

The Precision of Size Constancy

SUZANNE P. McKEE,* LESLIE WELCH†

Received 15th July 1991; in revised form 28 January 1992

The precision of objective size judgments, made when target disparity changed at random from trial-to-trial, was compared to the precision of angular size judgments made under the same condition. Subjects judged incremental changes in the vertical distance separating a pair of horizontal lines. For the objective judgments (in cm), the angle subtended by the target separation decreased with increasing depth consistent with the natural geometry of physical objects. For the angular judgments (in arc min), the angular separation did not change with disparity. For separations subtending an angle < 10 arc min, objective thresholds were considerably higher than angular thresholds, indicating that size constancy does not function well at small scales. At larger scales (> 20 arc min), the Weber fractions for angular and objective thresholds were nearly equal (~6%) for two of the three subjects. These same two subjects also learned to judge "objective size" when angular subtense systematically increased with increasing depth in an exact inversion of the natural relationship. Although their "anti-constancy" judgments were less precise (~9%) than their constancy judgments, the fact that subjects could learn this task with little practice suggests that constancy itself may be a learned response. Angular thresholds for targets presented only in the fixation plane were significantly lower than the angular thresholds measured with random changes in disparity, showing that observers with normal stereopsis do not have direct access to information about the angle subtended at the retina.

Objective size Size constancy Hyperacuity Spatial localization Stereopsis

INTRODUCTION

The apparent size of an object is relatively unaffected by substantial changes in viewing distance, despite the dependence of retinal image size on the distance between the object and the observer. This stability in perceived size is commonly known as "size constancy" and is thought to reflect an almost automatic re-scaling of size with perceived distance (Andrews, 1964; Gregory, 1966, 1987; Morgan, 1992). One consequence of this rescaling process is that the physical size of unfamiliar objects can be effortlessly judged over a large range of distances. While this ability may seem somewhat remarkable, humans are equipped with a number of sensory mechanisms for estimating relative distance. A machine that could measure angular subtense and estimate relative distance could certainly calculate objective size, relative to a reference size viewed at a known distance. In the case of the human observer, the reference size could be a part of the body, like the size of the hands viewed at arm's length (Morgan, 1989). Knowledge of objective size is useful for survival, and the brain has the information needed to calculate objective size, so "size constancy" is a fairly predictable aspect of human visual processing.

In traditional studies of size constancy, observers were shown an object at some faraway distance and were asked to adjust the size of an adjacent object until it

matched the distant object (Thouless, 1931; Holway & Boring, 1941). Sometimes, the aim of these studies was to determine what was actually seen by the observer—the objective size or the angular size? In other cases, the intent was to explore limitations on size constancy, e.g. over what distances observers could match objective size before perhaps regressing to match based on angular size. In one of the most interesting of these studies, Gilinsky (1955) found that observers were able to match either the retinal or the objective size of the test object, depending on the instructions given by the experimenter. Gilinsky's results, subsequently verified in other laboratories (Carlson, 1960, 1977; Leibowitz & Harvey, 1969), demonstrate that matching is a weak guide to the cognitive (or neural) operations underlying size constancy. One cannot determine whether the observer perceives retinal size, and then corrects this percept by some measure of depth to estimate objective size, or vice versa.

There is a psychophysical tool that could reveal the coding sequence. Instead of asking observers what they perceive, we can ask about the *precision* of their judgments of angular and objective size. What is the smallest detectable change in objective size? What is the threshold for discriminating differences in angular size? The precision of psychophysical thresholds is usually limited by noise in the neural pathways coding the stimulus dimensions—more noise means less precision. Let us assume that the calculation of objective size involves the simple combination of two independent neural measurements (angular size and depth). If the

*Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, CA 94115, U.S.A.

†University of California at San Diego, La Jolla, CA 92093, U.S.A.

