

On the Scaling of Size Judgements by Orientational Cues

MICHAEL J. MORGAN*

Received 15 July 1991; in revised form 6 February 1992

Observers carried out multiple, concurrent size discriminations with a range of size standards. The task was to classify each stimulus as larger or smaller than the appropriate standard size for the set to which it belonged. The set to which each stimulus belonged was indicated by its orientation, or in different experiments, by its spatial location. Observers were able to maintain appropriate discrimination, both when there were four concurrent standards and when there were eight. Both angle and position functioned as effective cues. The size of the orientational cue appeared to make little difference to the efficiency of discrimination. However, when the relationship between standard size and orientation was random, rather than regular, performance got worse. The analogy between such discrimination and size constancy is pointed out, and the results are discussed in relation to Andrews, D. P.'s [(1964) *Quarterly Journal of Experimental Psychology*, 16, 104–115] account of perceptual calibration.

Size Orientation Discrimination Constancy Calibration

INTRODUCTION

In everyday perception, it is a commonplace observation that the perceived size of objects depends not only upon the size of their retinal images, but upon the relation of their image size to their distance, a phenomenon referred to as “size constancy” or “size distance invariance.” In the laboratory, size constancy has been studied by a variety of techniques including matching, though seldom by obtaining full psychometric functions in forced-choice discrimination tasks. Psychometric functions, obtained by the method of constant stimuli or by one of its modifications, have the advantage that they can measure not only the central tendency in the observer's responding, and any accompanying biases, but also the sensitivity to changes along the underlying continuum (e.g. Morgan, Hole & Glennerster, 1990). A psychometric function would allow us, for example, to determine whether we are as consistent in comparing the sizes of objects at different distances as we are when they are at the same distance. Little information is available on this point.

An exception is a study by Burbeck (1987) which compared spatial frequency discrimination for a pair of grating targets presented at a single viewing distance to that for the identical grating targets at two different viewing distances. There was no difference in the just-noticeable-difference (jnd) for frequency between these two conditions, which implies that size constancy adds no significant extra uncertainty to the comparison process. More recently, Welch and McKee (1990) and

McKee and Welch (1992) measured thresholds for size discrimination when the stimuli were presented stereoscopically at varying distances. Following an earlier method described by Westheimer and McKee (1977), the observer's task was to decide whether the separation between a pair of lines was larger or smaller than a standard interval, the standard being the mean (and the mode) of the stimulus set. The ability of observers to compare stimuli to a purely internal standard has been well documented, and is clearly relevant to the question of the underlying metric used in size judgements. In these most recent studies, McKee and Welch used a standard that was scaled in angular size appropriately to its stereoscopic disparity. Thus, a stimulus that was correctly judged “large” at one disparity, would be judged correctly as “small” at a disparity that caused it to be seen as nearer to the observer. Performance under this “objective size” condition was compared to that under an “angular size” condition where the observer had simply to ignore the disparity and respond on the basis of the angular size of the stimulus alone. Performance under these two conditions was similar, provided that the mean stimulus size was 20 arcmin or greater. For smaller stimuli, angular jnds were smaller than objective size jnds.

The finding that highly accurate size discriminations can be influenced by size constancy supports the view that such discriminations can involve flexible decision rules. Is size constancy unique in this respect, or is it a manifestation of a much deeper underlying process of stimulus scaling? Some time ago, Andrews (1964) argued that constancy-like processes were part of a more general scheme for sensory calibration. The spirit of Andrew's argument is captured by the following example. Suppose

*Laboratory for Neuroscience, University of Edinburgh, 1 George Square, Edinburgh, Scotland.

that on average the retinal images of objects were smaller in the upper visual field than in the lower, and that the visual system had some means of registering this fact. It is extremely unlikely that the real size of objects changes as a function of their position on the retina, so the assumption can be made that this statistical difference is an artifact of viewing geometry. It would therefore be advantageous to correct for the statistical difference and to rescale the size of objects by their position in the image so that the average perceived size of objects is position invariant.

In the particular case just described, there is an obvious connection with size constancy, since the retinal image size of objects does tend to diminish when they are further away along the ground plane and thus higher in the image. But the scheme could apply quite generally. To take a further example, if horizontally-oriented lines were statistically larger (in the image) than vertically-oriented lines, we would expect an association between size and orientation to be formed. The idea that constancy depends upon learning was basic to the Empiricist philosophy of Perception, originated by John Locke and developed into a subtle theory of subjective geometry by Helmholtz (for an historical review see Morgan, 1977). If Andrews is correct, and constancy is an instance of a much more general process of sensory calibration, we might expect to be able to obtain comparable results to those of McKee and Welch with other correlated dimensions, in addition to the familiar case of size and distance. Here we investigate the correlation between size and orientation. It should be possible for the observer to correct for a systematic association between angle and size just as they can correct for an association between size and retinal disparity. If the sensory calibration scheme is general it should not matter that the association between two variables is an arbitrary one. In fact this conclusion is already prefigured by one of McKee and Welch's results: they found that observers could maintain a consistent size scale when the association between size and disparity was reversed from its normal direction.

In the experiments described here, the basic task carried out by the observers was a spatial-interval discrimination, in which they had to decide whether a pair of lines was more or less separated than the appropriate

standard. However, the standard varied according to the orientation in which the pair of lines was presented. A stimulus that would be correctly classified as "wider" in one orientation might be correctly classified as "narrower" in a nearby orientation (Fig. 1). Response feedback was given to aid the observers in selecting their responses, and separate psychometric functions were collected for each orientation in the experiment.

In this paper we ask the following questions: (1) Can multiple discriminations be carried out with an accuracy comparable to that of single discriminations? (2) Can multiple discriminations be maintained with their appropriate standards, without biases arising from neighbouring sets? (3) Does the magnitude of the orientational cue affect the ability to maintain separate discriminations of spatial interval? And finally, (4) is it important that the relation between angle and the size of the standard interval appropriate to that angle be monotonic, or will any one-to-one metrical relationship suffice?

GENERAL METHODS

Apparatus

Stimuli were generated by fast point plotting on a computer controlled Hewlett-Packard high-resolution display oscilloscope (Morgan & Watt, 1982, 1989) and were viewed from a distance of 1.8 m. On each trial the stimulus consisted of two parallel bars, each of length 12.8 arcmin, and of a specified orientation and separation. Each bar consisted of points spaced at 0.22 arcsec, and refreshed at a rate of 67 Hz. The bars were presented on an otherwise blank screen, with a luminance at least 1 log unit above detection threshold. A dim veiling luminance ($< 1 \text{ cd/m}^2$) was provided by leaving the door to the experimental room slightly open. One observer (DK) made his observations with the room lights on (veiling luminance of 20 cd/m^2). Stimuli were presented for 1 sec only, and were immediately followed by a random-dot pattern mask to prevent visible persistence.

Psychophysical method

For each standard stimulus independently, we measured the minimum detectable change in separation (δs) by the method of single stimuli. That is, on each trial the observer saw a single stimulus consisting of a pair of

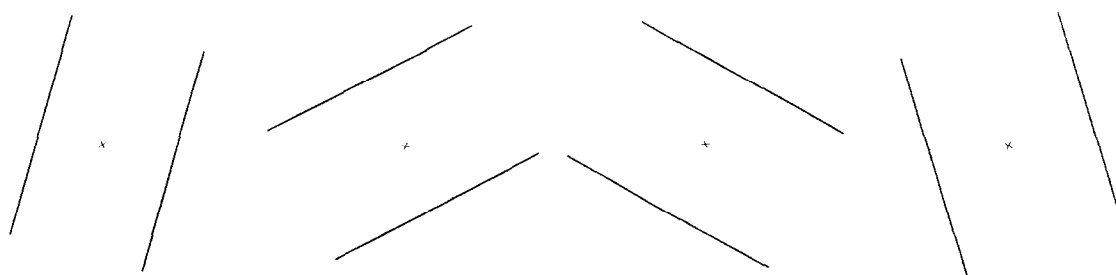


FIGURE 1. The figure shows a schematic representation of the four different size standards used in the experiments. The observer's task was to compare each stimulus presented in the experiment to the appropriate one of these four standards and to classify it as large or small. The appropriate standard of the stimulus was indicated by its orientation on the screen. The stimuli were presented in the same region of the screen, symmetrically placed around a fixation point (represented by a dot in the figure). In the case illustrated, the four standards are 50 deg apart in orientation, and differ in size by 7% of the smallest stimulus.

lines and had to judge whether the separation between the lines was larger or smaller than the appropriate standard, which had to be memorised. The actual size of the stimulus on each trial was determined by Watt and Andrews' (1981) method of adaptive probit estimation (APE). On each trial APE chooses the actual stimulus value randomly from four possible values placed symmetrically around the mean (standard) stimulus value. The minimum spacing of the stimuli used to determine the psychometric function was equal to 1/60 of the mean value of the stimulus set (i.e. for a 6 arcmin standard, the spacing was 6 arcsec). However, the actual spacing between these values in force at any time is chosen adaptively, and is increased as a function of the estimate of the observer's threshold, based upon the data collected so far during the session.

