Independence is elusive: Set size effects on encoding precision in visual search

Helga Mazyar* Baylor College of Medicine, Houston, TX, USA

Ronald van den Berg* Baylor College of Medicine, Houston, TX, USA

Robert L. Seilheimer* Baylor College of Medicine, Houston, TX, USA

Wei Ji Ma Baylor College of Medicine, Houston, TX, USA

Looking for a target in a visual scene becomes more difficult as the number of stimuli increases. In a signal detection theory view, this is due to the cumulative effect of noise in the encoding of the distractors, and potentially on top of that, to an increase of the noise (i.e., a decrease of precision) per stimulus with set size, reflecting divided attention. It has long been argued that human visual search behavior can be accounted for by the first factor alone. While such an account seems to be adequate for search tasks in which all distractors have the same, known feature value (i.e., are maximally predictable), we recently found a clear effect of set size on encoding precision when distractors are drawn from a uniform distribution (i.e., when they are maximally unpredictable). Here we interpolate between these two extreme cases to examine which of both conclusions holds more generally as distractor statistics are varied. In one experiment, we vary the level of distractor heterogeneity; in another we dissociate distractor homogeneity from predictability. In all conditions in both experiments, we found a strong decrease of precision with increasing set size, suggesting that precision being independent of set size is the exception rather than the rule.

Introduction

In natural environments, animals constantly have to attend to a large number of stimuli simultaneously. A central question whose history spans the history of psychology is how humans divide their attention across multiple stimuli. An important aspect of this question is whether the quality of the representation of an individual stimulus is independent of the number of attended stimuli (Broadbent, 1958; Townsend, 1974). The Scottish philosopher Sir William Hamilton (1788–1856) clearly articulated his belief that it is not: "The greater the number of objects among which the attention of the mind is distributed, the feebler and less distinct will be its cognizance of each" (Hamilton, 1859). More than a century later, rigorous psychophysical studies reached the opposite conclusion: For example, Palmer's work on simple visual search demonstrated that human performance is well described by a mathematical model in which the quality (precision) with which a stimulus is encoded is independent of the number of stimuli (Palmer, Verghese, & Pavel, 2000). Recently however, we showed that Palmer's conclusion is not universal but depends on the statistics of the distractor stimuli (Mazyar, van den Berg, & Ma, 2012). This raises the question which answer—dependence or independence of precision on set size—is the rule and which the exception; we will explore this question here.

Visual search is an oft-used paradigm for measuring the effects of set size on precision. Telling whether a target object is present among a set of distractor objects generally becomes more difficult as the number of stimuli increases. This can be understood from the premise that stimuli are encoded in a noisy manner. As more distractors are present in a scene, the noise they contribute is "drowning out" the target signal. This effect has been formalized in Bayes-optimal (Ma, Navalpakkam, Beck, van den Berg, & Pouget, 2011; Mazyar et al., 2012) and other signal detection theory models (Eckstein, Thomas, Palmer, & Shimozaki, 2000; Nolte & Jaarsma, 1966; Palmer, Ames, & Lindsey, 1993; Palmer et al., 2000; Rosenholtz, 2001; Verghese, 2001; Vincent, Baddeley, Troscianko, & Gilchrist, 2009).

Citation: Mazyar, H., van den Berg, R., Seilheimer, R. L., & Ma, W. J. (2013). Independence is elusive: Set size effects on encoding precision in visual search. *Journal of Vision*, *13*(5):8, 1–14, http://www.journalofvision.org/content/13/5/8, doi:10.1167/13.5.8.

doi: 10.1167/13.5.8 Received June 21, 2012; published April 10, 2013 ISSN 1534-7362

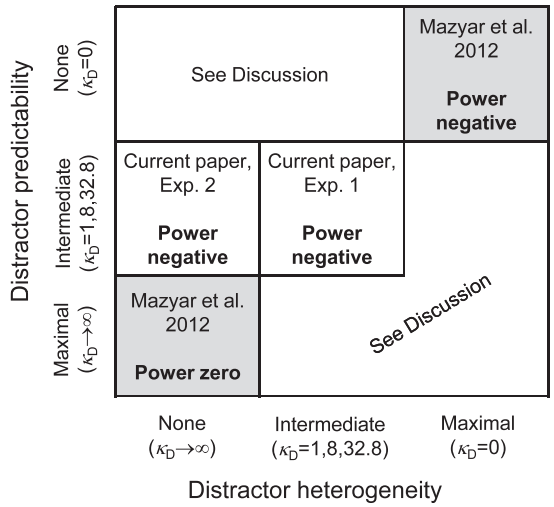

Figure 1. How does precision depend on set size? Each box in the diagram represents a distractor condition in single-target visual search. The gray boxes represent conditions for which set size effects on precision have been studied previously: homogeneous distractors that are predictable on every trial (bottom left) and maximally heterogeneous distractors that are also maximally unpredictable (top right). In this paper, we interpolate between these conditions (Experiment 1) and dissociate the factors (Experiment 2).

A second effect that might contribute to the decrease of performance with set size is the one mentioned above: The amount of noise in the encoding of a stimulus might increase with the number of stimuli, perhaps due to the spreading of an attentional resource (Palmer, 1990; Shaw, 1980). As a consequence, the precision of encoding an individual stimulus would decrease. This phenomenon—referred to as a "resource limitation" or also "limited capacity" (Townsend, 1974), although it is not a limit on the number of stimuli that is encoded—would cause performance to drop faster as a function of set size than it would due to the first effect alone. Many studies have reported, explicitly or implicitly, that the second effect is not needed to explain human performance, as long as no memory component is involved (Baldassi & Burr, 2000; Busey & Palmer, 2008; Palmer, 1994; Palmer et al., 1993; Palmer et al., 2000), thereby essentially proclaiming Hamilton wrong.

In 2012 we questioned the generality of this conclusion (Mazyar et al., 2012). We found that precision was independent of set size when the distractors within a given display were identical to each other (homogeneous) but not when they differed (heterogeneous). In the homogeneous condition, distractors were not only the same within a display, but also across trials. In the heterogeneous condition, distractors were drawn independently from a uniform distribution at every location and across trials, and were therefore both maximally heterogeneous and maximally unpredictable.

A way to conceptualize the space of distractor statistics is shown in Figure 1: One axis represents distractor heterogeneity within a display, the other distractor predictability across trials. In Mazyar et al. (2012), we contrasted two extremes in this space, represented by the bottom-left and top-right corners. Based on the results reported in that paper, we cannot determine whether the established conclusion—independence—or the new conclusion—dependence—is more generally valid. Therefore, we will here further explore this space. In Experiment 1, we test points along the diagonal by varying the degree of heterogeneity. In Experiment 2, we dissociate homogeneity from predictability by using unpredictable, homogeneous distractors. To anticipate our results, in all conditions tested, we find a decrease of precision with set size, suggesting that Hamilton's speculation states the rule rather than the exception.

Experimental methods

Subjects

Six subjects (four female, two male) participated in Experiment 1. Fifteen subjects participated in Experiment 2, but data from two subjects were excluded from the analysis because they performed at chance in one or more conditions. This left us with 13 subjects (six female, seven male). All subjects had normal or corrected-to-normal acuity and gave informed consent.

Apparatus and stimuli

Subjects viewed the stimuli on a 21-inch LCD monitor (60 Hz refresh rate) at a distance of approximately 60 cm. The background luminance was 33.1 cd/m^2. Stimuli consisted of Gabor patches with a spatial frequency of approximately 1.6 cycles/deg, a Gaussian envelope of approximately 0.29 deg, and a peak luminance of 122 cd/m^2.

