Visual discrimination of orientation statistics in crowded and uncrowded arrays

Joshua A. Solomon   City University London, London, UK

When required to identify the orientation of an item outside the center of the visual field, the mean orientation predicts performance better than the orientation of any individual item in that region. Here I examine whether the visual system also preserves the variance of orientations in these so-called "crowded" displays. In Experiment 1, I determined the separation between items necessary to prevent neighbors from interfering with discrimination between different orientations in a single, target item. In Experiment 2, I used this separation and measured the effect of orientation variance on discrimination between mean orientations in these consequently uncrowded displays. In Experiment 3, I measured the relationship between the just-noticeable difference in variance and the smaller of two orientation variances in uncrowded displays. Finally, in Experiments 4 and 5, I reduced the separation between items and measured the effect of crowding on mean and variance discriminations. When considered together, the results of all these experiments imply that the visual system computes orientation variances with both more efficiency and greater precision than it computes orientation means. Although crowding made it difficult for some observers to discriminate between small amounts of orientation variance, it had no other significant effect on visual estimates mean orientation and orientation variance.

Introduction

Under various constraints, observers must sacrifice acuity for capacity. Of these constraints, one of the best studied is viewing eccentricity. For example, in the parafovea, the addition of just two vertical Gabor patterns can cause the just-noticeable tilt of a third Gabor to double (Solomon, Felisberti, & Morgan, 2004; see Figure 1a). This dramatic loss of acuity is thought to reflect an obligatory statistical analysis by the visual system (Parkes, Lund, Angelucci, Solomon, & Morgan, 2001). Instead of reporting the target's tilt, observers report the mean.

To understand the effect of neighboring "flankers" on acuity for the orientation of a parafoveal target, it is first necessary to understand what limits acuity in their absence. The best model we currently have for orientation acuity is based on the idea of "equivalent noise." This idea was introduced by signal detection theorists (Nagaraja, 1964; Pelli, 1990) to describe random fluctuations in visual signals, which could be mimicked by random fluctuations in the stimuli that give rise to them.

The data in Figure 1b were collected by flashing a single Gabor for 0.1 s at 3.7° eccentricity. Consider the large data point, which represents targets tilted 1° anti-clockwise of vertical. When forced to choose, observer MM correctly identified the tilt as anti-clockwise on just 44 of 58 trials with these targets. The smooth curve shows a Gaussian distribution with standard deviation of 1.2°. It seems to fit the data pretty well. The implication is that there is a source of at least approximately Gaussian

noise somewhere in the visual system, which corrupts estimates of orientation. In theory, the mean orientation of three Gabors can have up to three times as much noise. That is why untilted flankers increase the target's just-noticeable tilt.

Outline of the paper

The experiments described below were motivated by a desire to know whether observers have access to only the mean orientation in a crowded region or whether other statistics like variance are also available. To satisfy this desire, I first established a display in which there was no crowding. Specifically, in Experiment 1, I verified that sufficiently distant flankers do not interfere with the identification of a pre-cued target. In an attempt to be systematic, I next measured accuracy for discriminating between demonstrably uncrowded Gabor arrays having different mean orientations. Those measurements were performed in Experiment 2. In Experiment 3, observers were asked to discriminate between arrays having different orientation variances. Finally, Experiments 4 and 5 were run to establish the effect of target–flanker spacing on mean and variance discriminations. Raw psychometric data from all these experiments are placed within the framework of an inefficient, noisy observer model (Solomon, 2009), the front end of which adds an independent sample of equivalent noise to the orientation of each Gabor the observer sees. The result of this modeling exercise is strong support for the notion that the visual system preserves all but the
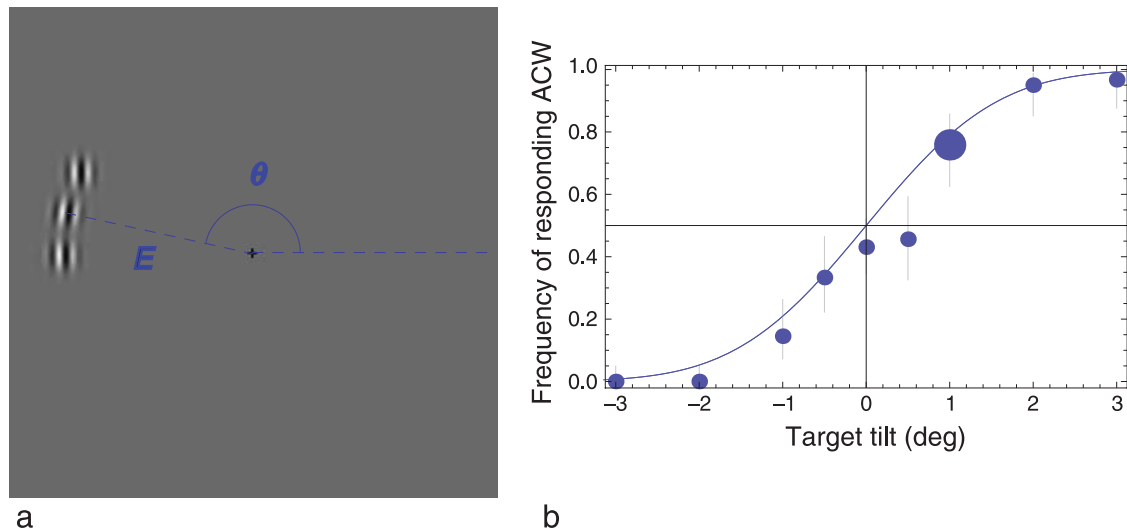
Figure 1. Orientation acuity for a parafoveal target is reduced when untilted flankers are present. (a) The "crowded" stimulus used by Solomon et al. (2004). All Gabor patterns were centered on a virtual, iso-eccentric circle with radius E = 3.7°. The tilted target had azimuth θ = 168° and the untilted flankers have azimuths $\theta \pm 12°$. (b) A psychometric function from that study. It was collected when the target appeared on its own (i.e., the flankers were absent). Each point shows how frequently MM classified the target as "anti-clockwise." Negative values indicate clockwise tilts. Error bars reflect 95% confidence intervals for the true probabilities. The solid curve is a cumulative Gaussian distribution with a standard deviation of 1.2°.

smallest variances in crowded displays. It seems to do so with even greater efficiency than it preserves the means.

Methodological considerations

Morgan, Chubb, and Solomon (2008) were the first to quantify orientation-variance discriminations within the framework of an inefficient, noisy observer model. (NB: They did not examine the effect of crowding.) Constraining the parameters of that model required measurements of just-noticeable differences (JNDs) in orientation variance. Within the context of Signal Detection Theory, Durlach and Braida (1969) identified two qualitatively different limitations on discrimination performance. The first type of limitation arises when observers compare two "sensory traces," each of which is corrupted by an independent sample of equivalent noise. The second type of limitation arises when observers compare stimuli one at a time with a remembered "context," which may also vary stochastically, during the course of an experiment.

As this latter, context-related variability is not a feature of the inefficient, noisy observer model, it seemed prudent to adopt Durlach and Braida's (1969) recommended technique for minimizing its contamination of our JNDs. Specifically, I used the two-alternative forced-choice (2AFC) paradigm, with a "roving pedestal." That is, on each trial, the observer had to select the larger of two orientation variances. The smaller of these variances is called the pedestal, and several different pedestals were randomly interleaved during the course of the experiment.[1]

To facilitate a direct comparison between variance discriminations and other types of discrimination, I used the 2AFC paradigm in all the experiments described below.

Experiment 1

Morgan and Solomon (2006) found that seven other Gabors, equally separated on a virtual, iso-eccentric circle, did not interfere with the identification of a pre-cued target. Here I simply verify that this result generalizes to the 2AFC paradigm.

Methods

There were four observers in total: myself (JAS), another professional psychophysicist (IM), an undergraduate with considerable psychophysical experience (HLW), and an undergraduate who was new to psychophysics (AS). Both of the latter were naive to the purposes of this experiment.

The experiment was conducted on a 15″ MacBook Pro computer, running the PsychToolbox (Brainard, 1997; Pelli, 1997; software available upon request). Display resolution was 1440 × 900 pixels. A comfortable viewing distance of approximately 0.5 m was maintained. At this viewing distance, each Gabor was centered 6 degrees of visual angle away from a central fixation spot (see Figure 2).
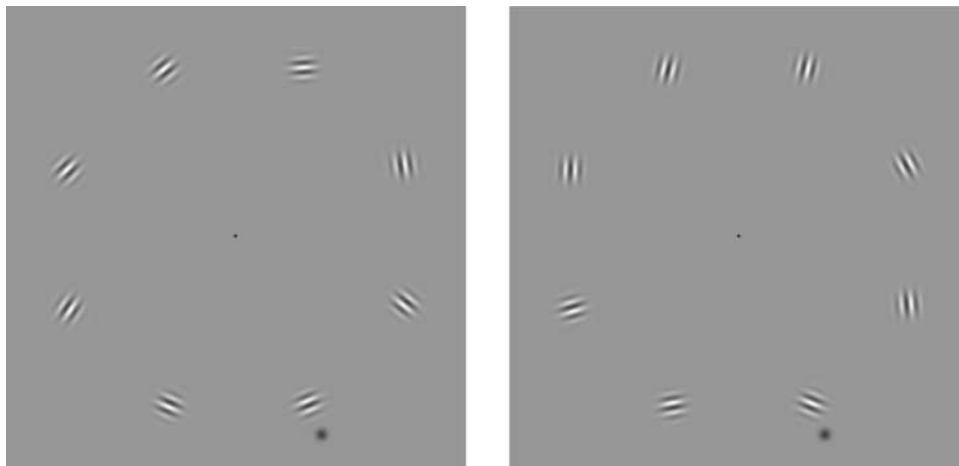
Figure 2. Demonstrably uncrowded Gabor arrays. The observer's task was to determine whether the second target Gabor (in the right panel) was oriented clockwise or anti-clockwise, with respect to the first (in the left panel). In this figure, the dark blobs responsible for cueing the target position have been superimposed onto the Gabor arrays themselves. In the actual experiment, blob onset preceded each Gabor array by 0.1 s.