Response feedback was given in the form of a tone following incorrect responses. An incorrect response is defined as a "wider" response to a stimulus that was actually narrower than the mean of its appropriate set, or vice versa. No feedback was given on trials when the stimulus was actually equal to the mean size of the set. To determine each psychometric function, a total of 64 trials was preceded by 8 trials in which feedback was given but the results were not included in the analysis. The standard deviation of the psychometric function was determined by Probit analysis (Finney, 1971) and this value was defined as the threshold (δs). The psychometric function was also analysed to determine the stimulus value at which the observer responded "wider" and "narrower" equally frequently [the point of subjective equality or P(50) point]. The difference between the point of subjective equality and the actual mean value of the set will be referred to as the bias for that set of stimuli.

In contrast to the conventional discrimination task, the standard interval against which the observer had to compare the current stimulus could vary from trial to trial. A second attribute of the stimulus, in addition

to the spatial separation of the two lines, indicated to the observer the standard size for that particular stimulus. The second attribute could be either the orientation of the stimulus, or its position on the screen. The number of concurrent discriminations was four in the early experiments, and eight in the later series. The values of the standard interval were spaced in steps of 30 arcsec. The association between angle and standard interval could either be monotonic (i.e. the size of interval increased systematically with angle away from the vertical) or pseudo-random. The latter will be referred to as the "disordered" condition.

Different psychometric functions were collected for each of the concurrent discriminations, and they were analysed separately at the end of each session. The different discriminations were randomly interleaved and were sampled without replacement until there had been 64 trials in each case. In a variant of this procedure referred to as the "blocked" condition, the thresholds for each of the discriminations was determined sequentially, with all the trials for discrimination 1 preceding those of discrimination 2, and so on. After each new block began, the observer was permitted to sample the new stimulus set by pressing on a third response button. These responses were not recorded, and since no decision was made, no feedback was given. The inspection button was also available during the randomly-interleaved condition, and observers were encouraged to use it freely at the start of each testing session.

In considering the precision of the observer's performance from these data, two issues must be logically separated. The distinction is explained graphically in Fig. 2 with the help of two highly schematic psychometric functions (PMFs). The first issue concerns the observer's bias. The left-hand function in Fig. 2 has no bias and when the stimulus has the same size as the standard (0 cue value on the horizontal axis) the observer responds "larger" 50% of the time. In the right-hand function the observer has a strong bias

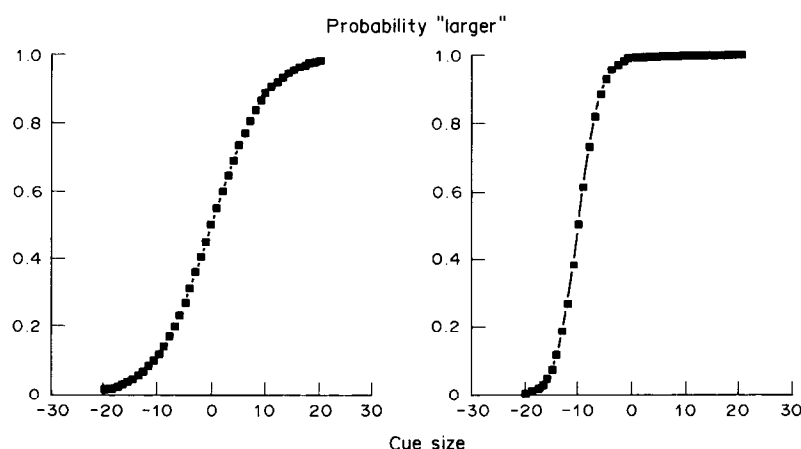


FIGURE 2. The figure shows two schematic psychometric functions to explain the difference between bias and slope. When the cue size is zero the stimulus has the same size as the standard; when it has positive values it is larger than the standard, and when it is negative it is smaller. The function on the right is sharper on the left and would yield a smaller jnd (defined by the standard deviation of the function). However, the function of the right also has a bias that shifts it leftwards. This increases the number of errors made by the observer, where an error is defined as a "larger" response to a smaller stimulus or vice versa. The function of the left is less steep and would yield a larger jnd than the one on the right, but it has no bias.

towards responding "larger" and the 50% point is shifted to smaller stimulus values. The bias is measured by the difference of the 50% point from zero: in this case, about 10 stimulus units. Note that this bias will introduce errors into the observer's responding.

Logically distinct from the question of bias is that of the observer's consistency. The left-hand function in Fig. 2 is more spread out than the one on the right, which means that there is a larger range of stimulus values over which the observer's responses are uncertain (in the information theoretic sense). A useful measure of spread of the PMF is its standard deviation, which we shall use as the measure of the jnd. Note that the jnd and bias are unrelated: in the example shown, the function with the larger bias also exhibits the smaller jnd.

A further point to be considered in relation to the PMFs is the influence of *misclassification* errors. The observer's task in these experiments is to assign each stimulus to its appropriate category, and then to decide whether the stimulus is "wide" or "narrow" in relation to the stimulus range for that category. If *all* the stimuli in a particular set were misclassified, for example by being assigned to the category nearest to the mean for the whole set, the effect will be to introduce a pure bias, without altering the shape of the PMF. If, on the other hand, stimuli in a category are sometimes correctly classified and at other times incorrectly classified, the effect will be to decrease the slope of the PMF and thus increase the threshold.

Subjects

The author (MM) served in all the experiments, and various other observers (HR, DK, AH) took part in different phases of the study. They were all psychophysically experienced, but had not previously taken part in experiments on spatial interval discrimination.

Order of experimental conditions

Different observers met different conditions in different orders, to control for the possible effects of long-term learning throughout the series. The order of training is given in detail in the Methods section for each experiment. MM experienced all the conditions, as well as many not described here, and had thus received much more extensive practice than any of the other observers.

EXPERIMENT 1

The aim of the first experiment was to investigate performance with four standards (7, 7.5, 8 and 8.5 arcmin) separated by an orientational difference of 1.5 deg. Previous experiments had shown that this was the smallest orientational difference for which MM could categorically identify the four angles without error. The four angles were spaced equally around the horizontal (90 deg) and the association between angle and interval was monotonic: thus the 7.0 interval set was associated with an orientation of 87.75 deg, the 7.5 set with an orientation of 89.25 deg, and so on (Fig. 1). Also investigated was a positional condition in which the four

different sets were presented in four different quadrants of the display screen (Fig. 1): in this case all the stimuli were presented at an angle of 90 deg.

Method

Observers and conditions. The two observers were MM and HR, whose experience with different conditions in the experiment was as follows. MM had carried out previous investigations with standards of 2, 4, 8 and 16 arcmin and with orientational separations varying from 0.5 to 50 deg. The effects of a positional rather than an orientational cue had also been investigated, as had the effects of reducing the number of concurrent stimuli in the range 1–4. Immediately before carrying out Expt 1, MM had carried out four sessions with the set 7.0, 7.5, 8.0 and 8.5 arcmin, with an orientational separation of 50 deg. The orientational separation was then decreased to 0.5 min for three sessions, and to 1.5 min for two sessions. The data reported in Expt 1 are the means for these last two sessions. MM then carried out six sessions with a disordered relation between angle and size, and then three sessions in the blocked conditions; the results reported in Expt 1 are the means over the last two of these sessions. MM then returned to an ordered relation with an orientational separation of 3 deg for four sessions. This was followed by four sessions with a disordered relation between angle and size, and then by six sessions with a positional cue. The positional data reported in Expt 1 are the means for the last two of these sessions.