Procedure

Experiment 1: Effect of level of heterogeneity

Each trial started with a fixation cross presented at the center of the screen (500 ms), followed by a display containing N stimuli (83 ms; five frames) (Figure 2a). Set size N was pseudorandomly chosen to be 1, 2, 4, or 8. Stimuli were placed 45° apart on an imaginary circle at the center of the screen with a radius of 5° of visual

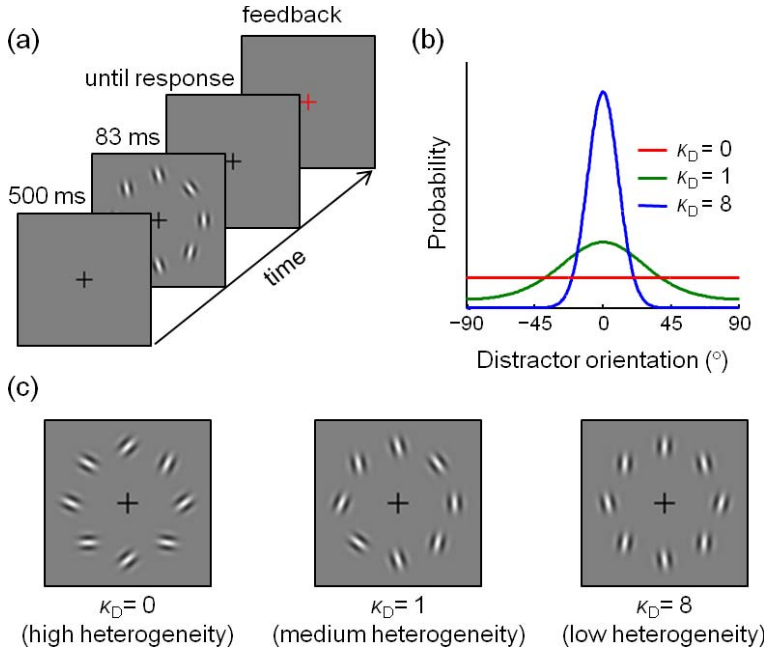

Figure 2. Experiment 1. (a) Time course of a single trial. (b) Von Mises distributions from which the distractor orientations were drawn in the three conditions (low, medium, and high heterogeneity). (c) Sample search displays in the three conditions. Displays are not to scale.

angle. On target-present trials, the stimulus set consisted of one target stimulus and $N-1$ distractor stimuli. On target-absent trials, it consisted of N distractor stimuli. The target stimulus was always vertical (denoted s_T and defined as 0°) and each distractor orientation s_i was independently drawn from a Von Mises distribution centered at s_T:

$$p(s_i|T_i=0) = \frac{1}{\pi I_0(\kappa_D)} e^{\kappa_D \cos 2(s_i - s_T)} \quad (1)$$

where $T_i = 0$ refers to a distractor being present at the i^{th} location, and I_0 is the modified Bessel function of the first kind of order 0. The concentration parameter of this distribution, κ_D, determined the level of heterogeneity in a display and was different across conditions: $\kappa_D = 0$ (high heterogeneity; uniform distribution), $\kappa_D = 1$ (medium heterogeneity), or $\kappa_D = 8$ (low heterogeneity; Figure 2b and c). The concentration parameter is inversely and monotonically related to the circular variance of the distractors. The observer reported by pressing a key whether or not the target was present and received immediate correctness feedback.

This design is based on that by Vincent et al. (2009), but we use different concentration parameters; in particular, we include a uniform distribution ($\kappa_D = 0$) to relate to our earlier results. Vincent et al. fixed set size at four, while our focus is on the effect of set size. We will also test a different set of models, specifically Bayes-optimal models and variable-precision models, which we will compare to Vincent et al.'s model.

Each subject completed six sessions, each featuring a single heterogeneity condition and consisting of four blocks of 175 trials. The order of the conditions used across the six sessions was high, medium, low, low, medium, high. At the beginning of each session, subjects were instructed in multiple ways about the distractor distribution they were going to experience. A plot of the distribution was shown on the screen along with 100 randomly drawn sample distractors; the meaning of the plot and the samples was explained verbally to the subject in the first three sessions. After seeing the plot and samples, the subject performed 100 practice trials; these were left out of our analyses, but were otherwise identical to the "real" trials.

Experiment 2: Heterogeneous versus homogeneous, unpredictable distractors

Experiment 2 consisted of two conditions, "heterogeneous" and "homogeneous" (Figure 7). The heterogeneous condition was identical to any one of the conditions in Experiment 1, except that the Von Mises distribution from which each distractor orientation was drawn had a concentration parameter of $\kappa_D = 32.8$ (corresponding to about 5°). The homogeneous condition was identical to the heterogeneous condition, except that on each trial a *single* distractor orientation was drawn from the $\kappa_D = 32.8$ Von Mises distribution and assigned to all distractors. Set sizes were 1, 2, 4, and 8 in the homogeneous condition and 1, 2, 3, and 4 in the heterogeneous condition; the latter was different from

Experiment 1 because the higher concentration parameter made this condition more difficult. Set size 1 trials had identical statistics in the two conditions. Each subject performed one session in the homogenous condition, followed by two sessions in the heterogeneous condition, followed by one session in the homogeneous condition. Each session consisted of four blocks of 175 trials.

The key difference with the homogeneous condition in our earlier work (Mazyar et al., 2012) is that in that study, the common distractor orientation was always 5° clockwise relative to the target, whereas in the current study, the common distractor orientation was unpredictable from trial to trial. In the earlier study, there were only eight different stimulus displays in the entire homogeneous condition (target present/absent, and four set sizes). By contrast, in the current experiment, the target-distractor orientation difference varies, making each trial unique and thus producing a richer data set.

Models

Experiment 1

Encoding model: Equal and variable precision

We model the observer's behavior as consisting of an encoding stage followed by a decision stage. In the encoding stage, stimuli are internally represented, or "measured," in a noisy manner. We denote by (s_1, …, s_N) the stimulus orientations on a given trial and by (x_1, …, x_N) the corresponding noisy measurements. We assume that the measurements are drawn from independent Von Mises distributions on [0, π),

$$p(x_i|s_i) = \frac{1}{\pi I_0(\kappa_i)} e^{\kappa_i \cos 2(s_i - x_i)} \quad (2)$$

The higher the concentration parameter κ_i, the less variable the measurements are for a given stimulus. In the limit of large κ_i, the Von Mises distribution becomes a Gaussian distribution with ${\sigma_i}^2 = 1/(4\kappa_i)$:

$$p(x_i|s_i) \propto e^{\kappa_i \cos 2(x_i - s_i)} \approx e^{\kappa_i (1 - \frac{1}{2}(2(x_i - s_i))^2)} \propto e^{-\frac{(x_i - s_i)^2}{2\sigma_i^2}}.$$

As in previous studies, we define encoding precision as Fisher information, denoted J, which is a general measure of the amount of information that a random variable (x_i in our case) carries about an unknown parameter upon which this variable depends (s_i in our case). Fisher information is related to the Von Mises concentration parameter through

$$J = \frac{\kappa I_1(\kappa)}{I_0(\kappa)},$$

where I_1 is the modified Bessel function of the first kind of order 1 (Keshvari, van den Berg, & Ma, 2012; Mazyar et al., 2012; van den Berg, Shin, Chou, George, & Ma, 2012).

Probabilistic models of perception typically assume that encoding precision is constant across stimuli and trials, as long as the physical conditions are held fixed. We recently showed, however, that precision in a visual search task varies across trials and stimuli (Mazyar et al., 2012). Therefore, we will here consider both an equal-precision (EP) and variable-precision (VP) model. In the EP model, all stimuli in a display are encoded with the same precision J, whose value we fit separately at each set size. In the VP model, we assume that J is a gamma-distributed random variable with a mean that we fit separately for each set size and a scale parameter τ that we assume to be constant across set size.

Optimal model

Our main model for the decision stage is a Bayesian model for heterogeneous visual search (Ma et al., 2011; Mazyar et al., 2012). Let $C = 0$ denote the state of the target being absent, $C = 1$ that of the target being present. The observer responds "target present" when, given the evidence in (x_1, …, x_N), the posterior probability of target presence is greater than that of target absence, $p(C = 1 \mid x_1, \ldots, x_N) > p(C = 0 \mid x_1, \ldots, x_N)$. A lengthy but straightforward derivation (see Appendix 1) reveals that this is equivalent to

$$\frac{p_{\text{present}}}{1 - p_{\text{present}}} \frac{1}{N} \sum_{i=1}^{N} d_i > 1 \quad (3)$$

where p_{present} is the observer's prior belief of the target being present and

$$d_i = \times \frac{I_0(\kappa_D) e^{\kappa_i \cos 2(x_i - s_T)}}{I_0\left(\sqrt{\kappa_i^2 + \kappa_D^2 + 2\kappa_i \kappa_D \cos 2(x_i - s_T)}\right)}$$

is the likelihood ratio of target presence at location i. To gain some intuition for Equation 3, consider the case that $p_{\text{present}} = 0.5$ and $\kappa_i = 0$ for all i; this means that no information is available. In this case, the posterior ratio (the left-hand side of Equation 3) equals 1. This is as expected because it means that the posterior probability that the target is present is 0.5. In another special case, $\kappa_D = 0$ (high-heterogeneity condition; uniform distractor distribution), the posterior ratio simplifies to

$$\frac{p_{\text{present}}}{1 - p_{\text{present}}} \frac{1}{N} \sum_{i=1}^{N} \frac{e^{\kappa_i \cos 2(x_i - s_T)}}{I_0(\kappa_i)},$$

which we studied before (Ma et al., 2011; Mazyar et al., 2012).