Each Gabor was the product of a sinusoidal luminance grating and a Gaussian blob. The grating had a spatial frequency of 3 cycles per degree and random spatial phase. The blob had a space constant (σ) of 0.25 degree of visual angle. Both grating and blob had maximum contrast.

One Gabor in each array was designated as the target. When other Gabors were present, they were equally spaced on an iso-eccentric circle. On each trial of the experiment, two arrays containing an identical number of Gabors were presented for 0.15 s each. For 1.5 s between these presentations, only the central fixation spot was visible. For 0.1 s, immediately prior to each of these presentations, the target was pre-cued by the appearance of a dark Gaussian blob ($\sigma = 0.125°$). The pre-cue had the same azimuth as the target, but its retinal eccentricity was greater: 6.5 degrees of visual angle. The 0.1-s cue–target onset asynchrony seems to be optimal for automatically attracting attention (Cheal & Lyon, 1991; Nakayama & Mackeben, 1989).

The orientation of each Gabor was selected at random, except that of the second target. The orientation of this latter Gabor was selected at random from the set $\{\theta_1 - 11.2°,\ \theta_1 - 5.6,\ \theta_1 + 5.6°,\ \theta_1 + 11.2°\}$, where θ_1 denotes the orientation of the first target. The observer's task was to report whether the second target was oriented clockwise or anti-clockwise of the first target. Pilot experimentation suggested that accuracy with these orientations would fall within the interval (50%, 99%). Yet another pilot experiment established accuracies in excess of 98% correct when arrays contained just one Gabor, and orientations were selected randomly from the set $\{\theta_1 - 45°,\ \theta_1 + 45°\}$.

For observers JAS, HLW, and AS, arrays contained either 1 or 8 Gabors. For observer IM, arrays contained either 1, 2, 4, 8, or 16 Gabors. For IM, HLW, and AS, both target Gabors had the same azimuth. For JAS, the second target had either the same azimuth as the first target, or its

azimuth was increased by 180°. For all observers, the azimuth of the first target was selected at random. After each observer ran a minimum of 100 practice trials, JAS ran a total of 800 and the others ran 500.

Results

When Signal Detection Theory (Green & Swets, 1966) is applied to 2AFC orientation discrimination, response frequencies reflect the comparison of two independent random variables. Each of these random variables represents the orientation of a target, after it has been perturbed by an independent sample of equivalent noise. As noted in the Introduction section, all the analyses described in this paper are predicated on the notion of Gaussian equivalent noise. Consequently, when flankers do not interfere, the psychometric function mapping the (signed) angle between two targets to the probability of an "anti-clockwise" response should take the form of a cumulative Gaussian distribution function, like that in Figure 1b.

The JNDs in Figure 3 are standard deviations of the cumulative Gaussian distributions that fit Experiment 1's response frequencies with maximum likelihood. For each observer, the JND with 8 Gabors per array is similar to the JND with 1 Gabor per array. Data from IM indicate that when there were 16 Gabors per array, the JND was larger. The JNDs derived from JAS's performances with targets on opposite sides of fixation (not shown) were also larger.

Discussion

The invariance of JND with up to 8 Gabors per array implies that the observers could discriminate between two successively displayed, differently oriented Gabors at 6°
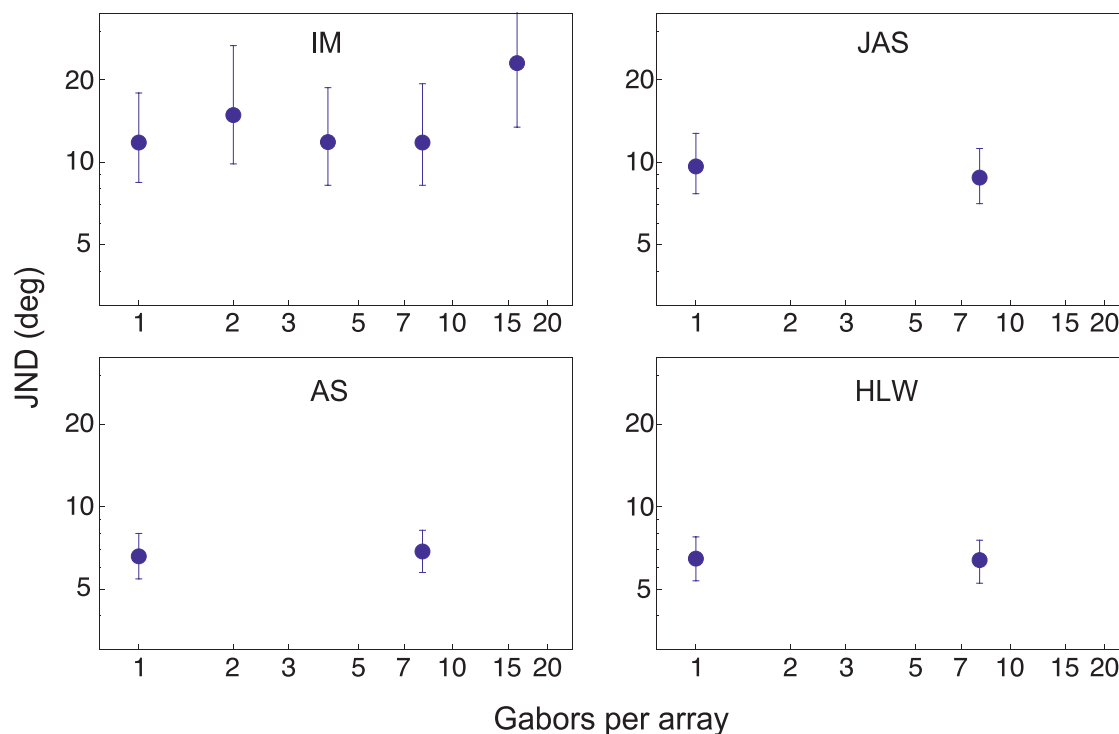
Figure 3. Just-noticeable differences (JNDs) between the orientations of two successively displayed, differently oriented Gabors at 6° viewing eccentricity. Error bars contain 95% confidence intervals. Whether there were eight equally spaced, iso-eccentric Gabors per array (as in Figure 2) or just one, JND is the same.

viewing eccentricity, without interference from seven iso-eccentric, randomly oriented distracters. Gabors in this configuration are said to be "demonstrably uncrowded," and thus were adopted for my examination of mean and variance discriminations in the absence of crowding (i.e., Experiments 2 and 3, below).

Experiment 2

Having established a stimulus configuration without crowding, I then proceeded systematically toward the goal of quantitatively analyzing variance discrimination under crowded and uncrowded conditions. Experiment 2 was designed to provide a baseline for comparison. It documents mean discrimination under uncrowded conditions.

Experiment 2 largely replicates the measurements and analysis performed by Dakin (2001). Nonetheless, there were two important differences between that study and the present one. Although Dakin systematically manipulated the density of his Gabor arrays, there were no conditions in which the individual Gabors were demonstrably uncrowded, according to the criteria described above. The second important difference between Dakin's method and the one used here is that Dakin used the vertical orientation as an implicit reference. That is, his observers reported whether

each mean was clockwise or anti-clockwise of vertical. Consequently, the stimuli used to measure just-noticeable departures from this reference are necessarily near vertical themselves.

Non-uniform distributions of stimulus orientation are fine when assessing mean discrimination, but they are not appropriate for assessing variance discrimination. They encourage observers to use energy at orientations orthogonal to the dominant one as a substitute for computing variance. For example, when observers know that most Gabors will have near-vertical orientations, high-variance arrays are simply those that activate visual mechanisms with a preference for horizontal stimuli. No other mechanisms would be required for variance discrimination. To prevent observers using this type of heuristic, it was necessary to adopt a uniform distribution of mean orientations for all of the experiments described in this paper.

Methods

The same general methods employed in Experiment 1 were reused in Experiment 2. However, observer IM was unavailable. Therefore, observations were solicited from another professional psychophysicist, MJM.

Experiment 2 also featured the same type of Gabor arrays used in Experiment 1. However, the pre-cue was no longer necessary. On each trial, the orientations of Gabors

in the first array were selected at random from a Gaussian distribution. The mean of the distribution $\bar{\theta}_1$, was selected randomly from the interval [0°, 180°). The Gaussian distribution of Gabors in the second array had the same standard deviation, but its mean $\bar{\theta}_1 \pm \Delta\theta$ was determined by one of a number of randomly interleaved, QUEST staircases (Watson & Pelli, 1983).

Each condition had its own staircase. There were twenty: five standard deviations of orientation ($\sigma_G \in \{1°, 2°, 4°, 8°, 16°\}$) were paired with four numbers of Gabor per array ($N \in \{1, 2, 4, 8\}$). The staircases were optimized to converge on the mean α of a (Weibull) psychometric function having the form

$$P(C) = \frac{1}{2} + \left(\frac{1}{2} - 0.01\right)\exp\left[-\left(\frac{\Delta\bar{\theta}}{\alpha}\right)^{\beta}\right], \quad (1)$$

where $P(C)$ represents the probability of a correct response. The value used for psychometric slope $\beta = 10$

was the arithmetic mean of maximum likelihood fits of Equation 1 to JAS's data from each condition in Experiment 1.

Good estimates of lapse rate are necessary when estimating model parameters from psychometric data (Wichmann & Hill, 2001). To obtain these estimates, staircases were ignored and $\Delta\bar{\theta}$ was set to 45° (22.5° for HLW) with a probability of 0.125 on every trial in Experiment 2.

Each observer completed 20 blocks of 100 trials each. New staircases were begun after the first 10 blocks.

