Vision Research 140 (2017) 144–156

Contents lists available at ScienceDirect

Vision Research

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

Set size manipulations reveal the boundary conditions of perceptual ensemble learning

Andrey Chetverikov[a,b,c,*], Gianluca Campana[d,e], Árni Kristjánsson[a]

[a] Laboratory for Visual Perception and Visuomotor Control, Faculty of Psychology, School of Health Sciences, University of Iceland, Reykjavik, Iceland
[b] Cognitive Research Lab, Russian Academy of National Economy and Public Administration, Moscow, Russia
[c] Department of Psychology, Saint Petersburg State University, Saint Petersburg, Russia
[d] Dipartimento di Psicologia Generale, Università degli Studi di Padova, Padova, Italy
[e] Human Inspired Technology Research Centre, Università degli Studi di Padova, Padova, Italy

ARTICLE INFO

Keywords:
Feature distributions
Summary statistics
Ensemble perception
Visual search
Priming of pop-out

ABSTRACT

Recent evidence suggests that observers can grasp patterns of feature variations in the environment with surprising efficiency. During visual search tasks where all distractors are randomly drawn from a certain distribution rather than all being homogeneous, observers are capable of learning highly complex statistical properties of distractor sets. After only a few trials (learning phase), the statistical properties of distributions - mean, variance and crucially, shape - can be learned, and these representations affect search during a subsequent test phase (Chetverikov, Campana, & Kristjánsson, 2016). To assess the limits of such distribution learning, we varied the information available to observers about the underlying distractor distributions by manipulating set size during the learning phase in two experiments. We found that robust distribution learning only occurred for large set sizes. We also used set size to assess whether the learning of distribution properties makes search more efficient. The results reveal how a certain minimum of information is required for learning to occur, thereby delineating the boundary conditions of learning of statistical variation in the environment. However, the benefits of distribution learning for search efficiency remain unclear.

How do observers represent the variation in the environment such as the colors in a moss-covered lava field or the brightness distribution in snow covered landscapes? Although we may think of moss as "green" and snow as "white", we clearly perceive more than a single feature value. On the other hand, encoding every feature at every location along with their conjunctions will require a lot of resources. The question is then how feature variation in the external world is translated into a representation, and the answer will likely be somewhere between the two extremes outlined above. Processing of such heterogeneous perceptual ensembles has been studied with texture segregation tasks (Julesz, 1981) but natural sets are typically not as regular as those studied by Julesz. Take color – color variation in natural environments is rarely uniform – and neither are the oriented edges available in natural statistical distributions. There is accumulating evidence that human observers can extract summary statistics such as the mean and standard deviation of a number of features, such as color, size, orientation and brightness, from stimulus sets having a certain variability (Alvarez, 2011; Ariely, 2001; Corbett & Melcher, 2014; Haberman & Whitney, 2012; Michael, de Gardelle, & Summerfield, 2014; Rosenholtz, Huang, Raj, Balas, & Ilie, 2012; Utochkin, 2015). Summary statistics provide a concise way of representing feature variation but they are still relatively coarse because two different ensembles might have the same statistics while coming from different distributions.

Our recent experiments have revealed that observers can represent more intricate feature variation than studies of simple statistical parameters have suggested. Chetverikov, Campana, and Kristjánsson (2017b) showed that after only a few trials observers can learn the properties of feature distributions of colored distractors in an odd-one-out visual search task over and above the mean and standard deviations, and in Chetverikov, Campana, and Kristjánsson (2016, 2017a), we found similar results for orientation. In those studies, assessed observers' representations by measuring their implicit expectations of upcoming stimulus distributions with response times (RTs) instead of explicit judgements of distribution properties. Namely, we measured effects of 'role-reversals' between targets and distractors on visual search performance (Kristjánsson & Driver, 2008). A role-reversal occurs when a target on a preceding trial becomes a distractor on the next trial, or vice-versa, which typically slows search (Becker, 2010). This effect is not limited to specific feature domains and seems to reflect

* Corresponding author at: Department of Psychology, University of Iceland, Saemundargata 2, 101 Reykjavík, Iceland.
E-mail address: andrey@hi.is (A. Chetverikov).

http://dx.doi.org/10.1016/j.visres.2017.08.003
Received 12 April 2017; Received in revised form 11 August 2017; Accepted 20 August 2017
0042-6989/

encoding of distractors in implicit short-term visual working memory (Carlisle & Kristjánsson, 2017; Lamy, Antebi, Aviani, & Carmel, 2008; Maljkovic & Nakayama, 1994). In a typical role-reversal study, the distractors are homogeneous. For example, in a color search observers would look for a red target among green distractors. After a few trials with repeated distractor colors observers encode the distractor features, and when the targets become green (among distractors of some other color), search is slowed. The key difference in our manipulation relative to previous studies was that distractors were heterogeneous and on a single trial their features were randomly drawn from a specific probability distribution. The distractors, in other words, formed a perceptual ensemble. Continuing with the example above, instead of a red target among green distractors, observers had to search for a red target among distractors of varying degrees of "greenness", akin to searching for a red berry within moss patches. Then, as these conditions were repeated for a few trials, a role-reversal to a greenish target resulted in slowed search. Importantly, the degree of slowing depended on the correspondence between target hue and the probability of that particular hue among previous distractors. This allowed us to assess observers' representations of ensembles encoded on previous search trials.

We assumed, in other words, that if a target falls within observers' representations of preceding distractor distributions it would cause role-reversal effects, that is, search would be slower because the features of the odd-one-out target would clash with representations of distractor distributions from previous trials (Chetverikov et al., 2016, 2017a,b). Using targets corresponding to different parts of previously learned distractor distributions allows us to infer the probabilistic representation of that distribution by assessing how much search is slowed. For example, following several odd-one-out search trials in the orientation domain with distractors drawn from a truncated Gaussian distribution with an orientation $\mu = 45^\circ$ and $\sigma = 15^\circ$ (range restricted to $45 \pm 30^\circ$), observers respond more slowly when a 45° odd-one-out target suddenly appears than when a 40° target appears, which, in turn, will yield longer response times than a 35° target, and search will be fastest for targets that fall outside the range of the previous distractor distribution. The search RTs will therefore be slow if observers expected this orientation to be from the distractor distribution of immediately preceding trials. The degree of slowing reflects encoded feature probability. By repeating blocks of learning and test trials with different test targets, we were able to "probe" observers' representations of feature distributions along the whole range of possible feature values and obtain detailed *continuous* estimates of these representations.

Importantly, we also found that even when two distributions have the same range or variance, observers' representations differ (Chetverikov et al., 2016). So in contrast to a Gaussian distribution, following learning of a uniform distribution with the same 45° mean and $\pm 30^\circ$ range, response times (RTs) for any target within $45 \pm 30^\circ$ degrees will be approximately the same. That is, even the shapes of the distributions (e.g., whether they are Gaussian, uniform, skewed or even bimodal) are encoded (Chetverikov et al., 2016, 2017a). Differences in the estimates for differently shaped distributions suggest that the precision of ensemble perception is much higher than was thought before.

1. Mechanisms of ensemble perception

How do observers obtain such precise ensemble representations from the stimuli presented on the screen? Recent studies involving explicit summary statistic judgements indicate that the aggregation is limited by the number of stimulus subsets rather than the number of stimuli within a subset (Attarha, Moore, & Vecera, 2014; Im & Halberda, 2013; Maule & Franklin, 2015; Utochkin & Tiurina, 2014; Utochkin & Yurevich, 2016). But the exact mechanisms of aggregation *within* subsets remain controversial. Several studies support the idea of limited sampling (Maule & Franklin, 2016; Myczek & Simons, 2008; Solomon, May, & Tyler, 2016; Tibber et al., 2015) with the number of sampled stimuli being below four. That is, observers can respond accurately when asked about summary statistics even if they analyse only a few exemplars from the stimulus set. Others have argued against this, however (Attarha & Moore, 2015; Attarha et al., 2014; Dakin, 2001; Im & Halberda, 2013; Tokita, Ueda, & Ishiguchi, 2016; Utochkin & Tiurina, 2014). Moreover, approximations involved in *explicit* averaging may differ from tasks where the use of statistics is not explicitly required but might nevertheless be useful or even necessary. Such tasks may include visual search (Rosenholtz et al., 2012), visual categorization (Utochkin, 2015), attentional selection (Im, Park, & Chong, 2015), or texture perception (Dakin, 2015). In particular, distribution learning in visual search (Chetverikov et al. (2016) is not required by the task and therefore allows the study of mechanisms involved in incidental use of summary statistics.

The use of explicit judgments about the properties of feature distributions in previous studies limits our understanding of the mechanisms leading to ensemble representations. It is possible that potential bottlenecks on the precision of such explicit judgments have little to do with distribution representations *per se*. There are a number of ways in which even if observers have highly precise representations of distributions, explicit judgments will still rely on only a few samples. For example, observers might use their representation to generate a limited sample for explicit judgements. That is, when asked to judge the mean, observers might simply sample the distributions they saw. Another option is that, during an averaging task, observers' might try to hold in working memory only the stimuli useful for the averaging they are asked to perform. Using tasks with incidental encoding would be helpful to understand whether limitations found in some studies for explicit averaging are related to ensemble encoding or simply reflect the use of explicit judgements.

Regardless of the mechanisms underlying explicit averaging, incidental distribution encoding within the present paradigm is of interest by itself. Previous results indicate that distribution representations that observers use in visual search are more precise than, for example, those that can be derived from forced-choice judgements (see review in Chetverikov et al., 2016). How this higher precision is obtained is an interesting question in and of itself, one we investigate here.

Our previous results indicate that distribution learning in visual search can occur rapidly (Chetverikov et al., 2017a). Sometimes only two trials seem to be needed to learn simpler distributions, while learning a more complex (bimodal) distribution required a larger number of search trials and involved a gradual change from a unimodal to a bimodal representation. This shows that distribution representations can be based on the accumulation of information coming from multiple samples – otherwise the representation would be the same regardless of trial number. But how many samples are needed from a single display is unknown. For example, on a given trial observers may sample three or four items (Maule & Franklin, 2016; Myczek & Simons, 2008; Solomon et al., 2016) and then integrate the samples from different trials.

Here we used set size manipulations to investigate the limits on processing of simultaneously presented information during feature distribution learning. If learning of distribution parameters is based on a few stimuli sampled from each trial, the learning should be equally efficient with small and large set sizes. On the other hand, if the learning is based on an aggregation (possibly, in parallel) of a large number of stimuli, larger set sizes would result in better learning.

