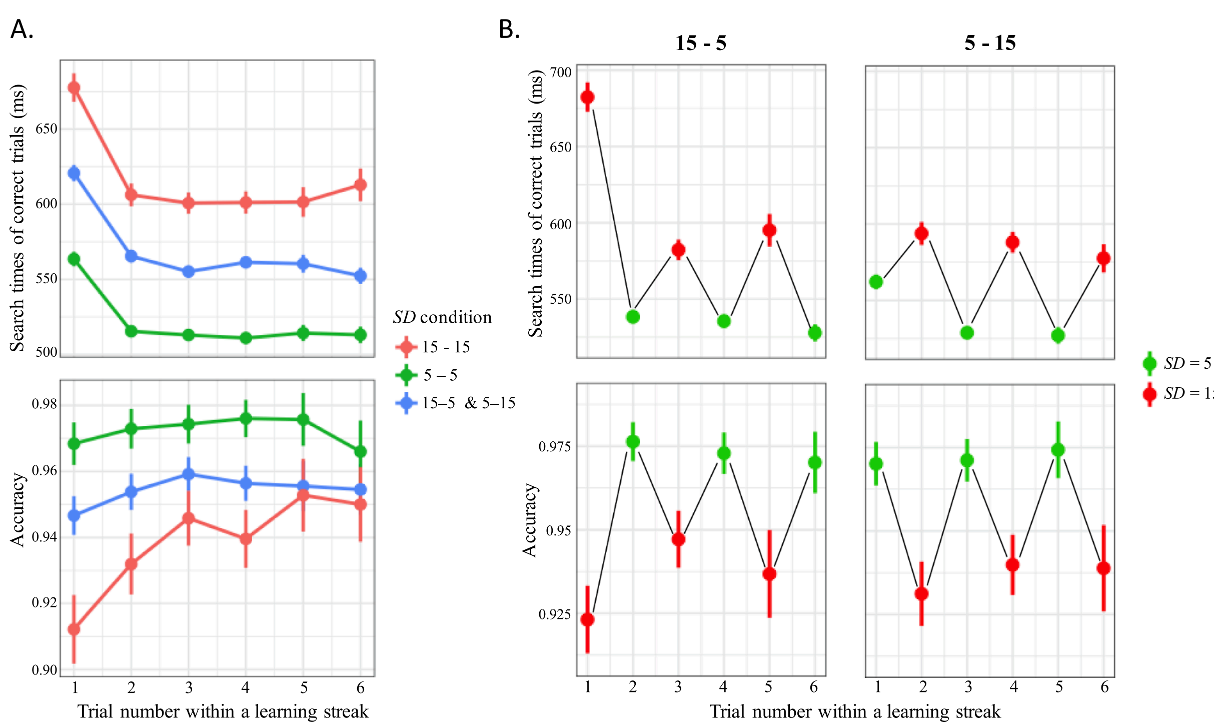

Fig. 9. A. Search times and accuracy as a function of trial number within learning streaks in Experiment 3. B. Search times and accuracy for the conditions where the SD of the distractor distributions alternates during the learning streak. The first column on the left shows the search times (top) and accuracy (bottom) for the condition "15–5", and the column on the right is for the condition "5–15". Error bars denote 95% confidence intervals.