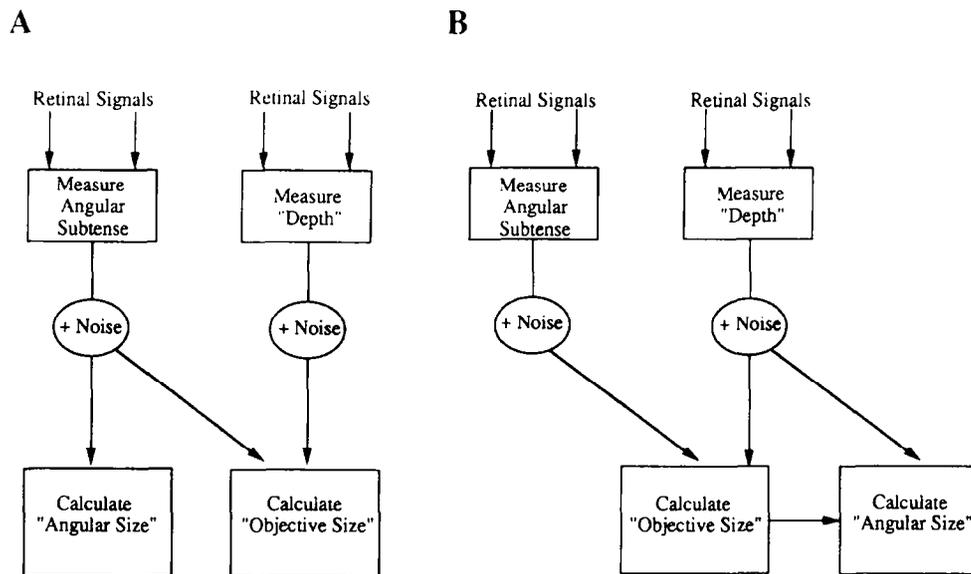


FIGURE 1. Flow chart diagramming sources in noise that limit the precision of objective and angular judgments. In (A) observer has direct access to information about angular subtense. Objective size is calculated from depth and angular size, and the additional noise from the depth estimate makes objective judgments less precise than angular judgments. In (B) observer does not have access to information about angular subtense and must calculate angular size indirectly from objective size and depth, thereby lowering the precision of angular judgments relative to objective judgments.

noise in the neural estimate of depth is roughly comparable, or substantially larger, than the noise in the estimate of angular size,* objective size judgments should be consistently less precise than angular size judgments, because the depth measurement would add noise to the calculation; as diagrammed in Fig. 1(A), there are two sources of noise in the objective size calculation and only one in the angular size calculation.

There are coding schemes in which objective size judgments would be *more* precise than angular judgments. For example, the brain may not have direct access to the signal for retinal or angular size (Gogel, 1969), but may instead estimate angular size indirectly by correcting the perceived objective size for an obligatory coupling between size and depth. As shown in Fig. 1(B), this correction should introduce additional noise, because the noise from the depth estimate enters the calculation twice—once in the calculation of objective size and again in the indirect calculation of angular size. Either of the models diagrammed in Fig. 1 would predict that observers can judge both angular and objective size, because all the information necessary is available for both calculations. The question is which do they do best.

In a recent study, Burbeck (1987a) used a measure of precision, spatial frequency discrimination, to analyze the coding sequence. She compared spatial frequency discrimination for a pair of grating targets presented at a single viewing distance to that for the identical pair of gratings (same object frequencies in c/cm) presented at two different distances; she found no significant

difference in performance for these two conditions, despite the difference in "retinal spatial frequency" (c/deg) necessarily introduced by the second viewing distance. In a second experiment, she showed that observers initially had great difficulty discriminating between targets presented at two different viewing distances on the basis of their "retinal frequencies" (c/deg), but were able to learn this "retinal frequency" discrimination with practice. Burbeck concluded that we do *not* have direct access to information about spatial frequency coded in angular units, and must instead estimate angular spatial frequency (and presumably also angular size) indirectly as diagrammed in Fig. 1(B). Not all angular dimensions are calculated indirectly. McKee and Welch (1989) found that the discrimination of angular velocities was decidedly superior to the discrimination of objective velocities.