2. Role of set size for search efficiency and inter-trial priming

Set size manipulations have played a key role in theories of visual attention. For easier searches where the target is easily found among distractors, RTs are constant with set size, or can even decrease (Bravo & Nakayama, 1992; Kristjánsson, 2015; Wang, Kristjánsson, & Nakayama, 2005; Wolfe & Horowitz, 2017). Using this classic manipulation may therefore also reveal whether and how distractor distribution learning affects search performance more generally.

Representing distractor distributions accurately may help observers make more optimal perceptual decisions, making search easier (Ma, Shen, Dziugaite, & van den Berg, 2015). Additionally, we expected that more precise estimates of distribution density would be more beneficial for Gaussian than uniform distributions. Chetverikov et al. (2016) found that even with the same range (or standard deviation) of distractor orientations, representations of uniform distributions are narrower for Gaussian than uniform ones. If this is true, search should be particularly easy within Gaussian compared to uniform distractor distributions when observers' estimates of distribution shape reflect the difference between these distributions.

We also assessed traditional target and distractor priming effects with different set sizes. Meeter and Olivers (2006) have argued that priming is weaker when there is less ambiguity in displays. Accordingly, they found that priming was strongest with fewer elements in the display. Rangelov, Müller, and Zehetleitner (2017), Rangelov, Zehetleitner, Muller, and Zehetleitner (2013) found that priming of pop-out is stronger with smaller set sizes that were also less dense. In other studies, however, priming effects were either the same or stronger with larger than smaller set sizes, however (Becker & Ansorge, 2013; Hodsoll, Humphreys, & Braithwaite, 2006; Kristjánsson & Driver, 2008; Wang et al., 2005; Wolfe, Butcher, Lee, & Hyle, 2003) or strong despite large set sizes (e.g., Ásgeirsson & Kristjánsson, 2011). Our aim was to see if set-size induced changes in priming of pop-out would mirror those observed for distribution learning. A positive answer to that question would suggest that the two effects might be governed by similar mechanisms. Interestingly, priming of pop-out in the orientation domain has been argued to mostly depend on distractors, not the target (Lamy, Yashar, & Ruderman, 2013). Given that we reliably found distribution learning in the orientation domain, it is then possible that distractor distribution learning is independent of target learning. However, this conclusion depends on the assumption that with the larger set sizes used in previous studies, target priming is absent as it was in Lamy et al. (2013).

3. Aims and hypotheses

We therefore had three aims, firstly to assess how much information is needed for encoding of perceptual ensembles by manipulating set size. For example is a certain minimum amount of information (or set size) required so that observers treat a display as a distribution in the first place? We assessed the precision of distribution learning by analyzing RTs as a function of the difference in orientation between the target and the mean of distractor distributions from preceding trials. In the absence of any learning, there would be no dependency between search times and target difference with distractors on previous trials; that is, RT curves would be flat. With relatively precise learning, RT curves would correspond to the probability density function (PDF) of previously learned distributions. Intermediate learning precision should result in non-flat RT curves that will be similar between distributions. We expected that if distribution learning is driven by sampling only a few items on each trial, learning should be equally precise with small and large set sizes. In contrast, if distribution learning is based on aggregating information from the whole display, then precision of distribution learning should increase with set size.

Secondly, we used set size to assess how distribution learning affects search efficiency. To this end, we analyzed whether evidence for distribution learning coincides with improved search efficiency. As described above, we expected in particular that RT differences between uniform and Gaussian distributions will increase for larger set sizes. Finally, we measured whether priming effects and distribution learning vary similarly with set size, which could indicate that the two reflect overlapping mechanisms. This was done by comparing priming effects from target or distractor set mean with distribution learning for different set sizes.

4. Experiment 1

4.1. Method

4.1.1. Participants

Eleven observers (seven female, age $M = 26.36$) participated in the experiment. All of them were staff or students at St. Petersburg State University, who participated without additional reward. The study was approved by the local ethics committee and carried out in accordance with the Declaration of Helsinki. The observers signed an informed consent form before participating.

4.1.2. Procedure

We used a task similar to our previous studies (Chetverikov et al., 2016, 2017a). Observers looked for an odd-one-out line among a set of lines differing in orientation. Stimuli were presented on an iiyama ProLite T2250MTS display (21.5″ with 1680 × 1050 resolution) using PsychoPy 1.82.01 (Peirce, 2007, 2009). Each line length was 1.41°. Observers indicated whether the target line was in the upper or the lower half of the screen by pressing the 'i' or 'j' keys on a standard keyboard. Trials were organized in intertwined prime and test 'streaks'. During prime streaks, distractors were randomly drawn from a uniform (range = 60 deg.) or a truncated Gaussian distribution (SD = 15 deg., range = 60 deg.). The distribution mean was the same within streak but chosen randomly between streaks. Both distributions had the same range on each trial, ensured by adding two distractors with orientation equal to the minimal and maximal values given the mean and range. Target orientation was selected randomly on each trial with the restriction that the distance between target orientation and distractor mean in feature space was 60 degrees at minimum. Within test streaks, distractor orientations were randomly drawn from a truncated Gaussian with SD = 10 deg. and range 20 deg. Each test trial had the same target orientation within a streak, while distractor mean was chosen randomly with a distance to target no less than 60 deg (as on prime trials).

The main difference from our previous experiments (Chetverikov et al., 2016, 2017a) was that set size during prime streaks (prime SS) varied between streaks (was constant within them). Four set sizes were used: 8, 14, 24, or 36 lines. The two smaller set sizes were fit into a 4 by 4 matrix, while the larger ones were fit into a 6 by 6 matrix, both with 3.2° cell size (Fig. 1). Line position within cells was jittered randomly by ± 0.5° on both coordinates. On test streaks, set size (test SS) was always 36. The mean distance between lines was constant.

Observers participated in three sessions of approximately 1040 trials. Each session had 208 prime and test streaks (4 set sizes * 2 prime distribution types * 26 repetitions). The first session was discarded, as participants had difficulties searching within the smallest set size: mean *RT* was very high ($M = 994$ ms), some of the participants had exceptionally high mean RTs (e.g., for one of them $M = 1727$ ms), while others had exceptionally low accuracy (M = 0.62 for one participant, M = 0.65 for another).

4.1.3. Data analyses

RTs as a function of set size and distribution type were analyzed using ANOVA or linear mixed effects regression (LMER; Bates, Mächler, Bolker, & Walker, 2015) where appropriate. RTs were log-transformed, errors were removed from RT analyses. Distribution learning was assessed by fitting models first to the overall data and then to data from individual observers (see Chetverikov et al., 2017b). The data from Experiments 1 and 2 is available at https://osf.io/3eunr/?view_only=f591c72e57b6417ea1b0d2aa02f411b2.

5. Results

5.1. Set size effects on prime trials

Fig. 2 shows that search for targets within the uniform distractor

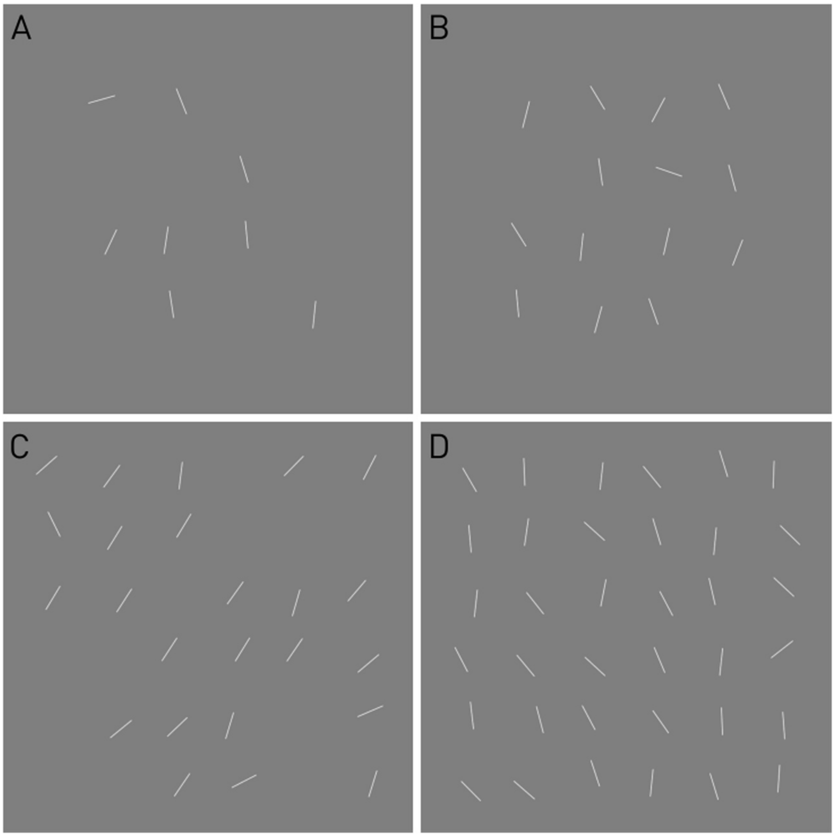

Fig. 1. Example stimuli with different set sizes, 8–36 lines from upper left to lower right.