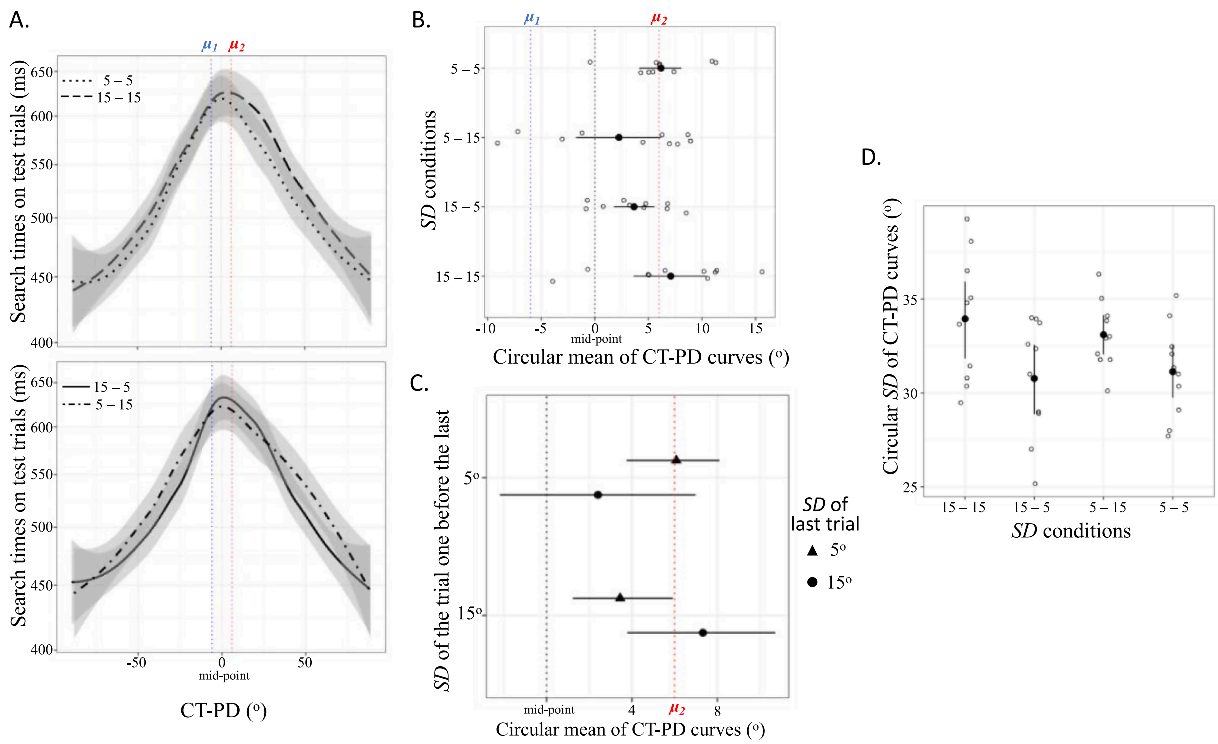

Fig. 10. A. An example of CT-PD curve (i.e. search times on test trials as a function of CT-PD distances) obtained from one participant in Experiment 3. The upper plot includes the CT-PD curves from the conditions where the SD of the two distractor distributions were equal, while the bottom one shows the conditions where SD of the distractor distributions alternated. The red (μ_2) and the blue (μ_1) dashed line indicate the mean of the distractor distribution on the last trial and on the one before the last learning trial, respectively. The grey fill around the curve indicates the 95% confidence interval of the LOESS (locally estimated scatterplot smoothing) regression fitted to the data. B. Circular means of the CT-PD curves for each condition of Experiment 3. The red (μ_2) and the blue (μ_1) dashed line indicate the mean of the distractor distribution on the last trial and the one before the last learning trial, respectively. The black filled circles indicate the aggregate data with error bars indicating 95% confidence intervals, whereas the smaller open circles correspond to individual data. C. Circular means of the CT-PD curves depending on the SD of the distractor distribution on the last or the one before the last learning trial. Error bars indicate 95% confidence intervals. The red dashed line indicates the mean of the distractor distribution on the last learning trial (μ_2), whereas the black one indicates the mid-point of the two distractor distributions. D. Circular SDs of the CT-PD curves obtained for each condition of Experiment 3. The black filled circles indicate the average circular SDs, whereas the smaller open circles indicate the circular SDs calculated from individuals. Error bars denote 95% confidence intervals. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

reversals was observed on the calculated *SD* of the CT-PD curves. When the *SD* on the last learning trial was 15°, the calculated *SDs* of CT-PD curves were much higher than the *SDs* of the CT-PD curves from learning streaks ending with distractor $SD = 5°$.

6. Analysis of aggregate data from all experiments

The main difference across the three experiments was the orientation distance (or the overlap) between the two distractor distributions used in the learning streaks. This was done in order to observe whether the distance between the distributions would influence their integration. Apart from this, all three experiments had a similar design in which there are two distractor distributions, one with a lower *SD* (5° and 8°) and one with a higher *SD* (15°). Given this similarity in their design, we combined the data from all three experiments to examine whether the *SD* of the last trial or the *SD* of the one before the last trial had an effect on the calculated circular means of the CT-PD curves.

Fig. 11 shows the combined data from all three experiments. In these plots, the reference point ($\mu = 0$) is taken as the mean of the distractor distribution on the last learning trial (μ_2, the red dotted line in Fig. 11A & B). Negative values indicate a bias towards the distractor distribution used on the trial before the last one. The labels “low” and “high” on the plots refer to distractor distributions with low (5° or 8°) or high *SD* (15°), respectively. We again ran a two-way Bayesian analysis to examine the effects of the two main factors; *SD* of the last trial and *SD* of the trial before the last. Participants and the experiment number were added to the model as random variables. For the effect of the *SD* of the last trial, there was positive evidence for the null model ($BF_{excl} = 4.8$). This indicates that the *SD* of the last trial did not have an effect on the circular means of the CT-PD curves. The evidence in favour of the null model was lower ($BF_{excl} = 2.3$) for the effect of the trial before the last one. However, there was positive evidence ($BF_{incl} = 4.1$) for the interaction of the two factors. This interaction is mostly due to the bias towards the distractor in the trial before the last, which was observed when the trial before the last one had a low *SD* and the *SD* of the last trial was high (Fig. 11B).