Results

The results of Experiment 2 were subjected to the same type of analysis as those of Experiment 1. In particular, the JNDs illustrated in Figure 4 represent the standard deviations of cumulative Gaussian distributions that fit Experiment 2's response frequencies with maximum likelihood. Numerical symbols indicate the size of each
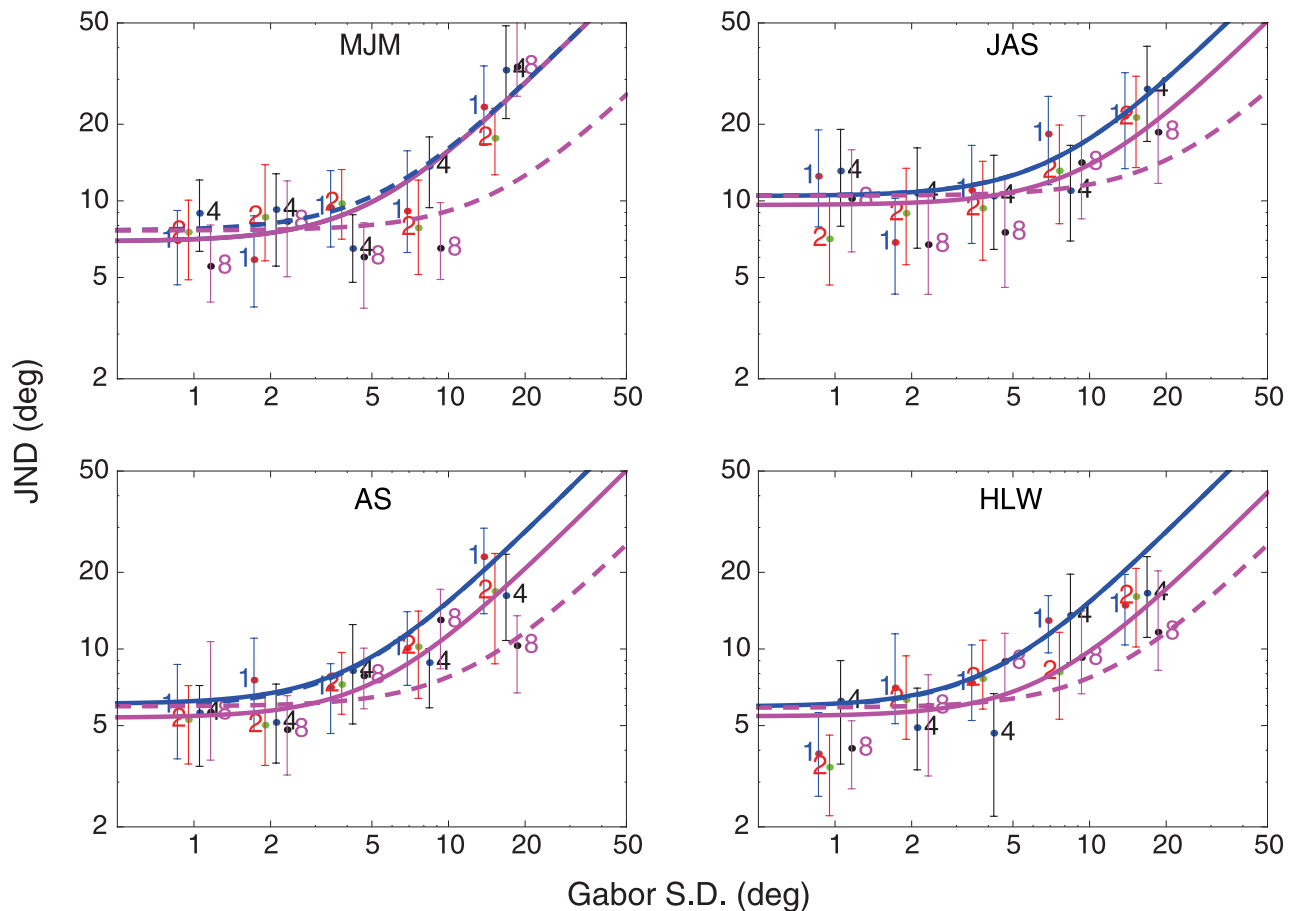

Figure 4. JNDs between the mean orientations of two successively displayed, uncrowded Gabor arrays. Numerals denote *N*, the number of Gabors per array. They have been nudged horizontally for better legibility. Error bars contain 95% confidence intervals. Solid blue and magenta curves illustrate the best 5-parameter fit to the data in each panel for *N* = 1 and *N* = 8, respectively. (Simultaneous fits for *N* = 2 and *N* = 4 are not shown.) These curves may be compared with the dashed blue and magenta curves, which show the best 2-parameter fits, i.e., when efficiencies were fixed at 100%. Note, however, that all of these latter fits were best when all the noise was late, i.e., $\sigma_E = 0$. All data were best fit with a sample size of one, two, or three Gabors per array (see text for details).

Gabor array, N. These symbols have been nudged horizontally for maximum legibility. All observers suffered an elevation of JND when the standard deviation of Gabor orientations was highest, but no systematic effect of array size is obvious from a visual inspection of the data.

Modeling

Following Dakin (2001) and Parkes et al. (2001), I conducted a more principled test for an effect of Gabor number by maximum likelihood fitting the data from each observer with an inefficient, noisy observer model. In its most general form, this model for mean discrimination has terms for both early noise, which independently perturbs the perceived orientation of each Gabor, and late noise, which perturbs estimates of sample mean. Within this framework, the proportion of "anti-clockwise" responses is given by the formula

$$P(\text{"ACW"})$$
$$= \gamma + (1 - \delta - \gamma)\Phi\left[\frac{\Delta\bar{\theta}}{\sqrt{\sigma_L^2 + 2(\sigma_E^2 + \sigma_G^2)/M_{\text{mean}}}}\right],$$
(2)

where σ_G is the standard deviation of the Gabor orientations, Φ is the cumulative Normal distribution function, and M_{mean}, σ_E, and σ_L are free parameters. M_{mean} represents the number of Gabors observers use when calculating the mean of each array. M_{mean} is necessarily an integer and $M_{\text{mean}} \leq N$. σ_E and σ_L represent the standard deviations of the early and late noises, respectively. Psychometric floor γ and ceiling $1 - \delta$ were set to reflect the empirical lapse rates: $\gamma = \delta = 0.03$ for HLW, $\gamma = \delta = 0.04$ for MJM and AS, and $\gamma = \delta = 0.05$ for JAS. NB: The fraction in Equation 2 has a 2 in its denominator because of the 2AFC paradigm; decisions are necessarily affected by orientation variance in both arrays.

When the sample size M_{mean} was allowed to vary with array size N, all data from observer MJM were nonetheless best fit when $M_{\text{mean}} = 1$. In this case (i.e., when M_{mean} does not vary with N), one of the two noise parameters is redundant. Therefore, I arbitrarily set $\sigma_L = 0$ and found that likelihood was maximized when $\sigma_E = 4.9°$. Fits to JAS's data were best when $\sigma_E = 4.0°$, $\sigma_L = 8.8°$, and $M_{\text{mean}} = \begin{cases} 1 & N = 1 \\ 2 & N \geq 2 \end{cases}$. For AS, fits were best when $\sigma_E = 2.9°$, $\sigma_L = 4.5°$, and $M_{\text{mean}} = \begin{cases} 1 & N = 1 \\ 2 & N \geq 2 \end{cases}$. For HLW: $\sigma_E = 2.1°$, $\sigma_L = 5.2°$, and $M_{\text{mean}} = \begin{cases} 1 & N = 1 \\ 2 & N = 2, 4 \\ 3 & N = 8 \end{cases}$. These fits are illustrated by the solid curves in Figure 4.[2]

All but the last of these maximum likelihood fits is consistent with the notion that observers have a capacity limit, whereby they are unable to use more than a maximum of M_{max} uncrowded Gabors, when attempting to compute their mean orientation. In other words, the fits satisfy the constraint $M_{\text{mean}} = \min\{M_{\text{max}}, N\}$, where M_{max} is an arbitrary integer, representing the maximum sample size. Imposition of this constraint results in an insignificantly poorer fit to HLW's data, as confirmed by the generalized likelihood ratio test (Mood, Graybill, & Boes, 1974).[3]

The dashed curves in Figure 4 illustrate the maximum likelihood fit of a noisy observer model that is not inefficient. That is, this observer has unlimited capacity (or $M_{\text{max}} \geq 8$). Its fit to MJM's data is clearly inferior, but its fit to the other observers' data at least does not appear entirely awful. To determine which capacities can be safely ruled out by each observer's data, I once again turned to the generalized likelihood ratio test. The results indicate that we can reject the hypothesis that $M_{\text{max}} > 1$ for MJM, we can reject the hypothesis that $M_{\text{max}} > 4$ for JAS, and we can reject the hypotheses that $M_{\text{max}} > 3$ for both AS and HLW.

Discussion

Despite Dakin's (2001) conclusion that $M_{\text{mean}} \approx \sqrt{N}$, the extremely limited capacities suggested by my modeling came as a surprise. On the basis of earlier work with very small Gabors (specifically Figure 3 of Parkes et al., 2001), I had expected sample sizes M_{mean} much closer to the array size N. As noted above, Experiment 2 differs from previous work by using a uniform distribution of mean orientations. If observers had different sampling efficiencies for different mean orientations, perhaps the small values of M_{mean} suggested by Experiment 2 were simply an artifact of collapsing across different mean orientations.

To test this idea, I reran Experiment 2, interleaving separate staircases for eight different "mean" orientations: 0°, 22.5°, 45°, 67.5°, 90°, 112.5°, 135°, and 155.5° anti-clockwise with respect to horizontal. One of these means was used for one presentation in each trial. A QUEST staircase determined the other presentation. Data appear below in Appendix A. They strongly suggest that there is no effect of mean orientation on sampling efficiency, and the randomization of mean orientation need not contaminate estimates of M_{mean}.

This does not explain MM's more efficient performance in Figure 3 of Parkes et al. (2001). Fortunately, that study's MM is this study's MJM, and he was available for an attempt at replicating that study's result with at least partially uncrowded Gabors. Data from this attempted replication appear below in Appendix B. They were somewhat equivocal. On the one hand, three of MJM's four JNDs formed a pattern similar to that reported by Parkes et al. On the other hand, the pattern of JAS's data seemed quite different. Further investigation of this issue was suspended due to MJM's unavailability and its merely

tangential relationship to this study's main goal, which is a quantitative analysis of *variance* discrimination not mean discrimination.

Experiment 3

In Experiment 3, I measured JNDs for 2AFC discrimination of orientation variance in demonstrably uncrowded arrays.

Methods

The same general methods employed in Experiment 2 were reused in Experiment 3, with the following exceptions. Only arrays of size $N = 8$ were used in this experiment. As before, the orientations in each array were selected from a Gaussian distribution. However, in this experiment, the mean of the Gaussian distribution was reselected at random for every array (not merely every trial). Within each trial, the distribution defining one array had a pedestal variance that was randomly selected from the set {1 deg^2, 4 deg^2, 16 deg^2}. The variance of the distribution defining the other array was determined by QUEST. It was the observer's task to select this latter array.