HR's history was simpler. She began with interval sets of 7.0, 7.5, 8.0 and 8.5 arcmin and stayed with these throughout all experiments. She began with three sessions in the "blocked" condition; the "blocked" data reported in Expt 1 are the means over the last two of these sessions. She then experienced the interleaved orientational 1.5 deg condition for seven sessions; the means of the last two of these are reported. HR then had two "blocked" sessions with a disordered relation between angle and size, followed by five sessions of randomly interleaved training in this condition. This was followed by a return to the "ordered" condition with an orientational cue of 3 deg for four sessions, and then by four sessions with a positional cue. The mean of the last two positional sessions is reported in Expt 1.

Results and discussion

Both observers were able to carry out the task, in the sense that the psychometric functions showed clear evidence for discrimination between the sets of stimuli. A typical set of psychometric functions from a single session is shown in Fig. 3. The important point to note is that the probability of a "wider" response depends not upon the absolute value of the stimulus, but also upon the set it belongs to. There are many cases where a given stimulus has two or even three different probabilities associated with it, depending on the set to which it belongs. Since the only cue to the set in these circumstances can have been the orientational cue, it follows that the observer was using this cue as well as the width

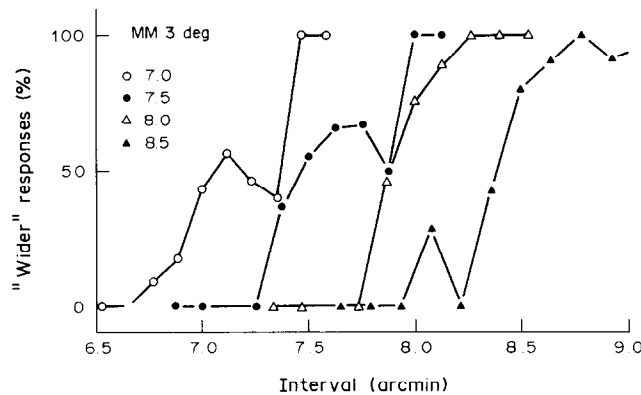


FIGURE 3. The figure shows representative psychometric functions from a single session with one observer (MM) in Expt 1. Each of the curves is derived from a different standard set, which was cued by its orientation. Because an adaptive method of stimulus presentation was used the number of observations in each data point are not equal, and there are many more observations in the centre of each function than towards its extremes. For this reason, points on the function with two or fewer observations have been eliminated. Note that the functions are distinct but overlapping on the horizontal axis: this means that same-sized stimuli were classified differently according to the set to which they belonged.

cue in arriving at a decision. The separation between the functions is not perfect; there are two cases where they almost touch, but none where they cross.

From functions like those in Fig. 3, thresholds and biases were determined, and these are shown separately

for each of the standard intervals in Fig. 4. Weber fractions for MM in the orientational-cue condition were in the region of 3%, which compares very favourably with independent reports of spatial interval acuity for simple, single discriminations (for a review see Morgan, 1990). This is impressive performance, bearing in mind the small orientational difference between the standards. Performance was not markedly better in the blocked condition, where the discriminations were carried out successively rather than in parallel, except with the 8.0 arcmin standard where an anomalously low Weber fraction of <2% was achieved.

Thresholds were considerably higher in the second observer (HR), possibly because of her lack of practice. Weber fractions with the orientational cue were in the region of 7%, but these were no worse than thresholds in the blocked condition. This observer's performance did improve considerably with the positional cue: thresholds in this condition were about 3%. For MM, on the other hand, thresholds with the positional cue were similar to those with the orientational cue. These differences between observers may reflect genuine differences in ability to use the orientational cue, or they may equally be due to differences in practice.

If observers had failed to use the orientational cue or the positional cue, and had instead simply judged each stimulus in relation to the mean for the whole set, they would have shown large biases in their psychometric

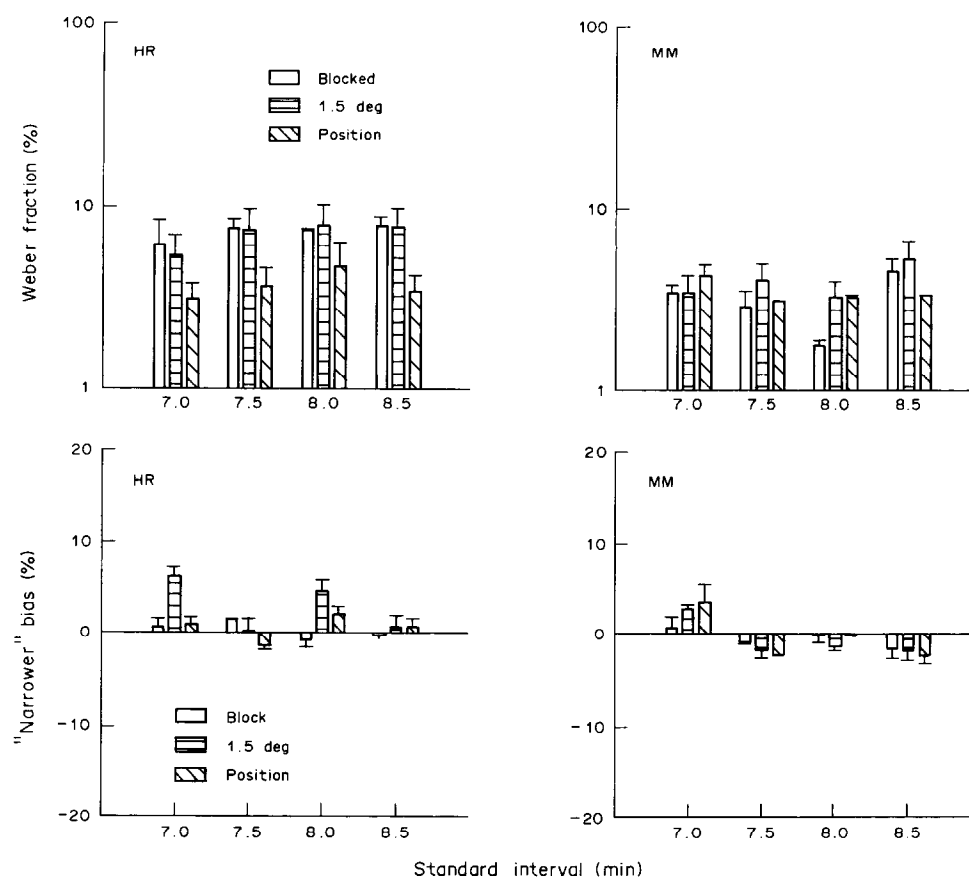


FIGURE 4. The four panels illustrate the results of Expt 1 for two observers: HR (left hand panels) and MM (right hand panels). The top panels show the thresholds (Weber fractions) under the orientational (1.5 deg) and positional cue conditions, and in a blocked condition where thresholds for the four stimulus standards were determined sequentially rather than being randomly interleaved. The bottom panels show the values of the 50% points (the PSEs) on the psychometric functions.

functions. Stimuli in the 7.0 arcmin set should mostly have been judged to be "narrow" while those in the 8.5 arcmin set should have been judged "wider". The mean of the whole set was 7.75 arcmin and thus the bias from indiscriminate performance would have been +0.75 arcmin (approx. +10%) for the 7.0 arcmin set and -0.75 (approx. -10%) for the 8.5 arcmin set. The bias data shown in Fig. 4 demonstrate that the biases were significantly smaller than these. HR showed only very small biases, which were mostly positive over the whole set. MM showed an increasing "wider" bias with a larger stimuli, but the effect was a small one. Further evidence on this point will be presented in the form of the psychometric functions in later experiments.

EXPERIMENT 2

The aim of the second experiment was to determine whether discrimination performance would be improved by increasing the orientational difference between the sets. The 1.5 deg orientational cue in Expt 1 was close to the threshold for orientational discrimination itself, which is in the region of 1 deg (Heeley & Buchanan-Smith, 1990). With such a small cue, there would have been considerable overlap between spatial-frequency tuned filters signalling "narrower" at one orientation, and those signalling "wider" at a neighbouring orien-

tation. Indeed, it is quite hard to see how the discrimination could be maintained at all if it depends on filters simultaneously tuned to width and to orientation. An alternative is that the stimuli are analysed categorically for orientation in order to select the appropriate spatial interval standard: if this is the case, it should matter little how separate the stimuli are in orientation, provided that the different orientations can be accurately labelled.

Performance with orientational cues of 1.5 and 3.0 deg were compared in observers MM and HR. A third observer (AH) was tested with orientation cues of 3, 6 and 50 deg. In all cases there were four interleaved spatial-interval standards, exactly as in the first experiment.