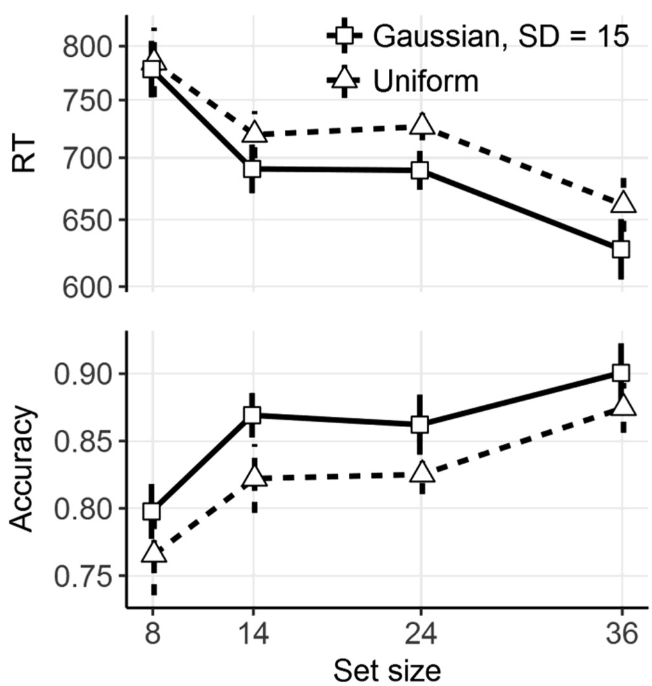

Fig. 2. Effects of set size on RT's during prime streaks for the different distribution types.

distribution was more difficult than within a Gaussian distribution with $SD = 15$. A repeated-measures ANOVA revealed main effects of set size ($F(3,\ 30) = 34.79$, $p < .001$, $\eta^2_G = 0.16$ for RT, $F(3,\ 30) = 27.95$, $p < .001$, $\eta^2_G = 0.24$ for accuracy) and distribution type ($F(1,\ 10) = 24.66$, $p < .001$, $\eta^2_G = 0.02$ for RT, $F(1, 10) = 29.93$, $p < .001$, $\eta^2_G = 0.07$ for accuracy), but no interaction. Note that ANOVAs ignore the fact that set size is a numeric variable. Analyses treating set size as numeric with LMER demonstrated a significant interaction for RT ($t(13418) = -2.72$, $p = .006$). Post-hoc comparisons showed that search was faster within a Gaussian distribution compared to a uniform distribution for all set size levels except for the smallest one ($B = -0.01$ (0.01), $p = .470$, for the rest of the comparisons $p < .01$). Consecutive comparisons (comparing set size 14 with 8, 24 with 14, and 36 with 24) separately for each distribution type showed that increasing set size decreased response time ($p < .001$) except for the comparison of set size 14 with 24 ($p > .9$).

5.2. Set size effects on test trials

Fig. 3 shows that set size during prime streaks affected search performance on test trials (*RT* increased, $F(3,\ 30) = 2.51$, $p = .113$, $\eta^2_G = 0.00$, and accuracy decreased, $F(3,\ 30) = 3.29$, $p = .046$, $\eta^2_G = 0.04$), but previous distribution type had no effect on test trial performance ($p > .12$). Post-hoc contrasts indicated that RTs were longer ($t(30) = 2.70$, $p = .031$) and accuracy was lower ($t(30) = -2.73$, $p = .029$) following the largest set size compared to the smallest one. No other contrasts were significant.

5.3. Priming of pop-out

We next analyzed priming effects (Kristjánsson & Campana, 2010) by the difference (in absolute degrees) between target orientations and between distractor mean orientations. We first analyzed priming for switches from prime to test streaks, because during switches between streaks both target and distractors changed, providing an opportunity to independently analyze their effects. The effects were analyzed with LMER by entering both target-to-target and distractor-mean-to-distractor-mean differences simultaneously into the regression model (the random effects for these variables grouped by participant were also added, controlling for between-subject variability). Fig. 4 shows the resulting regression coefficients showing the amount of slowing per

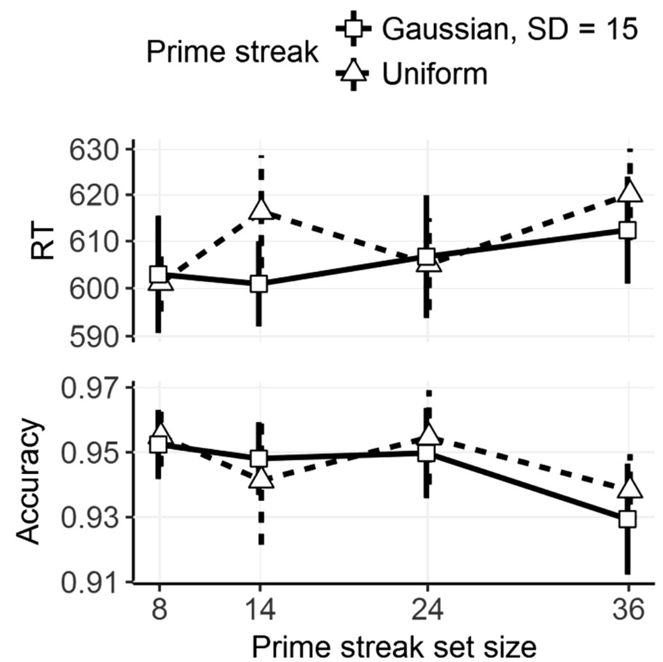

Fig. 3. Search performance on test trials as a function of prime streak set size and distribution type.

change of 1 degree in either target or distractor orientations. LMER indicated that with small SS neither the differences in target orientation between prime and test trials, nor differences in distractor means, had any effect ($t < 0.5$[1] with SS = 8 and $t < 1.5$ with SS = 14, Fig. 4, right panel). In contrast, with SS = 24 there was significant distractor priming ($t = 3.4$) but not target priming ($t = 1.7$), while for the largest SS both were present ($t = 2.0$ and $t = 4.9$ for distractor and target priming, respectively). Analyses of interactions further confirmed that target priming became more pronounced with increasing SS ($t = 2.7$) of preceding prime streak while distractor priming also increased but not quite significantly ($t = 1.9$).

We then did the reverse test, by analyzing RTs on the first trials during prime streaks as a function of the differences in targets and distractors with preceding test streaks (Fig. 4, left panel). The effect of distractor-mean-to-distractor-mean difference was significant for all tests while the effect of target orientation shift was not significant with the smallest SS ($t = -0.2$), close to significance with SS = 14 ($t = 1.9$) and significant for larger SS ($t > 2$). Analyses of interactions showed that priming from target increased with prime streak SS ($t = 2.4$) while distractor priming decreased ($t = -2.1$).

Finally, analyses of target differences within prime streaks (where the target varied between trials, but distractor mean stayed the same) showed that target priming increased with SS ($B = 0.0007$, 0.0008, 0.0009, 0.0010 with $t = 1.3$, 1.7, 2.1, and 2.6 for set sizes 8, 14, 24, and 36, respectively). Analyses of priming effects within test streaks were not run because test streak had constant SS.

5.4. Distribution learning

Fig.5 shows non-linear smoothed curves for RT as a function of the distance between targets on test trials and the means of preceding prime trial distractor distributions (CT-PD, current target – previous distractor), that, in other words, reflect role-reversals between targets and distractors (Kristjánsson & Driver, 2008; Becker, 2010). The distance was analyzed in absolute degrees as prime distributions were symmetric.

To test whether observers learned distractor distributions (that is: responded according to the learned probabilities) we fit the RT data from test trials using models based on actual distractor distributions (half-Gaussian, uniform, uniform with decrease) against simple linear and null models (see more formal description of the models in Chetverikov et al., 2017b. The null model assumed constant RT. The linear model assumes simple linear dependency of RT on CT-PD. The half-Gaussian model assumed that RTs are described by a half-Gaussian function with SD = 15. The uniform model assumed constant RT both within and outside the range of 30 degrees, but with different baselines for the two ranges. Finally, the uniform-with-decrease model assumed constant RT within a 30 degree range and linearly decreasing RTs outside this range. Fit quality was assessed with Bayesian Information Criteria (BIC). We further denote the comparisons between BIC models as ΔBIC, with ΔBIC_{NULL} showing the difference from the null model. $\Delta BIC > 2$ provide "positive" evidence, while $\Delta BIC > 6$ provide "strong" and $\Delta BIC > 10$ "very strong" evidence (Kass & Raftery, 1995).

For the combined data from all participants, the null model

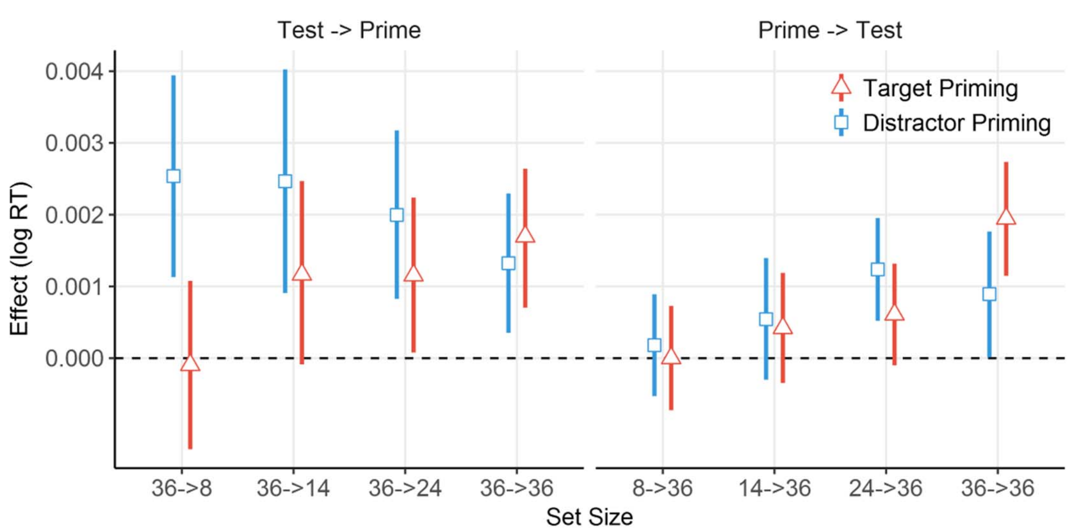

Fig. 4. Estimates of target and distractor priming effects in Experiment 1 on prime (left) and test (right) trials. Only the first trials after switch were analyzed, that is, for prime streaks the first trials after the switch from test streaks and vice versa. Vertical lines show 95% CI. If the effect is above zero (dashed line), the decrease in target-to-target similarity (target priming) or distractor-mean-to-distractor-mean similarity (distractor priming) slowed down the search.

[1] For the sake of brevity we do not report regression coefficients here; they are shown in Fig. 4.