Precision is related to measures of variability, so the standard deviations of the matches made in the early constancy studies supply information about the amount of noise involved in these judgments. Some studies do confirm Burbeck's conclusion. Leibowitz and Harvey (1969) found that the variability of the retinal size matches was greater than the objective size matches. On the other hand, Gilinsky's data indicate, that under optimum circumstances (large angular sizes viewed at distances < 200 ft), the standard deviations of the objective and angular matches were comparable. There are also persistent indications that size constancy fails at small angular sizes (Boring, 1943; Ross, Jenkins & Johnstone, 1980). Although Gibson (1950) believed that size constancy worked at all sizes, the variability of small-angle matches in his study was much greater than the variability of matches for objects subtending larger angles. This same variability at small angular sizes is apparent in Gilinsky's data, suggesting that the precision

*Comparable measurements of size and disparity discrimination indicate that one important depth signal, disparity coding, is a noisier physiological dimension than size coding (McKee, Levi & Bowne, 1990a).

of objective size judgments may be significantly degraded at small scales.

These early studies do not provide a clear picture of coding sequence, but that was not their intention. In these studies, the means and standard deviations were based on pooled data, obscuring systematic differences in the data of individual observers about the relative precision of their angular and objective judgments. The measurements were often made at large distances where binocular disparity information was approaching the stereoscopic limit, and the observers were given instructions, but not feedback about the correctness of their matches. With feedback, they might have been able to make highly precise judgments that exactly matched either the angular or objective size of the test target.

In the present study, we explored the generality of Burbeck's conclusions for a range of sizes and depths. We gave our subjects the difficult task of abstracting the objective size of a target despite random variations in its apparent depth, but we did provide feedback. In a companion study, we measured the ability to judge angular size despite random variations in target depth. We used a very simple, well-studied task to measure the precision of size judgments in which subjects were required to judge a single dimension, the vertical distance separating a pair of horizontal lines. We manipulated perceived depth by changing only the binocular disparity of the target, but we chose disparities that were easily discriminable and corresponded to physical distances ranging roughly from 1–2 m. Under these severely restricted circumstances, we compared the precision of angular and objective size judgments.

METHODS

Three different experimental conditions were used to assess the relative precision of angular and objective size judgments: (1) incremental judgments of *objective size* with randomly-chosen disparities; (2) incremental judgments of *angular size* with randomly-chosen disparities; (3) incremental judgments of size for targets presented only in the *fixation plane* (at one depth).

The target for all conditions is diagrammed in Fig. 2. It consisted of two short horizontal lines that defined a vertical separation (or size), a configuration chosen to minimize the effects of diplopia on the thresholds. The vertical separation was varied parametrically from 0.13 to 2.62 cm for the objective size judgments, and correspondingly, from 3 to 60 arc min for the angular size judgments. For a given experimental test run, we chose one particular separation (S), and measured the minimum detectable change in separation (ΔS) using the method of single stimuli. In this method, the subject is shown 1 of 7 possible stimuli chosen from a narrow range, e.g. 2.41, 2.48, 2.55, 2.62, 2.69, 2.76, or 2.83 cm, and is required to judge whether the presented sample is smaller or larger than the mean of the range, equal to 2.62 cm in this example. Thus, the subject judges

Test Target

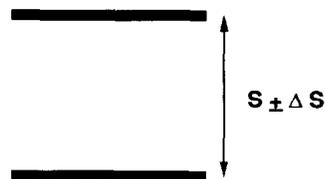


FIGURE 2. Diagram of target used for size thresholds. Subjects judged incremental changes in the vertical distance separating the horizontal lines relative to an implicit reference (the mean of the set of test stimuli). For the objective size judgments in cm, target disparity was varied at random from trial-to-trial, and the angular subtense was scaled with target disparity to be consistent with the apparent physical distance of the target. For the angular size judgments in min arc, target disparity was varied at random from trial-to-trial, but the mean angular subtense did not vary with target disparity.

the test sample against an implicit or remembered standard. While this type of judgment might seem difficult to perform, subjects have no trouble learning the task, and are able to establish an implicit standard with as few as 10 practice trials [see Westheimer & McKee (1977) and McKee, Welch, Taylor and Bowne (1990b) for previous examples of its use]. This method had an additional virtue for the present experiments; on our small CRT screens (subtense 3.4×4.2 deg), a visible standard equal to the mean separation (S) and presented adjacent to the test separation ($S \pm \Delta S$), might have introduced confounding cues that would not have had uniform effects in all experimental conditions.

