PERCEPTUAL ESTIMATION OF VARIANCE IN ORIENTATION AND ITS DEPENDENCE ON SAMPLE SIZE

by Steven A. Cholewiak

A thesis submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of requirements

for the degree of

Master of Science

Graduate Program in Psychology

Written under the direction of

Manish Singh

and approved by

____

____

____

____

New Brunswick, New Jersey

May, 2010

Abstract of the Thesis

Perceptual estimation of variance in orientation and its dependence on sample size

By Steven A. Cholewiak

Thesis director: Manish Singh

Recent research has shown that participants are very good at perceptually estimating summary statistics of sets of similar objects (e.g., Ariely, 2001; Chong & Treisman, 2003; 2005). While the research has focused on first-order statistics (e.g., the mean size of a set of discs), it is unlikely that a mental representation of the world includes only a list of mean estimates (or expected values) of various attributes. Therefore, a comprehensive theory of perceptual summary statistics would be incomplete without an investigation of the representation of second-order statistics (i.e., variance). Two experiments were conducted to test participants' ability to discriminate samples that differed in orientation variability. Discrimination thresholds and points of subjective equality for displays of oriented triangles were measured in Experiment 1. The results indicated that participants could discriminate variance without bias and that participant sensitivity (measured via relative thresholds, i.e., Weber fractions) was dependent upon sample size but not baseline variance. Experiment 2 investigated whether participants used a simpler second-order statistic, namely, sample range to discriminate dispersion in orientation. The results of Experiment 2 showed that variance was a much better predictor of performance than sample range. Taken together, the experiments suggest that variance information is part of the visual system's representation of scene variables. However, unlike the estimation of first-order statistics, the estimation of variance depends crucially on sample size.

Table of Contents

INTRODUCTION

Given the prevalence of ambiguity and noise in the environment, and the inductive nature of inferences in perception and cognition, most visual, visuo-motor, and cognitive tasks necessarily involve noisy or probabilistic representations. Whether the perceptual system is tasked with estimating scene parameters from images (Knill, Kersten & Yuille, 1996; Geisler & Kersten, 2002; Mamassian, Landy & Maloney, 2003) or to generate non-verbal numerical estimates (Burr & Ross, 2008; Cordes, Gallistel, Gelman & Latham, 2007), all representations involve uncertainty. Veridical representations of scene parameters would be a hindrance in a changing environment unless one could use past and present perceptual information in a dynamic manner, embracing the innate variability of the real-world. The fact that individuals have little difficulty in performing perceptual and cognitive tasks and can make judgments in a computationally optimal manner suggests that there may be explicit representations of uncertainty and variance in the perceptual system, which provides a wealth of information for planning future actions in an uncertain world.

All perceptual representations have local and global parameters associated with elements of the environment. Perceptual and cognitive systems need to be able to make computations with these representations in order to plan actions and to predict future events. Parameter estimates alone (whether individual values or means) are insufficient for most computations, which necessarily rely on variance information in a number of critical ways. For example, cue combination within and across modalities (e.g., visual/haptic integration) requires taking into account the variance in the systems' representations in order to combine noisy signals. Similarly, visually guided actions rely on perceptual estimates and need a representation of the distribution of relevant parameters in order to apply loss functions. Other examples where expected values of parameters alone are insufficient are source separation and categorization where the perceptual system attempts to determine whether there is a single underlying source for the observed samples, or multiple sources.

This highlights the question of what information perceptual systems should store. Should they store only the single “best estimate” value for each scene parameter, or should they store this best estimate along with the degree of uncertainty around this estimate? Given the necessity for variance in cue combination, visually guided actions (where different consequences have different “losses” associated with them), source separation, and categorization, a simple “best estimate” appears to be insufficient. However, this has not always been the view. A pictorial representation is superficially appealing because it would mean that we store parameters of a scene (e.g., object sizes, lengths, widths, orientations, etc.) for subsequent recreation and action. However, although estimating parameters could provide valuable information about a single scene/image, changing environments would require re-encoding of the parameters and significant processing and storage. The rigidity of the parameters and the inability to cope with changes in the scene make a pictorial representation of the environment an unappealing method of representation. Therefore, a more robust method of encoding is necessary.

A simple, universal axiom is that actions have consequences. In general, actions are planned in order to maximize some gain and minimize some loss. For example, a pitcher will throw a baseball with enough force so that it reaches the catcher’s mitt, but will adjust his/her trajectory so that the ball will cross through the “strike zone” and avoid hitting the batter and avoid straying far enough to be called a “ball”. To be able to throw a strike, the pitcher must have a representation of not only his motor variability, but also a loss function for the space of outcomes. If individuals only represented the mean, median, or mode of the parameters of the scene, they would not be able to apply loss functions to their decision spaces. This is because the entire probability distribution, not just the central tendency, is necessary for the use of loss functions. Therefore, one could only use loss functions if the distribution (or a way to recreate it) is stored.

Past research in decision making and action planning has found that participants have an implicit awareness of the uncertainty in planning motor actions and will utilize loss functions to maximize gain and minimize loss (Trommershäuser, Maloney, & Landy 2003; Landy, Goutcher, Trommershäuser, & Mamassian, 2007). Motor actions and perceptual decisions involve gains and losses associated with different outcomes and optimal performance thus depends on convolving a "loss function" with a noisy estimate / probabilistic representation of some variable. These experiments have shown that participants ' performance is close to optimal, which suggests that participants implicitly "know" their own perceptual/motor variability. If participants are aware that they are more variable in their motor actions, they will be more conservative in planning their actions to minimize loss. Therefore, individuals’ implicit awareness of their variability affects their motor strategies.

In addition to estimating and representing scene parameters pertaining to a single object / entity, the visual system must also deal with scenes that contain multiple similar elements. For example, cooks need to be quickly process the mean size, color, and shape of beans when sorting them to remove foreign objects (e.g., stones) and they need to rapidly assess variance when selecting root vegetables to ensure uniform size for even cooking. Similarly, quality control workers must be aware of not only the mean size of a product (say, eggs), but also how much each batch can vary in size, color, and quality before being classified as a higher/lower grade. Therefore, uncertainty or variance also arises naturally in situations where there are multiple similar elements with similar (but not identical) values along some property of interest (disc size, color, orientation, etc.). When a scene contains sets of multiple elements, as more objects are added to the set, more information must be processed and stored. While each parameter of an element within a scene has a lossy, variable encoding, there may also be representations of the distribution properties (i.e., summary statistics) of sets of multiple elements.

Summary statistics are representations of quantitative information about sets. They provide a way to represent large quantities of information along different dimensions. Research has shown that individuals can compute and represent the mean (Ariely, 2001; Chong & Treisman, 2003; 2005) for distributions of discs. However, as the previous examples demonstrate, probabilistic computations likely require not only the mean, but also the variance (and possibly higher-order statistics) of the relevant distributions.