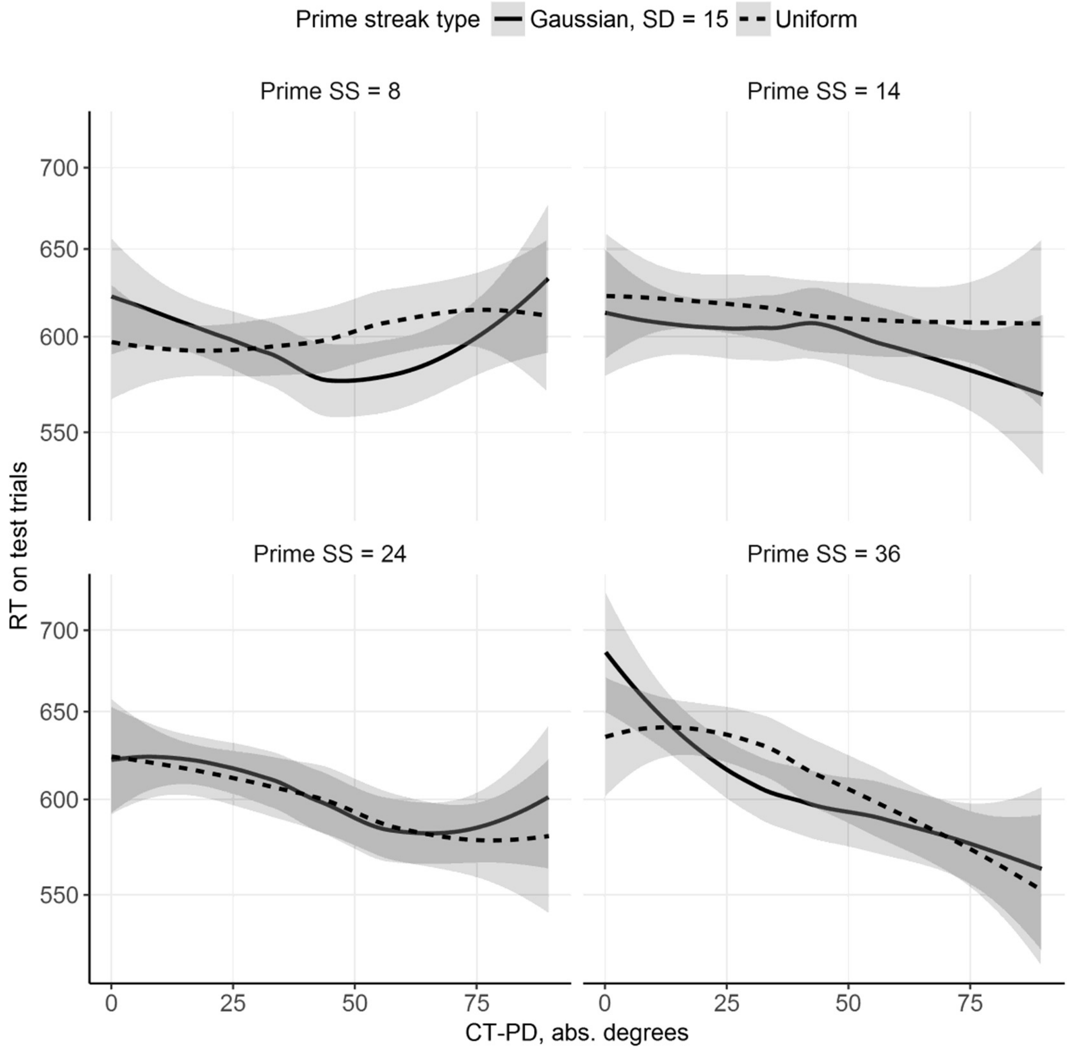

Fig. 5. Search times as function of the distance between target orientation on test streaks and the mean of preceding prime streak distractor distribution, separately for each prime SS.

Table 1
Model fits.

Prime SS	ΔBIC against null model				Best model	Model	2nd best
	Linear	HG	Uniform	UwD			ΔBIC
Prime Distribution: Gaussian, SD = 15							
8	−6.09	−3.12	−2.52	−6.69	Null	Uniform	2.52
14	−3.73	−5.60	−6.03	−3.56	Null	UwD	3.56
24	1.52	−1.82	4.12	1.50	Uniform**	Linear	2.59
36	17.01	18.44	6.95	9.43	HG*	Linear	1.43
Prime Distribution: Uniform							
8	−4.37	−6.70	−4.96	−3.95	Null	UwD	3.95
14	−5.92	−6.15	−5.38	−6.00	Null	Uniform	5.38
24	2.20	−0.58	−2.08	1.44	Linear*	UwD	0.76
36	14.86	2.33	9.69	17.53	UwD**	Linear	2.67

Note: HG – half-Gaussian, UwD – uniform with decrease. ΔBIC > 2 provide "positive" evidence, while ΔBIC > 6 provide "strong" and ΔBIC > 10 "very strong" evidence.
* Best model is better than the null model.
** Best model is better than the null and the 2nd best model.

provided the best fit following the two smallest set sizes independently of the distribution used (Table 1).

For the set size of 24, the uniform model was better than the null following a Gaussian distribution (ΔBIC_{NULL} = 4.12), while following the uniform distribution, the linear model was better than the null (ΔBIC_{NULL} = 2.20), although close to the uniform with decrease model (ΔBIC = 0.76). Note that the non-linear estimates of the RT function were quite similar with that set size (Fig. 5).

Only with the largest set size did a half-Gaussian model provide a good fit (Table 1; see also Fig. S1) that was significantly better than the null model (ΔBIC_{NULL} = 18.44), although close to the linear model (ΔBIC = 1.43). In contrast, but also in line with predictions, following a uniform distribution with the largest set size the best fit was provided by the uniform with decrease model (ΔBIC_{NULL} = 17.53), that was also better than the linear model (ΔBIC = 2.67).

To show that these results are not an artefact of aggregation, we analyzed slopes obtained for each observer with a simple two-segment model with a breaking point at the range of the prime distribution (± 30 deg.). If distributions are learned, RTs should mimic the shape of distribution PDF (i.e., Gaussian or uniform) with negative RT slopes within the range of Gaussian prime distribution and flat slopes within the range of the uniform prime distribution. The critical interaction between prime distribution type and slope type was significant only for prime SS = 36, $F(1, 10) = 6.05$, $p = .034$, $\eta^2_G = 0.12$ (Fig. 6).

6. Discussion

Experiment 1 shows that a large set size (36 lines) is necessary for distribution learning. Only then did we observe the expected difference in the shapes of RT functions on test trials: following Gaussian prime streaks, RT were best fit by half-Gaussian and linear models while

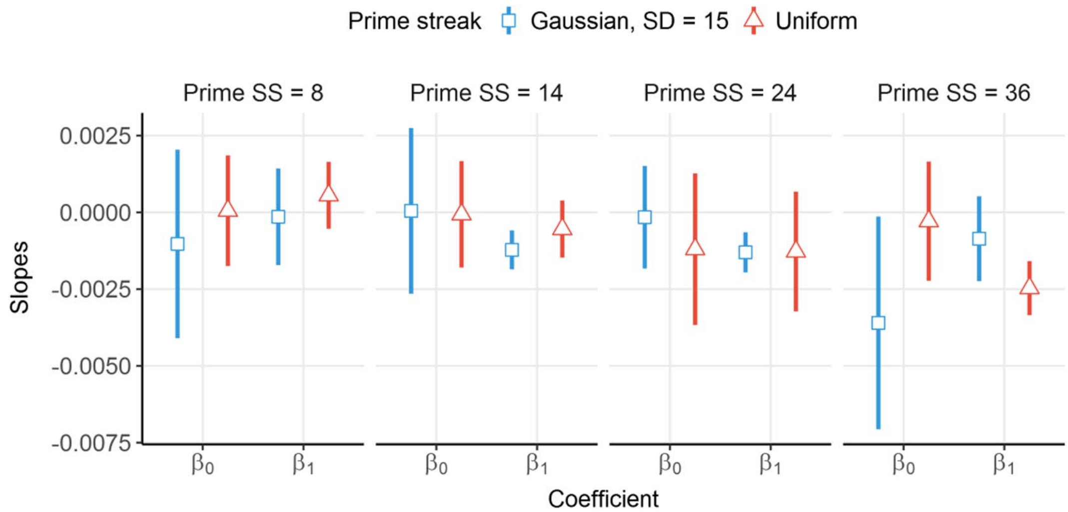

Fig. 6. Individual observers' slopes within (β_0) and outside (β_1) of the prime distribution range in Experiment 1. Lines show 95% CI. Slopes are obtained for log-transformed RTs.

following a uniform distribution the uniform-with-decrease model provided the best fit. For the second to largest set size (24 lines) some learning was observed but the distribution shape was not preserved. Following the two smallest set sizes, the null model provided the best fit for both distributions, indicating no learning of prime distributions. In fact, as Fig. 5 shows, RT functions were mostly flat, showing that RTs were similar regardless of the distance between targets on test trials and distractor mean on prime trials. This shows that the precision of distribution learning increases with set size, from no learning at all with the smallest set sizes through some learning (that does not account for distribution shape) with SS = 24 to accurate learning with SS = 36. Observers seemingly learn information about distribution properties by aggregating information over the whole display rather than sampling only a few display elements.

Our second aim was to assess the relationship between search efficiency and distribution learning. To this end, we measured differences in RTs and accuracy between uniform and Gaussian distributions as a function of set size. We found that set size increased search efficiency. However, distribution learning was only present for the large set sizes (Table 1, Fig.5), but the difference in search efficiency between two distribution types was already visible (and of similar magnitude) with set sizes of 14 and 24 (Fig.2). We therefore cannot conclude that distribution learning improves search.

We also assessed effects of distribution learning on priming of pop-out effects. Target priming became more pronounced with increasing prime streak set size (where the target varied while distractor distribution was stable). Similar results were found for switches between prime and test trials or between test and prime trials. Distractor priming also increased with set size on preceding trials (in analyses of prime to test switches) but decreased with increased set size of the current trial (in analyses of test to prime switches). Note that in this experiment two variables, set size and the difference in set sizes between trials are confounded (this is addressed in Experiment 2, see below). Nevertheless, the data show that both target and distractor priming were present with larger set sizes (it was observed in analyses of test trials only with the largest set size and in analyses of prime streaks only for the two largest set sizes). For the smallest set sizes, we found no priming effects. Overall, the results suggest that priming becomes stronger with larger set size, except for decreasing distractor priming with increased set size of the current trial. Our results therefore seemingly contradict earlier findings suggesting that with lower set sizes or sparser displays, priming effects are more pronounced (Meeter & Olivers, 2006; Rangelov, Müller, et al., 2013; Rangelov, Zehetleitner, et al., 2017) or that only distractor priming is observed within the orientation domain (Lamy et al., 2013), and support previous studies showing similar or stronger priming with larger set sizes (Becker & Ansorge, 2013; Hodsoll et al., 2006; Kristjánsson & Driver, 2008; Wang et al., 2005; Wolfe et al., 2003). The increase in the magnitude of priming effects with set size on preceding trials mirrors increases of precision for distribution learning indicating that the two may involve similar mechanisms.