Two QUEST staircases were randomly interleaved for each pedestal, one converging on an accuracy of $P(C) = 0.67$, the other converging on an accuracy of $P(C) = 0.84$, assuming Weibull psychometric functions like that defined in Equation 1 and substituting the increment in variance $\Delta \text{var}\theta$ (in deg^2), for the increment in mean $\Delta\bar{\theta}$. An appropriate value for psychometric slope $\beta = 3$ was determined in a pilot experiment.

The frequency of easy trials, on which QUEST was ignored and $\Delta \text{var}\theta$ was set to a value of 900 deg^2, was 0.10. This seemed sufficient for a reasonable estimate of lapse rate.

Each observer ran six blocks of 110 trials each. QUEST was reinitialized at the beginning of the fourth block.

Results

Although Gaussian noise mandates Gaussian psychometric functions for mean discrimination, no such mandate exists for variance discrimination. Therefore, I have taken an atheoretical approach to defining JNDs for variance discrimination as the means of cumulative Weibull distributions (e.g., Equation 1), maximum likelihood fit to the raw psychometric data. This approach was also taken by Morgan et al. (2008).

Previously, I have argued for the importance of plotting JNDs in variance rather than standard deviation. One benefit of plotting data in this way is that sensory thresholds

are implicated by any negative masking or "dip" in the curve relating JND to pedestal variance (Solomon, 2009).[4] Data from IM clearly do not suggest any dip, but data from JAS do (see Figure 5). At least, they contain the hint of a dip: the JND for a 4-deg^2 pedestal is slightly lower than that for 1-deg^2 pedestal. Confidence intervals suggest that this dip may not be significant, but I applied the sensory threshold model described by Morgan et al. (2008) just to make sure.

Modeling

Morgan et al. (2008) derived the inefficient, noisy observer model for 2AFC variance discrimination. Equation 3 extends their model to include the possibility of lapse rates $\delta \neq 0$. Proportion correct is given by the formula

$$P(C) = (1 - 2\delta)\left(1 - F\left[\frac{\text{var}\theta + \sigma_E^2}{\text{var}\theta + \Delta\text{var}\theta + \sigma_E^2}\right]\right) + \delta,$$

(3)

where F is the F-distribution, with degrees of freedom $M_{\text{var}} - 1$ and $M_{\text{var}} - 1$. As in Equation 2, here M_{var} represents the number of Gabors observers use when calculating the variance of each array. M_{var} is necessarily an integer and $M_{\text{var}} \leq N$. σ_E represents the standard deviation of the early noise, and δ is the empirically determined lapse rate (0 for all observers in Experiment 3, except AS, for whom it was 0.014).

Note that this formula does not include a term for late noise, which could further perturb estimates of sample variance. Were such a term added, its distribution would have to be convolved with the F-distribution. Obtaining maximum likelihood fits with such a complicated psychometric function would not only be computationally intractable, it would necessarily over-fit the data from Experiment 3, which used just one array size, and thus cannot simultaneously constrain the variances of early and late noises.

Maximum likelihood fits to IM's data were obtained with $M_{\text{var}} = 7$ and $\sigma_E = 2.5°$. For JAS, the best-fitting parameter values were $M_{\text{var}} = 6$ and $\sigma_E = 2.0°$. For AS, they were $M_{\text{var}} = 6$ and $\sigma_E = 3.5°$, and for HLW, they were $M_{\text{var}} = 8$ and $\sigma_E = 3.0°$. A comparison of these values with those derived in Experiment 2 suggests that discriminations of orientation variance enjoy both greater sampling efficiency and less noise than discriminations of mean orientation, at least with uncrowded arrays.

Curves illustrating these model fits appear in Figure 5. Although they are strictly increasing with pedestal variance, JAS's data do not. As noted above, this "dip" may indicate a sensory threshold, below which all variances would be indiscriminable. Morgan et al. (2008) described how the inefficient, noisy observer model could be extended to include such a threshold. I fit that extended
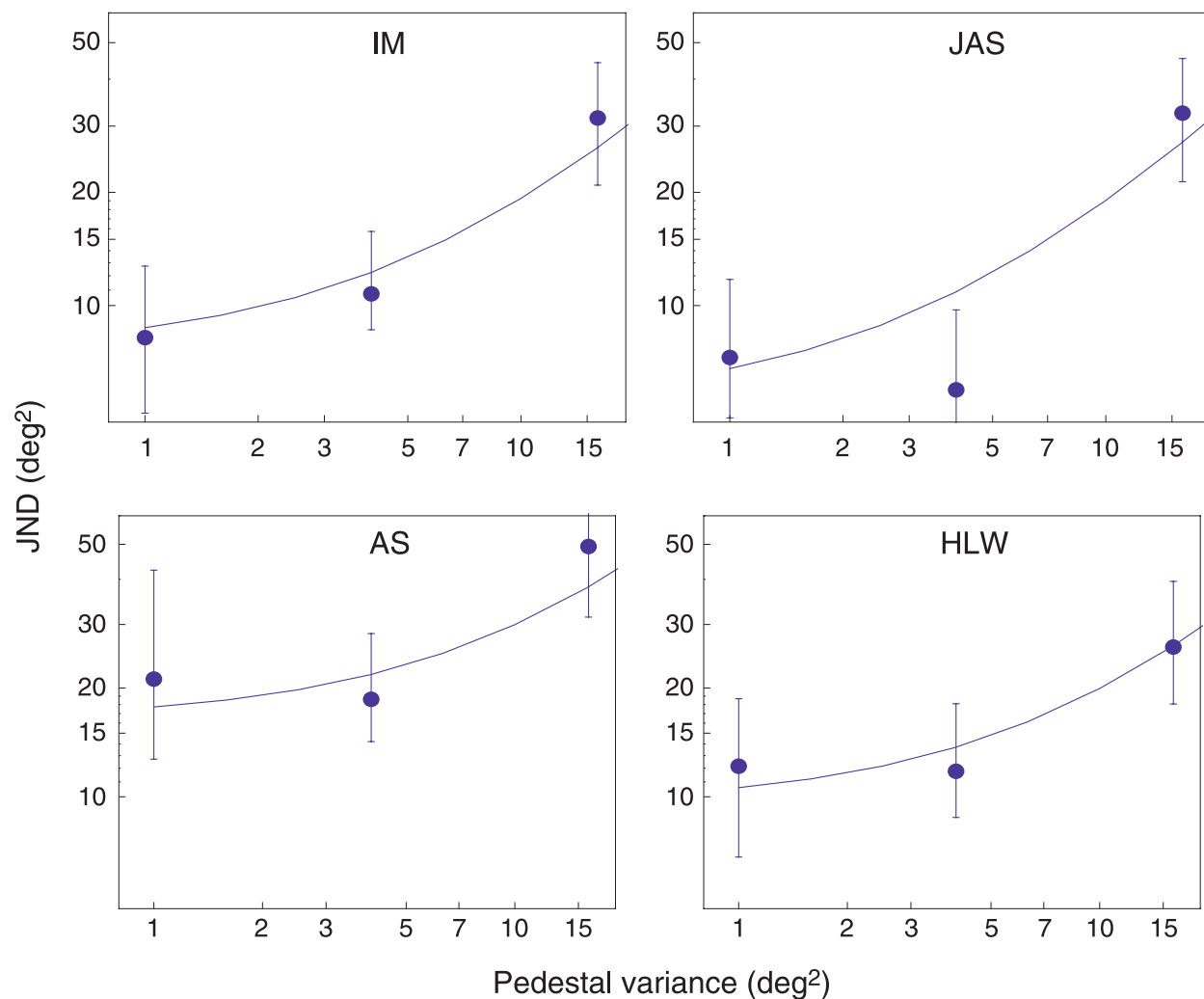
Figure 5. JNDs between orientation variance in two successively displayed arrays containing eight uncrowded Gabors each. Error bars contain 95% confidence intervals. Curves illustrate maximum likelihood fits of the inefficient, noisy observer model. All data were best fit with a sample size between 6 and 8 Gabors per array. Adding a sensory threshold to the two-parameter model did not significantly improve any fit.

model to each observer's results. In each case, I found the improvement too modest to reject the null hypotheses of the more restrictive, 2-parameter model.

The availability of mean discrimination data and variance discrimination data from the same observers in otherwise identical conditions affords a statistical evaluation of other hypotheses within the framework of an inefficient, noisy observer. In its most general form, this model has four free parameters, with which to fit all of the responses with eight-Gabor arrays: σ_E, σ_L, M_{mean}, and M_{var}. After finding the maximum likelihood fit with all four parameters free to vary (JAS: $\sigma_E = 2.0°$, $\sigma_L = 8.5°$, $M_{mean} = 2$, and $M_{var} = 6$; AS: $\sigma_E = 3.5°$, $\sigma_L = 3.6°$, $M_{mean} = 1$, and $M_{var} = 6$; HLW: $\sigma_E = 3.0°$, $\sigma_L = 3.4°$, $M_{mean} = 1$, and $M_{var} = 8$), I independently tested the hypotheses of no late noise (i.e., $\sigma_L = 0$) and equal sampling efficiencies for the two tasks (i.e., $M_{mean} = M_{var}$). Data from each observer (except IM, who did not participate in Experiment 2) allowed the rejection of both hypotheses.

Discussion

The preceding modeling confirms that, in the absence of crowding, orientation-variance discriminations are more efficient than mean orientation discriminations. This may seem paradoxical, because ideal estimates of variance are based on individual deviations from the estimated mean. However, our result is not a logical impossibility. Several sub-ideal estimates of variance do not require estimates of the mean and yet can be highly efficient. See the General discussion section for further thoughts on this issue.

The high ratio of late noise to early noise, as estimated within the context of the inefficient, noisy observer model, is also noteworthy. This suggests that orientation discrimination in all but the most varying arrays is primarily limited by late noise.