Most work investigating the representation of summary statistics has used discs of varying radii as stimuli (Ariely, 2001; Chong & Treisman, 2003; 2005). Discs have been appealing because they provide a simple geometric stimulus, but they would not be ideal for investigating variance perception. The reason for this is that the relationship between physical “size” (which could mean area or radius/diameter for circles) and perceived size is highly non-linear, and moreover the precise mathematical form of this relationship is not fully known (but see Teghtsoonian, 1965). Therefore, in the current experiments, element orientation (rather than size) was used to investigate the representation of variance. One could use oriented line segments, but because orientation is a periodic parameter, stimuli could only be manipulated between 0° and 180° and any stimuli with variability would thus need to have a relatively small variance. Therefore, we chose to use oriented ("pointy") triangles, in order to maximize the range of distinct orientations (0° - 360°).

The current study thus addressed the following questions:

1. How well can participants estimate variance? Specifically, how precise is the representation of variance?
2. When subjects estimate dispersion in orientation, are they in fact using variance, or a simpler statistic such as sample range?

Two experiments were conducted to address these questions. In Sections 1 and 2, the methods and findings of Experiments 1 and 2 are presented, respectively. A general discussion follows in Section 3.

EXPERIMENT 1: SENSITIVITY TO SET VARIANCE

In order to determine if participants could generate and process representations of second-order summary statistics, Experiment 1 was designed to evaluate participants' abilities to compare samples of oriented isosceles ("pointy") triangles that differed in orientation variance. Experiment 1 specifically asked how sensitive participants were in estimating the variance of sets of oriented elements, and how their sensitivity varied with sample size. In addition, a post-hoc analysis of the variance in the experimental stimuli (i.e., the sampling distributions) was contrasted with participant performance in order to estimate the internal noise of participant's orientation variance representations.

I. Methods

a. Participants

Six graduate students (4 males and 2 females) at Rutgers University took part in the experiment. Participants S1, S4, and S6 had previous experience with psychophysical experiments. Participants S2, S3, and S5 were naïve participants who received compensation for their experimental time. The participants gave their written consent to the experimental protocol that had been approved by the Institutional Review Board at Rutgers University.

b. Stimuli

The stimuli were composed of a number of isosceles triangles with orientations drawn from von Mises distributions. Each triangle subtended 1.23° degrees of visual angle (DVA) and had a 5:1 (altitude:base) aspect ratio. The triangles were randomly placed in a circular area with diameter = 11.87° DVA, and were placed such that no two triangles overlapped (with a 0.09° DVA buffer between elements). For each stimulus, the sample of orientations was drawn from a von Mises distribution using the randraw MATLAB function (Bar-Guy, 2005) with a probability density function of:

$$f(x|\mu, \kappa) = \frac{e^{\kappa * \cos(x-\mu)}}{2\pi I_0(\kappa)}$$

where μ is the mean orientation of the distribution, κ is the concentration parameter, and I_0 is the modified Bessel function of order 0. The von Mises distribution was chosen for this experiment because it is considered the "circular normal distribution" and the concentration parameter (κ) closely approximates the reciprocal of the variance of a Gaussian distribution:

$$\lim_{\sigma \to 0} \sigma^2 = \frac{1}{\kappa}$$

In order to test the influence of sample size on variance perception, a 3x3 design was employed with baseline SD and sample size as independent variables. Three baseline, or standard, SDs (10°, 20°, and 30°) and three sample sizes (N = 10, 20, and 30) were used to create 9 different conditions. For each baseline SD, 8 comparison SDs (±10%, ±30%, ±50%, and ±70% of baseline SD) were used in the method of constant stimuli.

c. Procedure

A two-interval, forced-choice procedure was employed. On each trial, the participants received two presentations of stimuli (two intervals) and were asked to choose the stimulus with the higher orientation variance. On each trial within a condition, where baseline SD and N were constant, one of the stimuli (randomly chosen to be the 1st or 2nd) was drawn from a distribution with $\mu_{Baseline} = \text{rand} * 2\pi, \sigma_{Baseline} = Baseline\ SD$ and the other stimulus was drawn from a distribution with $\mu_{Test} =$

$\mu_{Baseline}, \sigma_{Test} = Test\ SD$, where the comparison, or test, SD was randomly chosen from the 8 possible comparison SD values.

Hence there were 9 conditions (3 baseline SDs × 3 sample sizes), and each condition consisted of 400 trials (8 comparison SDs × 50 repetitions), for a total of 3600 trials per participant. Each condition took participants approximately 20-25 minutes to complete and participants completed a maximum of 2 such sessions a day with at least a 15 minute break between conditions. Before each condition, participants were reminded that the triangles were placed randomly and that they should make their decisions based solely upon the variance in the triangles' orientations.

The participants were comfortably seated before a computer screen and a keyboard, and used a chin rest placed 1 meter from the screen. Practice trials were provided, where a series of representative stimuli were presented to familiarize the participant with distributions of orientations of triangles that had a fixed mean orientation but different variances. The example practice trials included stimuli with exaggerated variance differences to illustrate the task. Participants were encouraged to ask questions and to raise any issues with the experimenter. After the practice trials, the experimenter would initialize the experiment and leave the room.

The first interval in a trial was initiated by a keystroke on the computer keyboard. A fixation cross was flashed for 1000 ms, followed by the first stimulus for 300 ms, a mask for 50 ms, an inter-stimulus interval of 100 ms, the second stimulus for 300ms, a mask for 50ms, and then a black response screen requesting the participant to respond (see Figure 1). No feedback was provided. The participants were allowed to take a break half-way through each experimental session.

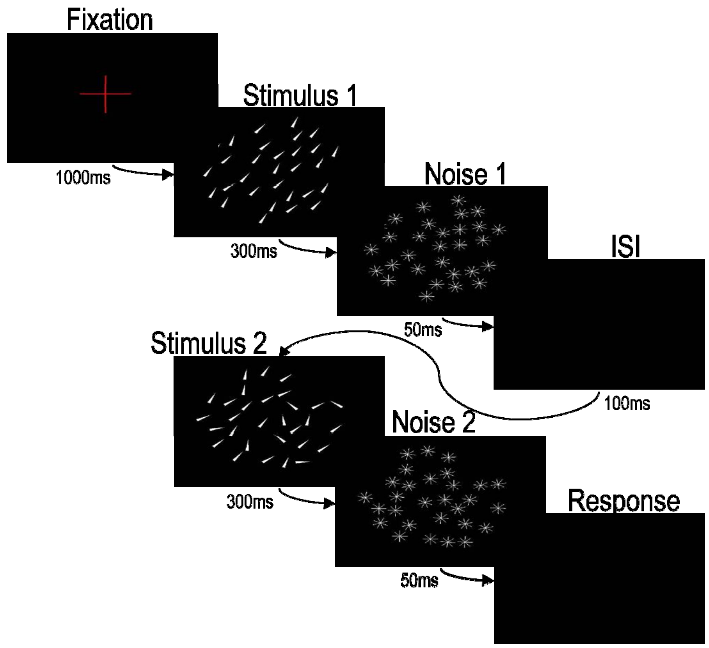


Figure 1. Example flow diagram for a single trial. Note that the second stimulus display has higher variance than the first. Note that the difference in variance between the two displays has been exaggerated in this Figure for illustrative purposes.

d. Data analysis

Stimulus parameters and participant responses were analyzed using the psignifit toolbox version 2.5.6 for Matlab R2008b, which implements the maximum-likelihood method and confidence interval bootstrap methods described by Wichmann and Hill (2001a; 2001b). Analyses were conducted to address the following questions:

- How sensitive/precise are participants in estimating the variance of a set of oriented elements?
- How does this sensitivity vary with sample size?