Our design had some limitations. Firstly, set size varied on prime trials while for test trials set size was always the largest (36). It is possible that switches to larger set sizes decrease effects of previously obtained knowledge. Second, large spacing between elements ("gaps") in stimulus matrices might lead to segmentation of stimuli into subsets with separate statistics. Thirdly, observers may not have been able to utilize distribution learning because the information was too scarce. Low set sizes and only a few repetitions (3–4 trials on prime streaks) result in an overall low number of distractors seen for each streak, and this might not be sufficient for building a representation of distractor distribution. More information could be provided by increasing trial number during learning streaks – and in Experiment 2 we therefore increased trial number during learning streaks.

7. Experiment 2

In Experiment 1 we found only learning of distractor distributions for learning phases with large set sizes. A potential reason is that observers did not receive enough information - that the total number of observed distractors (determined by set size and trial number within prime streaks) was too low for distribution learning. Another possibility is that there is something special about large set sizes, and sparser displays may not contain the critical aspects that induce distribution learning.

To investigate these issues, in Experiment 2 we used two set sizes (16 and 36) both on test and prime trials and used stimulus matrices without "gaps" in both cases. We also increased the number of repetitions on prime trials to increase the amount of information available for observers with lower set sizes. Finally, we increased the sample size to add power to assessments of potential learning effects.

7.1. Method

7.1.1. Participants

Sixteen observers (ten female, age M = 26.88) participated. All were staff or students at St. Petersburg State University, participating without additional reward. The study was approved by the local ethics committee and carried out in accordance with the Declaration of Helsinki. The observers signed informed consent before participating.

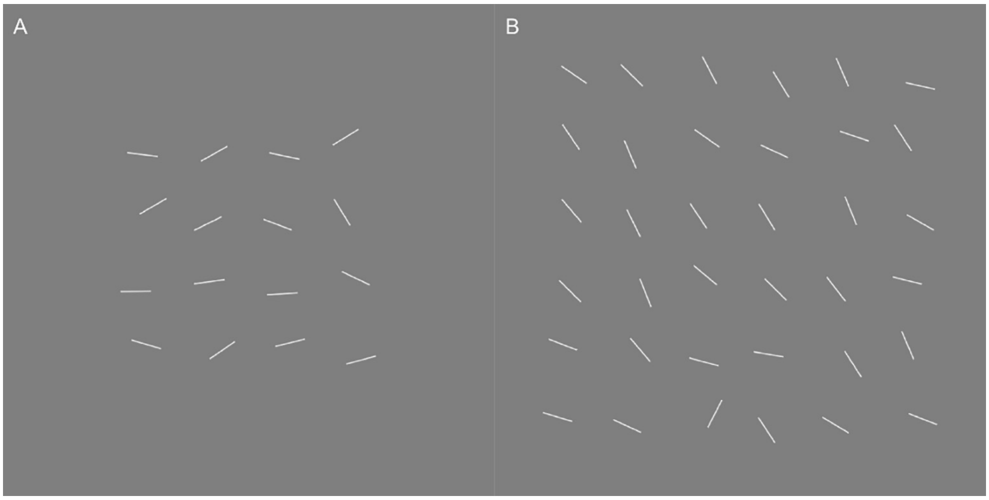

Fig. 7. Example stimuli in Experiment 2 with set sizes 16 (panel A; target is in the third row from bottom, fourth column from the left) and 36 (panel B; target is in the bottom row, third column).

7.1.2. Procedure

Experiment 2 followed the same procedure as Experiment 1 with the following changes. First, only two set sizes were used (16 and 36) and importantly both were used on both prime and test trials. Stimuli on trials with 16 lines were fit into a 4 × 4 matrix and 36 lines were fit into a 6 × 6 matrix with the same spacing as in Experiment 1 (Fig. 7). The prime trials in each block were now 5–6 trials per block (test streaks still had 1 or 2 trials). Participants underwent two sessions with 1460 trials each (≈150 test trials for Prime SS × Test SS combination).

8. Results

8.1. Set size effects on prime trials

A repeated-measures ANOVA revealed main effects of distribution type both for RT ($F(1, 15) = 44.17$, $p < .001$, $\eta^2_G = 0.01$) and accuracy ($F(1, 15) = 36.18$, $p < .001$, $\eta^2_G = 0.07$). Interestingly, the main effect for set size was significant only for accuracy ($F(1, 15) = 71.77$, $p < .001$, $\eta^2_G = 0.14$) but not RT ($F(1, 15) = 0.17$, $p = .682$, $\eta^2_G < 0.01$), which is explained by a significant interaction between the two factors ($F(1, 15) = 6.21$, $p = .025$, $\eta^2_G < 0.01$; the interaction is not significant for accuracy, $F(1, 15) = 0.60$, $p = .451$, $\eta^2_G < 0.01$). As Fig. 8 (left panel) shows, increasing set size decreased RT for the Gaussian but not the uniform distribution. This may reflect some benefit from learning the distribution across trials during the prime streak (see discussion).

8.2. Set size effects on test trials

A repeated-measures ANOVA using prime SS, prime distribution type, and test SS showed that RTs depended on prime SS ($F(1, 15) = 15.21$, $p = .001$, $\eta^2_G < 0.01$) and test SS ($F(1, 15) = 9.78$, $p = .007$, $\eta^2_G < 0.01$). No other factors nor interactions were significant. Accuracy analyses showed significant effects of test SS ($F(1, 15) = 20.51$, $p < .001$, $\eta^2_G = 0.06$) interacting with prime SS ($F(1, 15) = 5.50$, $p = .033$, $\eta^2_G = 0.03$). Post-hoc tests indicated that participants were more accurate when switching from prime SS = 16 to test SS = 36 than when switching to test SS = 16, $M = 0.93$ [0.91, 0.94] vs. $M = 0.96$ [0.95, 0.97], $t(15.0) = -4.32$, $p < .001$. The test distribution following switches from prime SS = 36 did not affect accuracy ($M = 0.93$ [0.92, 0.95] vs. $M = 0.94$ [0.92, 0.96], $t(15.0) = -0.91$, $p = .379$).

8.3. Priming of pop-out

The analyses of priming effects were similar to Experiment 1. Fig. 9 (right panel) shows that on the first trials of test streaks, the magnitude of priming from distractors and targets was similar (all effects significant, $t > 2$, expect for distractor priming with a switch from SS = 16 to SS = 36, which was borderline significant, $t = 1.9$). On prime trials (Fig. 9, left panel), target effects were generally weaker than distractor effects and were significant only with a prime and test of SS = 16. The difference was especially strong with a switch from test SS = 36 (distractor effect $t = 6.0$ and $t = 7.7$ with a switch to SS = 16 and 36, respectively; corresponding target effect t's = 1.3 and 0.4). The only significant interaction, however, reflected the effect of the previous set size on a switch from test to prime streaks ($t = 2.6$).[2]

Within prime streaks (where target varied, but distractors were constant), target priming was present and similar in magnitude regardless of SS ($B = 0.0010$, $t = 4.2$ and $B = 0.0009$, $t = 3.6$ for SS = 16 and 36, respectively). Within test streaks (where distractors varied while targets were constant), distractor priming was stronger with larger SS ($B = 0.0018$, $t = 3.8$ vs. $B = 0.0008$, $t = 1.5$).

8.4. Distribution learning

As in Experiment 1, we fit half-Gaussian, uniform-with-decrease, uniform, linear, and null models to the test streak RTs (Fig. 10) for each combination of prime and test streak set sizes. Table 2 shows that, as in Experiment 1 there was distribution learning following prime streaks of SS = 36 but not SS = 16. The uniform-with-decrease model provided the best fit following the uniform distribution while the linear model was best following the Gaussian distribution but only for prime SS = 36 (see Fig. S2 for a depiction of model fits). For prime SS = 16, uniform-with-decrease and linear models provided similar fits, and for the Gaussian prime streak, neither of them was better than the null model when the data from two test SS were analyzed separately.

We then compared individual observers' slopes within and outside the range of prime distractor distribution (mean ± 30 deg.) with repeated-measures ANOVA. Within the prime distribution, there was a significant interaction between prime SS and prime distribution type ($F(1, 15) = 8.51$, $p = .011$, $\eta^2_G = 0.06$). With prime SS = 36, slopes within the prime distribution (β_0) were flat following a uniform distribution and negative following the Gaussian distribution, independently of test SS (Fig. 11, right panel). In contrast, for the prime SS = 16, they were similar regardless of prime distribution type (Fig. 11, left panel). Analysing slopes outside the prime distribution (β_1), only the main effect of test SS was significant ($F(1, 15) = 5.33$,

[2] The same results were obtained by computing individual regression coefficients for target and distractor priming for each subject in each combination of conditions and then analyzing them with a repeated-measures ANOVA.

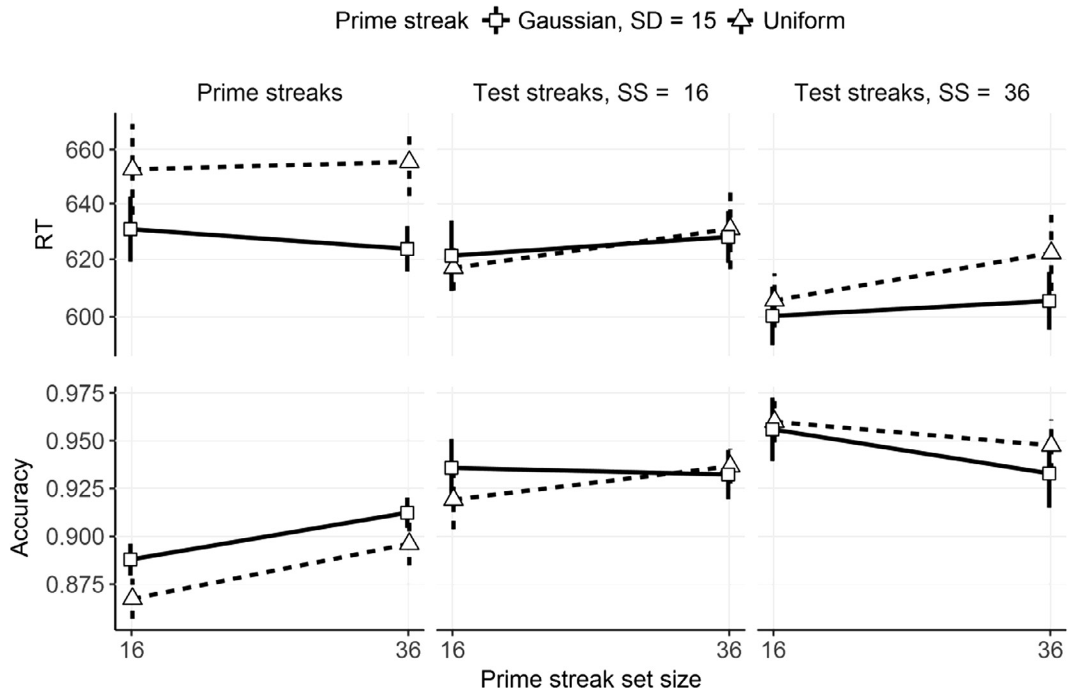

Fig. 8. Set size effects during prime and test streaks in Experiment 2.

p = .036, $\eta^2_G = 0.03$). This reflects that with test SS = 16 ($M = -0.001$ ($SD = 0.002$)) slopes were slightly more negative than with test SS = 36 ($M = -0.001$ ($SD = 0.002$)). Finally, comparisons within prime SS levels indicated that with prime SS = 16 there were no significant effects while with prime SS = 36 there was a significant effect of prime distribution type ($F(1, 15) = 13.98$, $p = .002$, $\eta^2_G = 0.06$) and an interaction between prime distribution type and slope type ($F(1, 15) = 5.59$, $p = .032$, $\eta^2_G = 0.05$). Within the prime distribution range, slopes were negative following Gaussian prime and flat following a uniform prime distribution while outside the prime distribution they were equally negative.