In order to further examine this interaction, we ran one-sample Bayesian hypothesis testing in which we compared a point null model ($\mu = 0°$, the red dashed line on Fig. 11B) to an alternative model of $\mu \neq 0°$. The evidence in favour of the alternative model was negligible when the *SD* of the last trial was low ($BF_{10} = 1.3$) or high ($BF_{10} = 1$). However, there was positive evidence ($BF_{10} = 5$) in favour of the alternative model when the *SD* of the trial before the last one was low. In contrast, when the *SD* of the trial before the last one was high, there was positive evidence in favour of the null model ($BF_{01} = 3$). These results indicate that the bias towards the trial before the last one can appear if the *SD* on that trial was low and the *SD* on the last trial was high. In other words, this slight bias was observed when uncertainty in the last trial was high (i.e. low reliability) *and* uncertainty was low (i.e. high reliability) in the trial before the last one.

Fig. 11C shows the circular *SDs* calculated from the CT-PD curves for each condition. The same two-way analysis was run in order to observe

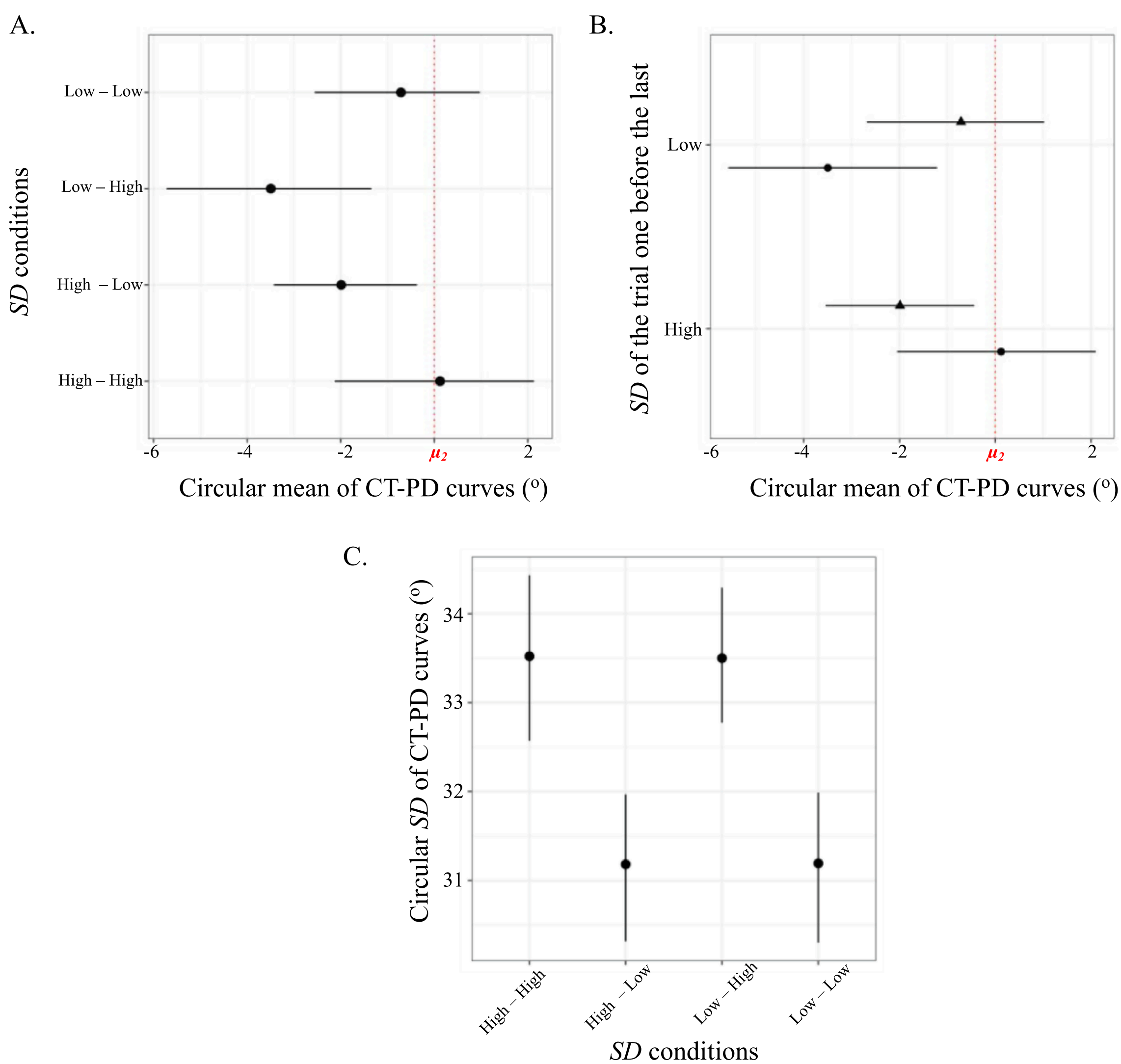

Fig. 11. A. Circular means of the CT-PD curves obtained from combining the data from all three experiments. The labels "low" and "high" refer to the SD of distractor distributions. Negative values indicate a bias towards the distractor on the trial before the last one. The reference point μ_2 (the red dashed line) indicates the mean of the distractor distribution on the last learning trial. Error bars indicate 95% confidence intervals. B. Same data plotted depending on the SD of the distractor distribution on the last or the one before the last learning trial C. Circular SDs of the CT-PD curves obtained by combining the data from all three experiments. The labels "low" and "high" refer to the SD of distractor distributions. Error bars denote 95% confidence intervals. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