II. RESULTS

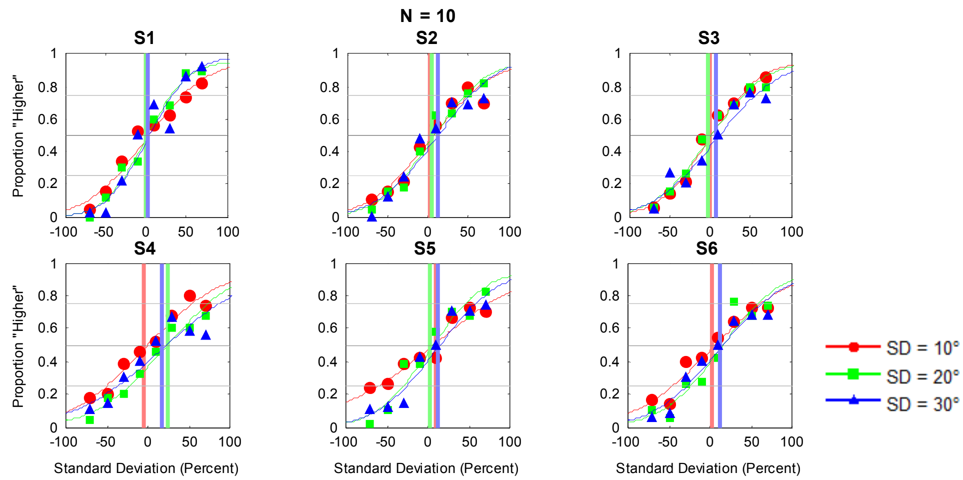
a. Bias

Confidence intervals indicated that most participant PSEs did not significantly differ from the baseline SDs for the 9 conditions. The baseline SD did not fall out outside of the 95% confidence intervals for S1 or S3. Participant S2 had 1 positively biased PSE, S5 had 2 positively biased PSEs, S4 had 4 positively biased PSEs, and S6 had 6 positively biased PSEs. Although most participants had few, if any, biases for individual conditions, only one, S6, had a consistently positive bias.

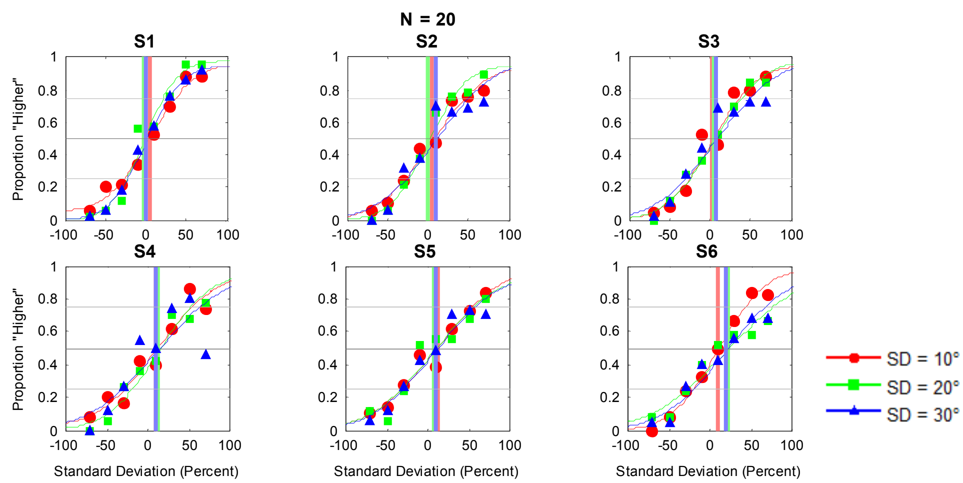
b. Effect of Baseline Variance

In order to determine the effect of baseline SD on observer performance, psychometric curves were fit to the participants' data using psignifit (Wichmann & Hill, 2001a; 2001b) and performance was compared within-participant between baseline SD conditions, keeping sample size constant. If variance perception, as measured by this experimental task, conforms to Weber's Law, then psychometric curves with constant N should have similar shapes (sigmas) and there should be no effect of baseline variance. The plots in Figure 2 suggest that this is approximately true. We will return to this question when we report the difference thresholds.

A.



B.



C.

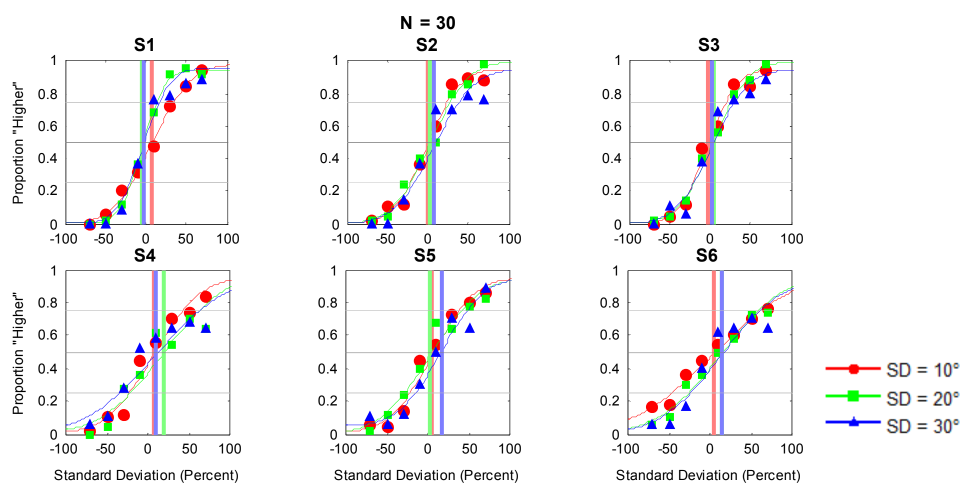
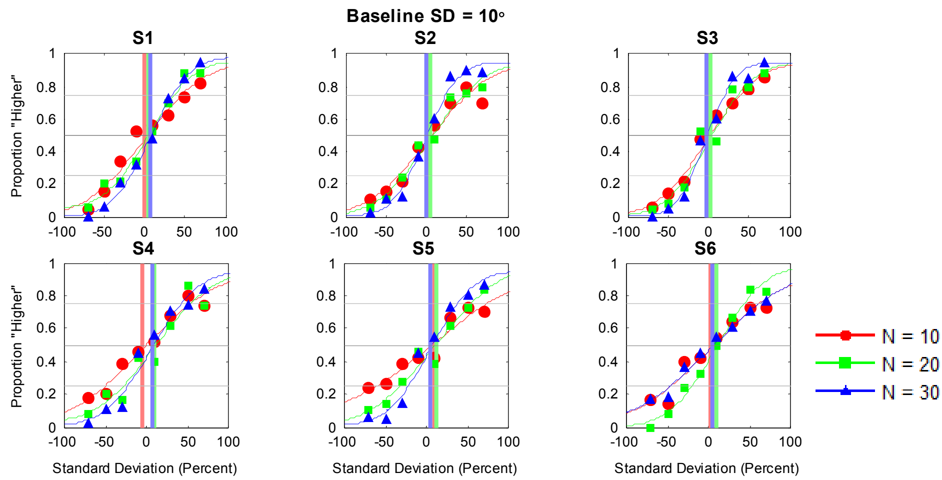