9. Discussion

Experiment 2 yielded two main results. First, distribution learning requires large prime set size regardless of set size on test trials, even when spacing between elements was matched. Distribution learning is therefore unlikely to depend on sampling of a limited number of stimuli as can be the case with explicit judgments of summary statistics (Maule & Franklin, 2016; Myczek & Simons, 2008; Solomon et al., 2016; Tibber et al., 2015). If that were the case, observers should be able to sample the same amount of information and learn the distribution shape with the smaller set sizes. In contrast, distribution learning is likely to reflect parallel processing of a large number of presented stimuli.

Second, distribution learning seemingly affects visual search efficiency. While search times were similar for uniform distributions regardless of set size, for the Gaussian distributions larger set size resulted in faster search suggesting that observers can use distribution shape to guide their search. The Gaussian distribution we used is relatively narrow and knowing that distractors belong to it may be useful for outlier detection. The benefits of larger set sizes for a Gaussian distribution with SD = 10 used on test trials were even larger (Fig. 8). This comparison should be treated with caution, however, as repetition

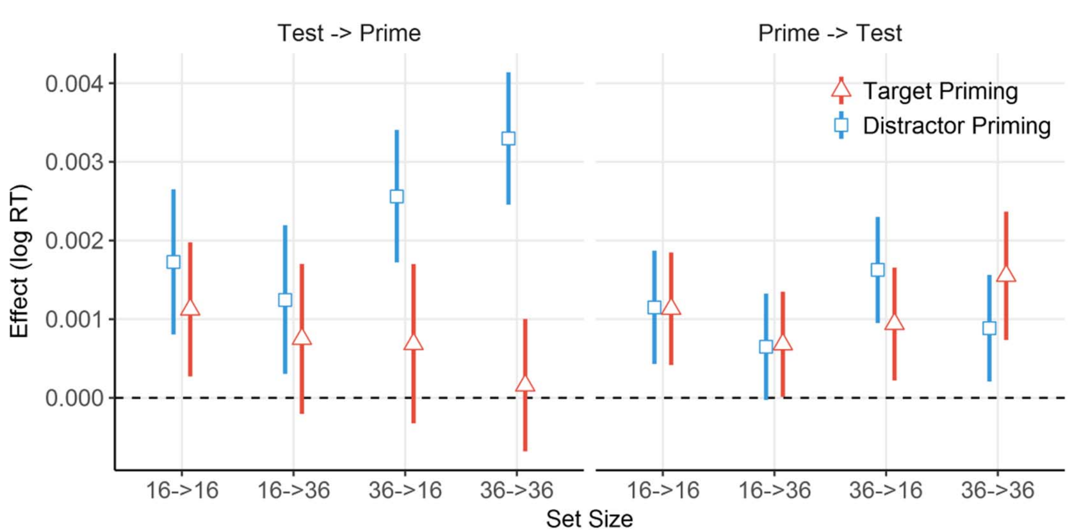

Fig. 9. Estimates of target and distractor priming effects in Experiment 2 on prime (left) and test (right) trials. Only the first trials after switch were analyzed, that is, for prime streaks the first trials after the switch from test streaks and vice versa. Lines show 95% CI.

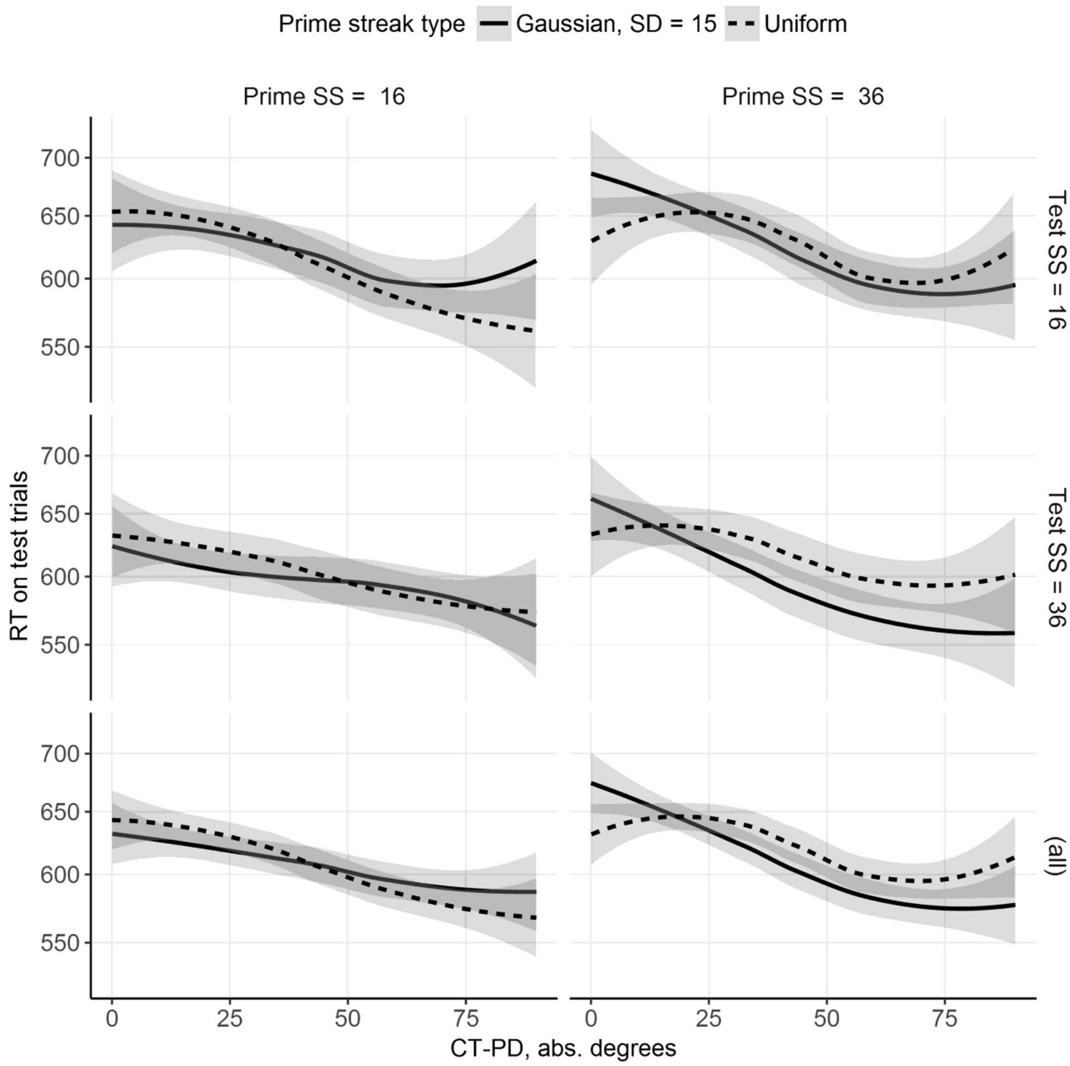

Fig. 10. RTs as a function of the distance between target orientation on test streaks and the mean of preceding prime streak distractor distribution, separately for each combination or prime SS and test SS, in Experiment 2. The bottom row shows the data for two prime SS aggregated over test SS.

Table 2
Model fits for test streaks RT in Experiment 2.

Prime SS	Test SS	ΔBIC against null model				Best model	Model	2nd best
		Linear	HG	Uniform	UwD			ΔBIC
Prime Streak: Gaussian, SD = 15								
16	16	0.80	−1.36	−2.81	0.27	Linear	UwD	0.53
	36	−1.42	−3.54	−3.70	−1.97	Null	Linear	1.42
	All	5.30	0.98	−0.35	4.36	Linear*	UwD	0.94
36	16	21.57	15.79	18.20	17.24	Linear**	Uniform	3.37
	36	26.73	21.21	22.13	21.00	Linear**	Uniform	4.61
	All	52.88	41.81	45.13	43.14	Linear**	Uniform	7.75
Prime Streak: Uniform								
16	16	23.86	11.15	10.27	23.92	UwD*	Linear	0.05
	36	5.21	0.37	1.19	4.48	Linear*	UwD	0.74
	All	33.49	17.05	17.15	32.33	Linear*	UwD	1.15
36	16	2.22	−3.66	2.55	3.96	UwD*	Uniform	1.41
	36	2.58	−1.83	−0.58	3.17	UwD*	Linear	0.59
	All	11.11	0.81	8.33	13.42	UwD**	Linear	2.31

Note: HG – half-Gaussian, UwD – uniform with decrease. ΔBIC > 2 provide "positive" evidence, while ΔBIC > 6 provide "strong" and ΔBIC > 10 "very strong" evidence.
* Best model is better than the null model.
** Best model is better than the null and the 2nd best model.