This experiment also included (for observer MM only) a control condition in which the orientational cue was entirely absent. All the stimuli in this condition were presented at the same orientation (horizontal). There was thus no opportunity to separate the classes of stimuli, and the observer had to judge all the stimuli as "wider" or "narrower" than an arbitrarily chosen internal standard.

Method

Observers and conditions. The order in which HR and MM met the 1.5 and 3.0 deg orientation conditions was explained in the Methods section of Expt 1. Means were

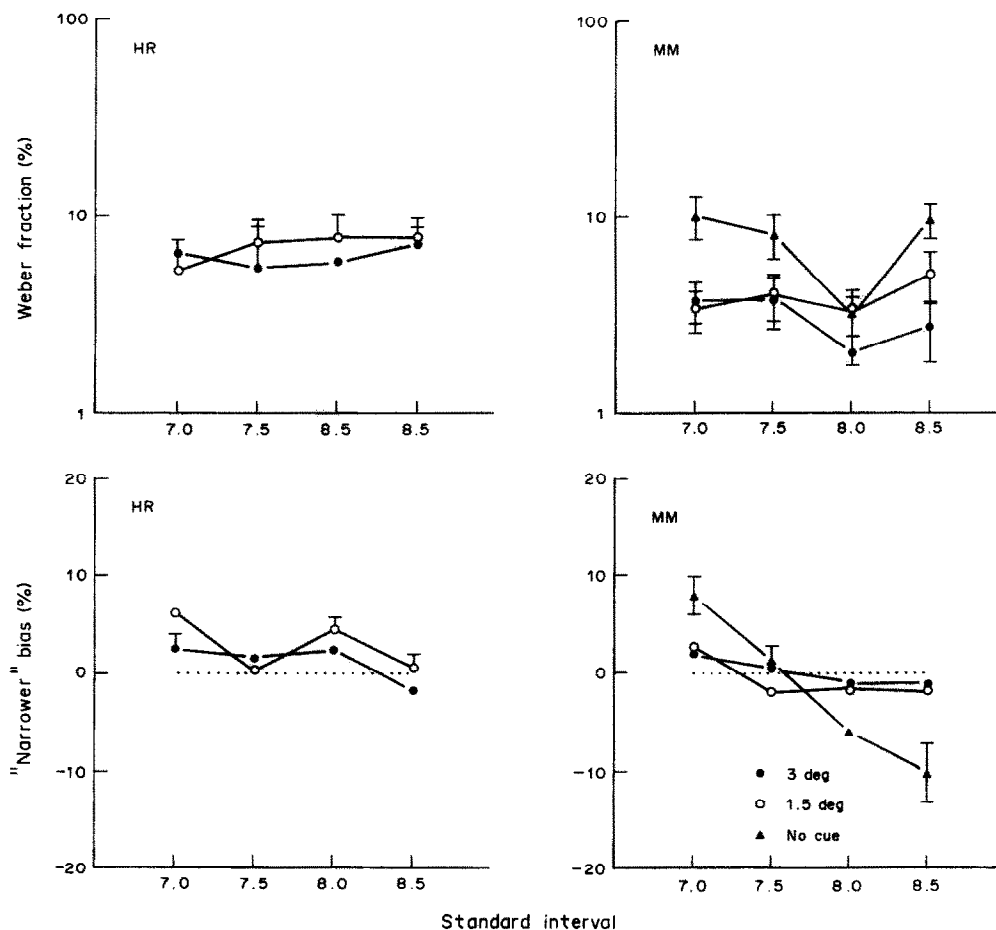


FIGURE 5. The figure illustrates the results from Expt 2, in which the size of the orientational cue was increased to 3 deg. The 1.5 deg data are derived from Expt 1. The size of the orientational cue appears to make little systematic difference to performance.

taken from the last two sessions in each condition. MM experienced the 50 deg condition shortly before the 3 deg condition: training was given for four sessions, and means were taken from the last two. The third observer, AH, had relatively little experience with the task. He began with four sessions using a 50 deg cue; this was followed by two sessions with a positional cue, three sessions with a 3 deg cue; two more sessions with a positional cue, and finally, three sessions with a 6 deg cue. Data (means) were taken from the last two sessions in each condition.

Results and discussion

Thresholds and biases for angles of 3 and 1.5 deg are compared in Fig. 5. In general, there was little effect of the different orientational cue size. Observer HR achieved Weber fractions in the region of 7% with both angles and the effects on her biases were insignificant. MM showed an improvement in acuity with the 3 deg orientational cue at the two larger separations, but not with the smaller. Biases were unaffected, and were negligible with both sizes of orientational cue.

In the "no cue" condition MM had significantly elevated thresholds and a large systematic bias. The bias is in the direction expected if each stimulus is compared to the mean of the whole set rather than to its appropriate standard. In theory, the observer could choose an arbitrary standard from somewhere within the set and classify all stimuli as "wider" or "narrower" than that

standard, irrespective of response feedback. The result would be a unitary psychometric function, and the threshold for each of the four standards would be calculated from different segments of this unitary function. In these circumstances it would be possible for one of the stimulus sets, placed near to the actual centre of the psychometric function, to have a threshold close to normal values. Something like this seems to have happened for the 8.0 arcmin condition, where the threshold in the "no cue" condition was similar to that with a cue present. Further evidence on the nature of performance without an orientational cue will be presented in Expt 5.

Results for the 50 deg condition vs 3 deg are shown in Fig. 6. There was little systematic effect of cue size upon threshold in either observer. The less-practised observer (AH) achieved Weber fractions of about 6–7%, quite similar to those of HR with a 3 deg cue. There was no evidence for lower thresholds with the large orientational cue: if anything, the data for MM point in the opposite direction.

As we found in Expt 1, biases were much smaller than would have been expected if observers had classified stimuli by the mean of the overall set. For observer AH there was an overall "narrower" bias, which was reduced in the wider stimuli when the orientational cue was small, but not when it was large.

Some additional data on the effects of orientational separation upon acuity and bias are shown for one

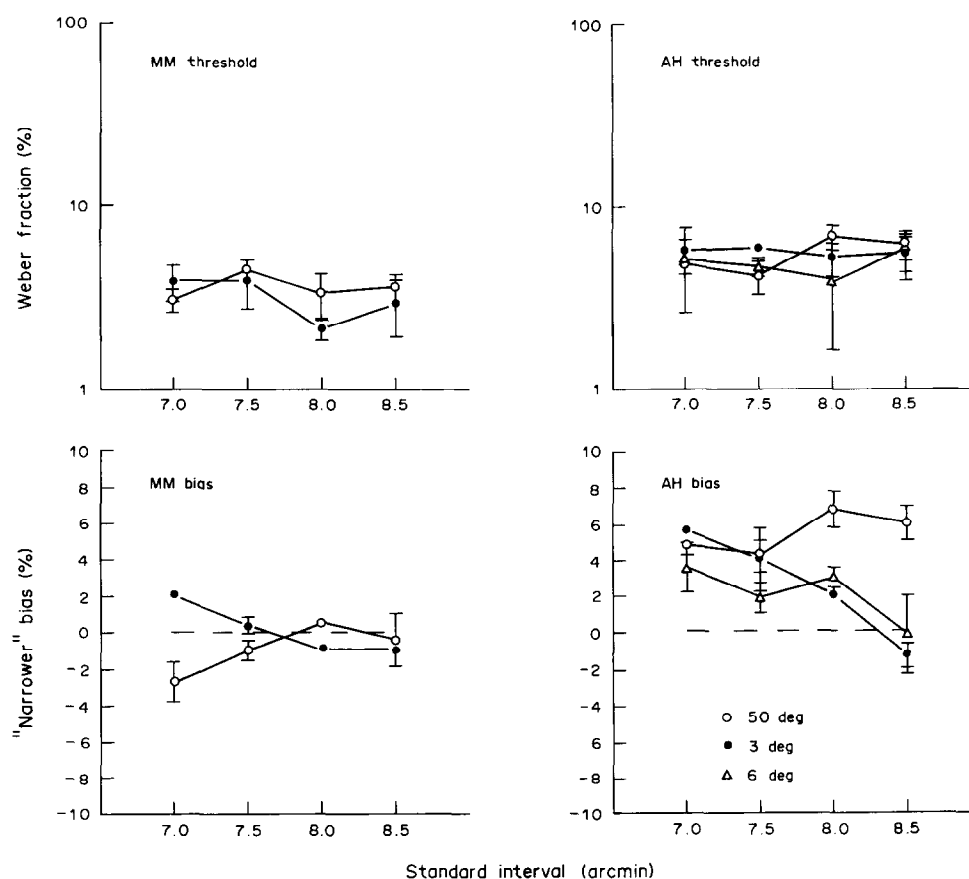


FIGURE 6. The figure compares size discrimination performance at 50 and 3 deg of orientational separation between the sets (MM and AH), with additional observations for a 6 deg separation for one observer (AH).