Figure 2. Within-participant psychometric curves for 6 participants with sample sizes of 10 (A), 20 (B), and 30 (C) elements held constant and varying baseline SD: 10° (red), 20° (green), and 30° (blue).

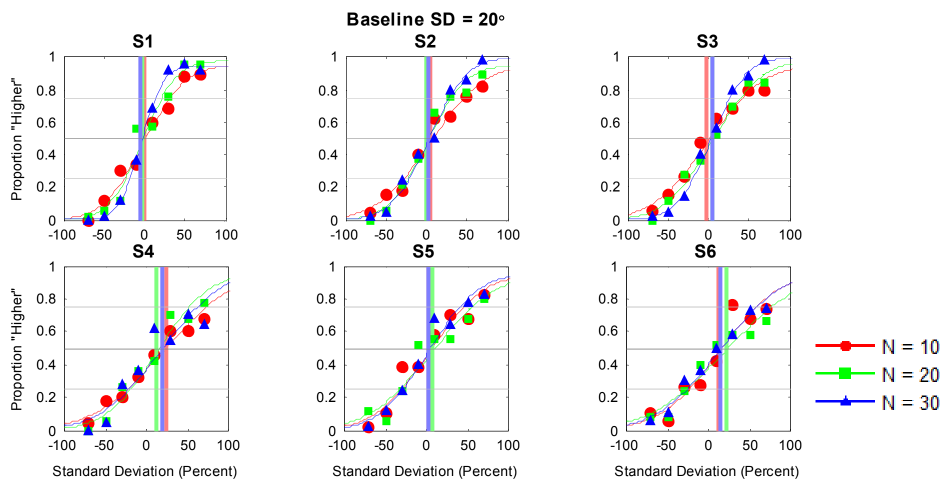
c. Effect of Sample Size

Figure 3 shows the same data and psychometric curves rearranged to show the influence of sample size. Interestingly, the sigmas of the psychometric curves for participant responses, which are related to the participants' sensitivity, appeared to increase as sample size increased (see Figure 3).

A.



B.



C.

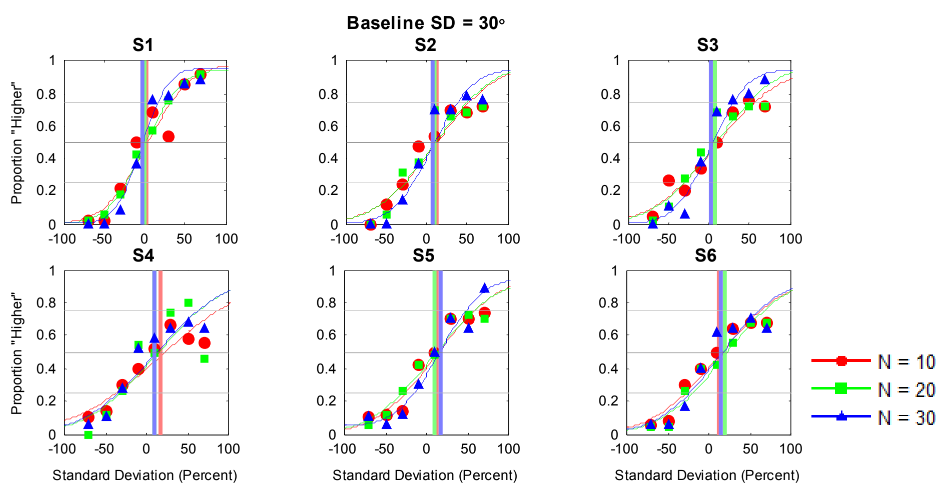


Figure 3. Within-participant psychometric curves for 6 participants with baseline SDs of 10° (A), 20° (B), and 30° (C) held constant and varying sample size: N = 10 (red), 20 (green), and 30 (blue) elements.

d. Sensitivity

The apparent trend of increasing sigma of the psychometric fits with increasing sample size warranted additional analysis, so difference thresholds, using Gaussian CDF best-fits, were estimated using psignifit's curve fitting toolbox (Wichmann & Hill, 2001a; 2001b). Difference thresholds were defined as the PSE-75% "higher" threshold, estimated using the psychometric curve. Note that the difference threshold is a raw, unnormalized threshold in degrees for each participant for each condition. For each participant, as baseline standard deviation increased, so did the difference threshold (see Figure 4). For most participants and most conditions with constant sample size and increasing baseline SD, this increase was linear, as would be expected from Weber's Law. In addition, for most participants, larger sample sizes resulted in lower difference thresholds.

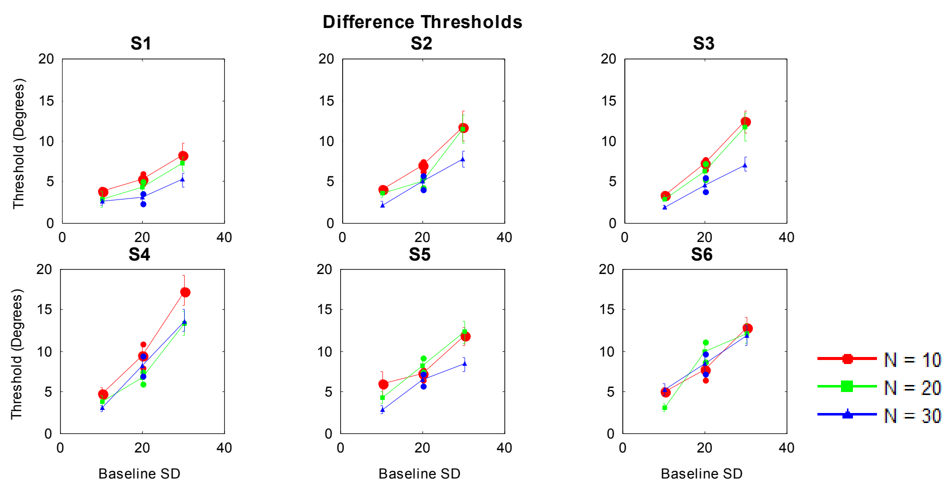


Figure 4. Difference thresholds for 6 participants for 9 conditions: N = 10 (red), 20 (green), 30 (blue) elements × baseline SD = 10°, 20°, 30°.