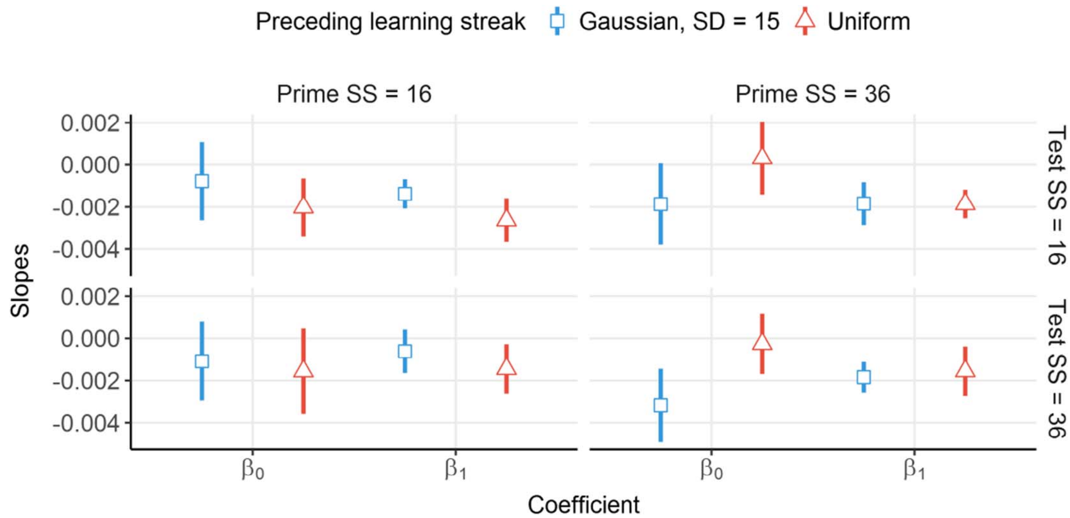

Fig. 11. Observers' slopes within (β_0) and outside (β_1) of prime distribution range in Experiment 2. Lines show 95% CI. Slopes are obtained for log-transformed RT.

number differed between test and prime streaks and in the absence of different test streaks distributions we cannot determine whether observers were able to learn the distribution shape.

The results also elucidate the role of target and distractor priming with different set sizes. A comparison of priming effects from test and prime streaks showed that target and distractor priming were similar in magnitude during switches from prime to test trials while for shifts from test to prime trials distractor priming was stronger. This could reflect three factors: streak length, stability of target or distractors, and search difficulty. Prime streaks were longer, more difficult, and had stable distractor distributions while test streaks were short, easier, and had stable targets. It is nevertheless clear that both target and distractors are learned in orientation search and affect search efficiency on subsequent trials in contrast to the claims of Lamy et al. (2013).

We did not find any significant effect of set size on target priming. However, distractor priming was significantly more pronounced after a switch from test to prime trials when the test streak had larger set size (36 lines), and distractor priming was also stronger with larger set size within test streaks. In sum, this experiment supports the results of Experiment 1, more often revealing larger priming effects with increasing set size, in contradiction to some previous results (Meeter & Olivers, 2006; Rangelov, Müller, et al., 2013; Rangelov, Zehetleitner, et al., 2017). The increased distractor priming follows increases in the of precision in distractor distribution learning and further supports the idea that the two might reflect similar mechanisms.

10. General discussion

Observers can learn surprisingly complex distractor distributions and quickly learn to distinguish between Gaussian and uniform distractor distributions. We infer this from inhibitory effects of role-reversals – if the target falls within the represented distractor distribution, search is slowed, and by how much depends on the distribution shape (Chetverikov et al., 2016, 2017a,b).

But our current findings also show that this learning has limits – that a certain minimum amount of information is required on each trial for distribution learning to occur. The smaller set sizes that we used did not lead to any learning of the characteristics of distractor distributions while there was strong learning of the shape of distractor distributions for the largest set size. This clearly shows that distribution encoding does not simply reflect sampling of only a few display items. In contrast, observers seem to aggregate information over the whole display. This result supports accounts of ensemble perception that do not rely on a limited sampling mechanism (Attarha & Moore, 2015; Attarha et al., 2014; Dakin, 2001; Im & Halberda, 2013; Tokita et al., 2016; Utochkin & Tiurina, 2014). Moreover, this shows that studies finding strong limits of information integration in explicit averaging judgments (Maule & Franklin, 2016; Myczek & Simons, 2008; Solomon et al., 2016; Tibber et al., 2015) might reflect bottlenecks related to explicit judgements rather than to encoding of perceptual ensembles.

But the results also pose a question for further studies: why does distribution encoding diminish or even disappear with smaller set sizes? There are at least two possible reasons for this. The simplest explanation is that there is not enough information about the distribution within displays with the smaller set sizes. Another possibility is that the smaller set sizes are simply not considered examples of a distribution by the visual system – the displays are simply too sparse for the system to treat them in this way.

Previously, we found that observers are able to grasp distribution shape from one or two repetitions with 36 stimuli (Chetverikov et al., 2017a), amounting at most to the encoding of 70 distractor lines. In Experiment 2 with prime SS = 16, observers had 75–90 examples of distractors in prime streaks. So why does no learning take place with small set sizes? It is unlikely to be related to area as in Experiment 1 the two largest set sizes were positioned on the same matrix. Neither can the density of the display explain this, as this was controlled for in Experiment 2. The first explanation we suggested (lack of information) would then imply that there is a heavy loss of information in-between trials. This may occur because priming effects stop accumulating quickly (Brascamp, Pels, & Kristjánsson, 2011; Martini, 2010). However, there is also evidence that learning of distribution shape might occur with sequences longer than 6 trials (Chetverikov et al., 2017a, Exp. 3).

The second explanation suggested above – that some displays are not treated as distributions – does not imply any loss of information. Rather the absence of distribution learning might reflect that the stimuli are processed in small chunks ("serial" processing, though not necessarily limited to single stimuli, but perhaps to small groups). Speculatively, if there is a limited number of stimuli on the screen, observers may look for target using local cues (e.g., angles between lines). With larger displays that strategy might be less useful, resulting in more global processing. This calls for studies of attentional distribution and display segmentation in distribution learning.

Experiment 2 suggests that learning distributions may affect search efficacy. Search for stimuli from a Gaussian prime distribution with SS = 36 compared to SS = 16 was faster, while no such difference was observed with the uniform prime distribution. Distribution learning was observed only with SS = 36 and the additional search time benefits for the Gaussian distribution could be attributed to learning. But in Experiment 1, RT differences between distributions were the same for different set sizes, except for the smallest one, suggesting a role of other factors (display size, target eccentricity) as well.

One potential factor affecting both search times and distribution learning, and partially dependent on set size, might be the total area of stimuli. With increased area, larger parts of the search display might appear in the periphery. Recently, Rosenholtz and colleagues suggested that peripheral visual representations might lose detail while keeping a range of statistics intact (Balas, Nakano, & Rosenholtz, 2009; Chang & Rosenholtz, 2016; Rosenholtz et al., 2012). Such statistics include, among others, pairwise correlations of responses between oriented wavelets across different orientations, spatial locations, and spatial scales. These statistics are detailed enough for approximate encoding of orientation PDF. It is therefore possible that increasing the total area might increase precision of distribution encoding due to, paradoxically, less detailed peripheral perception. Note, however, that in Experiment 1 two pairs of set sizes (8 and 14, 24 and 36) had the same total area – the lines were positioned in a matrix of 4 × 4 or 6 × 6 lines, respectively. Larger area therefore cannot be the only factor providing the advantage in distribution learning: set sizes 24 and 36 in Experiment 1 had the same total area, yet the search times were lower and the distribution learning was more precise in the latter than the former case. Notably, our results also show that, even if target and distractor priming occurs, this does not necessarily mean that distribution shape is learned. While target priming was generally weaker than distractor priming, it was nevertheless observed with the smaller set sizes (16 lines) in Experiment 2. The magnitude of distractor priming was also similar between set sizes. At the same time, no distribution shape learning was found for this set size. Note, however, that the fact that distribution shape is not learned accurately does not mean that observers do not learn some information about distributions. In Experiment 2, with the smaller set size, RT functions on test trials were best fit by linear and uniform-with-decrease models regardless of distribution shape. This means that observers approximated distractor distributions similarly. In contrast, in Experiment 1 RT functions on test trials following the two smallest set sizes (8 and 16) were best fit by null models. But there were also no significant priming effects with these set sizes. In sum, where there was no distribution learning at all, no priming of pop-out effects were observed. When there was at least some priming of pop-out, there was also some evidence of distribution learning, although it was not necessarily particularly accurate. We believe that this indicates that priming effects and distribution shape learning may be governed by the same learning mechanism, but the precision of that learning varies.

Finally, the results show a complicated pattern of findings surrounding the interaction of set-size with the target and distractor priming of pop-out. In general, the results indicate stronger priming of pop-out with larger set size. This goes against findings of some previous studies (Rangelov, Müller, et al., 2013; Rangelov, Zehetleitner, et al., 2017) and the idea that uncertainty plays a domineering role in priming of pop-out (Meeter & Olivers, 2006) while supporting other studies that found similar or larger priming with larger set sizes (Becker & Ansorge, 2013; Hodsoll et al., 2006; Kristjánsson & Driver, 2008; Wang et al., 2005; Wolfe et al., 2003). Some previous studies may have confounded spatial density of stimuli with set size as we did in Experiment 1. For example, in a typical visual search display with items arranged in a circle that was used both by Becker and Ansorge (2013) and Rangelov et al. (2017, Exp. 3 and 4), increasing the number of items automatically means decreasing their density. When the spatial density is similar despite different set sizes (as in our Experiment 2), the priming effects seem to depend less on set size, although they are still larger with larger set sizes. Note, however, that in Experiment 1 we also found positive correlations between priming effects and set size despite covariation of density with set size. Search difficulty, is, on its own, also unlikely to explain the differences in the results. In both experiments reported here search difficulty decreased with set size while priming effects became stronger. Yet, with the displays used by Rangelov and colleagues (2017), search difficulty similarly decreased with set size but priming effects became weaker. A possible explanation is that search difficulty lies in different boundaries – in our experiments the average error percentage varied approximately from 75% to 95% while in the data reported by Rangelov and colleagues (2017) the lowest accuracy was about 88%. This might also hint at the importance of another variable – positional variation of individual stimuli. In our experiments, stimuli were jittered which might reduce the usefulness of local cues, while in the studies of Rangelov and colleagues, the stimuli were arranged on a fixed grid which might entice participants to use such cues for search. Accordingly, when they are used, specific target and distractor parameters may become less relevant. In general, the results indicate that many questions remain unanswered in the priming of pop-out literature despite it's relatively long history.