In the objective size condition, the subject's task was to judge small incremental changes in the *objective* size (measured in cm), despite random variations in disparity. On each trial, the target was presented in 1 of 9 depth planes chosen at random. The nine planes spanned the range from +40 min of arc crossed disparity to -40 min uncrossed disparity in equally-spaced intervals 10 arc min apart; this disparity range corresponded to physical distances ranging from 1.16 to 2.12 m (see diagram on left side of Fig. 3). There is only a small calculated difference between angular and objective size at small disparities (± 20 arc min), so, to enhance the difference between these two conditions, we increased the probability that the target was presented at a large disparity; the target was twice as likely to appear in the 4 extreme planes (± 30 arc min and ± 40 arc min) as in the fixation plane and the 4 nearer planes. The targets were actually presented on the screens of an electronic stereoscope at a fixed distance of 1.5 m, and target disparity was manipulated to produce changes in perceived depth (right side of Fig. 3). The mean objective size (in cm) of the vertical separation between the target lines (S) was constant for any experimental test run. When the three-dimensional distance to a real object is increased, the retinal angle subtended by the object necessarily decreases. To simulate the natural relationship between angular subtense and objective size, we

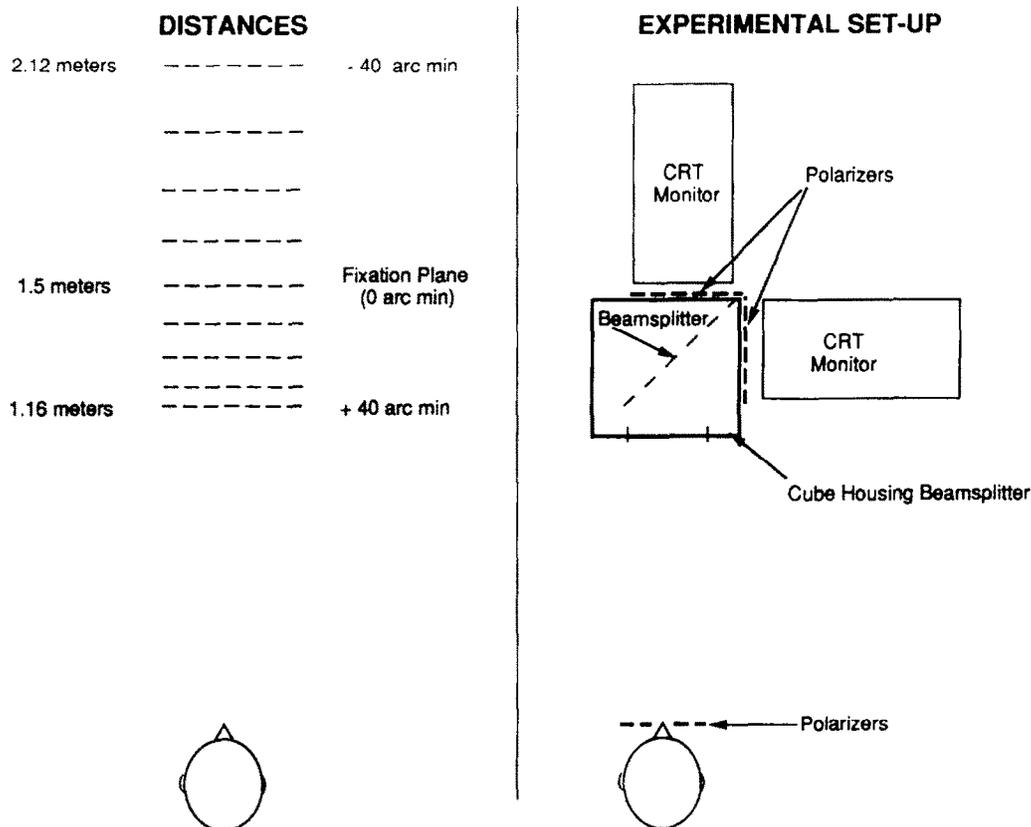


FIGURE 3. Diagram of experimental set-up. Left side of figure shows the apparent distances associated with the disparities used to measure angular and objective size thresholds. Right side shows the actual electronic stereoscope used to present stereoscopic display. See Methods for details.