When fitting the model, we leave open the possibility that p_{present} is not equal to 0.5, reflecting that the observer may not use the true frequencies of target-absent and target-present trials during inference. A value of p_{present} greater than 0.5 will make the observer respond "target present" in some cases when the evidence points towards target absence. This aspect of the model makes it strictly speaking suboptimal, and large deviations from 0.5 will make the observer stray far from performance maximization. However, since the observer otherwise accurately takes into account the structure of the generative model, we will for convenience still refer to this Bayesian model as optimal.

Max$_d$ model

Vincent et al. (2009) proposed an alternative model in which the observer computes the likelihood ratio of target presence at each individual location (for the i^{th} location, this ratio would be the summand in Equation 3), and responds "target present" if the largest of these ratios exceeds a fixed decision criterion k:

$$\underset{i}{\text{argmax}}\ d_i > k \tag{4}$$

with d_i as defined above. Contrary to the claims of Vincent et al. (2009), this is not an optimal model, because the decision rule is not equivalent to Equation 3 with $p_{\text{present}} = 0.5$. It is not clearly Bayesian either, because it is not obvious that there exists a generative model for which the decision rule maximizes performance (Ma, 2012). However, it is a very plausible model for human behavior.

Lapse rate

Due to lapses of attention, a certain proportion of trial responses may have been random guesses. We allow for this possibility by including a lapse rate in the models: on each trial, there is a probability λ that the observer produces a random guess instead of responding according to the decision rule of the respective model. This parameter can also capture unintended key presses.

Summary of models

We have defined two encoding models—EP and VP—and two decision models—optimal (O) and max$_d$ (M). This gives rise to a total of four models, which we will name EPO, EPM, VPO, and VPM. These models have six, six, seven, and seven free parameters, respectively: four (mean) precision parameters, a lapse rate parameter, a scale parameter in the VP models, a prior parameter in the optimal models, and a decision criterion in the max$_d$ models.

Experiment 2

In Experiment 2, we only consider the winning model for Experiment 1, namely the VPO model. In the heterogeneous condition, the VPO model is identical to that in Experiment 1. In the homogenenous condition, the model is different because a common distractor orientation is drawn from a Von Mises distribution on each trial. The optimal decision rule is then to respond "target present" when

$$\frac{p_{\text{present}}}{1-p_{\text{present}}}\frac{1}{N}\sum_{i=1}^{N} e^{\kappa_i \cos 2(x_i - s_T)}\frac{I_0(\kappa_{C=1,i})}{I_0(\kappa_{C=0})} > 1 \tag{5}$$

where

$$\kappa_{C=1,i} = \sqrt{\left(\sum_{j\neq i}\kappa_j \cos 2x_j\right)^2 + \left(\sum_{j\neq i}\kappa_j \sin 2x_j\right)^2}$$

$$\kappa_{C=0} = \sqrt{\left(\sum_{j=0}^{N}\kappa_j \cos 2x_j\right)^2 + \left(\sum_{j=0}^{N}\kappa_j \sin 2x_j\right)^2}, \tag{6}$$

(see Appendix 1). Here, all sums start at 0, and we have defined $x_0 = s_T$ and $\kappa_0 = \kappa_D$. When $p_{\text{present}} = 0.5$ and $\kappa_i = 0$ for all i, the posterior ratio again equals 1.

Based on our findings in Experiment 1, we further specified the VPO model in two ways. First, we describe mean precision as a function of set size by a power law, $\overline{J} = \overline{J}_1 N^{\alpha}$. Second, we fixed the lapse rate to 0. We fitted the homogeneous and heterogeneous conditions together, assuming that p_{present} and $\overline{J}_1$ are shared between conditions. For $\overline{J}_1$, we make this assumption because at set size 1, the stimulus statistics are identical in the two conditions. The power α and the scale parameter τ can vary between the conditions. Thus, in this experiment, the VPO model has six free parameters to fit both conditions.

Results

Experiment 1

In this experiment, we tested the effect of the level of distractor heterogeneity within a display on the relationship between precision and set size in detection of a single target. Figure 3 shows subjects' perfor-

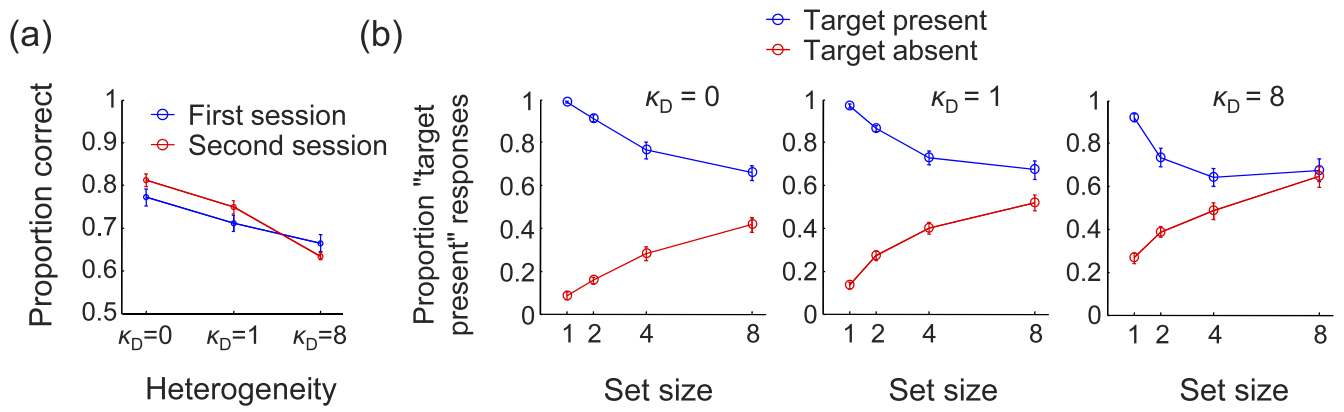

Figure 3. Data from Experiment 1. Here and elsewhere, error bars indicate 1 *SEM*. (a) Mean performance for the three heterogeneity conditions. Each condition was presented in two out of six sessions. (b) Hit and false-alarm rates as a function of set size for each condition.

mance. A two-way repeated-measures ANOVA showed a significant main effect of session heterogeneity level (κ_D) on performance, $F(2, 10) = 19.5$, $p < 0.001$, but no effect of session number, $F(1, 5) = 2.73$, $p = 0.16$, and no interaction, $F(2, 10) = 2.06$, $p = 0.18$ (Figure 3a). Moreover, at every level of heterogeneity, hit rate decreased (one-way repeated-measures ANOVA, $F(3, 15) > 14.7$, $p < 0.001$) and false-alarm rate increased ($F(3, 15) > 21.2$, $p < 0.001$), as a function of set size (Figure 3b).