how the calculated circular *SD* from CT-PD curves was influenced by the *SD* of the last trial and of the trial before the last. There was strong evidence for the effect of the *SD* of the last trial ($BF_{incl} = 3 \times 10^6$). When the *SD* of the last trial was high (15°), the calculated *SD* of the CT-PD curve was significantly higher compared to when the *SD* of the last trial was low. There was positive evidence in favour of the null model for the effect of the *SD* of the trial before the last one ($BF_{excl} = 5.2$), as well as for the interaction of the two factors ($BF_{excl} = 3.7$). These results demonstrate that the *SD* of CT-PD curve was only influenced by the distractor *SD* of the last trial: higher *SD* on the last trial yielded higher *SD* in CT-PD curves. Overall, the results indicate that observers were strongly biased towards the statistical properties of the most recent trial within the learning streak.

7. General discussion

Chetverikov and colleagues have shown how the visual system can encode feature distributions in surprising detail (Chetverikov et al., 2016), and how this learning develops over time (Chetverikov et al., 2017b). We wondered to what extent and in what ways these encoded feature distributions can be used by the visual system, particularly in integrating probabilistic visual input over time. In three experiments we addressed the question of whether the visual system optimally integrates two different distributions that alternated on adjacent visual search trials. We expected that the visual system would learn the feature distributions in light of our previous findings but the main question was how the two distributions would be integrated. Our question reflected whether the visual system would be capable of optimal integration in accordance with classic Bayesian integration principles (e.g., Knill, 2007; Knill & Richards, 1996; Körding & Wolpert, 2006) suggesting that observers will weigh more reliable information more strongly, in our case the distribution with the lower variance. We therefore expected the visual system to integrate the information by weighing the distribution with lower *SD* (and therefore more homogeneity) more highly in the integration process (this general logic is schematized in Fig. 2.)

However, we found little evidence of integration of visual feature distributions across learning trials. The observers' representations measured with the *FDL* method mostly reflected the statistics from the immediately preceding trial. The pattern of results we obtained was similar to the prediction depicted on the rightmost plot in Fig. 2B.

However, our overall results indicate a recency effect as opposed to a complete lack of integration. In certain conditions of each experiment, as well as in the combined data from all experiments, a bias towards the distractor distribution used on the trial before the last one can be observed. For example, the interaction observed between the *SD* of the last trial and of the trial before the last (Fig. 11B) was in line with the Bayesian principles of integration. The trial before the last one exerted an attractive bias only if it had lower uncertainty (i.e. was more reliable), *and* if the last trial had higher uncertainty (i.e. less reliable). However, there was positive evidence for the lack of an effect of the distractor *SD* on the last trial. This indicates that the influence of the *SD* of the last trial was mostly overridden by the strong recency effect. Even though the evidence to assess the main effect of the *SD* of the trial before the last one was weak, the results from the one sample tests on the combined data yielded positive evidence for a bias towards the midpoint when the trial before the last one had low *SD*, as well as positive evidence for no effect if it had high *SD*. This differential influence of the *SD* of the trial before the last one indicates that the visual system does not completely dismiss the information presented before the last learning trial (as depicted in the rightmost plot of Fig. 2B), and does take into account the uncertainty in the distribution parameters, but also suggests a dominant recency effect that tends to override the influence of the previous trials.

Even though the *SD* of the distractor distribution on the last trial had almost no effect on the circular means of CT-PD curves, it had a strong effect on their calculated circular *SD*s. When the distractor *SD* was high on the last trial, the *SD*s of the CT-PD curves were significantly higher than when the last distractor *SD* was low (Fig. 11C). This is in line with the recency effect we have observed in the circular means of the CT-PD curves, as well as with previous studies on FDL (e.g., Chetverikov et al., 2016, Exp. 1) where larger distractor *SD*s resulted in wider CT-PD curves. In addition to this, the calculated circular *SD*s in all conditions were notably higher ($>30^\circ$) than the actual *SD*s ($<15^\circ$) of the distractor distributions. Search times in FDL studies as a function of CT-PD distances are strongly proportional to the actual probability distribution function of the distractors. Given this, the high circular *SD*s obtained from the CT-PD curves suggest that observers' representation of the distractor distribution incorporates higher variance or uncertainty compared to the actual variance of the underlying distractor distribution. This could also speculatively be one of the reasons for why the *SD* of the distractor distributions had a minimal effect on the integration of the distractor distributions. The *SD* contrast between the two distractor distributions could be much lower than the *SD* contrast between the representations of these two distractor distributions.