varied the angular size presented on the screens of the stereoscope systematically for each depth plane as though the presented objective size ($S \pm \Delta S$) were being viewed at the physical distance specified by the target disparity, i.e. 2.62 cm was set equal to 1.29 deg at a distance of 1.16 m (equal to 40 min of crossed disparity with respect to the fixation plane), and to 0.71 deg at a distance of 2.12 m (40 min of uncrossed disparity). Note that the only cue to depth in our experiments was binocular disparity.

The angular size condition was designed to be exactly parallel to the objective size condition. The subject's task was to judge small incremental changes in *angular* size, despite random changes in disparity. On each trial, the target was again presented in 1 of the 9 depth planes chosen at random, but the mean vertical target separation in angular units did *not* vary with target depth. In the third experimental condition, the subject judged small incremental changes in the vertical distance (or size) separating the target lines for targets presented only in the fixation plane. In all three conditions, the size of the incremental steps were increased proportionally with the size of the mean separation because we anticipated a proportional increase in threshold consistent with Weber's law.

In a variant of the objective size condition, the "*anti-constancy*" condition, subjects made incremental judgments of size when changes in the angular subtense of the target were completely inverted from

the natural arrangement described above for the objective size judgments—the angular size of the target increased systematically as the apparent physical distance to the target increased. For example, an objective size of 2.62 cm was set equal to 0.71 deg at a distance of 1.16 m, and to 1.29 deg at a distance of 2.12 m, exactly the opposite of the compensation for apparent physical distance used for objective size condition. Subjects were required to judge incremental changes in "objective" size, while compensating for the inverted changes in angular subtense associated with the depth plane of the target. In short, they were forced to learn a new association between angular subtense and target disparity in order to minimize the error feedback. When the target appeared far away, they judged its size against a large implicit standard, and when it appeared close, they judged its size against a small implicit standard.

The targets were composed of thin, bright lines, drawn by computer-generated signals on the screens of two Hewlett-Packard 1332A monitors, each equipped with a P-4 phosphor. To insure best performance for larger sizes, e.g. 60 arc min, that necessarily stimulated parafoveal loci, we increased line length with target eccentricity. The length of the target lines was increased parametrically with parametric increases in the vertical distance separating the target lines (S), according to the hyperacuity scaling function originally described by Levi, Klein and Aitsebaomo (1985): $\text{length} = 9(1 + E/0.8)$ where E is the eccentricity given in deg,

and 9 arc min is the length of the foveal targets. Thus, the horizontal length of the target lines ranged from 9 to 14.6 arc min. However, the angular length of the horizontal target lines was *not* changed with changes in disparity. In a preliminary study, we found that systematic alterations in line length consistent with changes in an objective length did not produce a significant improvement in the objective size increment thresholds possibly because target length was obscured by diplopia at the extreme disparities.

The images on the two monitors were superimposed by a beam-splitting pellicle. Orthogonally-oriented polarizers placed in front of the monitors and the subject's eyes insured that only one screen was visible to each eye. Before the presentation of each test stimulus, a fixation pattern was presented for 800 msec, and then turned off 60 msec before the presentation of the test target. The fixation pattern, presented binocularly, consisted of four corner brackets forming an implicit square 60 arc min on a side with a bright point in the center of the square. The lines forming the test configuration were presented symmetrically arrayed around the center of the fixation pattern. As the fixation pattern might have provided an additional alignment cue for the largest angular test size (60 arc min), it was increased to 90 arc min for that size alone. Generally, the invisible fixation pattern would *not* have provided a better reference than the visible test lines. Hyperacuity studies which have compared positional judgments for simultaneous presentation of test lines to successive presentation have found a marked decrement in thresholds for successive presentation, particularly for test targets separated by small distances (Foley, 1976; McKee *et al.*, 1990b; White, Levi & Aitsebaomo, 1992). McKee and Levi (1987) tested the efficacy of a recently-extinguished fixation pattern as a reference system for positional judgments; they found that the threshold with only the invisible fixation pattern as a reference was significantly higher than the thresholds in the presence of a visible target, except when the visible targets were separated by 100 arc min or more.