Moreover, the ratios implied above ($\sigma_L/\sigma_E > 1$) must be understood as a lower bound on the ratio for uncrowded mean discriminations that would have been obtained with

a non-zero late noise for variance discrimination. To fit mean and variance discriminations as well as they were fit above, the presence of such noise would have to be offset by a reduction in the amount of early noise that was theoretically present for both mean and variance discriminations, and this reduction in turn would have to be offset by yet more late noise for mean discriminations.

Experiments 4 and 5

The final step required for a quantitative analysis of variance discrimination with crowded stimuli is a comparison between the inefficient, noisy observer's behavior with demonstrably uncrowded stimuli (discussed above) and its behavior in crowded situations. The latter was the subject of investigation in Experiments 4 and 5.

Methods

In general, the same methodology and analytical techniques developed for use with uncrowded stimuli were also deployed here. Crowded Gabor arrays were created by reducing the separation between neighboring Gabors. Examples are shown in Figure 6. As above, all Gabors were arranged on a virtual circle. The center of this circle in Experiments 4 and 5 had the same retinal eccentricity as the individual Gabors in Experiments 1–3.

The procedure used in Experiment 4 was identical to that used in Experiment 2, with the exception that only arrays of size $N = 8$ were used. On each trial, the first array was centered at a random azimuth and 6 degrees of visual angle. Observers were encouraged to maintain fixation, so that Gabors in the second array would have the same retinal positions as Gabors in the first array. There were a total of nine viewing conditions: three standard deviations of orientation ($\sigma_G \in \{1°, 8°, 16°\}$), fully crossed with

three Gabor spacings. In the most tightly crowded condition, the distance between the center of each Gabor and the center of its array was 1.5 degrees of visual angle. Nearest neighbors thus had a center-to-center spacing of 1.2 degrees of visual angle. Nearest neighbor spacings for the lesser crowded conditions were 2.3 and 4.7 degrees of visual angle. This latter spacing, which was 0.92 times that used in Experiments 1–3, was the maximum with which I could be certain to avoid losing parts of Gabors at the top and bottom of the computer screen.

In Experiment 4, each observer completed 12 blocks of 165 trials each. New staircases were begun after the first 6 blocks. As in Experiment 3, the frequency of easy trials, on which QUEST was ignored and $\Delta\theta$ was set to a value of 45°, was set to 0.10.

The procedure used in Experiment 5 was identical to that used in Experiment 3, except for the configuration of Gabors, which was the same as that used in Experiment 4. Each observer completed 24 blocks of 165 trials each. New staircases were begun after every 6 blocks.

Results

Like Experiment 2, the results of Experiment 4 (left-hand panels in Figure 7) clearly demonstrate an increase in the JND for mean orientation when there is an increase in the variance (or S.D.) of the orientations in the stimulus. With one exception, a visual inspection of these left-hand panels reveals no systematic effect of crowding. That one exception can be found with the 8° Gabor S.D. in Figure 7e, where crowding seems to elevate the JND.

Crowding also appears not to have affected the JND in orientation variance when the stimulus has high variance to begin with. This can seen by looking at the right-hand sides of the right-hand panels in Figure 7. Thus crowding certainly does not adversely effect all variance discriminations. On the other hand, Figures 7b and 7d provide evidence that it can elevate the JND in orientation
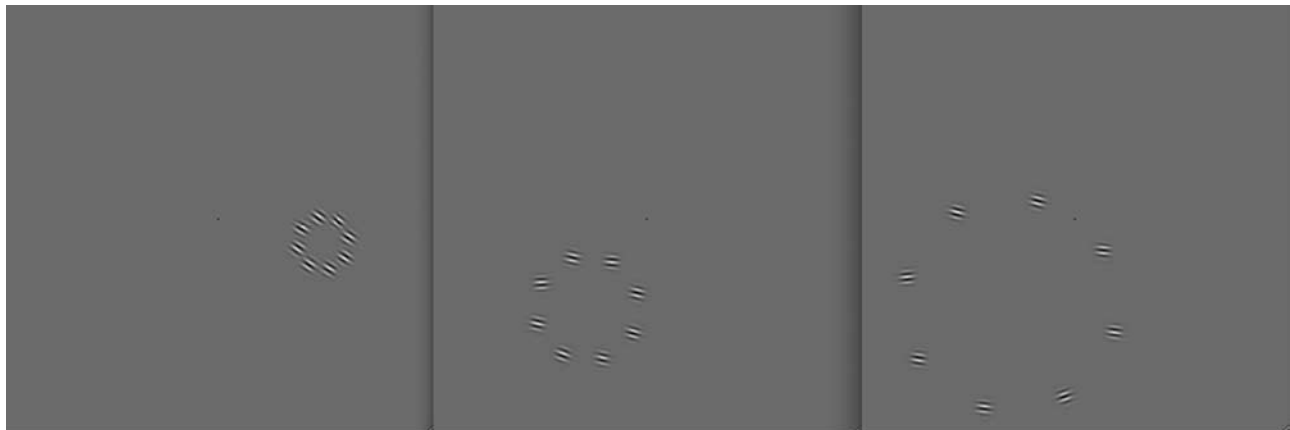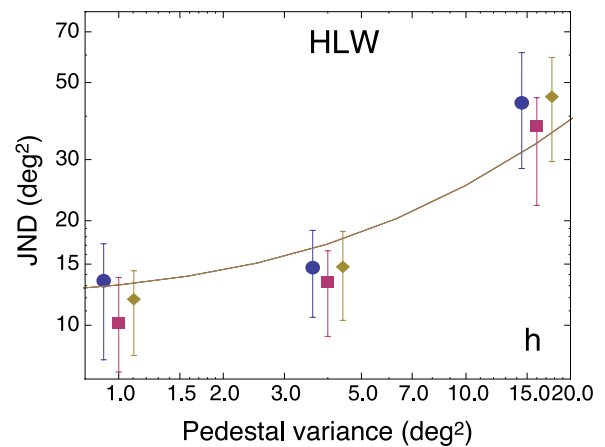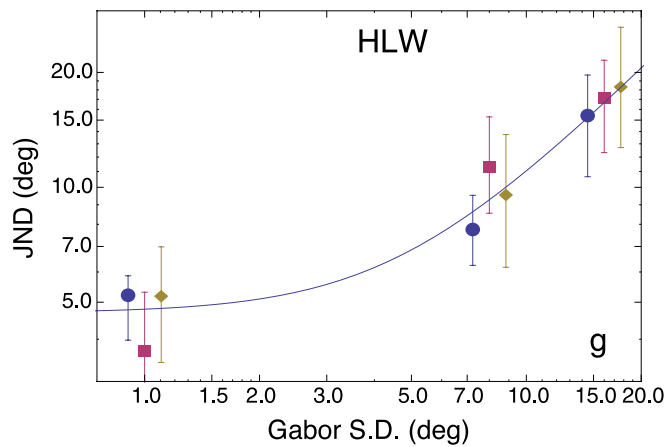

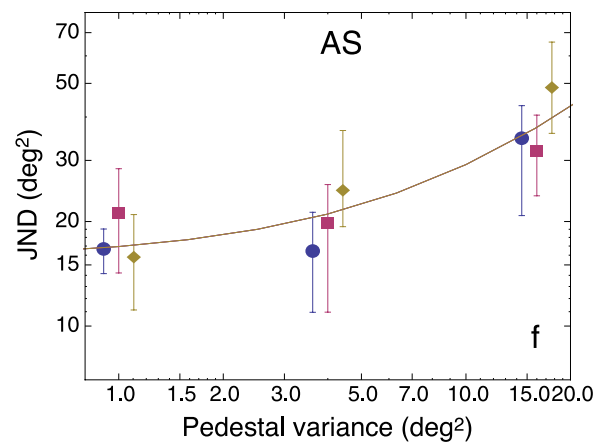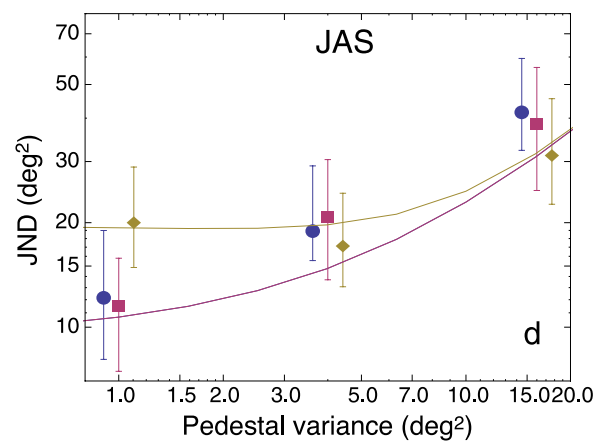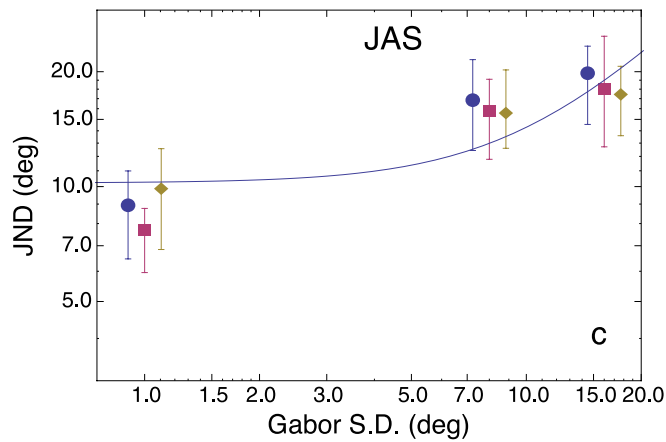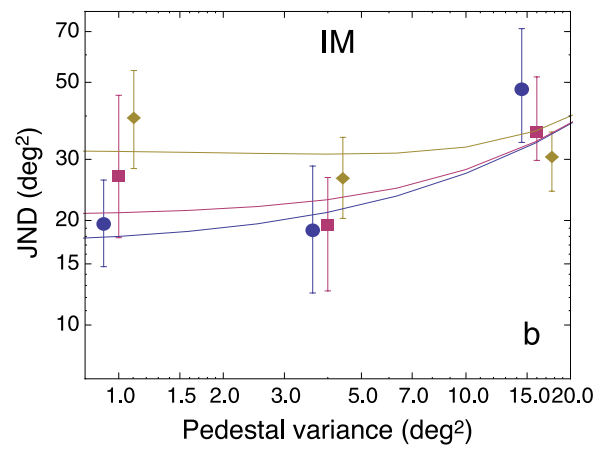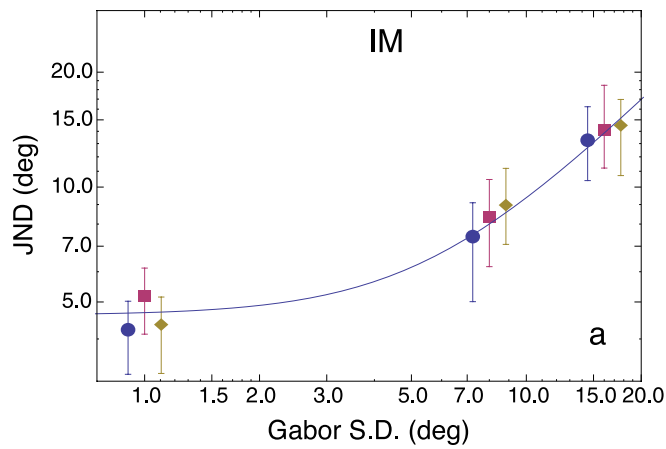Figure 6. Sample Gabor arrays used in Experiments 4 and 5. There were two independent variables in each experiment: the distance between nearest neighbors and the variance of their orientations.
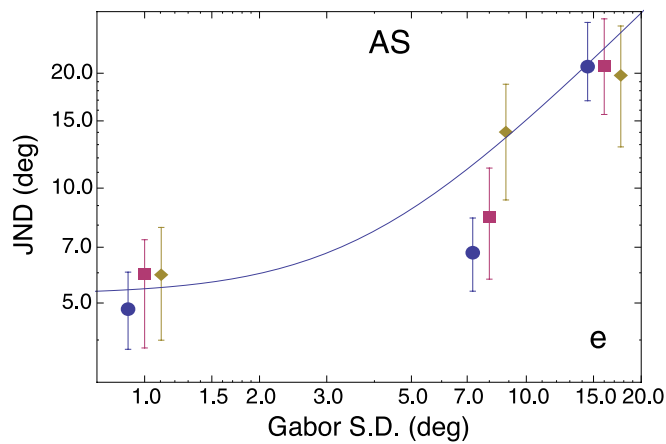
variance when the pedestal variance is low, at least for some observers (IM and JAS).