The visual system is less sensitive to ensemble information of visual features when the variance among the relevant feature increases (Fouriezos, Rubenfeld, & Capstick, 2008; Haberman, Lee, & Whitney, 2015; Im & Halberda, 2013; Solomon, Morgan, & Chubb, 2011). Even though the variance of the distractor distribution within a single search display in our experiments could be low (e.g., 5°), the temporal variation between different trials of a learning streak was high (the mean of the distractor distribution on each learning trial was different than the preceding one because the distributions alternated). Moreover, in certain conditions the *SD*s of distractor distributions alternated as well, between 5° and 15°. We speculate that all this temporal uncertainty on the learning trials made the visual system rely only on the statistical information extracted from the last learning trial. This may reflect that the system weighs recent information very highly because the input is unreliable due to the alternation between different distractor distributions and prefers not to integrate unreliable information over time.

Repetition benefits on search times and accuracy (Kristjánsson & Ásgeirsson, 2019) were mostly observed after the first learning trial only. This is most likely because of a sudden change of target and distractor features due to the start of a new block of learning trials. Statistical stability over successive search trials can significantly improve search performance (Chetverikov et al., 2016), even if the statistical stability is irrelevant to the search feature (Corbett & Melcher, 2014). Search performance did not improve much after the first trial in our experiments, which indicates that the inter-trial priming effects during the learning streak were minimal. This is in line with the very small integration observed within a learning streak.

Temporal integration of visual features into ensembles has been observed in many studies (Chong & Treisman, 2003, Haberman et al., 2009; Albrecht & Scholl, 2010; Whiting & Oriet, 2011; Hubert-Wallander & Boynton, 2015; Oriet & Hozempa, 2016). Moreover, integration efficiency of temporally presented visual items can be even higher than, or at least equal to, integration efficiency of spatially presented items (Florey, Dakin, & Mareschal, 2017; Gorea, Belkoura, & Solomon, 2014). Our results with the FDL method are in seeming contrast with these studies. One of the main differences between our study and previous ones is that implicit measures not requiring explicit perceptual judgments were used to assess observers' representations of feature ensembles. This methodological difference is the most likely reason for the discrepancy between our findings and those previous findings on temporal integration. While our observers spatially integrated the orientation information within a single visual search trial to encode the distractor orientation distribution, there was little temporal integration of orientation distributions across successive trials. Recency effects have been observed in temporal integration of certain visual features into ensembles, such as size, facial expression and motion direction (Hubert-Wallander & Boynton, 2015). Our results point to a similar dominant influence of more recent information in the temporal integration of ensembles.

As mentioned before, the orientation distance between the two distractor distributions could influence the integration process. It is possible that there might be an optimum distance that could facilitate integration. However, the orientation difference that would facilitate integration can change depending on task and integration type. For example, Fischer and Whitney (2014) observed that the strength of serial dependence during orientation judgments weakens when the difference in orientation between the previous and current item exceeds approximately 28°. Utochkin and Yurevich (2016) found evidence of integration of different orientations into a single ensemble when the orientation difference between the distractor lines was 9°, but not when it was 22.5°. We varied the orientation distance between the two distributions means from 30° to 12°. Yet, the constant result patterns over the three experiments indicates that this distance was not the factor preventing integration within the limits we tested.

In a similar vein, whether separate distributions are perceived to be generated from the same source can also influence how they are treated[4]. In Bayesian cue combination terms, the visual system will integrate two sources of information, if they are treated as two independent estimates of the same visual object. This can explain our results and it seems likely that the visual system would only integrate information if it facilitates efficient visual processing. For example, serial dependence is considered an adaptive strategy to deal with noise and uncertainty in a world where natural scenes are generally stable and temporally continuous (Kiyonaga, Scimeca, Bliss, & Whitney, 2017). But temporal integration of perceptual information may not be advantageous when the visual world changes from moment to moment. Taubert, Alais, and Burr (2016) observed that serial dependence only occurred for facial features that are generally stable overtime (e.g., gender or

[4] For example, if you are judging the ripeness of berries on a branch, you would base your judgment on the color distribution of the visible berries on that branch. However, in order to make a more accurate judgment, you can glance at the branch from a different location, so the color distribution your visual system extracts will change since different berries will become visible to you. Your visual system would integrate the two color distributions for a better estimate of the ripeness of the berries only if the visual system knows that the two color distributions were generated from the same branch.

identity), but not for more changeable features (e.g., facial expression). Speculatively, our results suggest a similar effect of temporal stability on priming.