We fitted the models using maximum-likelihood (ML) estimation, for each subject separately (see Appendix 2 for an overview of parameter estimates). To examine how well the models fit the subject data, we computed predicted hit and false-alarm rates under each model using the ML parameter estimates. All four

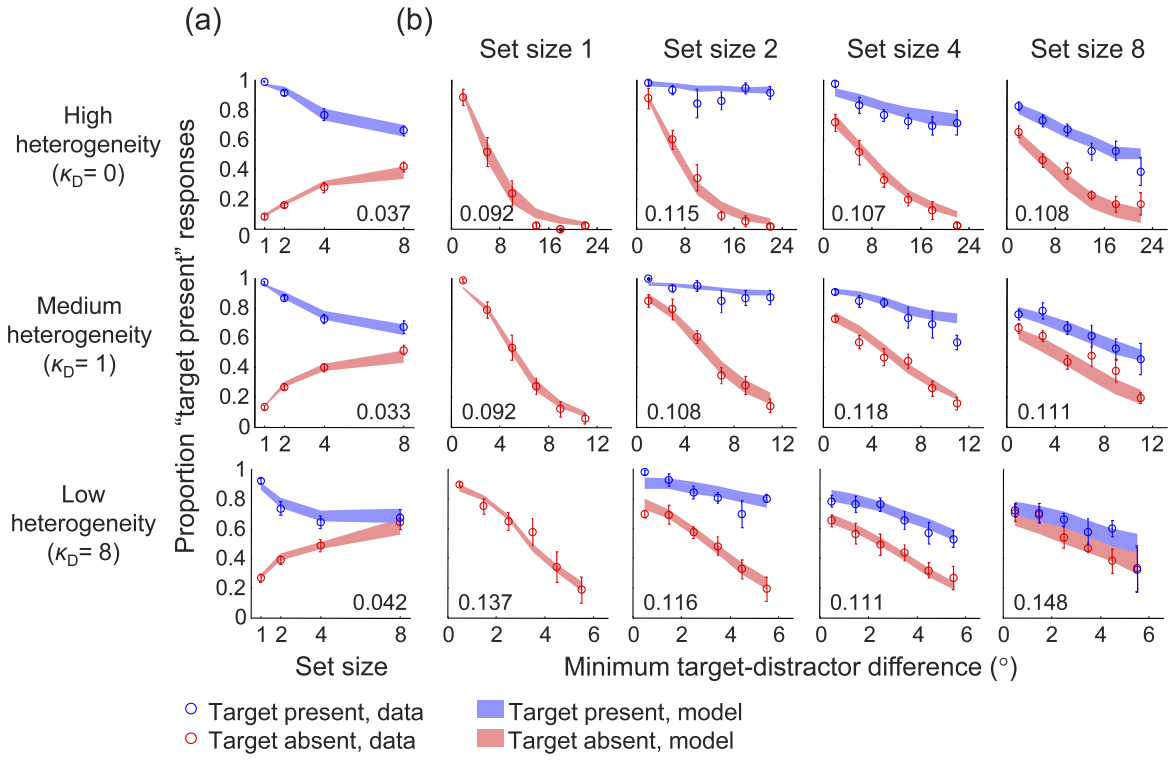

Figure 4. VPO model fits in Experiment 1. (a) Hit and false-alarm rates in the three heterogeneity conditions. Here and elsewhere, shaded areas indicate 1 *SEM* in the model, and numbers indicate the root-mean-square error between model and data, averaged over subjects. (b) Proportion "target present" responses as a function of the minimum target-distractor difference, separately for target-present (blue) and target-absent (red) trials. For set size 1, there are no distractors on target-present trials. Note that the values on the *x*-axes differ between rows.

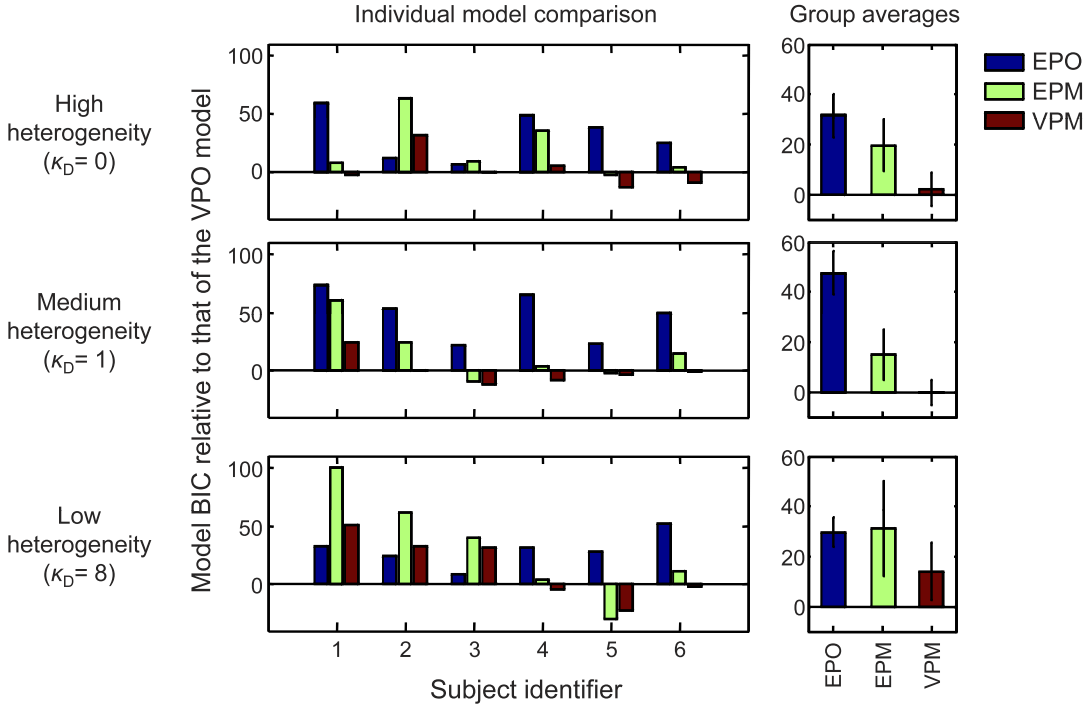

Figure 5. Model comparison results for Experiment 1. Shown are BIC values of the EPO, EPM, and VPM models relative to the VPO model, for each subject (left) as well as the group averages (right). Each row corresponds to a heterogeneity condition. Higher BIC values mean worse fits.

models provide good fits to the hit and false-alarm rates (Figure 4a and Supplementary Figures S1 through S3a). A more detailed view of the data and model predictions can be obtained by plotting the proportion "target present" responses as a function of the minimum orientation difference between the target and any distractor (Figure 4b and Supplementary Figures S1 through S3b). These plots show that the proportion of "target present" responses decreases as the minimum target-distractor difference increases, both for target-present and target-absent trials. This is expected, because the larger the minimum target-distractor difference, the more dissimilar to the target the distractors tend to be.

To quantify the goodness of fit of the four models, we computed the Bayesian Information Criterion (BIC; Schwartz, 1978) for each subject at every level of heterogeneity (Figure 5). Out of 18 comparisons (six subjects × three conditions), the VPM model wins nine, the VPO model eight, the EPM model one, and the EPO model zero times. Hence, both EP models perform poorly on these data, but we cannot distinguish the VPO and VPM models. Therefore, we will consider both VP models in our examination of the relation between encoding precision and set size.

Figure 6 shows the estimates of mean precision ($\bar{J}$) as a function of set size in both VP models. For each level of heterogeneity, the relationship between set size and mean precision is fitted well by a power-law function with a power close to −1 in both models (−0.95 ± 0.26, −1.08 ± 0.20, and −0.91 ± 0.08 for $\kappa_D = 0$, 1, and 8, respectively in the VPO model, and −1.06 ± 0.17, −1.17 ± 0.13, and −1.34 ± 0.06 in the VPM model; mean and *SEM* across subjects), similar to our previous results (Mazyar et al., 2012). We also tested a variant of the VPO model with the prior probability p_{present} fixed to 0.5. Paired t tests showed no significant differences between the powers obtained with this model and those obtained with the "fixed prior" model ($p = 0.41$, $p = 0.48$, and $p = 0.71$ for $\kappa_D = 0$, 1, and 8, respectively).

Experiment 2

In Experiment 2, we aimed to dissociate distractor homogeneity from distractor predictability in set size effects on precision. We did this by making distractors homogeneous, but with a common orientation that varied across trials (Figure 7). The control condition was identical except that distractors were heterogeneous.