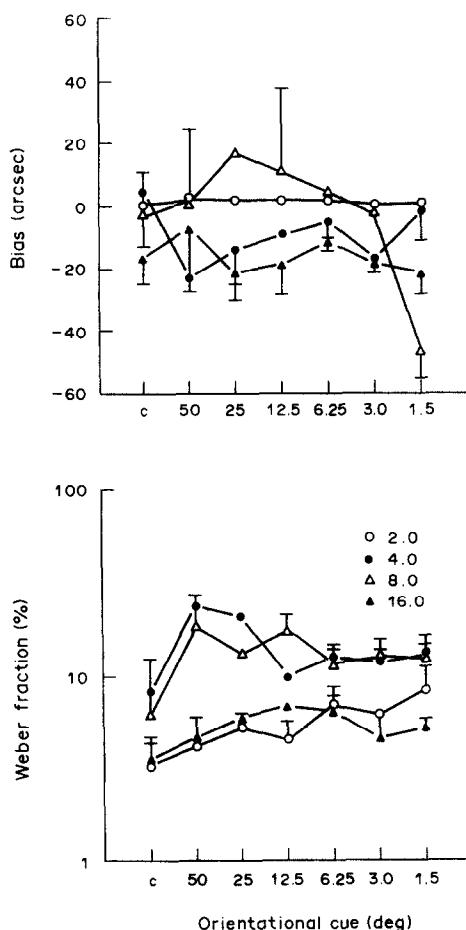


FIGURE 7. The figure shows the effects of differing orientational separation between the discrimination sets for one observer (MM). The four spatial interval standards in this experiment were 2, 4, 8 and 16 arcmin, and results are shown separately for these four standards by different symbols (see key). The point labelled "c" refers to a control procedure in which the different standards were presented separately, as in the "blocked" condition of Expt 1, rather than being randomly interleaved.

observer (MM) in Fig. 7. These data were collected with a stimulus set comprising standards of 2, 4, 8 and 16 arcmin. These data were collected early on in the study, before those for the smaller stimulus range (7–8.5 arcmin). The pattern of data is quite complex and not easy to understand, but it is clear that there is no general improvement in discrimination due to increasing the size of the orientational cue. The trend for the two larger stimuli is opposite to that for the two smaller, but in neither case is the extent of the trend very marked. The biases are likewise small (< 20 arcsec) and unaffected by the size of the orientational cue. Note that biases in this graph have been plotted in absolute units (arcsec) rather than in percentage terms, in order to emphasise the fact that they are numerically quite similar for the different sizes of standard.

Overall, the three sets of data reveal little systematic effect of the size of the orientational cue upon ability to carry out the concurrent interval discriminations. This is evidence for a two-stage process in discrimination: first the stimulus is categorically assigned to its orientation class, and then the appropriate standard for that orientation is selected. It is difficult to reconcile the absence

of an orientational effect with a one-stage model, in which discrimination is accomplished by the distribution of activity over filters jointly tuned in orientation and spatial frequency.

EXPERIMENT 3

The purpose of this experiment was to investigate the effects of disturbing the monotonic relation between angle and standard size used in the previous experiments. Hitherto, the greater the clockwise rotation of the stimulus, the larger has been the standard size associated with it. Observers may learn a discrete association between rotation and standard, or they may learn a continuous adjustment of the standard in relation to angle. If they have learned a continuous adjustment, performance will fail when the relation between size and angle is non-monotonic. To see whether this would occur, we compared performance in the normal "ordered" condition with performance in a "disordered" condition, in which the four angles, in increasing clockwise order, were associated with intervals of 8.0, 7.5, 8.5 and 7.0 arcmin respectively. The orientational difference between the sets was 1.5 deg, as in Expt 1.

Method

Observers and conditions. The two observers were HR and MM. The order in which they experienced the ordered and disordered conditions was specified in the Methods section for Expt 1. For both observers, the ordered condition was tested first. MM experienced 6 sessions of disordered training and HR had 5 sessions. In both cases, mean data were taken from the last two testing sessions.

The procedure in the disordered condition was identical to that in the ordered, except for the relation between angle and mean interval. In particular, feedback for incorrect responses was given, so that the observers had the opportunity to learn the new association between angle and mean spatial interval.

Results and discussion

The results shown in Fig. 8 may be summarised as follows. There was no evidence that thresholds increased in the disordered condition; indeed, in the case of one observer (HR) Weber fractions in the disordered condition were reduced to 5%. However, biases were clearly greater for both observers in the disordered condition, in the direction that would be expected if discrimination between the sets were partially lost. In HR the difference in bias between the extremes of the sets was almost 20%. This is still not so great a bias as we should expect if there were no discrimination between the sets, but it is quite considerable.

EXPERIMENT 4

The purpose of Expt 4 was to extend the number of concurrent sets from four to eight. The mean values of the eight sets varied from 6.0 to 9.5 arcmin in steps of

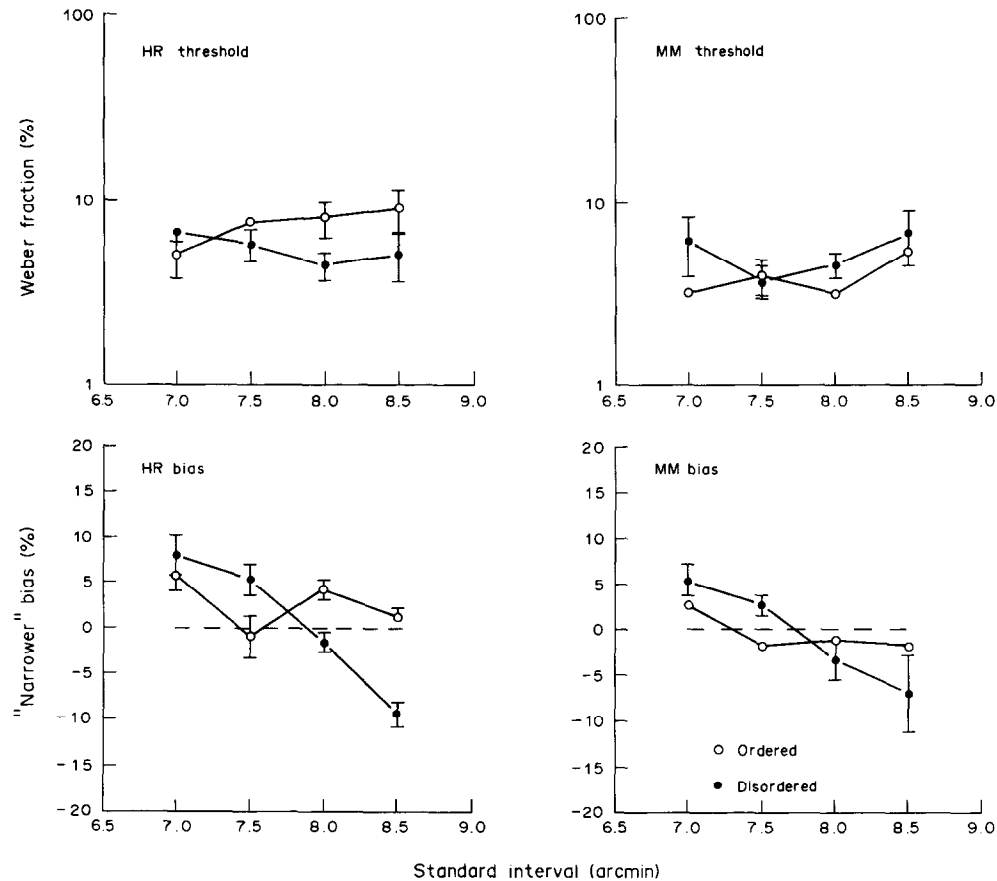


FIGURE 8. The figure shows the results of Expt 3, which investigated the effects of an ordered vs disordered relation between the orientational and size dimensions. In the ordered case, larger standards were associated with more clockwise rotation angles for the stimulus. In the disordered case, the association was non-monotonic. Note that biases are greater in the disordered condition: the observer is tending to classify the stimuli by their absolute size not by their size relative to the appropriate standard for that orientation.