Temporal integration of visual feature distributions across successive visual search displays has been observed before using the FDL method – for example a bimodal distribution learned for 2–3 trials is encoded as a uniform distribution with no drop between the two peaks but after 8–11 trials observers learn the bimodality of the distribution (Chetverikov et al., 2017b). When there is no temporal uncertainty within a learning streak, the visual system seems to be able to integrate feature distributions that are even much more complex than the ones we have used in this study. The only methodological difference in our study was the alternation in distractor distributions within the learning trials. After observing, in Experiment 1, the negative influence of this temporal uncertainty within a streak, in our second and third experiments we decided to decrease this uncertainty within a streak by increasing the overlap between the two distractors. The magnitudes of Bayes factors obtained in each individual experiment were low, but the observed effects were all in the same direction. Furthermore, when we combined the data from all three experiments, accounting for the between-study variability, we saw more conclusive evidence. The results from this mini-meta-analysis should be interpreted cautiously since it includes data from three different experiments in which the orientation difference between the distractor distributions were different. This can add heterogeneity to the aggregate data, which could, in principle, bias the results in favor of null models. Nevertheless, all things considered, these results still provide a very strong contrast with previous FDL studies where significant temporal integration was easily observed. The most straightforward conclusion is that the visual system does not integrate information across trials in cases where the visual input exhibits temporal unreliability, which induces a strong reliance on the most recent input.

Our results suggest that there are strict limits on feature distribution learning and that unreliable variance strongly affects the learning. We have previously shown limits on what can be learned through FDL methods (Chetverikov et al., 2017c) where a certain minimum amount of input (in this case set-size) determines whether learning occurs. Similarly, a certain amount of continuity and reliability could be another prerequisite for detailed encoding of feature distributions during sequential presentation (e.g., Chetverikov et al., 2016; Chetverikov et al., 2017a; Chetverikov et al., 2017b; Chetverikov et al., 2017c; Chetverikov et al., 2020; Hansmann-Roth et al., 2019).

However, apart from indicating what sort of learning the FDL paradigm can achieve, our results may also reveal how the visual system may use the encoded information. Sensory history can be used by the visual system to adjust and calibrate itself to uncertain and dynamic environments. For example, passive exposure to visual features can increase efficiency in search for outliers. (Kompaniez-Dunigan, Abbey, Boone, & Webster, 2015; McDermott, Malkoc, Mulligan, & Webster, 2010; Wissig, Patterson, & Kohn, 2013). However, inter-trial priming (Kristjánsson & Campana, 2010; Maljkovic & Nakayama, 1994), depends on the role assigned to the features (target vs. distractors). This indicates a more active interaction between the visual system and the encoded visual features with respect to their roles given a visual task. Moreover, the detailed feature distributions encoded during learning streaks in FDL studies determine which target observers pick when the search display includes two targets instead of one (Chetverikov, Campana, & Kristjánsson, 2020). In other words, previous studies on priming and role reversals (for a review, see Kristjánsson & Campana, 2010) and on FDL paradigm (for a review see Chetverikov et al. 2019) strongly suggest that encoded feature distributions are utilized by the visual system. Here, we tested with FDL methodology whether detailed probabilistic feature distributions can be integrated when two different orientation distributions are presented in an alternating order. Our results showed that the visual system relied heavily on the distractor distribution on the last trial with this current method, indicating that the temporal integration observed in previous FDL studies is limited to more reliable and temporally stable sensory input.

8. Conclusions

Our results reveal limits to what can be learned about distractor distributions during sequential presentation of search arrays. The visual system clearly prefers more reliable input than was available in our study. We speculate that the mechanisms revealed in previous FDL studies help with learning environmental statistics in stable environments and trial-by-trial variability can interfere with the learning, suggesting that the visual system generally assumes that the statistics within natural environments are stable.

CRediT authorship contribution statement

Ömer Dağlar Tanrıkulu: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Visualization, Project administration. **Andrey Chetverikov:** Conceptualization, Methodology, Formal analysis, Writing - review & editing, Funding acquisition. **Arni Kristjánsson:** Conceptualization, Resources, Writing - review & editing, Supervision, Funding acquisition.

Acknowledgments

We thank Sabrina Hansmann-Roth for valuable feedback on the data analysis and the interpretation of the data. ODT and AK were supported by grant IRF #173947-052 from the Icelandic Research Fund. AC is supported by Radboud Excellence Fellowship.

References

Albrecht, A. R., & Scholl, B. J. (2010). Perceptually averaging in a continuous visual world: Extracting statistical summary representations over time. *Psychological Science, 21*(4), 560–567.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences, 15*(3), 122–131.

Atchley, P., & Andersen, G. J. (1995). Discrimination of speed distributions: Sensitivity to statistical properties. *Vision Research, 35*(22), 3131–3144. https://doi.org/10.1016/0042-6989(95)00057-7.

Banks, M. S., & Backus, B. T. (1998). Extra-retinal and perspective cues cause the small range of the induces effect. *Vision Research, 38*(2), 187–194.

Becker, S. I. (2008). The stage of priming: Are intertrial repetition effects attentional or decisional? *Vision Research, 48*(5), 664–684.

Blake, A., Bülthoff, H. H., & Sheinberg, D. (1993). Shape from texture: Ideal observers and human psychophysics. *Vision Research, 33*(12), 1723–1737.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436.

Chetverikov, A., Campana, G., & Kristjánsson, Á. (2016). Building ensemble representations: How the shape of preceding distractor distributions affects visual search. *Cognition, 153*, 196–210.