The target duration for most of these experiments was 150 msec, too brief to permit a voluntary shift in convergence. Since the targets were presented at disparities symmetrically arrayed around the fixation plane, e.g. ± 40 arc min, we may assume that the subjects kept their eyes close to the fixation plane for most of the experimental session, and that additional proprioceptive cues from convergence and accommodation to depth planes other than the fixation plane were not available at these brief durations. For some of the experiments, duration was increased to either 1000 or 1500 msec, a time sufficient to allow convergence to the depth plane occupied by the target. If convergence were accompanied by accommodation, as is probable, the targets off the fixation plane would have been slightly out of focus in this long duration condition (maximally $\sim \frac{1}{3}$ D).

Line luminance was measured with a Pritchard photometer. The test pattern for these measurements was a long vertical line created with the same timing and

intensity characteristics as the experimental target lines. The measurements for the line on each screen were made in the dark, through the pellicle, at two different distances (147 and 258 cm) and with two different apertures (1 deg and 20 arc min) for accuracy. The photometer measurements (which are measured in cd/m^2) were converted to *line* luminances using the formula: $L = (\text{meter reading})\pi d/4$ where d is the diameter of the measuring aperture. The line on each screen measured approx. 0.22 cd/m . It should be noted that these small *line* intensities produce bright, easily visible lines; if we filled the screen with such lines, one line every 1 mm, the mean luminance of the screen would equal 22 cd/m^2 . The background luminance measured 0.52 cd/m^2 . Overhead fluorescent lighting located about 2 m from the CRT monitors supplied indirect illumination of the experimental setting at a moderate photopic level. Room furniture and experimental equipment were clearly visible. The only immediate reference frame for the targets was the 14×11.5 cm opening in the cube that contained the beamsplitting pellicle; the opening was 45 cm in front of the fixation plane. The edges of the CRT screens and the pellicle were very dim and partially obscured by the cube housing the pellicle. Moreover, the two CRT screens were not superimposed optically, so they appeared to float at an indistinct distance. The only target providing good information for convergence in the fixation plane was the bright square fixation pattern. In a control experiment for our earlier study of constancy (McKee & Welch, 1989), we found that making these measurements in total darkness did not change the pattern of results from those obtained when room furniture was visible.

Each of the increment thresholds presented in this paper is based on at least 300 trials, usually from two separate test runs of 150 trials each. Additional experimental sets were taken for some sizes when there was a substantial difference between the thresholds for the first and second sets; all sets were summed to estimate the threshold. The thresholds were estimated from the psychometric functions generated by plotting the percentage of trials on which the observer responded that the presented separation ($S \pm \Delta S$) was larger than the mean separation (S) as a function of the distance separating the target lines. A cumulative normal function was fitted to the resulting function by Probit analysis. Threshold was defined as that incremental change in size that produced a change in response rate from the 50% to the 75% level, equal to a d' of 0.675. Each experimental test run began with 10–20 practice trials. All measurements were made using audible error feedback.

For those cases where we wished to measure the perceived size as a function of target disparity, data from 900–1000 trials were accumulated at a single criterion size, e.g. $1.2 \text{ cm} \pm \Delta$ for the objective size judgments or $30 \text{ arc min} \pm \Delta$ for the angular size judgments, and sorted by disparity into separate bins. A point of subjective equality (PSE) (stimulus corresponding to the 50% response value on the psychometric function) was estimated from the psychometric function generated by the

data in each of the separate bins. The standard errors of the PSEs hovered between 1 and 2%, except for one case, the "anti-constancy" PSE corresponding to a disparity of 10 arc min, where erratic performance and a small number of trials combined to make the PSE indeterminate. An increment threshold was also estimated from the data from the 900-1000 trials pooled over all disparities.