Lapse rates, estimates of which are required for maximum likelihood modeling, were $\gamma = \delta = 0.01$ for each observer in Experiment 4, except AS: $\gamma = \delta = 0.03$. In Experiment 5, the lapse rates (δ) were IM: 0.02, JAS: 0.02, AS: 0.01, and HLW: 0 (sic).

Modeling

To determine how (and if so, to what degree) crowding affected the observers' discriminations of mean orientation and orientation variance, I once again modeled them as inefficient and noisy. As noted above, most of the left-hand panels in Figure 7 suggest no effect of Gabor spacing on the JND in mean orientation. A rigorous evaluation of this suggestion was conducted by comparing maximum likelihood fits of nested models to each observer's data from Experiment 4. In the more general model, sample size and equivalent noise were allowed to vary with Gabor spacing. In the nested model, they were not. The ratio between these maximum likelihoods was greatest (2.5 natural log units) for observer AS but still not large enough to reject the null hypothesis that Gabor spacing had no effect on these parameters.

Consequently, I did not bother attempting to maximize the likelihood of the most general model that could be constrained by Experiments 4 and 5 when their data were pooled. Instead, I assumed that a single sample size M_{mean} would be sufficient to describe all mean discriminations and maximized the likelihood of a slightly less general model. In that model, 11 parameters were allowed to vary. They were the sample size M_{mean}, the late noise that perturbs estimates of mean orientation σ_L, and three parameters for each Gabor spacing: the standard deviation of the early noise σ_E, the sample size used for variance discriminations M_{var}, and the threshold below which all variances would be indistinguishable c.

Although higher efficiencies necessarily manifest as lower JNDs, lower values of early noise do not. A large effect of Gabor spacing on σ_E would not seriously affect mean discriminations that were primarily limited by late

noise. Therefore, I refit each observer's data with a nested model, in which σ_E remained invariant with spacing, and submitted the reduction in likelihood to the chi-square test. The reductions in likelihood were not significant. Therefore, none of the observers' data suggest an effect of Gabor spacing on σ_E.

What about other effects of spacing? Clearly, spacing is having some effect on the JNDs in Figures 7b and 7d. Crowding could produce a sensory threshold, thereby obliterating small differences between Gabor orientations, or it could affect efficiency, forcing observers to use fewer Gabors in their estimates of variance. To test each of these hypotheses, I refit each observer's data with further nested models, one in which c remained invariant, one in which M_{var} remained invariant, one in which both remained invariant, and one in which M_{var} remained invariant with spacing and the sensory threshold c was fixed at zero.

For AS and HLW, in all cases, the reductions in likelihood were insignificant. Forcing M_{var} to remain invariant with Gabor spacing also resulted in insignificant likelihood reductions when the inefficient, noisy observer model was fit to the data from IM and JAS; however, the other aforementioned nested models were significantly less likely. Therefore, the data from these two observers are consistent with the conclusion that crowding did produce a sensory threshold, but it did not affect the efficiency with which orientation variances were estimated.

The data from each observer were subjected to two final tests, one for the hypothesis that $\sigma_L = 0$ and one for the hypothesis that $M_{var} = M_{mean}$. These constraints also significantly reduced the maximum likelihood. Therefore, all of the data are consistent with the conclusion that orientation-variance discrimination is both less noisy and more efficient than mean orientation discrimination, regardless of the spacing between Gabors.

Curves in Figure 7 were produced with the following parameter values. For IM, $\sigma_E = 4.0°$, $\sigma_L = 3.2°$, $M_{mean} = 3$, $M_{var} = 8$, and sensory thresholds effectively equated all variances below 160 and 240 deg^2, for moderately spaced and tightly crowded Gabors, respectively. For JAS, $\sigma_E = 2.6°$, $\sigma_L = 9.9°$, $M_{mean} = 2$, $M_{var} = 6$. Only those data collected with tightly crowded Gabors supported a sensory threshold for this observer, the best-fitting one effectively equated all variances below 89 deg^2. For AS, $\sigma_E = 3.4°$, $\sigma_L = 2.2°$, $M_{mean} = 2$, $M_{var} = 6$, and for HLW, $\sigma_E = 2.9°$, $\sigma_L = 3.7°$, $M_{mean} = 2$, $M_{var} = 6$. Neither of these latter two data sets supported a sensory threshold.

Figure 7. JNDs between (a, c, e, g) the mean orientations or (b, d, f, h) orientation variances of two successively displayed arrays containing eight Gabors each. Blue circles, magenta squares, and yellow diamonds represent results with widely separated, moderately separated, and tightly crowded arrays, respectively, Circles and diamonds have been nudged horizontally for better legibility. Error bars contain 95% confidence intervals. Color-coded curves illustrate the behavior of inefficient, noisy observers. Data from two observers (IM and JAS) were significantly better fit when a sensory threshold squelched small variances in the more tightly crowded arrays. However, no other parameters of the inefficient, noisy observer model were significantly affected by crowding.

Discussion

The preceding modeling confirms that orientation-variance discriminations are both less noisy and more efficient than mean orientation discriminations, even when stimuli are crowded. Crowding's only effect seemed to be the imposition of a sensory threshold. This made it hard

for some observers to discriminate small variances from one another, but it had no effect on large variances or estimates of mean orientation.

General discussion

If human observers really do behave like noisy, inefficient versions of the ideal observer, then all of the current results suggest that they are more efficient, i.e., they use greater sample sizes when computing orientation variances. Stimulus configuration seems to make no difference. The estimates of sampling efficiencies derived from the demonstrably uncrowded configuration used in Experiments 2 and 3 were similar to those derived from the various configurations used in Experiments 4 and 5. In all cases, $M_{var} > M_{mean}$.

Whereas previous research has proven that crowding can drastically impair the identification of individual items, the current results are consistent with the notion that it preserves a good deal of statistical information. The only effect of crowding that can be seen in these results is the elevation of some observers' JNDs for discriminating between small variances. I have attributed this elevation to a sensory threshold. If this model is correct, then small orientation variances are sometimes squelched in crowded situations. This notion is consistent with several recent demonstrations that crowded targets appear tilted in the direction of slightly tilted flankers (Kapadia, Ito, Gilbert, & Westheimer, 1995; Mareschal, Morgan, & Solomon, 2010; Solomon et al., 2004). Note that both of these conditions seem to be necessary. "Small-angle assimilation" is not seen when the targets are not crowded (Solomon & Morgan, 2006) and it is not seen when the flankers have a substantially different tilt (Mareschal et al., 2010; Solomon et al., 2004).

Without specifically considering the effect of crowding, Morgan et al. (2008) suggested that the squelching of small orientation variances would be useful for eliminating early noise from higher level visual analyses. Why should this happen only for crowded orientation signals? I would like to speculate that the reason is that crowded signals typically emanate from the same object, and orientations within an object have a much greater likelihood of being parallel than those emanating from separate objects. In other words, I believe that the present results are consistent with the Bayesian notion that our perceptions result from the combination of the noisy data (i.e., likelihoods) our visual systems can gather from the environment and pre-existing knowledge (i.e., priors) about the environment, which may simply be hardwired into the visual system. The recent profusion of computational models inspired by this notion (e.g., Knill, Kersten, & Yuille, 1996) suggests an alternative approach (not dissimilar to that used by Tomassini, Morgan, & Solomon, 2010) to modifying the equivalent noise model so that it might account for the present data.