Chetverikov, A., Campana, G., & Kristjánsson, Á. (2017d). Learning features in a complex and changing environment: A distribution-based framework for visual attention and vision in general. In *Progress in brain research* (pp. 97–120). Elsevier.

Chetverikov, A., Campana, G., & Kristjánsson, Á. (2017a). Representing color ensembles. *Psychological Science, 28*(10), 1510–1517.

Chetverikov, A., Campana, G., & Kristjánsson, Á. (2017b). Rapid learning of visual ensembles. *Journal of Vision, 17*(2), 21. https://doi.org/10.1167/17.2.21.

Chetverikov, A., Campana, G., & Kristjánsson, Á. (2017c). Set size manipulations reveal the boundary conditions of perceptual ensemble learning. *Vision Research, 140*, 144–156.

Chetverikov, A., Campana, G., & Kristjánsson, Á. (2020). Probabilistic rejection templates in visual working memory. *Cognition, 196*, 104075. https://doi.org/10.1016/j.cognition.2019.104075.

Chetverikov A., Hansmann-Roth S., Tanrıkulu Ö.D., Kristjánsson Á. (2019) Feature Distribution Learning (FDL): A new method for studying visual ensembles perception with priming of attention shifts. In: Neuromethods. Humana Press.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research, 43*(4), 393–404.

Corbett, J. E., & Melcher, D. (2014). Stable statistical representations facilitate visual search. *Journal of Experimental Psychology: Human Perception and Performance, 40*(5), 1915–1925.

Crawford, L. E., Corbin, J. C., & Landy, D. (2019). Prior experience informs ensemble encoding. *Psychonomic Bulletin & Review, 26*(3), 993–1000. https://doi.org/10.3758/s13423-018-1542-6.

Dakin, S. C., & Watt, R. J. (1997). The computation of orientation statistics from visual texture. *Vision Research, 37*(22), 3181–3192. https://doi.org/10.1016/S0042-6989(97)00133-8.

Dakin, S. C. (2015). Seeing statistical regularities: Texture and pattern perception. In J. Wagemans (Ed.), The oxford handbook of perceptual organization (pp. 150–167).

Dosher, B., & Lu, Z.-L. (2017). Visual perceptual learning and models. *Annual Review of Vision Science, 3*(1), 343–363.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*(6870), 429–433.

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience, 17*(5), 738–743.

Florey, J., Dakin, S. C., & Mareschal, I. (2017). Comparing averaging limits for social cues over space and time. *Journal of Vision, 17*(9), 17–17.

Fouriezos, G., Rubenfeld, S., & Capstick, G. (2008). Visual statistical decisions. *Perception & Psychophysics, 70*(3), 456–464.

Goolsby, B. A., & Suzuki, S. (2001). Understanding priming of color-singleton search: Roles of attention at encoding and "retrieval". *Perception & Psychophysics, 63*(6), 929–944.

Gorea, A., Belkoura, S., & Solomon, J. A. (2014). Summary statistics for size over space and time. *Journal of Vision, 14*(9), 22–22.

Haberman, J., Harp, T., & Whitney, D. (2009). Averaging facial expression over time. *Journal of Vision, 9*(11), 1.

Haberman, J., Lee, P., & Whitney, D. (2015). Mixed emotions: Sensitivity to facial variance in a crowd of faces. *Journal of Vision, 15*(4), 16–16.

Haberman, J., & Whitney, D. (2012). Ensemble perception: Summarizing the scene and broadening the limits of visual processing. In J. Wolfe, & L. Robertson (Eds.), *From perception to consciousness: Searching with Anne Treisman* (pp. 339–349). Oxford: Oxford University Press.

Hansmann-Roth, S., Chetverikov, A., & Kristjánsson, Á. (2019). Representing color and orientation ensembles: Can observers learn multiple feature distributions? *Journal of Vision, 19*(9), 2. https://doi.org/10.1167/19.9.2.

Hansmann-Roth, S., Kristjánsson, Á., Whitney, D. W., & Chetverikov, A. (2021). Dissociating implicit and explicit ensemble representations reveals the limits of visual perception and the richness of behavior. *Scientific Reports, 11*, 3899. https://doi.org/10.1038/s41598-021-83358-y.

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision, 4*(12), 1–1.

Hubert-Wallander, B., & Boynton, G. M. (2015). Not all summary statistics are made equal: Evidence from extracting summaries across time. *Journal of Vision, 15*(4), 5.

Im, H. Y., & Halberda, J. (2013). The effects of sampling and internal noise on the representation of ensemble average size. *Attention, Perception, & Psychophysics, 75*(2), 278–286.

Kiyonaga, A., Scimeca, J. M., Bliss, D. P., & Whitney, D. (2017). Serial dependence across perception, attention, and memory. *Trends in Cognitive Sciences, 21*(7), 493–497.

Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). What's new in Psychtoolbox-3? 785. *Perception, 36*(ECVP Supplement 14).