In order to directly test for Weber-type behavior, the difference thresholds were converted into relative thresholds (Weber fractions), which are the difference thresholds normalized by the baseline SDs. Figure 5 shows that the relative thresholds are largely independent of baseline SD, as would be expected from Weber's Law. However, relative thresholds tended to decrease with sample size: the larger the sample size, the more precisely participants can estimate the variance.

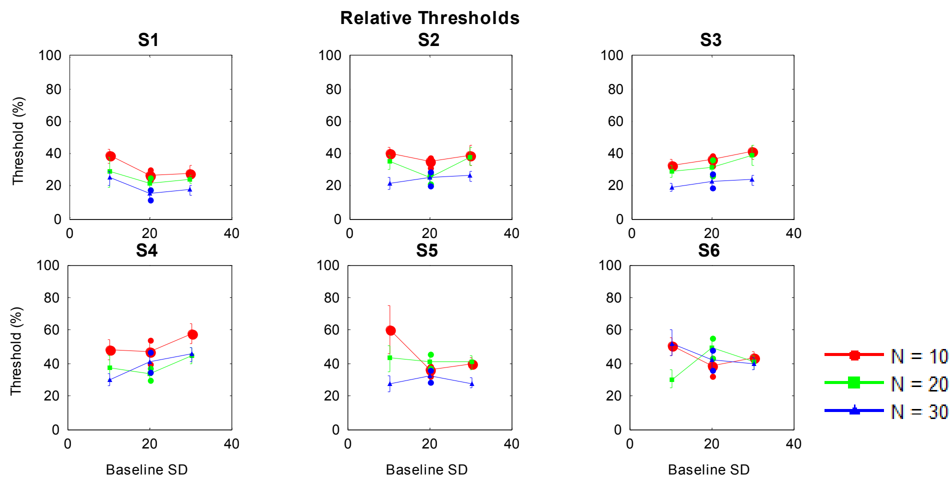


Figure 5. Relative thresholds for 6 participants for 9 conditions: N = 10 (red), 20 (green), 30 (blue) elements × baseline SD = 10°, 20°, 30°.

e. Internal vs. External Noise

The encoding of external stimuli is an inherently noisy process. As illustrated in the process flow diagram from Gallistel and King (2009), scene perception can be distorted by external physical noise sources that act upon the distal stimulus or internal biophysical noise sources that can alter the sensory signal (see Figure 6).

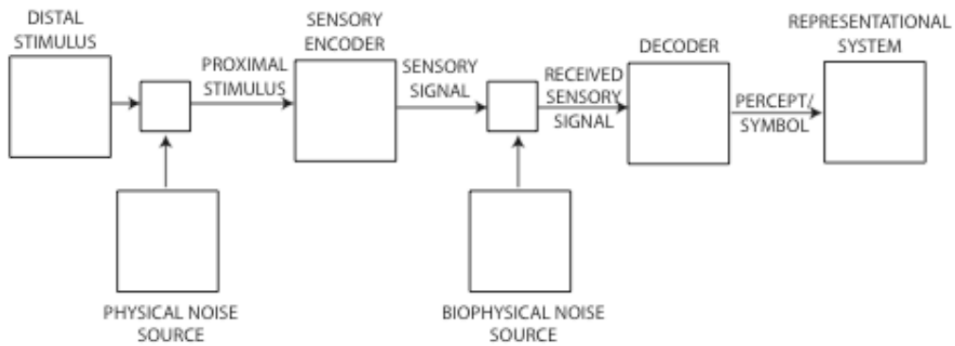


Figure 6. Process flow diagram of the encoding of a distal stimulus into a percept (Gallistel & King, 2009).

In the Experiment 1, physical noise was controlled, such that a single sample of orientations presented repeatedly would generate the same proximal stimulus, assuming other environmental factors (distance from the stimuli, fatigue, etc.) are fixed. However, note that from trial to trial the samples of orientations are drawn from a probabilistic distribution, the von Mises distribution and, therefore, there is variance from one sample (on a particular trial) to the next. When orientations are drawn from an underlying distribution with a given population mean and population variance, the mean and variance of the sample will vary around the population parameters. These sampling distributions have a mean and variance (the means of the sampling distributions of the mean and variance, and the variances of the sampling distributions of the mean and variance) and can thus be used to create observer models.

An ideal computational model would have an absence of external physical noise or internal noise sources and would respond with perfect precision and accuracy for every trial. However, due to sample-to-sample variance from samples being drawn from a probabilistic distribution, some trials may have a stimulus drawn from a higher population variance condition, but have a lower sample variance. This variance is manifested as variability in the sample's variance around the mean variance (i.e., the baseline SD) for the condition and is quantified across the course of the experimental condition as the standard error of the variance (i.e., the standard deviation of the sampling distribution of variance). If the ideal computational model then made its noise-free decisions, then the model's psychometric curve would have a sigma (i.e., standard deviation of a Gaussian psychometric fit) equal to the standard error of the variance.

However, humans are not noise free and have numerous internal sources of biophysical noise. Although the performance would be predicted by the standard deviation of the variance for a perceptual system with no internal noise, the human visual system is noisy and noisy representations will ultimately decrease the performance. When representations of the samples of orientations (the distal stimuli) are generated, there are two primary possible sources of noise:

1. The representation of the orientation of individual elements.
2. The representation of the orientation variance of the sample.

Each orientation will thus be represented with some variability and so will the holistic representation of their variance in the scene, so human performance could only be lower than would be predicted by the standard error of the variance.

The sigmas of the psychometric curves (Figure 7) are a measurement that includes both the internal and external sources of noise. It is the parameter fit to the individuals' psychophysical data that represents an individual's variance in their own representation of the variance. Consistent with the difference threshold analysis (Figure 4), as baseline SD increased, so did the sigma of the psychometric curves.