I cannot be certain why estimates of mean orientation are so much less efficient and less precise than estimates of orientation variance. However, it is by no means certain, or even likely, that human observers really do behave like noisy, inefficient versions of the ideal observer. There is no limit to the alternative statistics observers could compute and still produce reasonable psychometric functions. Consider, for example, an observer whose variance discriminations are based on orientation *range* (i.e., the smallest angle containing all 8 orientations). This range observer's
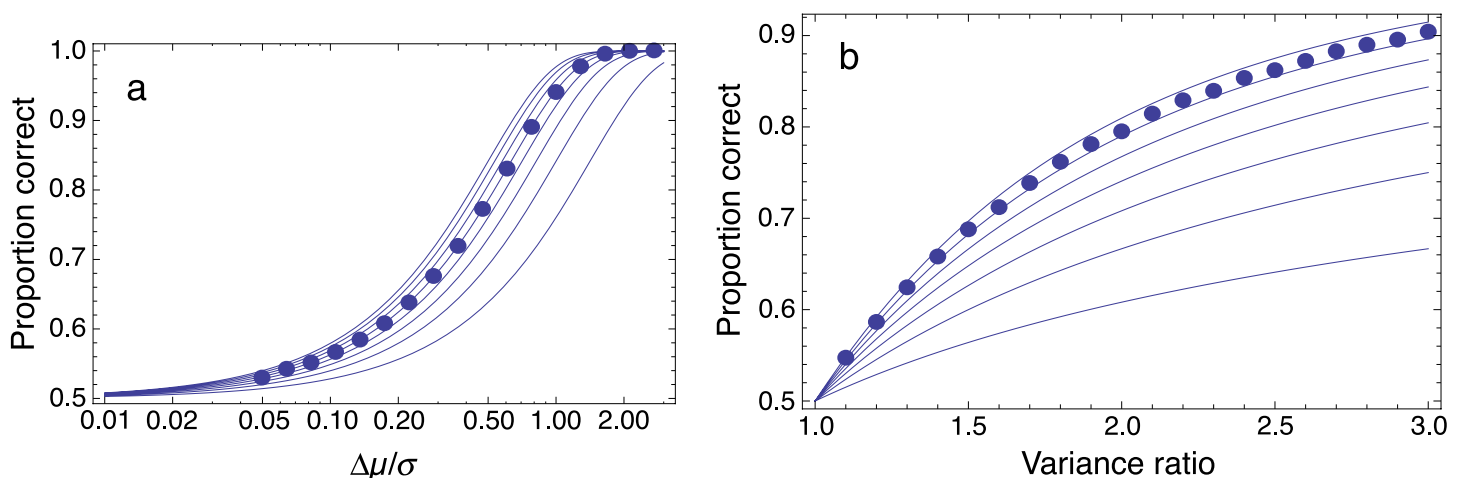

Figure 8. Comparing the performances of an otherwise ideal observer whose responses are based solely on the range of orientations in an 8-Gabor array (blue circles) and the performances of otherwise ideal observers having variously limited efficiencies (curves). (a) Psychometric functions for mean discrimination. Values on its horizontal axis represent the ratio of the difference between mean orientations to the S.D. of orientations in each interval. (b) Psychometric functions for variance discrimination. Values on its horizontal axis represent the ratio of orientation variances in the two intervals. In both panels, the highest curve represents 100% efficiency (i.e., $M = 8$). Successively lower curves represent integral decreases in sample size (i.e., $M = 7$, $M = 6$, etc.). The range observer is more efficient at variance discrimination ($M \approx 7$ in (b)) than mean discrimination ($M \approx 5$ in (a)).

performance is illustrated by the circular symbols in Figure 8b. For comparison, an observer whose mean discriminations are based on the middle of this range is illustrated by the circular symbols in Figure 8a.

The top curve in each panel represents the performance of a truly ideal observer, who must first estimate the mean orientation, and then estimate the sum of squared deviations away from it. The first (second, third, etc.) curve below the top represents the performance of an otherwise ideal observer, which ignores 1 (2, 3, etc.) randomly selected element(s). Figure 8b shows that the range observer's performance is similar to that of an ideal that ignores just 1 element in a 2AFC variance discrimination task.

However, when discriminating between mean orientations, the range observer's performance is similar to that of an ideal that ignores 3 elements. Thus, the behavior of the range observer is not qualitatively different from that of our human observers. Both exhibit greater efficiency when discriminating different variances than they do when discriminating different means.

Whereas mean orientation is a cyclical quantity, orientation variance is not. Even with successive displays of a single Gabor, our observers often exhibited some hesitancy when deciding whether the second was a clockwise or an anti-clockwise rotation of the first. Of course, both directions were correct; it is just that one rotation subtended smaller angle than the other. Variance discriminations have no such inherent ambiguity, and neither do mean discriminations on non-cyclic dimensions, such as size. It remains to be determined whether all feature means are estimated

with less efficiency and less precision than their corresponding variances, or whether the inferiority of mean estimates is specific to those features defined on cyclic dimensions, like orientation.

Appendix A

Supplementary Experiment 1

This experiment was designed to explore whether there might be an effect of mean orientation on the sample size for mean discrimination M_{mean}.

Methods

The methods were identical to those used in Experiment 2; however, JAS was the only observer. Staircases for eight different "mean" orientations were randomly interleaved. One of these means was used for one presentation in each trial. There were a total of 48 different conditions, each with its own staircase: eight mean orientations × three standard deviations of orientation ($\sigma_G \in \{1°, 8°, 16°\}$) × two numbers of Gabor per array ($N \in \{1, 8\}$).

Results

Results appear in Figure A1. The first thing to appreciate in these data is the oblique effect (Appelle, 1972). JNDs in
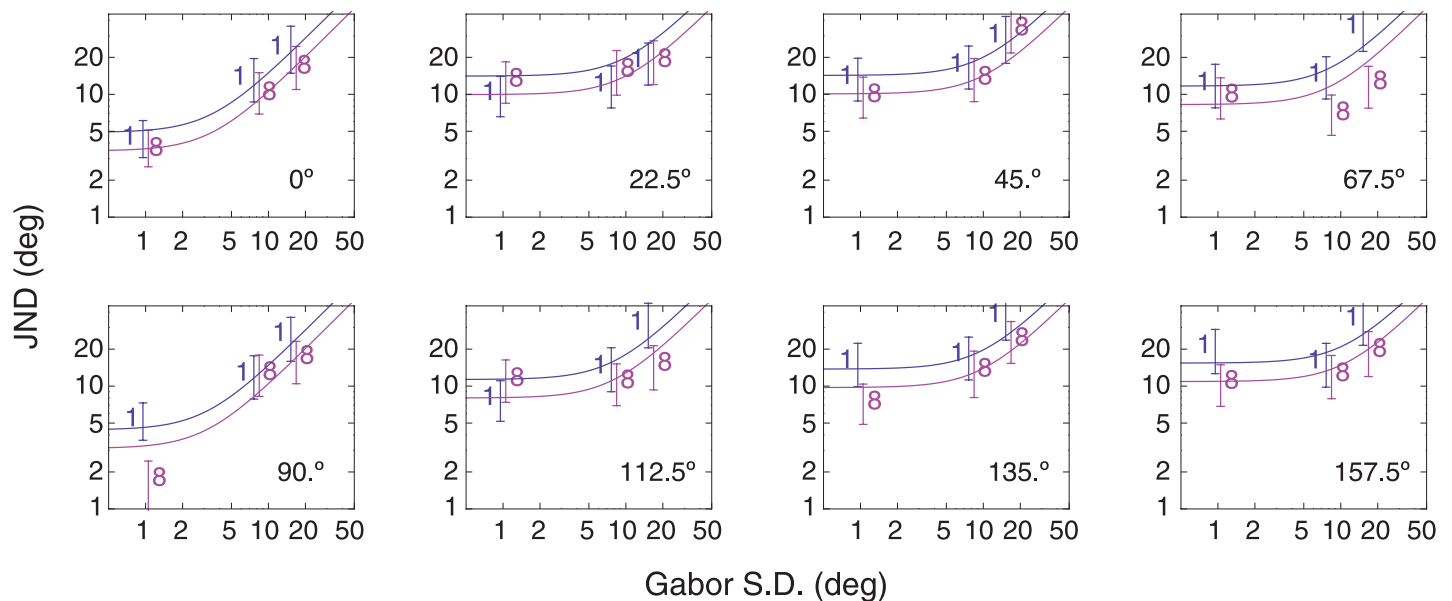

Figure A1. JNDs between the mean orientations in a pair of successively displayed, uncrowded Gabor arrays. The orientations of Gabors in one array of each pair were drawn from a Gaussian distribution having the mean orientation indicated (with respect to horizontal) in the lower right-hand corner of each panel. Other numerals denote the number of Gabors per array. They have been nudged horizontally for better legibility. Error bars contain 95% confidence intervals. Curves illustrate maximum likelihood fits of two-parameter equivalent-noise models. Eight-Gabor-array data were best fit with a sample size of two Gabors per array (magenta curves). All other data are best fit with a sample size of one Gabor per array (blue curves).

the upper left (horizontal mean) and lower left (vertical mean) panels are lower than those in the other panels. The second thing to notice is that there usually is not too much difference between the $N = 1$ JNDs and the $N = 8$ JNDs.

Modeling

In order to estimate both the equivalent noise and the efficiency of an otherwise ideal observer, it is necessary to have at least one JND on the left-hand (i.e., essentially flat) side of the JND-vs.-noise curve and another on the right-hand side (which has an asymptotic gradient of 1). On the other hand, the model described in Equation 2 is not strictly valid on cyclic dimensions, because (with what are consequently Wrapped Normal distributions) the sample and population variances become increasingly different as the latter increases. This problem of "wrap-around" becomes especially acute when the probability density of orientations orthogonal to the population mean is non-negligible.

To avoid wraparound, I adopted the seemingly conservative criterion of discarding any data for which 45° (the upper limit of the ordinates in Figure A1) is contained in the 95% confidence interval for threshold. The remaining data were fit, separately for each mean, with the two-parameter, equivalent noise model.

In 6 of 8 cases, maximum likelihood fits were obtained when $M_{mean} = \begin{cases} 1 & N = 1 \\ 2 & N = 8 \end{cases}$. (The exceptions were mean = 22.5°: $M_{mean} = 1 \forall N$; and mean = 67.5°: $M_{mean} = \begin{cases} 1 & N = 1 \\ 3 & N = 8 \end{cases}$.) Forcing $M_{mean} = \begin{cases} 1 & N = 1 \\ 2 & N = 8 \end{cases}$ for all mean

orientations and refitting the data resulted in an insignificant decrease in joint likelihood.

Furthermore, when these data were collapsed across mean orientation, the best fit of the equivalent noise model once again was obtained when $M_{mean} = \begin{cases} 1 & N = 1 \\ 2 & N = 8 \end{cases}$ (and $\sigma_E = 8.6°$).