The two authors and a third experienced observer served as subjects for these experiments; all three had 20/20 visual acuity for the viewing distance of 1.5 m and good stereoacuity, and all had had much practice on size or separation judgments in previous studies. Our original intent was to assess the natural capabilities of well-trained adult observers in judging angular and objective sizes. These experiments were *not* designed to measure the ability to learn novel stimulus criteria. Subjects LW and SM spent a brief period (roughly 300 trials) practising with both types of judgments before the data presented here were collected; neither author noted great difficulty in performing either task, but both had participated in an earlier study in which judgments of angular and objective size for a single size had been a control condition for velocity constancy (McKee & Welch, 1989). The main function of practice was to reduce confusion between the "objective size" conditions, and the "angular size" conditions. Subjects tended to take "objective" thresholds, and "angular" thresholds in blocks, although sometimes they interspersed conditions as needed to complete at least 300 trials for all tested sizes in all experimental conditions. When switching from one condition to another, a subject would often take some practice (40-150 trials) at one size to be reminded of the appropriate criterion for that block of thresholds; data from these designated practice sets were discarded. Questions arising from our initial results lead us to design the anti-constancy experiment. Clearly, "anti-constancy" is not a natural part of human experience, so for these measurements, subjects LW and SM practised for 600 trials before collecting the data shown in Table 2 and Fig. 9. The average Weber fractions from the practice sets were 0.125 for LW and 0.10 for SM—slightly higher than the Weber fractions for "anti-constancy" shown in Table 2 which were based on accumulated data from several runs (900 trials for LW and 1050 trials for SM), taken after completing the 600 practice trials. These thresholds do not necessarily represent asymptotic performance; with continued practice, performance may have improved.

The third subject (WAM) did not seem to have any natural ability to estimate objective size from disparity at a short duration (150 msec) despite abundant practice. He felt that he might be able to perform this task better at a longer duration, so we repeated these size measurements at a duration of 1000 msec. His data for both durations, taken after considerable practice in the indicated conditions, are presented in all appropriate figures. For those interested in the genetic contribution to these perceptual abilities, subject WAM shares half his genetic

endowment with subject SM; his father declined to participate in these experiments.

RESULTS

The first experiment compared objective size judgments, made when target disparity changed randomly from trial-to-trial, to angular size judgments made under the same condition. The resulting increment thresholds, each based on the pooled data from all 9 depth planes (300 trials), are shown in Fig. 4 for a range of sizes. Thresholds for targets presented only in the fixation plane are also plotted in the same figure. For ease of comparison, all thresholds are presented in a common framework based on angular units. The separation plotted on the horizontal axis refers to the mean angular size. By design, the angle subtended by the mean objective size at 1.5 m (the fixation plane) equals this mean angular size. The minimum detectable increment in objective units (cm) was first translated into the dimensionless Weber fraction, and then plotted as an incremental change in the mean angular size corresponding to the mean objective size.

Two things are apparent from Fig. 4. First, the objective size thresholds are decidedly less precise than the angular size thresholds at small separations (≤ 10 arc min), but are nearly equal to the angular thresholds at separations > 10 arc min. Second, the random disparity angular thresholds are much less precise than thresholds measured in a single depth plane, indicating that, when observers view a target binocularly, they do not have access to a pure angular or retinal estimate of size uncontaminated by depth signals.

Size constancy at small angular subtense

Although the objective thresholds were converted into angular units, the responses in the objective size condition were, of course, scored according to whether the observer correctly judged the physical size in cm, *not* in arc. If the observer were unable to compensate for the changes in apparent depth associated with target disparity and could only judge the target according to its angular subtense, he or she would make many errors, because a small size (cm) in the nearest plane would subtend a larger angle than a large size (cm) in the farthest plane. Thus, the increasing discrepancy between the objective and angular thresholds at small scales may be due to the failure of size constancy for features subtending small angles (Boring, 1943; Ross *et al.*, 1980).