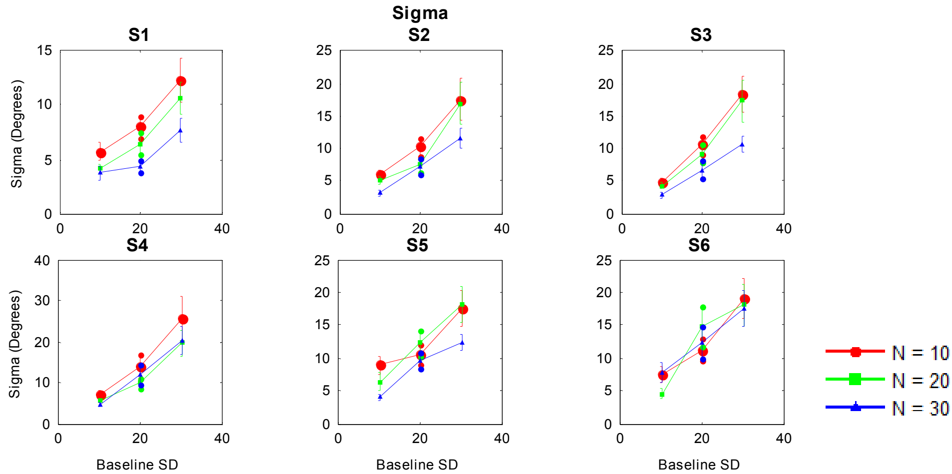


Figure 7. Psychometric curve sigmas for 6 participants for 9 conditions: N = 10 (red), 20 (green), 30 (blue) elements × baseline SD = 10°, 20°, 30°.

In order to determine the contribution of internal sources of noise (representation of the orientation of individual elements, and representation of the sample variance), participant performance was compared for each condition to a model with no internal noise. By plotting the sigma of the Gaussian psychometric curve fits against the standard error of the variance for each condition, one can estimate the contribution of the internal sources of variability (see Figure 8). Assuming internal noise affects performance in a constant manner, a linear best fit line was fit to the data to estimate how well participants performed compared to an ideal noise-free observer model. Participants appeared to have representations of the sampling distributions that are approximately 1/3 as precise as an ideal model due to the internal biophysical noise sources (S1: 55% of ideal, S2: 31% of ideal, S3: 29% of ideal, S4: 21% of ideal, S5: 33% of ideal, S6: 27% of ideal). This suggests that, on average, two-thirds of the noise/imprecision in participants' performance is due to internal sources of noise, and only one-third is due to external sources (sampling variability).

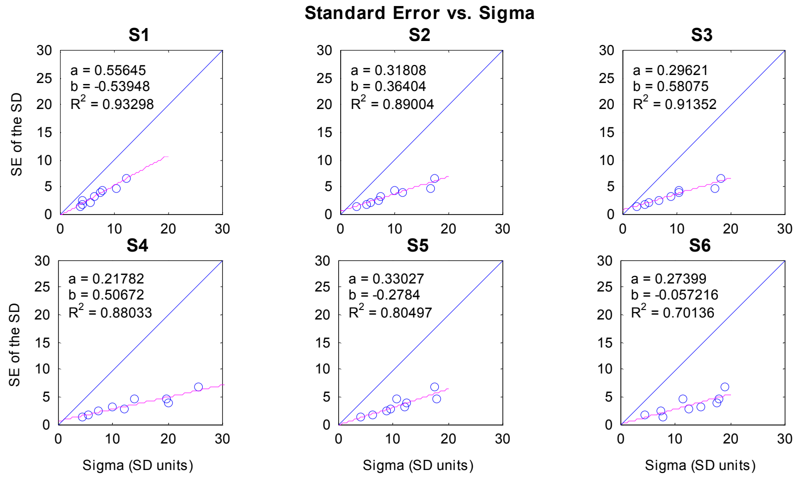


Figure 8. Comparing the psychometric curve sigmas versus the standard error of the standard deviation for 6 participants for 9 conditions. Pink lines are linear fits.

III. Discussion

The results of Experiment 1 showed participants are sensitive to orientation variance and that their relative thresholds (difference thresholds normalized by baseline SD) were largely independent of baseline SD for the majority of participants. No bias would be predicted for participant responses given the two-interval, forced-choice higher/lower paradigm employed.

A more interesting result was that the general trend for relative thresholds was to increase with decreasing sample size. That is to say, as sample sizes became smaller, participants needed a larger difference in the baseline and comparison variances before they could detect a difference. Participants relative thresholds (Weber fractions) ranged from approximately 15% for n = 30 and up to > 40% for n=10. Therefore, participant sensitivity to the variance summary statistic was dependent upon the number of elements in the scene. This result contrasts sharply with previous research on first-order summary statistics: the estimation of the mean disc size in a set has been shown to be largely unaffected by set size (in the range of 1-12 elements used by Chong & Treisman, 2003; and for sample sizes ranging from 4-16 elements for Ariely, 2001).

Further analysis of the participants' psychometric curves revealed that participants appear to have representations of the sampling distributions that are approximately 1/3 as precise as ideal representations, which assume no noise in the representation of individual orientations and representation of sample variance.

EXPERIMENT 2: SAMPLE RANGE VS. VARIANCE

The results from Experiment 1 clearly showed that participants are sensitive to the dispersion of orientations and could judge elements drawn from a higher variance population distribution to be different from samples drawn from a lower variance population. However, these results do not necessarily show that participants are sensitive to variance per se. A natural, and simpler, alternative in performing this task and judging the dispersion would be to use sample range. This is due to the fact that the range statistic is diagnostic of the variability (i.e., higher range is correlated with higher variance). Indeed, as long as the baseline and comparison samples are drawn from the same distribution, it is difficult to tell whether observers are using SD or sample range. Therefore, in order to experimentally distinguish between the strategies of using sample range or variance, one must set up a situation in which the two statistics would yield different predictions. We do this by drawing samples from different population distributions for the baseline and comparison stimuli displays.

The strategy for this experiment was then to use Uniform distribution for baseline, or standard, samples and Gaussian distribution for comparison, or test, samples, such that samples drawn from the two distributions would have the same sample variance, but different sample ranges. Participants would be asked to compare the variability of the baseline and comparison samples and if participants chose to use range instead of variability, one would predict a bias in their psychometric curves due to the differing sample ranges.