We found a significant effect of set size on hit rate (Figure 8a) in both the homogeneous, $F(3, 36) = 15.7$, $p < 0.001$, and heterogeneous, $F(3, 36) = 30.2$, $p < 0.001$, conditions. There was also a significant effect of set size on the false alarm rates in both conditions (homoge-

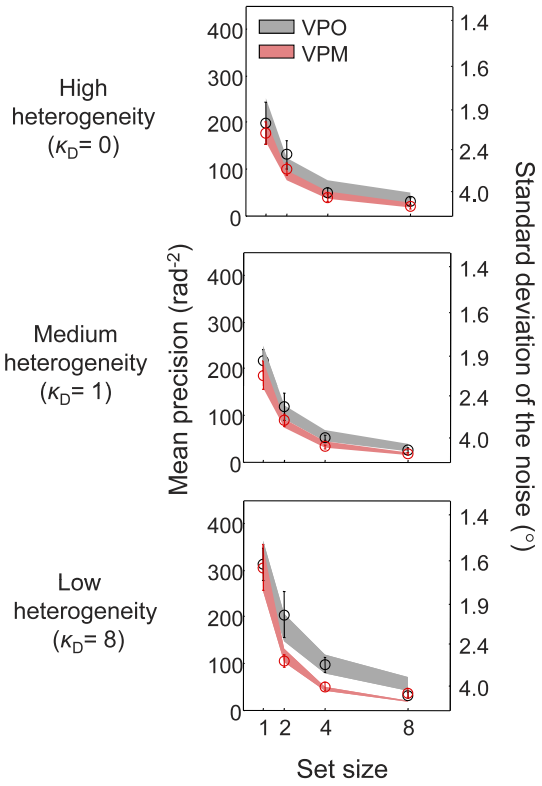

Figure 6. Dependence of precision on set size in Experiment 1. Estimates of mean precision parameter at each set size in the VPO (black) and VPM (red) models. The right *y*-axis shows the corresponding standard deviation of the Gaussian noise distributions, computed using the mapping $\sigma^2 = 1/(4\bar{J})$ (see Models). The shades represent the best-fitting power law (mean over subjects; width indicates 1 *SEM*).

neous: $F(3, 36) = 3.73$, $p < 0.05$; heterogeneous: $F(3, 36) = 6.90$, $p < 0.01$).

We fitted the VPO model to both conditions simultaneously. It provides good fits to hit and false-alarm rates (Figure 8a) and somewhat less good fits to the proportion of "target present" responses as a function of the target-distractor difference (Figure 8b). In those responses, both in the data and in the model, there is a dip in the target-present curves, especially noticeable at set sizes 4 and 8. This dip results from two opposing effects. First, if the distractor orientation is close to that of the target, the measurements of the distractors contribute to evidence for target presence; this effect diminishes with increasing target-distractor difference. Second, if the distractor orientation is very different from the target, it is easy to infer that one stimulus is different from all others and must therefore be the target; this effect diminishes with decreasing target-distractor difference. In the model, these counteracting effects are seen in the two factors in the summand of Equation 5: $e^{\kappa_i \cos 2(x_i - s_T)}$ accounts for the "target similarity effect," whereas $I_0(\kappa_{C=1,i})/I_0(\kappa_{C=0})$ reflects the "oddball effect." The former is obvious: The smaller the difference between the measurement x_i and the target orientation s_T, the larger $e^{\kappa_i \cos 2(x_i - s_T)}$. The latter requires a bit more thought. If we represent the i^{th} measurement as a vector in the plane with angle $2x_i$ and length κ_i, then $\kappa_{C=0}$, given by the second line of Equation 6, is the length of the vector sum of all measurements, and $\kappa_{C=1,i}$, given by the first line, is the length of the vector sum of all measurements except for the i^{th} one. These lengths are greater when the vectors in the sum are more aligned with each other, that is, when the measurements are closer to each other. Therefore, the factor $I_0(\kappa_{C=1,i})/I_0(\kappa_{C=0})$ is the answer to the question: If I were to leave out the i^{th} measurement, how strongly aligned would the remaining measurements be relative to the original alignment of all measurements? If the i^{th} measurement is different from all others, as will often be the case if it was produced by the target, then alignment would increase by leaving it out, and the factor $I_0(\kappa_{C=1,i})/I_0(\kappa_{C=0})$ would be greater than 1. Therefore, this ratio measures the strength of the evidence that the i^{th} stimulus is an odd element in

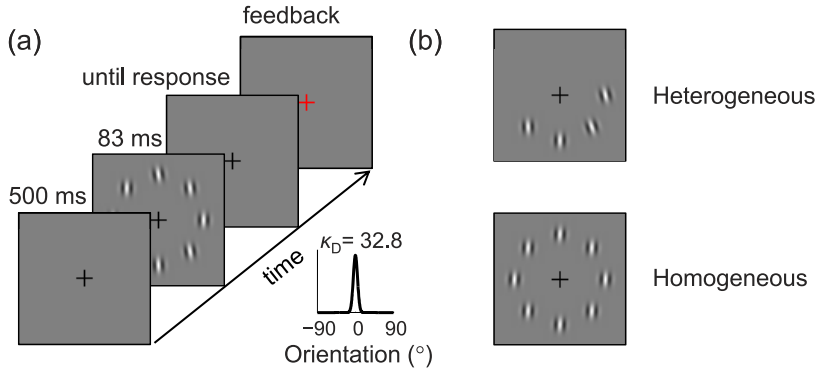

Figure 7. Experiment 2. (a) Time course of a trial (heterogeneous condition). Distractor orientations were drawn from a Von Mises distribution with a concentration parameter $\kappa_D = 32.8$ (corresponding to 5°; inset). Subjects judged whether a vertical target was present among the stimuli. (b) In the heterogeneous condition, distractor orientations were drawn independently. In the homogeneous condition, a common distractor orientation was drawn on each trial and assigned to all distractors on that trial.

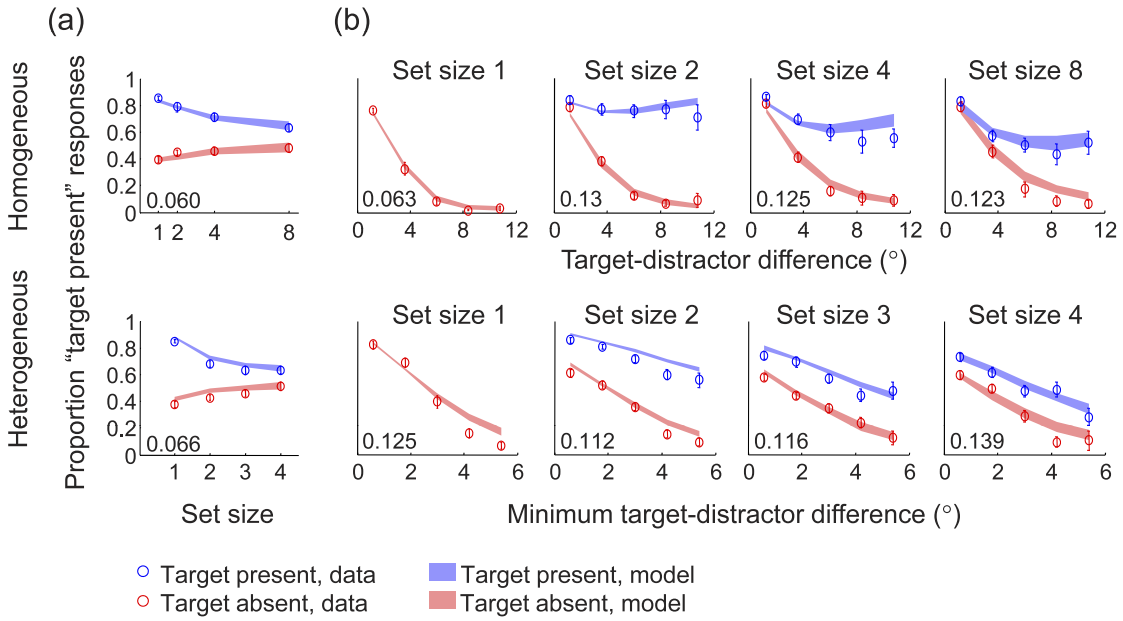

Figure 8. Data and VPO model fits from Experiment 2. Top row: homogeneous; bottom row: heterogeneous. (a) Hit and false alarm: data (circles and error bars) and VPO model fits (shaded areas). (b) Proportion "target present" responses as a function of the target-distractor difference, separately for target-present (blue) and target-absent (red) trials. For set size 1, there are no distractors on target-present trials. Note that the values on the *x*-axes and the set sizes are different in the homogeneous and heterogeneous conditions.

the display. In Experiment 1, since the distractors always differed amongst each other, there is no such factor.