0.5 arcmin. The orientational difference between the sets was 3 deg. A disordered condition was also investigated, in which the association between angle and mean size was non-monotonic.

Method

Observers and conditions. The two observers were DK and MM. The more experienced observer (MM) carried out these observations after those described in the earlier experiments. He had two sessions with the ordered condition and an orientational cue, one session with a positional cue, and one with both an orientational cue and a positional cue. The data described below are from these single sessions, or from the second session in the case of the orientational cue. DK received no training before three sessions with the ordered condition, data from the last of which is reported here, followed by two sessions with the disordered condition, the last of which is reported.

The standards spanned the range 6.0–9.5 arcmin in steps of 0.5 arcmin, and an associated orientational range of 79.5–100.5 deg in steps of 3 deg. In the disordered condition the size of the standards in order of orientation (increasing CW) was as follows: 8.0, 9.0, 7.0, 9.5, 8.5, 6.0, 6.5 arcmin. By chance, there was modest but nonsignificant negative correlation between size and angle, which accounted for 22.7% of the variance.

Except for the different range of sizes and angles, the procedure was exactly as in the previous experiments. Some additional experiments were carried out by MM using a positional rather than an orientational cue. In this case, the stimuli occurred at different positions around a notional circle on the screen according to their standard size. Considering the notional circle as a clock face, the standards were presented 45 deg apart, with the first stimulus (standard size = 6 arcmin) at 22.5 deg CW of vertical. All the stimuli had a vertical orientation, as opposed to their different positions. In a final condition, the orientational and positional cues were combined, so that each stimulus had an orientational cue increasing in 3 deg steps with its position (CW) around the circle. The orientational range was the same as in the pure orientational-cue condition described earlier.

Results

Psychometric functions from the ordered and disordered condition for one observer (MM) are shown in Fig. 9. It will be seen that good separation was maintained between the functions at different angles in the case of the ordered condition, but not in the disordered. Weber fractions and biases are shown in Fig. 10. It is clear that Weber fractions tended to be considerably higher in the disordered condition, and also there were consistent biases in the disordered condition. These

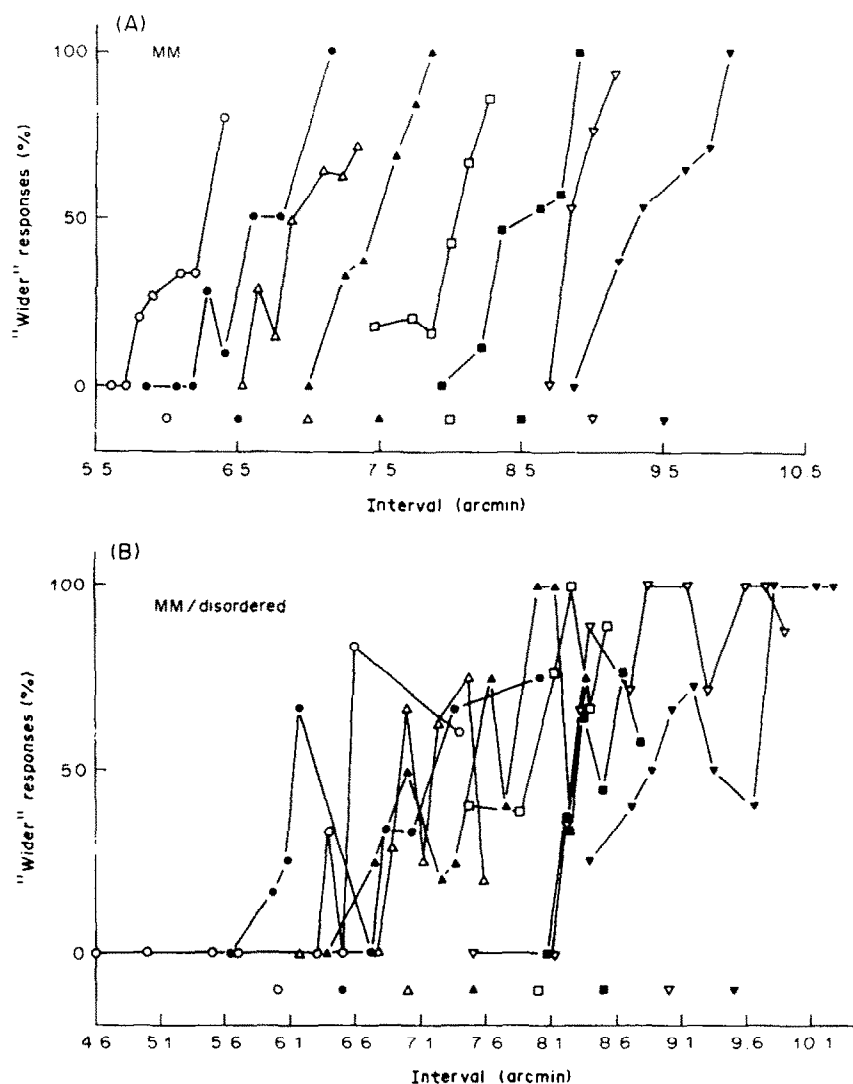


FIGURE 9. The figure shows representative psychometric functions for one observer (MM) in Expt 4, where there were 8 concurrent size standards. In the panel the relation between standard size and orientation was regular (larger sizes were associated with more CW rotation); in the bottom panel the relation was random. The size of the standards is indicated by the position of the symbols on the horizontal axis, and the psychometric function for each standard is plotted with the appropriate symbol to its standard. As in Fig. 4, data points with two or fewer observations have not been plotted.

biases took the form of an increased probability of classifying the absolutely smaller stimuli as "narrower" irrespective of their size relative to the mean of their appropriate set, as signalled by their orientation. This is what would be expected if discrimination between the different sets was breaking down. The drift of the biases is in the same direction in the ordered condition, but is much less marked. For both observers the bias is greater for the smallest stimulus in the set than for the others: the reason for this is not known.

It is clear from these data that the observers were able to use the orientational cue in the ordered (monotonic) condition to select the appropriate reference standard. The fact that this was easier in the ordered than in the disordered case argues that continuous correction was applied rather than a categorical judgement. This accords with the observer's subjective impressions in performing the task. The number of angles was too great to permit an easy absolute identification of the kind that could occur when there were only four

angles; instead, there was a general impression of whether the target "looked" large or small with respect to its degree of rotation. It is rather hard to state whether there was an actual change in perceived size as a function of angle, just as it is hard in size constancy experiments to distinguish between the perceived size of the stimulus and its angular subtense (Thouless, 1931).

Some further results for one observer (MM) comparing the effects of an orientational and a positional cue, as well as the combination of orientational and positional cues, are shown in Fig. 11. In general, there seems to be little difference between these conditions.

EXPERIMENT 5

The final series of observations reported here are concerned with the effects of increasing the difference between the members of the set of standard intervals. In most of the previous experiments that difference has