In order to compute the magnitude of the predicted bias, a Monte Carlo simulation was used to calculate the expected PSE if participants used sample range instead of variance. The basic idea was to draw samples from the 2 distributions with identical sample ranges and to measure the sampling distributions' standard deviations to compare predicted PSEs (see Figure 9). 100,000 samples were drawn from separate Uniform population distributions with population means of 0, sample sizes of 10, 20, 30, and 40 elements, and population ranges of 30, 60, 90, and 120 degrees (for a total of 16 sampling distributions, 4 sample sizes × 4 ranges). From these Uniform samples, sampling distributions were derived and the sample means, sample ranges, and standard deviations were estimated. Gaussian sample distributions with constant sample sizes were successively drawn while increasing the population standard deviations from 0 in increments of 0.01° until the Gaussian mean sample ranges were equal or greater than of the Uniform mean sample ranges. This resulted in Uniform and Gaussian sampling distributions with mean sample ranges within 0.5° of each other, but with differing sample variances (see Table 1 for the results). Samples with $n = 40$ led to the greatest difference between the sample standard deviations between the Uniform and the Gaussian distributions. Therefore, samples with 40 elements were chosen to maximize the predicted bias for this experiment. For samples with 40 elements, the mean Gaussian sample SD were, on average, ~23% lower than the mean Uniform sample SD with the same mean sample range. Therefore, the prediction would be that if participants use sample range when making judgments of dispersion, their PSEs would be 23% lower than the Uniform distributions' mean sample variances.

$$
\begin{array}{ccc}
\overline{W}_{Uniform} & \leftrightarrow & \overline{sd}_{Uniform} \\
|| & & \\
\overline{W}_{Gaussian} & \leftrightarrow & \overline{sd}_{Gaussian}
\end{array}
$$

Figure 9. Commutative diagram illustrating the strategy for Experiment 2, where the distributions' mean sample ranges ($\overline{W}_{Uniform}$ and $\overline{W}_{Gaussian}$) were equated, leaving the mean sample standard deviations ($\overline{sd}_{Uniform}$ and $\overline{sd}_{Gaussian}$) to vary. See Table 1 for calculations.

Table 1.
Sample statistics from Uniform and Gaussian sample distributions with N = 100,000 samples.

Sample Size	Uniform Population Range	Uniform Sample SD	Uniform Sample Range	Gaussian Sample Range	Gaussian Sample SD	1 – (Gaussian SD / Uniform SD)
10	30	8.54	24.53	24.62	7.78	0.0890
10	60	17.08	49.08	49.24	15.57	0.0884
10	90	25.62	73.62	73.80	23.32	0.0899
10	120	34.15	98.16	98.27	31.04	0.0911
20	30	8.613	27.15	27.30	7.21	0.1627
20	60	17.21	54.27	54.54	14.41	0.1630
20	90	25.84	81.45	81.84	21.61	0.1638
20	120	34.44	108.59	108.61	28.70	0.1668
30	30	8.63	28.06	28.17	6.84	0.2077
30	60	17.26	56.13	56.43	13.69	0.2067
30	90	25.89	84.20	84.48	20.51	0.2079
30	120	34.50	112.26	112.33	27.26	0.2099
40	30	8.64	28.54	28.97	6.66	0.2293
40	60	17.28	57.07	57.43	13.21	0.2357
40	90	25.92	85.63	85.92	19.76	0.2378
40	120	34.56	114.16	114.58	26.33	0.2381

Because the baseline distribution (Uniform) and the comparison distribution (Gaussian) had sample variances equated, if participants actually used variance when judging the dispersion of the samples, one would predict a bias of 0% of the baseline SD. However, if participants instead used range when judging the dispersion, although they would be explicitly instructed to use the variability, then one would predict a bias of -23% of the baseline SD.

I. METHODS

a. Participants

The six graduate students from Exp. 1 participated in Exp. 2. The participants gave their written consent to the experimental protocol that had been approved by the Institutional Review Board at Rutgers University.

b. Stimuli

As in Exp. 1, stimuli were composed of a number of isosceles triangles. Each triangle had a 5:1 (altitude:base) aspect ratio and subtended 1.23° degrees of visual angle (DVA). The triangles were randomly placed in a 13.69° DVA circular area, with the same density as the N=30 conditions in Exp. 1, and were placed such that no two triangles overlapped.

For each trial, samples were drawn from Uniform and Gaussian distributions using the randraw MATLAB function (Bar-Guy, 2005).

As described in the discussion of the Monte Carlo simulation, in order to maximize the expected bias, sample size was fixed at n = 40. Three baseline SDs (baseline SD = 10°, 20°, and 30°) drawn from Uniform distributions with were used. For each baseline SD condition, 9 comparison SDs (±10%, ±30%, ±50%, ±70%, and -90% of baseline SD) were used to create psychometric curves. The comparison SDs were chosen to straddle the two possible predictions (0% PSE bias if sample variance was used, -23% PSE bias if sample range was used) and to provide ample sampling space for individual differences is dispersion perception.

c. Procedure

As in Exp. 1, a two-interval, forced-choice procedure was employed. On each trial, the participants saw two presentations of stimuli (two intervals) and were asked to choose the stimulus with the higher orientation variance. On each trial within a condition, where baseline SD was constant, one of the stimuli (randomly chosen to be the 1st or 2nd) was drawn from a the Uniform distribution with $\mu_{Baseline} = rand * 2\pi, \sigma_{Baseline} = Baseline\ SD$ and the other stimulus was drawn from a Gaussian distribution with $\mu_{Test} = \mu_{Baseline}, \sigma_{Test} = Test\ SD$, where the comparison SD was randomly chosen from the 9 possible comparison SD values.

Initial training was provided and participants were encouraged to ask questions and to raise any issues with the experimenter. After practice, the experimenter would initialize the experiment and leave the room.

Trial structure and timing were the same as in Exp. 1.

d. Data analysis

Stimuli and participant responses were analyzed with MATLAB using psignifit. Of specific interest, as stated in the Exp. 2 introduction, were the participants' biases (difference between the respective PSEs and baseline variance value).

II. Results

Psychometric curves were fit to the participants' data and participants' PSEs were compared to the predicted biases (0% if participants used variance, -23% if participants used range). As illustrated in Figure 11, every participant responded with essentially no bias. When they deviated from zero, biases were small and positive.

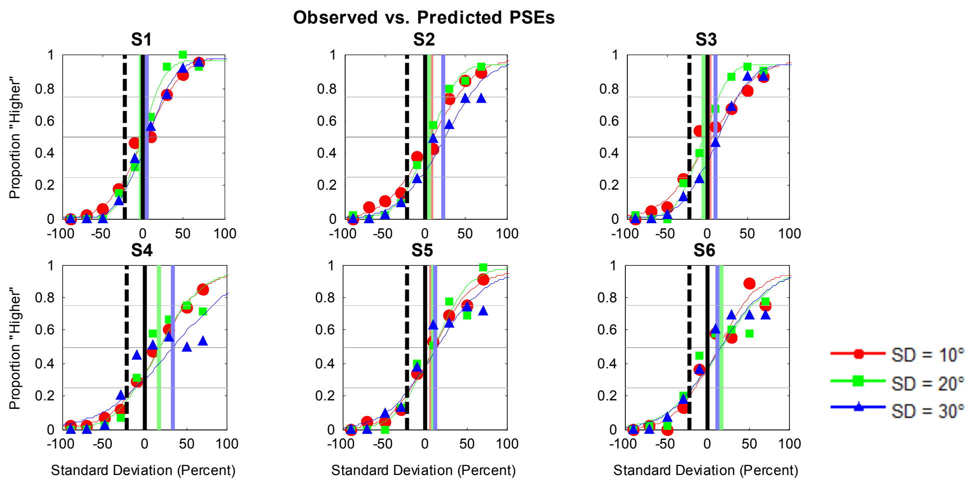


Figure 10. Within-participant psychometric curves for 6 participants. Each curve represents a baseline SD of 10° (red), 20° (green), or 30° (blue). The colored vertical lines represent the observed PSEs, the black dashed line represents the predicted PSE if participants used range, and the black solid line represents the predicted PSE if participants used variance.