Discussion

The foregoing model fits make a strong case against an "oblique effect" on the sampling size for mean discrimination M_{mean}. The fits also strongly suggest that Experiment 2's randomization of mean orientation in no way contaminated the estimates of sampling efficiency that were inferred from its results.

Appendix B

Supplementary Experiment 2

This experiment was a straightforward attempt to replicate a result originally described by the squares in Figure 3 of Parkes et al. (2001).

Methods

The only differences between Supplementary Experiment 2 and the methods described by Parkes et al. (2001)
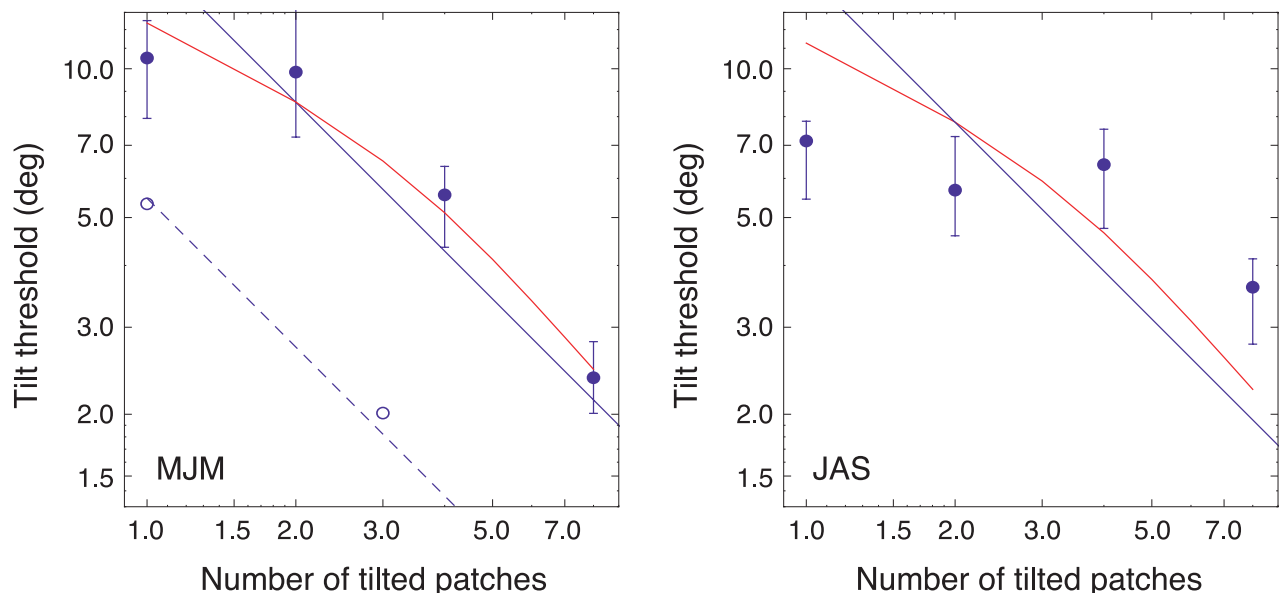

Figure B1. The tilt of a parallel subset required for discriminating clockwise from anti-clockwise in otherwise vertical, 8-Gabor arrays in the center of the visual field. Solid symbols summarize data from Supplementary Experiment 2. Open symbols were taken from Figure 3 of Parkes et al. (2001). Solid blue and dashed lines with gradients of −1 are those that best fit each set of data, respectively. Red curves illustrate the behavior of the best-fitting "max model."

were (i) I used 8 Gabors instead of 9 and (ii) I used QUEST (two 100-trial staircases per data point).

Results and modeling

The left panel in Figure B1 shows data from MJM (the sole observer in the original experiment). The right panel shows data from JAS. Each observer achieved 100% correct on 96 randomly interleaved "bonus" trials with targets tilted 45°. Thus, neither observer seems to randomly ignore any of the eight Gabors. Nonetheless, only the rightmost three of MJM's four data points are consistent with a gradient of -1, which is the signature of orientation averaging, and only the rightmost two of JAS's data points are consistent with that gradient. In fact, MJM's data are more consistent with the "max model," in which decisions are based on the direction of the most oblique Gabor, after its orientation has been perturbed by an independent sample of Gaussian noise (Baldassi & Verghese, 2002; Morgan & Solomon, 2005; Parkes et al., 2001). However, not even the max model can provide a satisfactory fit to the data from JAS.

Discussion

These results do not fit easily within the equivalent noise framework. Neither decisions based on an average orientation nor those based on the maximally tilted one would be ideal for this task (Solomon & Morgan, 2001). The only thing clear from these data is the seemingly qualitative difference between the observers' performances. JAS was better with a single tilted target, and MJM was better when all eight were tilted.

Acknowledgments

Commercial relationships: none.
Corresponding author: Joshua A Solomon.
Email: J.A.Solomon@city.ac.uk.
Address: City University London, London, EC1V 0HB, UK.

Footnotes

[1]To appreciate the necessity of a roving pedestal, consider what might happen without one. If trials were blocked by pedestal variance, observers could learn to make a reasonable response on each trial without ever seeing the second alternative.

[2]These fits were obtained when five parameters were allowed to vary: σ_E, σ_L, and a separate value of M_{mean} for $N = 2$, 4, and 8. (By definition, $M_{mean} = 1$ when $N = 1$.) Since the displays used in this experiment were demonstrably uncrowded, it seems safe to assume that σ_E remains invariant with N. However, it is conceivable that σ_L does not. I have not adopted the assumption that σ_L remains invariant with N in any subsequent modeling, but here it is necessary to avoid "over-fitting" these data with too many free parameters.

[3]A comparison between the generalized likelihood ratio and the chi-square distribution was used to statistically test all hypotheses in this study. For all of these tests, the critical region α was 0.05. Do not confuse this α, preferred by Mood et al. to the somewhat more popular P-value, with the Weibull mean in Equation 1!

[4]That paper distinguishes (hard) sensory thresholds, below which all variances are indistinguishable, from the softer thresholds produced by an accelerating transducer. This distinction will not be maintained here. Both hard and soft thresholds effectively diminish the differences between small variances, and both produce negative masking. A multiple-alternative ranking paradigm (e.g., two-response/four-alternative forced choice; Solomon, 2007) would be required to distinguish between hard and soft thresholds.

References

Appelle, S. (1972). Pattern and discrimination as a function of stimulus orientation: The "oblique effect" in man and animals. *Psychological Bulletin, 78,* 266–278.

Baldassi, S., & Verghese, P. (2002). Comparing integration rules in visual search. *Journal of Vision, 2*(8):3 559–570, http://www.journalofvision.org/content/2/8/3, doi:10.1167/2.8.3. [PubMed] [Article]

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Cheal, M., & Lyon, D. R. (1991). Central and peripheral precuing of forced-choice discrimination. *Quarterly Journal of Experimental Psychology, 43,* 859–880.

Dakin, S. C. (2001). Information limit on the spatial integration of local orientation signals. *Journal of the Optical Society of America A, 18,* 1016–1026.

Durlach, N. I., & Braida, L. D. (1969). Intensity perception: I. Preliminary theory of intensity resolution. *America, 46,* 372–383.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron, 15,* 843–856.

Knill, D. C., Kersten, D., & Yuille, A. (1996). Perception as Bayesian inference. In D. C. Knill & W. Richards (Eds.), *Perception as Bayesian inference* (pp. 1–21). Cambridge, UK: Cambridge University Press.

Mareschal, I. M., Morgan, M. J., & Solomon, J. A. (2010). Cortical distance determines whether flankers cause crowding or the tilt illusion. *Journal of Vision, 10*(8):13, 1–14, http://www.journalofvision.org/content/10/8/13, doi:10.1167/10.8.13. [PubMed] [Article]

Mood, A. M., Graybill, F. A., & Boes, D. C. (1974). *Introduction to the theory of statistics.* McGraw-Hill.

Morgan, M., Chubb, C., & Solomon, J. A. (2008). A 'dipper' function for texture discrimination based on orientation variance. *Journal of Vision, 8*(11):9, 1–8, http://www.journalofvision.org/content/8/11/9, doi:10.1167/8.11.9. [PubMed] [Article]

Morgan, M. J., & Solomon, J. A. (2005). Capacity limits for spatial discrimination. In L. Itti, G. Rees, & J. Tsotsos (Eds.), *Neurobiology of attention* (pp. 8–10). San Diego, CA: Elsevier.

Morgan, M. J., & Solomon, J. A. (2006). Attentional capacity limit causes spatial neglect in normal observers. *Vision Research, 46,* 1868–1875.

Nagaraja, N. S. (1964). Effect of luminance noise on contrast thresholds. *Journal of the Optical Society of America, 54,* 950–955.

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research, 29,* 1631–1647.

Parkes, L., Lund, J., Angelluci, A., Solomon, J., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience, 4,* 739–744.

Pelli, D. G. (1990). The quantum efficiency of vision. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 3–24). Cambridge, UK: Cambridge University Press.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Solomon, J. A. (2007). Intrinsic uncertainty explains second responses. *Spatial Vision, 20,* 45–60.

Solomon, J. A. (2009). The history of dipper functions. *Attention, Perception & Psychophysics, 71,* 435–443.

Solomon, J. A., Felisberti, F. M., & Morgan, M. J. (2004). Crowding and the tilt illusion: Toward a unified account. *Journal of Vision, 4*(6):9, 500–508, http://www.journalofvision.org/content/4/6/9, doi:10.1167/4.6.9. [PubMed] [Article]

Solomon, J. A., & Morgan, M. J. (2001). Odd-men-out are poorly localized in brief displays. *Journal of Vision, 1*(1):2 9–17, http://www.journalofvision.org/content/1/1/2, doi:10.1167/1.1.2. [PubMed] [Article]

Solomon, J. A., & Morgan, M. J. (2006). Stochastic re-calibration. *Proceedings of the Royal Society B, 273,* 2681–2686.

Tomassini, A., Morgan, M. J., & Solomon, J. A. (2010). Orientation uncertainty reduces perceived obliquity. *Vision Research, 50,* 541–547.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33,* 113–120.

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