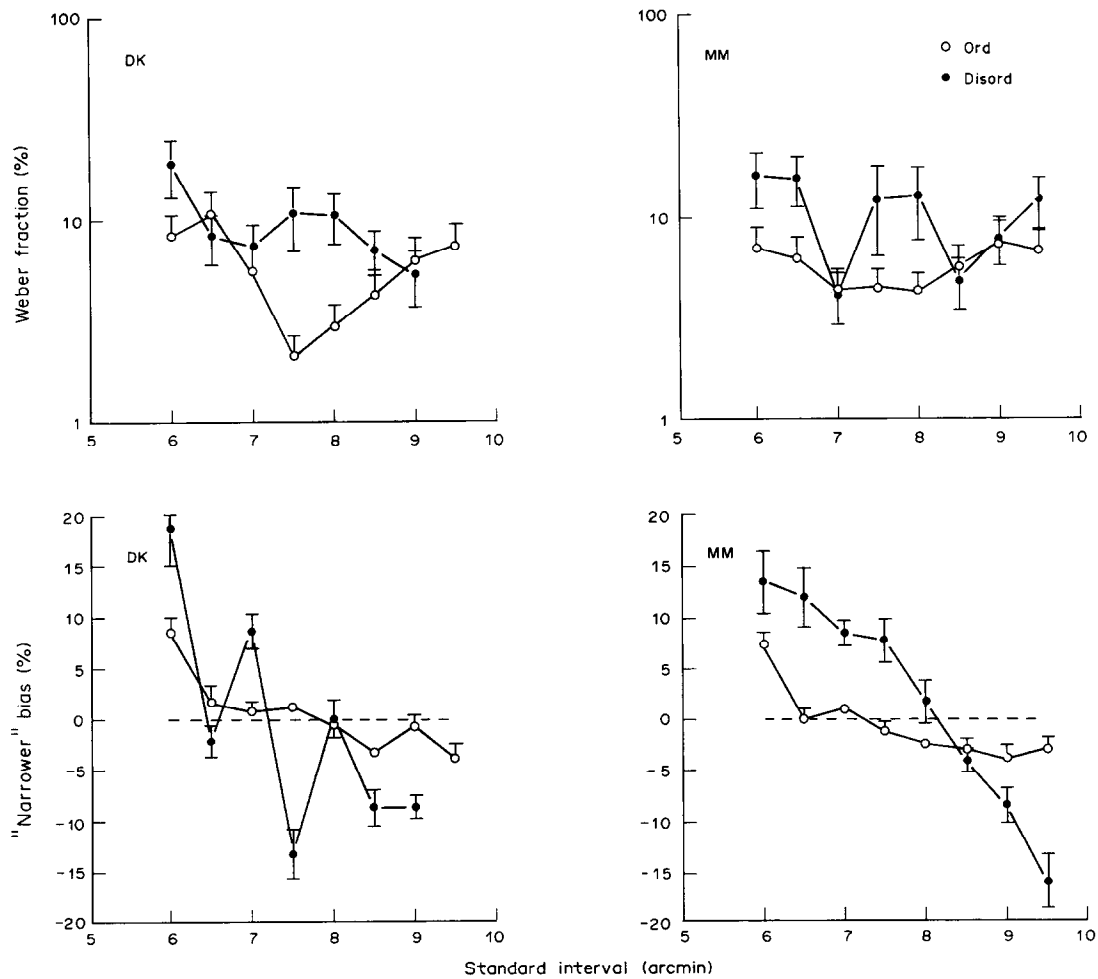


FIGURE 10. The figure shows Weber fractions and biases (50% points on the psychometric functions) in Expt 4, where there were 8 concurrent size standards. The relation between standard size and orientation was either ordered or disordered (see the caption to Fig. 9). Note that biases are larger in the disordered condition.

been 0.5 arcmin. It is possible that the ability to carry out multiple, concurrent discriminations depends upon the standards for the discriminations being close together in value. To evaluate this possibility, the ability to discriminate was tested with 5 different ranges of standard stimuli.

Method

The methods were the same as those in Expt 1-4 except for the different range of standards. The values of the 4 standards used for each of the 5 ranges is given in Table 1. The smallest range, 1.5 arcmin between the

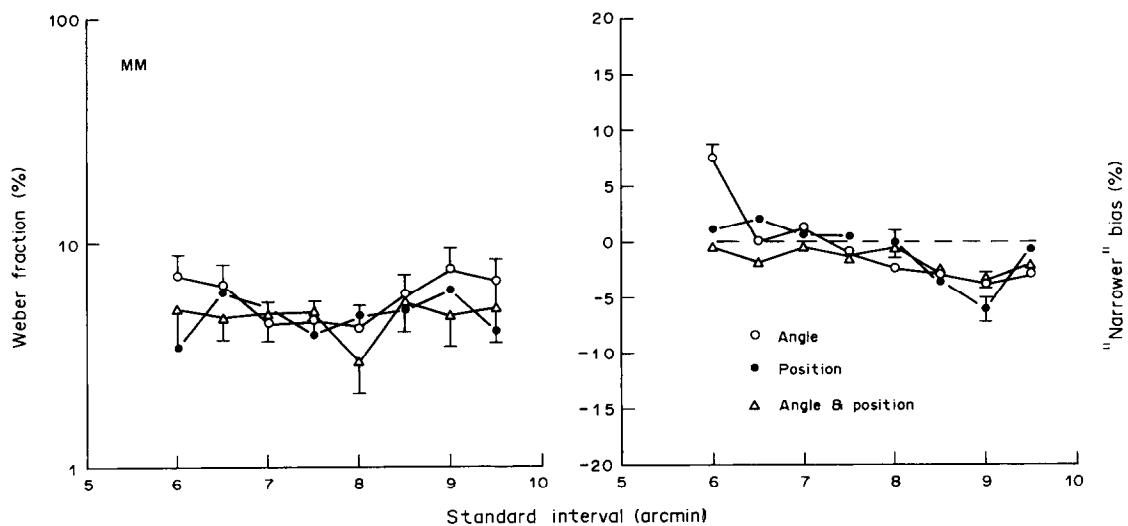


FIGURE 11. The figure shows results from Expt 4 for one observer (MM) comparing the effects of an orientational cue, a positional cue, and a combination of the two. There appears to be little difference between these conditions.

TABLE 1

| Range | Standard sizes (arcmin) | | | |
|-------|-------------------------|-----|----|-----|
| | S1 | S2 | S3 | S4 |
| 14 | 2 | 4 | 8 | 16 |
| 8.5 | 3.5 | 5 | 8 | 12 |
| 5.0 | 5 | 6.5 | 8 | 10 |
| 3.0 | 6 | 7 | 8 | 9 |
| 1.5 | 7 | 7.5 | 8 | 8.5 |

largest and smallest standards, was the one used in previous experiments; the largest was a range of 14 arcmin (2–16 arcmin). For each range of standards, three different conditions of orientation cueing were used: a 50 deg cue, a 1.5 deg cue, and a no-cue control. Each of the 60 thresholds in the experiment was determined at least twice, and the results given below refer to the second determination. Because of the very time-consuming nature of this study, only the author (MM) took part.

Results

The results are presented in Fig. 12. Considering first the threshold data, the most general finding is that thresholds decrease as the range of the standards is reduced. This was found in both sizes of orientational cue (50 and 1.5 deg). The same trend was found in the no-cue

control condition, although thresholds there were generally elevated as well. The relationship between threshold and the standard range is initially quite puzzling, since one would have expected a larger difference between the members of a set to make them more discriminable. The most probable explanation is that confusions between the sets will have a more deleterious effect upon Weber fractions when the spacing between sets is large than when it is small. Consider, for example, the case where the stimulus range is largest (14 arcmin), and where a stimulus of 3 arcmin is presented from the 2 arcmin standard set. If this stimulus is misclassified as "smaller" because it is attributed to the 4 arcmin set, the effect will be to stretch out considerably the psychometric function for the 2 min set. On the other hand, the misclassification of a stimulus midway between two sets of 7.0 and 7.5 arcmin sets will have a much smaller effect on the slope of the psychometric function.

GENERAL DISCUSSION

These experiments show that size judgements can be appropriately scaled by either an orientational cue or by a positional cue. The main evidence for this is that distinct psychometric functions can be obtained for the different size standards, depending upon the level of the