In sum, we found that distribution learning should not be thought of as being based on a limited sampling mechanism. Small set sizes and sparse displays nevertheless hamper learning, and this provides clues regarding the boundary conditions of distractor distribution learning.

Acknowledgments

This paper was supported by the Russian Foundation for Humanities (#15-36-01358A2) and the Icelandic Research Fund (#152427). The authors are grateful to Alena Begler for help with the data collection.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.visres.2017.08.003.

References

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences, 15*(3), 122–131. http://dx.doi.org/10.1016/j.tics.2011.01.003.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12*(2), 157–162. Retrieved from http://pss.sagepub.com/content/12/2/157.short.

Ásgeirsson, Á. G., & Kristjánsson, Á. (2011). Episodic retrieval and feature facilitation in intertrial priming of visual search. *Attention, Perception & Psychophysics, 73*(5), 1350–1360. http://dx.doi.org/10.3758/s13414-011-0119-5.

Attarha, M., & Moore, C. M. (2015). The capacity limitations of orientation summary statistics. *Attention, Perception, & Psychophysics, 77*(4), 1116–1131. http://dx.doi.org/10.3758/s13414-015-0870-0.

Attarha, M., Moore, C. M., & Vecera, S. P. (2014). Summary statistics of size: Fixed processing capacity for multiple ensembles but unlimited processing capacity for single ensembles. *Journal of Experimental Psychology: Human Perception and Performance, 40*(4), 1440–1449. http://dx.doi.org/10.1037/a0036206.

Balas, B. J., Nakano, L., & Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. *Journal of Vision, 9*(12), 13.1–18. http://dx.doi.org/10.1167/9.12.13.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1) https://doi.org/10.18637/jss.v067.i01.

Becker, S. I. (2010). The role of target-distractor relationships in guiding attention and the eyes in visual search. *Journal of Experimental Psychology: General, 139*(2), 247–265. http://dx.doi.org/10.1037/a0018808.

Becker, S. I., & Ansorge, U. (2013). Higher set sizes in pop-out search displays do not eliminate priming or enhance target selection. *Vision Research, 81*, 18–28. http://dx.doi.org/10.1016/j.visres.2013.01.009.

Brascamp, J. W., Pels, E., & Kristjánsson, Á. (2011). Priming of pop-out on multiple time scales during visual search. *Vision Research, 51*, 1972–1978.

Bravo, M. J., & Nakayama, K. (1992). The role of attention in different visual-search tasks. *Perception & Psychophysics, 51*(5), 465–472. http://dx.doi.org/10.3758/BF03211642.

Carlisle, N. B., & Kristjánsson, Á. (2017). How visual working memory contents influence priming of visual attention. *Psychological Research Psychologische Forschung, 1–7*. http://dx.doi.org/10.1007/s00426-017-0866-6.

Chang, H., & Rosenholtz, R. (2016). Search performance is better predicted by tileability than presence of a unique basic feature. *Journal of Vision, 16*(August), 1–18. http://dx.doi.org/10.1167/16.10.13.doi.

Chetverikov, A., Campana, G., & Kristjánsson, Á. (2016). Building ensemble representations: How the shape of preceding distractor distributions affects visual search. *Cognition, 153*, 196–210. http://dx.doi.org/10.1016/j.cognition.2016.04.018.

Chetverikov, A., Campana, G., & Kristjánsson, Á. (2017a). Rapid learning of visual ensembles. *Journal of Vision, 17*(21), 1–15. http://dx.doi.org/10.1167/17.2.21.

Chetverikov, A., Campana, G., & Kristjánsson, Á. (2017b). Representing color ensembles.

Psychological Science. http://dx.doi.org/10.1177/0956797617713787.
Corbett, J. E., & Melcher, D. (2014). Stable statistical representations facilitate visual search. *Journal of Experimental Psychology: Human Perception and Performance, 40*(5), 1915–1925. http://dx.doi.org/10.1037/a0037375.
Dakin, S. C. (2001). Information limit on the spatial integration of local orientation signals. *Journal of the Optical Society of America A, 18*(5), 1016–1026. http://dx.doi.org/10.1364/JOSAA.18.001016.
Dakin, S. C. (2015). Seeing Statistical Regularities: Texture and Pattern Perception. In J. Wagemans (Ed.), *The Oxford Handbook of Perceptual Organization* (pp. 150–167).
Haberman, J., & Whitney, D. (2012). Ensemble perception: Summarizing the scene and broadening the limits of visual processing. In J. M. Wolfe, & L. Robertson (Eds.). *From perception to consciousness: Searching with Anne Treisman* (pp. 339–349). Oxford: Oxford University Press.
Hodsoll, J. P., Humphreys, G. W., & Braithwaite, J. J. (2006). Dissociating the effects of similarity, salience, and top-down processes in search for linearly separable size targets. *Perception & Psychophysics, 68*(4), 558–570. http://dx.doi.org/10.3758/BF03208758.
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. http://dx.doi.org/10.3758/s13414-012-0399-4.
Im, H. Y., Park, W. J., & Chong, S. C. (2015). Ensemble statistics as units of selection. *Journal of Cognitive Psychology, 27*(1), 114–127. http://dx.doi.org/10.1080/20445911.2014.985301.
Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature, 290*(5802), 91–97. http://dx.doi.org/10.1038/290091a0.
Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773–795. http://dx.doi.org/10.1080/01621459.1995.10476572.
Kristjánsson, Á. (2015). Reconsidering visual search. *i-Perception, 6*(6), 1–14. http://dx.doi.org/10.1177/2041669515614670.
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. http://dx.doi.org/10.3758/APP.72.1.5.
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. http://dx.doi.org/10.1016/j.visres.2008.02.007.
Lamy, D. F., 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. http://dx.doi.org/10.1016/j.visres.2007.10.009.
Lamy, D. F., Yashar, A., & Ruderman, L. (2013). Orientation search is mediated by distractor suppression: Evidence from priming of pop-out. *Vision Research, 81*, 29–35. http://dx.doi.org/10.1016/j.visres.2013.01.008.
Ma, W. J., Shen, S., Dziugaite, G., & van den Berg, R. (2015). Requiem for the max rule? *Vision Research, 116*, 179–193. http://dx.doi.org/10.1016/j.visres.2014.12.019.
Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition, 22*(6), 657–672. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7808275.
Martini, P. (2010). System identification in priming of pop-out. *Vision Research, 50*(21), 2110–2115. http://dx.doi.org/10.1016/j.visres.2010.07.024.
Maule, J., & Franklin, A. (2015). Effects of ensemble complexity and perceptual similarity on rapid averaging of hue. *Journal of Vision, 15*(4), 6. http://dx.doi.org/10.1167/15.4.6.doi.
Maule, J., & Franklin, A. (2016). Accurate rapid averaging of multihue ensembles is due to a limited capacity subsampling mechanism. *Journal of the Optical Society of America A, 33*(3), A22–A29. http://dx.doi.org/10.1364/JOSAA.33.000A22.
Meeter, M., & Olivers, C. N. L. (2006). Intertrial priming stemming from ambiguity: A new account of priming in visual search. *Visual Cognition, 13*(2), 202–222. http://dx.doi.org/10.1080/13506280500277488.
Michael, E., de Gardelle, V., & Summerfield, C. (2014). Priming by the variability of visual information. *PNAS, 111*(21), 7873–7878. http://dx.doi.org/10.1073/pnas.1308674111.
Myczek, K., & Simons, D. J. (2008). Better than average: Alternatives to statistical summary representations for rapid judgments of average size. *Perception & Psychophysics, 70*(5), 772–788. http://dx.doi.org/10.3758/PP.70.5.772.
Peirce, J. W. (2007). PsychoPy–Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1–2), 8–13. http://dx.doi.org/10.1016/j.jneumeth.2006.11.017.
Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2*. http://dx.doi.org/10.3389/neuro.11.010.2008.
Rangelov, D., Müller, H. J., & Zehetleitner, M. (2017). Failure to pop out: Feature singletons do not capture attention under low signal-to-noise ratio conditions. *Journal of Experimental Psychology: General*. http://dx.doi.org/10.1037/xge0000284.
Rangelov, D., Zehetleitner, M., Muller, H. J., & Zehetleitner, M. (2013). Visual search for feature singletons: Multiple mechanisms produce sequence effects in visual search. *Journal of Vision, 13*(2), 1–16. http://dx.doi.org/10.1167/13.3.22.doi.
Rosenholtz, R., Huang, J., Raj, A., Balas, B. J., & Ilie, L. (2012). A summary statistic representation in peripheral vision explains visual search. *Journal of Vision, 12*(4), http://dx.doi.org/10.1167/12.4.14 14–14.
Solomon, J. A., May, K. A., & Tyler, C. W. (2016). Inefficiency of orientation averaging: Evidence for hybrid serial/parallel temporal integration. *Journal of Vision, 16*(1), 13. http://dx.doi.org/10.1167/16.1.13.
Tibber, M. S., Anderson, E. J., Bobin, T., Carlin, P., Shergill, S. S., & Dakin, S. C. (2015). Local and global limits on visual processing in schizophrenia. *PLoS One, 10*(2), 1–17. http://dx.doi.org/10.1371/journal.pone.0117951.
Tokita, M., Ueda, S., & Ishiguchi, A. (2016). Evidence for a global sampling process in extraction of summary statistics of item sizes in a set. *Frontiers in Psychology, 7*(MAY), http://dx.doi.org/10.3389/fpsyg.2016.00711.
Utochkin, I. S. (2015). Ensemble summary statistics as a basis for rapid visual categorization. *Journal of Vision, 15*(4), 8. http://dx.doi.org/10.1167/15.4.8.
Utochkin, I. S., & Tiurina, N. A. (2014). Parallel averaging of size is possible but range-limited: A reply to Marchant, Simons, and De Fockert. *Acta Psychologica, 146*, 7–18. http://dx.doi.org/10.1016/j.actpsy.2013.11.012.
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. http://dx.doi.org/10.1037/xhp0000203.
Wang, D., Kristjánsson, Á., & Nakayama, K. (2005). Efficient visual search without top-down or bottom-up guidance. *Perception & Psychophysics, 67*(2), 239–253. http://dx.doi.org/10.3758/BF03206488.
Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 483–502. http://dx.doi.org/10.1037/0096-1523.29.2.483.
Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour, 1*(3), 58. http://dx.doi.org/10.1038/s41562-017-0058.