The maximum-likelihood estimates of the power in the assumed power law relationship between mean precision and set size were -0.83 ± 0.11 (homogeneous) and -0.85 ± 0.13 (heterogeneous), which are not significantly different from each other ($t(12) = 0.10$, $p = 0.92$).

Mean precision and standard deviation of the noise are shown as a function of set size in Figure 9a. Power estimates across all conditions of Experiments 1 and 2 are compared in Figure 9b, presenting a consistent picture: All powers are clearly negative. Parameter estimates for Experiment 2 are listed in Appendix 2.

Discussion

Understanding how sensory precision depends on set size is essential for understanding how humans divide their attention over multiple objects, especially in split-second decisions. Here, we extended a previous study to examine this relationship under two manipulations of distractor statistics: changing the degree of heterogeneity and using homogeneous but unpredictable

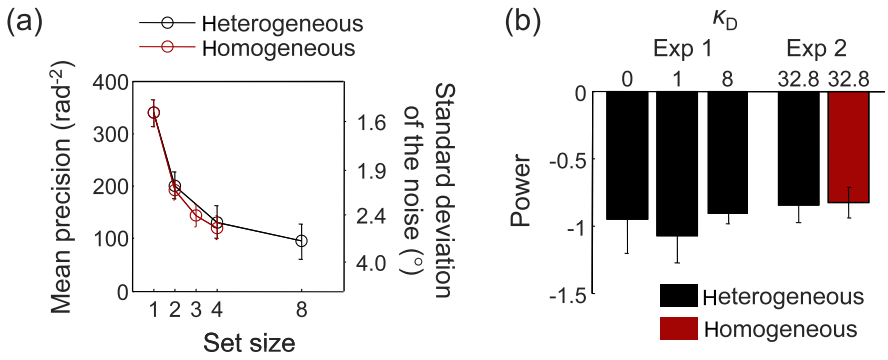

Figure 9. Comparison between conditions of the relationship between mean precision and set size in the VPO model. (a) Mean precision estimates for the heterogeneous (black) and homogeneous (red) conditions of Experiment 2. (b) Estimates of the power in the relationship between mean precision and set size for all conditions in Experiments 1 and 2.

distractors. To our surprise, we found a steep decrease of mean precision with set size in all conditions, with powers in the range of −1.3 to −0.8. This suggests that our earlier interpretation that the critical factor is whether distractors are heterogeneous or homogeneous (Mazyar et al., 2012) must be revised, and leads us to a new hypothesis: Precision is independent of set size ("capacity is unlimited") only when distractors are *predictable*. Stated otherwise, unless visual displays are largely predictable across trials, the spreading of visual attention has detrimental effects on the quality of encoding of each stimulus.

We should examine whether previous literature is consistent with the new hypothesis. Many earlier studies (Baldassi & Burr, 2000; Busey & Palmer, 2008; Mazyar et al., 2012; Palmer, 1994; Palmer et al., 1993; Palmer et al., 2000) reported no effect of set size on precision at least in some conditions. Except for two, these studies all used completely predictable distractors. The first exception was experiment 4 in Palmer (1994), in which distractors were "T" shapes that could be rotated 0°, 90°, 180°, or 270°. The task-relevant feature was the point at which the two line segments intersected. The distractors always had one line intersecting the other perfectly in the middle (thus the "T" shape) while the target could have one line intersecting the other line anywhere from one third the distance from the end of the line (offset "T" shape) to the end of the line ("L" shape). Thus, although the rotations caused variability, the distractors were predictable in the task-relevant feature. The second exception was experiment 3 of Mazyar et al. (2012), where the target orientation was on every trial drawn from a uniform distribution and identified to the subject through a pre-cue. Distractors were homogeneous and always tilted 5° clockwise with respect to the target orientation. Although the distractor orientation thus varied from trial to trial, observers did have enough information to predict this orientation after seeing the target cue. Altogether, we conclude that earlier studies are consistent with the predictability hypothesis.

In the above-mentioned studies that reported independence, distractors were not only predictable and homogeneous but also linearly separable between the target orientation and the set of possible distractor orientations (Bauer, Jolicoeur, & Cowan, 1996; D'Zmura, 1991). An account in which precision is independent of set size whenever distractor orientations—either homogeneous or heterogeneous—are drawn from a distribution that has mass only on "one side" of the target, as introduced by Rosenholtz (2001), is conceivable. Based on her results, however, one might expect a decrease of precision with set size as well. A concern is that such distractor distributions might be very difficult for subjects to learn.

In Experiment 2 we used homogeneous, unpredictable distractors to dissociate homogeneity from predictability. Further dissociation is possible by using heterogeneous, predictable distractors. In such a paradigm, which would correspond to the below-diagonal squares in the diagram in Figure 1, there would be variability in distractor orientations within a display but little or none across trials. In the extreme, distractors would be drawn from a uniform distribution but remain the same across trials. This seems an unnatural form of search but it would provide an independent test of our hypothesis.

A full description of the relationship between precision and set size requires filling two more squares in the diagram in Figure 1, in the top row. The top left square represents a condition similar to the homogeneous condition in Experiment 2 but with distractor orientations drawn from a uniform distribution, making the task too easy for subjects without additional modifications. When subject performance is near ceiling, models will generally be difficult to distinguish. The top center square could be realized by partially correlating the distractor orientations within a given display: One can think of the correlations between the distractor orientations within a display increasing within each row from zero on the diagonal (independent draws) to maximal on the left edge (homogeneous distractors). This case requires a study of its own, since it involves the question of whether humans can learn to take into account partial stimulus correlations in visual search.

If the hypothesis that precision is independent of set size only when distractors are predictable is confirmed, an important question is why this would happen. One reason we can imagine is that under complete predictability, the observer can form an internal template consisting of all distractors before the search display appears, and simply detect deviations from this template. This might be neurocomputationally more efficient than scrutinizing all stimuli, and this efficiency might translate to set-size independent estimated precision. Alternatively, it could be that every appearance of a set size effect on precision is the guise of an underlying form of suboptimal inference that the observer performs on measurements that are not subject to a set size effect (Beck, Ma, Pitkow, Latham, & Pouget, 2012). While this notion is conceptually appealing, it remains to be seen whether a plausible suboptimal inference model can be formulated that accounts for the data. Moreover, it would have to be explained why this form of suboptimality affects predictable displays much less than unpredictable ones.

Keywords: visual search, visual attention, capacity limitations, precision, Bayesian inference

Acknowledgments

WJM is supported by award R01EY020958 from the National Eye Institute and award W911NF-12-1-0262 from the Army Research Office. RLS is supported by award 5T32 GM007330-35 from the National Institute for General Medical Sciences.

*HM, RVDB, and RLS contributed equally to this article.

Commercial relationships: none.
Corresponding author: Wei Ji Ma.
Email: wjma@bcm.edu.
Address: Baylor College of Medicine, Houston, TX, USA.

References

Baldassi, S., & Burr, D. C. (2000). Feature-based integration of orientation signals in visual search. *Vision Research*, *40*, 1293–1300.

Bauer, B., Jolicoeur, P., & Cowan, W. B. (1996). Visual search for colour targets that are or are not linearly separable from distractors. *Vision Research*, *36*(101), 1439–1465.

Beck, J. M., Ma, W. J., Pitkow, X., Latham, P. E., & Pouget, A. (2012). Not noisy, just wrong: The role of suboptimal inference in behavioral variability. *Neuron*, *74*(1), 30–39.

Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon.

Busey, T., & Palmer, J. (2008). Set-size effects for identification versus localization depend on the visual search task. *Journal of Experimental Psychology: Human Perception & Performance*, *34*(4), 790–810.

D'Zmura, M. (1991). Color in visual search. *Vision Research*, *31*, 951–966.

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts effects of set size on visual search accuracy for feature, conjunction, triple conjunction and disjunction displays. *Perception & Psychophysics*, *62*, 425–451.