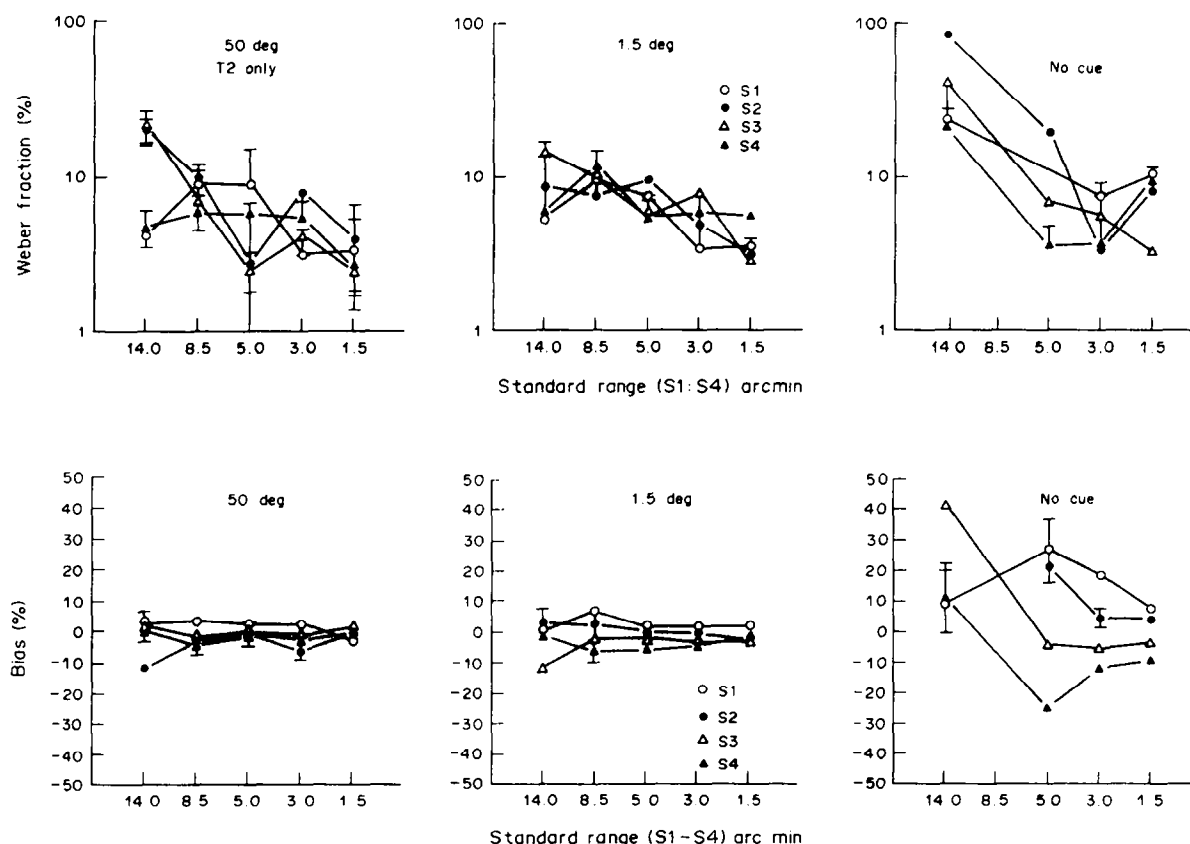


FIGURE 12. The figure shows results for a single observer (MM) in Expt 5, which investigated the effects of the size difference between the four standard intervals used in the discrimination. The magnitude of the size difference is shown on the horizontal axis in terms of its range: further details are given in Table 1. The cue indicating which stimulus was appropriate on a particular trial was either a 50 deg difference between sets (left hand panels); a 1.5 deg difference (middle panel); or there was no cue at all (right hand panel). The results for each of the four sets are shown separately by different symbols; S1 refers to the smallest standard in the set, S2 to the next smallest, and so on. The top set of panels shows thresholds (Weber fractions) from the experiment, and the bottom panels show the biases (50% point of the psychometric functions).

orientational or positional cue. When there are only four standard categories it appears to matter little whether the relation between orientation and size is monotonic or disordered: this suggests a categorical decision process rather than an analogue calibration. Further evidence for a categorical decision process is that the magnitude of the orientational cue does not appear to matter, so long as the categories are discriminable without error. However, with a larger number of standards, the task is easier when the size and orientational cues are monotonically related than when the relationship is random. Whether extensive learning could have established a non-monotonic relation we are unable to say.

The purpose of the experiment was to test the theory that size constancy is only one example of a general metrical scaling process that tends to recalibrate dependent stimulus dimensions. Other explanations of the findings will now be considered.

First, it is proper to question whether any real perceptual recalibration occurred as a result of the discrimination procedure. Observers certainly adopted appropriate standards to the different sets for deciding whether stimuli were "large" or "small", but this does not prove that physically different stimuli were seen as the same, or that same ones were seen as different. It would not be wise to venture too deeply into this philosophical territory; whether systematic decisions can be distinguished ultimately from perceptions is not easy to answer. For what it is worth, the subjective impressions of observers suggest that it was still possible to distinguish the differently-sized sets to some extent by their perceived size even after extensive training. This does not mean that no perceptual adjustment had occurred: training may very well have partially reduced the perceived difference between the sets. In any case, size constancy does not imply a complete failure to distinguish between nearby and farther objects of the same size: it is well known that a residual difference can be seen, even when the objects are judged to be the same size. The same is very likely to be true of arbitrary correlations, such as those studied here.

Another objection to the rescaling explanation is that the observers may simply have learned a set of independent discriminations, which had nothing to do with one another. It would not be surprising to find that an observer could perform concurrent discriminations of vernier acuity, curvature discrimination, orientation discrimination, and spatial interval discrimination. (In fact, in separate experiments we have found that such multiple, independent discriminations can be carried out quite easily and with little loss of acuity). If this is the explanation of multiple-interval performance, however, it is difficult to see why performance broke down when there were eight standards and the relationship between size and orientation was disordered rather than monotonic. The reply to this might be that the disordered relation made the appropriate standards harder to remember, but this would be to concede that the discriminations were not really independent. None of these arguments is conclusive. The best test would be to see if discrimination could be maintained with a truly con-

tinuous association between size and angle over a certain domain of the manifold. If it could, the possibility of discrete, independent discriminations would be ruled out.

The experiments and conclusions leave questions unanswered. It will be important to establish whether size is special case, or whether the rescaling process described here is general. The possibility that size is a special case can be tested by making other dimensions, such as orientation and curvature, the targets of the putative rescaling process. The connection between rescaling and adaptation requires investigation, by determining whether there are conventional aftereffects of multiple stimulus training with correlated dimensions. Finally, the evident conceptual link (Barlow, 1989) between scaling and the class of McCollough effects could be investigated by using motion and colour both as scaling and as target dimensions in multiple discriminations.

REFERENCES

- Andrews, D. P. (1964). Error-correcting perceptual mechanisms. *Quarterly Journal of Experimental Psychology*, 16, 104-115.
- Barlow, H. B. (1989). A theory about the functional role and synaptic mechanisms of visual after-effects. In Blakemore, C. B. (Ed.), *Vision: Coding and efficiency*. Cambridge: Cambridge University Press.
- Burbeck, C. A. (1987). On the locus of spatial frequency discrimination. *Journal of the Optical Society of America A*, 4, 1807-1813.
- Finney, D. J. (1971). *Probit analysis*. Cambridge: Cambridge University Press.
- Heeley, D. & Buchanan-Smith, H. M. (1990). Recognition of stimulus orientation. *Vision Research*, 30, 1429-1437.
- Levi, D. M. & Westheimer, G. (1987). Spatial-interval discrimination in the human fovea: What delimits the interval? *Journal of the Optical Society of America A*, 4, 1304-1313.
- McKee, S. P. & Welch, L. (1992). The precision of size constancy. *Vision Research*, 32, 1447-1460.
- Morgan, M. J. (1977). *Molyneux's question*. Cambridge: Cambridge University Press.
- Morgan, M. J. (1990). Hyperacuity. In Regan, D. (Ed.), *Spatial vision*. London: Macmillan.
- Morgan, M. J. & Watt, R. J. (1982). The modulation transfer function of a display oscilloscope. *Vision Research*, 22, 1083-1085.
- Morgan, M. J. & Watt, R. J. (1989). The Weber's Law for distance is not an artefact of eccentricity. *Vision Research*, 29, 1457-1462.
- Morgan, M. J., Hole, G. & Glennerster, A. (1990). Biases and sensitivities in the geometrical illusions. *Vision Research*, 30, 1793-1990.
- Morgan, M. J., Ward, R. M. & Hole, G. J. (1990). Evidence for positional encoding in hyperacuity. *Journal of the Optical Society of America*, 7, 297-304.
- Thouless, R. (1931). Phenomenal regression to the real object. *British Journal of Psychology*, 21, 339-359.
- Watt, R. J. & Andrews, D. P. (1981). APE: Adaptive probit estimation of psychometric functions. *Current Psychological Reviews*, 1, 205-214.
- Welch, L. & McKee, S. P. (1990). Position judgements are corrected for disparity at large scales. *Investigative Ophthalmology and Visual Science*, 31, 411.
- Westheimer, G. & McKee, S. P. (1977). Spatial configurations for visual hyperacuity. *Vision Research*, 17, 941-947.

Acknowledgements—I thank Suzanne McKee and Roger Watt for fruitful discussion, and Roger Watt for the use of his laboratory. The experiments were performed in The Centre for Cognitive and Computational Neuroscience at the University of Stirling while the author was supported by a grant from the Carnegie Trust. The work was also supported by the SERC Image Interpretation Initiative. The two anonymous referees provided many helpful suggestions.