Knill, D. C., & Richards, W. (Eds.). (1996). *Perception as Bayesian Inference*. Cambridge University Press.

Knill D.C. (2007) Bayesian models of sensory cue integration. In: Kenji Doya, Shin Ishii, Alexandre Pouget, Rajesh P. N. Rao, Bayesian brain: probabilistic approach to neural coding, The MIT Press, Cambridge, 189-206.

Kohn, A. (2007). Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology, 97*(5), 3155–3164.

Kompaniez-Dunigan, E., Abbey, C. K., Boone, J. M., & Webster, M. A. (2015). Adaptation and visual search in mammographic images. *Attention, Perception, & Psychophysics, 77*(4), 1081–1087.

Körding, K. P., & Wolpert, D. M. (2006). Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences, 10*(7), 319–326.

Kristjánsson, Á., & Ásgeirsson, Á. G. (2019). Attentional priming: Recent insights and current controversies. *Current Opinion in Psychology, 29*, 71–75.

Kristjánsson, Á., & Campana, G. (2010). Where perception meets memory: A review of repetition priming in visual search tasks. *Attention, Perception & Psychophysics, 72*(1), 5–18.

Kristjánsson, Á., & Driver, J. (2008). Priming in visual search: Separating the effects of target repetition, distractor repetition and role-reversal. *Vision Research, 48*(10), 1217–1232.

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.

Larsson, J., & Smith, A. T. (2012). fMRI repetition suppression: Neuronal adaptation or stimulus expectation? *Cerebral Cortex, 22*(3), 567–576.

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

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

McDermott, K. C., Malkoc, G., Mulligan, J. B., & Webster, M. A. (2010). Adaptation and visual salience. *Journal of Vision, 10*(13), 17–17.

Morey, R. D., & Rouder, J. N. (2018). BayesFactor: Computation of Bayes Factors for Common Designs. *R package version*, 0.9.12-4.2. https://CRAN.R-project.org/package=BayesFactor.

Oriet, C., & Hozempa, K. (2016). Incidental statistical summary representation over time. *Journal of Vision, 16*(3), 3.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology., 56*(5), 356–374.

Sasaki, Y., Nanez, J. E., & Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. *Nature Reviews Neuroscience, 11*(1), 53–60.

Shurygina, O., Kristjánsson, Á., Tudge, L., & Chetverikov, A. (2019). Expectations and perceptual priming in a visual search task: Evidence from eye movements and behavior. *Journal of Experimental Psychology: Human Perception and Performance, 45*(4), 489.

Sigurdardottir, H. M., Kristjánsson, Á., & Driver, J. (2008). Repetition streaks increase perceptual sensitivity in brief visual search displays. *Visual Cognition, 16*, 643–658.

Solomon, J. A., Morgan, M., & Chubb, C. (2011). Efficiencies for the statistics of size discrimination. *Journal of Vision, 11*(12), 13–13.

Tanrıkulu, Ö. D., Chetverikov, A., & Kristjánsson, Á. (2020). Encoding perceptual ensembles during visual search in peripheral vision. *Journal of Vision, 20*(8). https://doi.org/10.1167/jov.20.8.20.

Tanrıkulu, Ö. D., Hansmann-Roth, S., Chetverikov, A., & Kristjánsson, Á. (2020). What kind of empirical evidence is needed for probabilistic mental representations? An example from visual perception. Manuscript submitted for publication; Preprint available: 10.31234/osf.io/4ap3f.

Taubert, J., Alais, D., & Burr, D. (2016). Different coding strategies for the perception of stable and changeable facial attributes. *Scientific Reports, 6*, 32239.

Tower-Richardi, S. M., Leber, A. B., & Golomb, J. D. (2016). Spatial priming in ecologically relevant reference frames. *Attention, Perception, & Psychophysics, 78*(1), 114–132.

Utochkin, I. S., & Yurevich, M. A. (2016). Similarity and heterogeneity effects in visual search are mediated by "segmentability". *Journal of Experimental Psychology: Human Perception and Performance., 42*(7), 995–1007.

Walther, C., Schweinberger, S. R., Kaiser, D., & Kovács, G. (2013). Neural correlates of priming and adaptation in familiar face perception. *Cortex, 49*(7), 1963–1977.

Wang, D., Kristjansson, A., & Nakayama, K. (2005). Efficient visual search without top-down or bottom-up guidance. *Perception & Psychophysics, 67*(2), 239–253.

Webster, M. A. (2015). Visual adaptation. *Annual review of vision science, 1*(1), 547–567.

Whiting, B. F., & Oriet, C. (2011). Rapid averaging? Not so fast! *Psychonomic Bulletin & Review, 18*(3), 484–489.

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology, 69*(1), 105–129.

Wissig, S. C., Patterson, C. A., & Kohn, A. (2013). Adaptation improves performance on a visual search task. *Journal of vision, 13*(2), 6. https://doi.org/10.1167/13.2.6.