95% confidence intervals indicated that all participant PSEs were significantly different from the -23% bias predication for all conditions, which rules out the usage of sample range as a response method for all participants (see Figure 11).

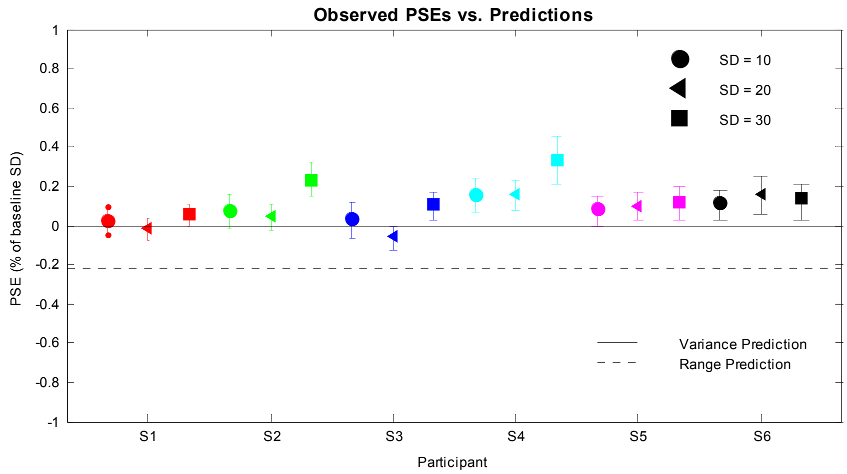


Figure 11. Participant PSEs versus predicted PSEs for usage of Variance or Range in variance judgement tasks. The black dashed line represents the predicted PSE if participants used range and the black solid line represents the predicted PSE if participants used variance. Error bars are 95% confidence intervals.

III. Discussion

The results of Experiment 2 provide supportive evidence that participants were using the variance, rather than the range, when making judgments about the dispersion of orientations in the experimental stimuli. All of the observed PSEs were significantly different from the range prediction and were consistent with the prediction for usage of sample variance. Experiment 1 suggested that participants were sensitive to variance and Experiment 2's results agree. Reassuringly, observers appear to be using the variance when instructed to use variance and not sample range.

Participants' PSE were precise; that is, the PSEs for different baseline SDs did not vary to a significant degree. As with Experiment 1's constant N results, there was no significant effect for Experiment 2's PSEs.

CONCLUSIONS

The world is highly stochastic and variable, with pervasive ambiguity and noise, and so humans have adapted to cope with an ever-changing environment. Every organic and inorganic object has a range of properties that vary from object to object. Some varying in imperceptible ways, such as the differing sizes of grains of sand on the beach, while others can be so great that the sheer magnitude boggles the mind (e.g., sizes of planets, stars, and galaxies). For most sets of similar objects that we encounter in our everyday lives (e.g., eggs, kiwis, basketballs, etc.), there are limited ranges of perceptibly differing properties. These properties, such as orientation, size, and color, provide cues as to how we should interact, handle, and relate the sets of objects. And because our perceptual systems exist to not only provide information about what is happening in the present, but to guide us as we plan our future decisions and actions in an uncertain world, these properties can help us identify how we should interact with the sets, from choosing to medicate a seemingly sick herd of sheep to making the decision to purchase a ripe basket of strawberries. Because statistical summary information is so valuable, it is beneficial to have explicit representations of uncertainty and variance in the perceptual system.

Previous research has investigated the representation of summary statistics, but has focused primarily on the representation of the mean size of sets of discs (Ariely, 2001; Chong & Treisman, 2003; 2005). Variance perception has been an overlooked, but critically important, summary statistic that allows for perceptual processes to categorize and produce goal-directed actions. It allows for cue combination within and across modalities and for loss functions to be applied to visually guided actions.

The experiments in this study specifically addressed the question of whether orientation variance is represented in the brain and how sensitive humans are to changes in orientation variance. Experiment 1 asked: How well can participants estimate variance and how precise is the representation of variance? And Experiment 2 investigated whether participants used a “simpler” summary statistic, range, rather than variance when estimating dispersion in orientation.

Experiment 1 clearly showed that not only was the difference threshold a relatively constant proportion of base SD for the range of SDs tested, but interestingly that the relative thresholds (Weber fractions) were sample size dependent. And although the visual system is clearly sensitive to variance, the sensitivity appears to be not as high as sensitivity to disc size, where discrimination thresholds are ~6-8% (Chong & Treisman, 2003). However, for a direct comparison to perception of the mean, we will need to test the sensitivity to the mean for orientation. In order to directly compare the sensitivity of variance perception to perception of the mean for sets of elements, one needs to compare sensitivities in the same domain (e.g., disc size, orientation, etc.). Although ideally this would mean extending the methodologies and analyses to the domain of disc size, the mapping between physical disc size and perceived size is less clear (as discussed in the Introduction).

The results of Experiment 2 showed that participants were not using range when they were instructed to make their judgments of dispersion based upon variance. These results support the methodology and conclusions of Experiment 1.

The results of Experiments 1 and 2 support the conclusion that the human perceptual system has representations of the variances of sets of elements and that it can provide a wealth of information for actions and decision making.

REFERENCES

Ariely, D. (2001). Seeing sets: representation by statistical properties. *Psychological Science, 12,* 157-162.

Bar-Guy, A. (2005). Randraw: Efficient random variates generator. Retrieved August 8th, 2008 from http://www.mathworks.com/matlabcentral/fileexchange/7309.

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

Chong, S.C. & Treisman, A. (2005). Statistical processing: computing the average size in perceptual groups. *Vision Research, 45,* 891–900.

Gallistel, C.R. & King, A.P. (2009). Memory and the computational brain: Why cognitive science will transform neuroscience. Hoboken, NJ: Wiley-Blackwell.

Landy, M.S., Goutcher, R., Trommershäuser, J., & Mamassian, P. Visual estimation under risk. *Journal of Vision, 7*, 1–15.

Teghtsoonian, M. (1965). The judgment of size. *American Journal of Psychology, 78,* 392-402.

Trommershäuser, J., Maloney, L.T., & Landy, M.S. (2003). Statistical decision theory and the selection of rapid, goal-directed movements. *Journal of the Optical Society of America, 20,* 1419-1433.

Wichmann, F. A. & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling and goodness-of-fit. *Perception and Psychophysics, 63,* 1293-1313.

Wichmann, F. A. & Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception and Psychophysics, 63,* 1314-1329.