Hamilton, W. (1859). *Lectures on metaphysics and logic* (vol. 1). Boston: Gould and Lincoln.

Keshvari, S., van den Berg, R., & Ma, W. J. (2012). Probabilistic computation in human perception under variability in encoding precision. *PLoS ONE*, *7*(6): e40216, doi:10.1371/journal.pone.0040216.

Ma, W. J. (2012). Organizing probabilistic models of perception. *Trends in Cognitive Science*, *16*(10), 511–518.

Ma, W. J., Navalpakkam, V., Beck, J. M., Van den Berg, R., & Pouget, A. (2011). Behavior and neural basis of near-optimal visual search. *Nature Neuroscience*, *14*, 783–790.

Mazyar, H., van den Berg, R., & Ma, W. J. (2012). Does precision decrease with set size? *Journal of Vision*, *12*(6):10, 11–16, http://www.journalofvision.org/content/12/6/10, doi:10.1167/12.6.10. [PubMed] [Article]

Nolte, L. W., & Jaarsma, D. (1967). More on the detection of one of M orthogonal signals. *Journal of the Acoustical Society of America*, *41*, 497–505.

Palmer, J. (1990). Attentional limits on the perception and memory of visual information. *Journal of Experimental Psychology: Human Perception & Performance*, *16*(2), 332–350.

Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research*, *34*(13), 1703–1721.

Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception & Performance*, *19*(1), 108–130.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, *40*(10–12), 1227–1268.

Rosenholtz, R. (2001). Visual search for orientation among heterogeneous distractors: Experimental results and implications for signal detection theory models of search. *Journal of Experimental Psychology: Human Perception & Performance*, *27*(4), 985–999.

Schwartz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464.

Shaw, M. L. (1980). Identifying attentional and decision-making components in information processing. In R. S. Nickerson (Ed.), *Attention and performance* (vol. VIII, pp. 277–296). Hillsdale, NJ: Erlbaum.

Townsend, J. (1974). Issues and models concerning the processing of a finite number of inputs. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 133–168). Hillsdale, NJ: Erlbaum.

van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision

accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences, USA*, *109*(22), 8780–8785.

Verghese, P. (2001). Visual search and attention: A signal detection theory approach. *Neuron*, *31*(4), 523–535.

Vincent, B. T., Baddeley, R. J., Troscianko, T., & Gilchrist, I. D. (2009). Optimal feature integration in visual search. *Journal of Vision*, *9*(5):15, 1–11, http://www.journalofvision.org/content/9/5/15, doi:10.1167/9.5.15. [PubMed] [Article]

Appendix 1: Optimal model

Optimal model

Deriving predictions of the optimal-observer model consists of three steps: defining the generative model, computing the observer's decision rule (to be applied on each trial), and computing predictions for response probabilities across many trials. The generative models of Experiments 1 and 2 are shown in Figure A1. Each variable is a node. Variables are as follows: C (target presence, 0 or 1), T_i (target presence at the i^{th} location, 0 or 1), s_i (orientation at the i^{th} location), and x_i (measurement of orientation at the i^{th} location). The homogeneous condition in Experiment 2 has an extra variable, s_{D} (common distractor orientation). We denote the vector $(T_1,\ldots, T_N)$ by $\mathbf{T}$, and similarly for $\mathbf{s}$ and $\mathbf{x}$. A probability distribution is associated with each variable in the generative model. The equations are given in Figure A1. Some distributions are common to both the heterogeneous and the homogeneous condition. In the order of Figure A1, these distributions reflect that there is a probability that the target is present (p_{present}); that if the target is absent, it is absent everywhere; that if the target is present, it is only at one location, selected with equal probability; that the target always has value s_{T}; that the noise corrupting a measurement is independent across locations; and that it follows a Von Mises distribution centered at the true orientation (Equation 2). A few distributions are specific to the experiment. For heterogeneous distractors, these reflect that given target presence at a location, the orientation at that location is not influenced by other locations, and that each distractor orientation is drawn from a Von Mises distribution centered at s_{T} (Equation 1). For homogeneous distractors, the distributions reflect that a common distractor orientation s_{D} is drawn from a Von Mises distribution centered at s_{T}, that locations are "coupled" through this common orientation, and that every distractor orientation is equal to s_{D}. In both conditions, the optimal decision rule is to respond "target present" if $p(C=1|\mathbf{x})/p(\ C=0|\mathbf{x}) > 1$. We will now work out this expression using the respective generative models.

Derivation of the optimal decision rule for the heterogeneous conditions

For the scenario of heterogeneous distractors independently drawn from a distribution $p(s_i \mid T_i = 0)$, a single target with value s_{T}, and a general noise distribution $p(x_i \mid s_i)$, the posterior ratio is equal to (Ma et al., 2011)

$$\frac{p(C=1|\mathbf{x})}{p(C=0|\mathbf{x})} = \frac{p_{\text{present}}}{1-p_{\text{present}}}\frac{1}{N}\sum_{i=1}^{N}\frac{p(x_i|s_i=s_{\text{T}})}{\int p(x_i|s_i)p(s_i|T_i=0)ds_i}.$$

The summand is the likelihood ratio of target presence at the i^{th} location. In our task, the integral in the denominator can be evaluated as

$$\begin{aligned}\int p(x_i|s_i)p(s_i|T_i=0)ds_i &= \int \frac{1}{\pi I_0(\kappa_i)}e^{\kappa_i\cos 2(x_i-s_i)}\frac{1}{\pi I_0(\kappa_{\text{D}})}e^{\kappa_i\cos 2(s_i-s_{\text{T}})}ds_i \\ &= \frac{I_0(\kappa_{\text{combined}})}{\pi I_0(\kappa_i)I_0(\kappa_{\text{D}})},\end{aligned}$$

where $\kappa_{\text{combined}} = \sqrt{\kappa_i^2+\kappa_{\text{D}}^2+2\kappa_i\kappa_{\text{D}}\cos 2(x_i-s_{\text{T}})}$. Thus, the posterior ratio is

$$\begin{aligned}\frac{p(C=1|\mathbf{x})}{p(C=0|\mathbf{x})} &= \frac{p_{\text{present}}}{1-p_{\text{present}}}\frac{1}{N}\sum_{i=1}^{N}\frac{\frac{1}{\pi I_0(\kappa_i)}e^{\kappa_i\cos 2(x_i-s_{\text{T}})}}{\frac{I_0(\kappa_{\text{combined}})}{\pi I_0(\kappa_i)I_0(\kappa_{\text{D}})}} \\ &= \frac{p_{\text{present}}}{1-p_{\text{present}}}\times\frac{1}{N}\sum_{i=1}^{N}\frac{I_0(\kappa_{\text{D}})e^{\kappa_i\cos 2(x_i-s_{\text{T}})}}{I_0(\kappa_{\text{combined}})}.\end{aligned}$$

Derivation of the optimal decision rule for the homogeneous condition

The posterior ratio is obtained by marginalizing over $\mathbf{T}$, $\mathbf{s}$, and s_{D}:

$$\frac{p(C=1|\mathbf{x})}{p(C=0|\mathbf{x})}$$

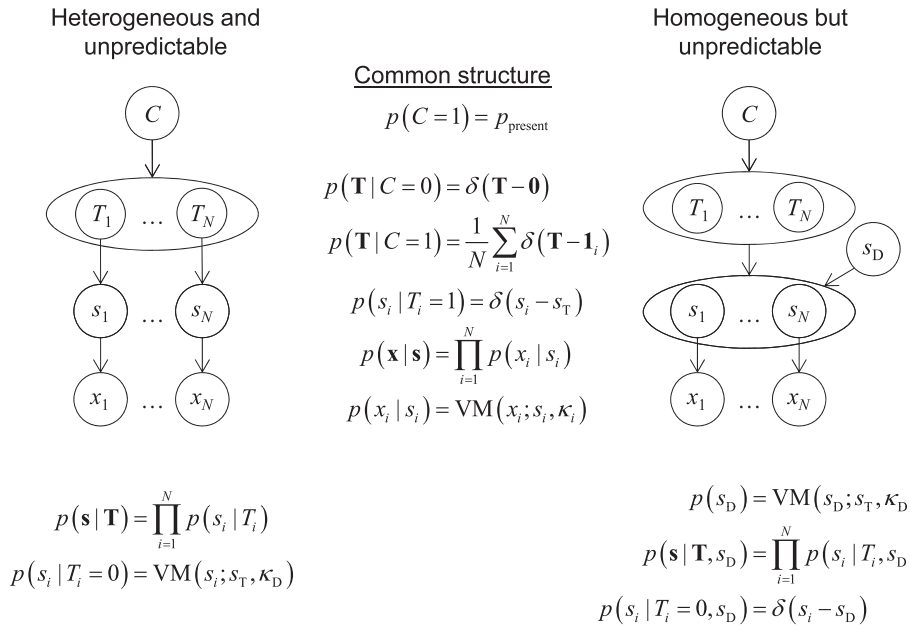

Figure A1. Generative models. These diagrams depict the dependencies between the variable of interest (target presence, *C*) and the measurements (**x**). Part of the statistical structure is shared (top), and part of it is specific to the experiment (bottom). **0** is the zero vector, and $\mathbf{1}_i$ is a vector of zeroes with a 1 in the i^{th} entry. VM stands for the Von Mises distribution on (−90°, 90°); in parentheses are its argument, its mean, and its concentration parameter.

$$= \frac{p_{\text{present}}}{1-p_{\text{present}}} \times \frac{\frac{1}{N}\sum_{i=1}^{N}\iint p(\mathbf{x}|\mathbf{s})p(\mathbf{s}|s_{\text{D}},\mathbf{T}=\mathbf{1}_i)p(s_D)d\mathbf{s}ds_{\text{D}}}{\iint p(\mathbf{x}|\mathbf{s})p(\mathbf{s}|s_{\text{D}},\mathbf{T}=\mathbf{0})p(s_{\text{D}})ds_{\text{D}}d\mathbf{s}}$$

$$= \frac{p_{\text{present}}}{1-p_{\text{present}}} \times \frac{\frac{1}{N}\sum_{i=1}^{N}p(x_i|s_i=s_{\text{T}})\int\left(\prod_{j\neq i}p(x_j|s_j=s_{\text{D}})\right)p(s_{\text{D}})ds_{\text{D}}}{\int\left(\prod_{i}p(x_i|s_i=s_{\text{D}})\right)p(s_{\text{D}})ds_{\text{D}}}$$

Note that the integral over s_{D} is outside the product, indicating that the s_{D} is common to all distractor locations. In Experiment 1, there was no such outer integral. In both the numerator and denominator we find an integral of a product of Von Mises distributions over s_{D}. For convenience, we define $x_0 = s_{\text{T}}$ and $\kappa_0 = \kappa_{\text{D}}$, so that $p(s_{\text{D}})$ formally becomes a factor in the product. Then the posterior ratio simplifies to

$$\frac{p(C=1|\mathbf{x})}{p(C=0|\mathbf{x})} = \frac{p_{\text{present}}}{1-p_{\text{present}}}\frac{1}{N}\sum_{i=1}^{N}e^{\kappa_i\cos2(x_i-s_{\text{T}})}\frac{I_0(\kappa_{C=1,i})}{I_0(\kappa_{C=0})},$$

with

$$\kappa_{C=1,i} = \sqrt{\left(\sum_{j\neq i}\kappa_j\cos2x_j\right)^2+\left(\sum_{j\neq i}\kappa_j\sin2x_j\right)^2},$$

$$\kappa_{C=0} = \sqrt{\left(\sum_{j=0}^{N}\kappa_j\cos2x_j\right)^2+\left(\sum_{j=0}^{N}\kappa_j\sin2x_j\right)^2}.$$

Model predictions

To produce predictions for human behavior, we applied the (optimal or $\max_d$) decision rule to a large number of simulated samples **x** drawn using the presented stimuli on a given trial. The result is a prediction for the probability of reporting "target present" on that trial for a given set of model parameter values. Maximum-likelihood parameter estimates were obtained by computing the joint response probabilities under a large number of parameter combinations (31 values per parameter). We verified that our results were insensitive to the range and discretization of the parameter space.

Appendix 2: Parameter estimates

Maximum-likelihood estimates of parameters in Experiment 1

The parameters in Tables 1-4 below are the mean precision at set sizes 1, 2, 4, and 8 (in rad^{-2}), the scale parameter of the gamma distribution over precision (in rad^{-2}), the observer's prior probability that the target is present, and the lapse rate.

Condition	J_1	J_2	J_4	J_8	p_{present}	λ
$\kappa_D = 0$	267 ± 99	64 ± 15	21.4 ± 4	27.2 ± 4.8	0.423 ± 0.024	$(6.2 \pm 1.2)\cdot 10^{-2}$
$\kappa_D = 1$	205 ± 73	71 ± 12	31.3 ± 3	45 ± 13	0.482 ± 0.021	0.110 ± 0.010
$\kappa_D = 8$	331 ± 75	133 ± 27	75 ± 17	28 ± 15	0.508 ± 0.013	0.123 ± 0.018

Table 1. Equal precision, optimal decision rule (EPO).

Condition	J_1	J_2	J_4	J_8	p_{present}	λ
$\kappa_D = 0$	$(31 \pm 12)\cdot 10$	53 ± 12	17.2 ± 3.6	5.44 ± 0.73	5.06 ± 0.32	$(1.22 \pm 0.36)\cdot 10^{-2}$
$\kappa_D = 1$	175 ± 44	54.8 ± 9.7	18.3 ± 2.7	10.4 ± 1.5	3.53 ± 0.22	$(1.22 \pm 0.58)\cdot 10^{-2}$
$\kappa_D = 8$	273 ± 49	92 ± 14	45.5 ± 9.4	37.2 ± 6	2.22 ± 0.12	$(1.11 \pm 0.54)\cdot 10^{-2}$

Table 2. Equal precision, max decision rule (EPM).

Condition	$\bar{J}_1$	$\bar{J}_2$	$\bar{J}_4$	$\bar{J}_8$	τ	p_{present}	λ
$\kappa_D = 0$	198 ± 45	131 ± 30	49 ± 11	31.2 ± 7.6	119 ± 28	0.508 ± 0.015	$(4.4 \pm 2.8)\cdot 10^{-3}$
$\kappa_D = 1$	218 ± 23	118 ± 30	51.7 ± 6.1	25.8 ± 4.9	120 ± 27	0.522 ± 0.011	$(1.7 \pm 1.1)\cdot 10^{-2}$
$\kappa_D = 8$	313 ± 33	205 ± 49	97 ± 16	31 ± 7	195 ± 66	0.525 ± 0.011	0 ± 0

Table 3. Variable precision, optimal decision rule (VPO).

Condition	$\bar{J}_1$	$\bar{J}_2$	$\bar{J}_3$	$\bar{J}_4$	τ	p_{present}	λ
$\kappa_D = 0$	178 ± 24	99 ± 13	39.1 ± 7.4	18.7 ± 3.4	64 ± 10	3.47 ± 0.34	$(7.8 \pm 2.0)\cdot 10^{-3}$
$\kappa_D = 1$	186 ± 30	88 ± 13	32.7 ± 3.6	17.1 ± 1.8	52.4 ± 6.1	2.56 ± 0.10	$(8.9 \pm 4.1)\cdot 10^{-3}$
$\kappa_D = 8$	306 ± 49	105 ± 13	49.7 ± 9.1	37.2 ± 8.2	51 ± 16	1.89 ± 0.14	$(1.1 \pm 1.1)\cdot 10^{-3}$

Table 4. Variable precision, max decision rule (VPM).

Condition	$\bar{J}_1$	β (power)	τ	p_{present}
Homogeneous $\kappa_D = 32.8$	339 ± 26	−0.83 ± 0.11	147 ± 38	0.506 ± 0.005
Heterogeneous $\kappa_D = 32.8$		−0.85 ± 0.13	208 ± 42	

Table 5.

Maximum-likelihood estimates of parameters in Experiment 2 (VPO model)

The parameters in Table 5 below are the mean precision at set size 1 (in rad^{-2}), the power in the power law dependence of mean precision on set size, the scale parameter of the gamma distribution over precision (in rad^{-2}), and the observer's prior probability that the target is present.