There is another explanation. Increment judgments at very small angular sizes are known to be relatively less precise than judgments at larger angular sizes. The Weber fraction for judging changes in a 1 min spatial interval is 0.2-0.3, while the Weber fraction for a 10 min spatial interval is about 0.03 (Westheimer & McKee, 1977; Klein & Levi, 1987). Our measurements of objective size thresholds are necessarily based on

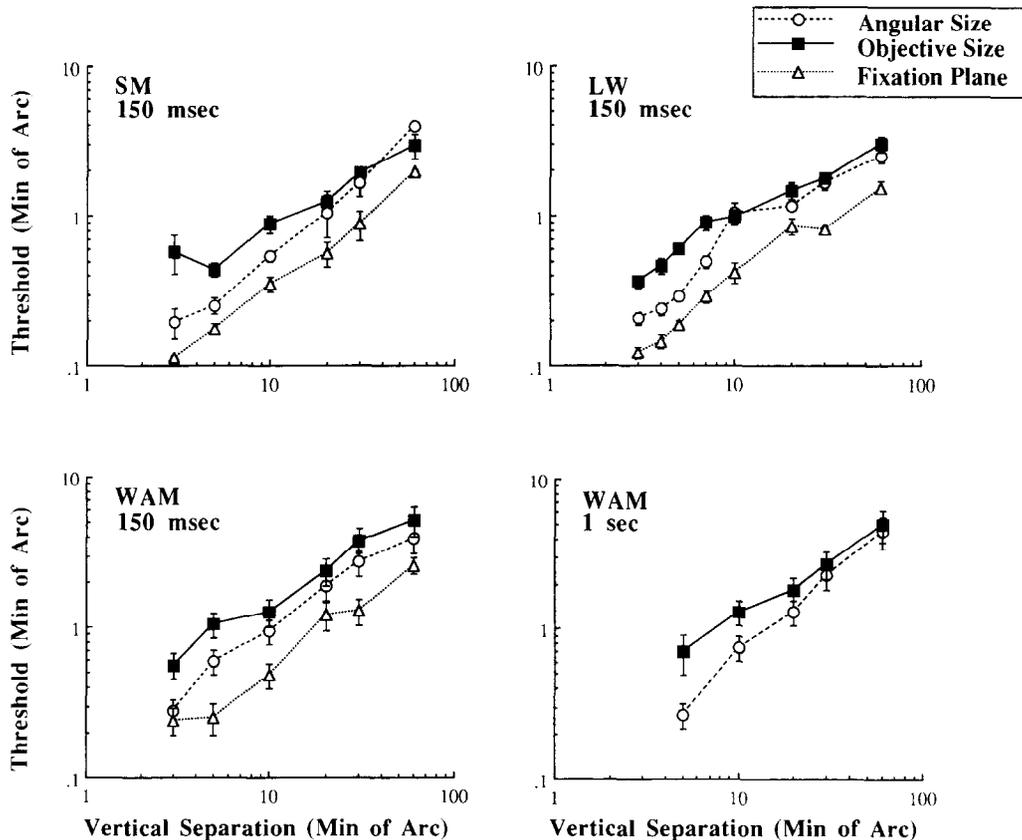


FIGURE 4. Increment thresholds for vertical separation for three experimental conditions: Objective size judgments with random trial-to-trial changes in disparity; angular size judgments with random trial-to-trial changes in disparity; size judgments for target presented only in fixation plane.

data from a range of angular sizes, consistent with the changes in disparity, so the objective thresholds for small scales might be elevated because they are based on a mixture of precise and imprecise angular signals. However, as long as we operate within the range of angular sizes where Weber's Law holds, our objective thresholds should not be affected by this pooling of large and small angular signals. For subjects LW and SM, Weber's Law holds for separations as small as 3 arc min (see their fixation plane data in Fig. 4), and the deviation from Weber's Law apparent in the fixation plane data of subject WAM is too small to account for the near doubling of his objective thresholds at the small scales. To check whether this mixture explanation accounts for the imprecision of the small-scale objective thresholds, we increased the number of trials for 0.2 cm, (mean angular subtense = 5 arc min) thereby obtaining adequate estimates of the increment thresholds associated with each of the 9 depth planes. The Weber fractions ($\Delta S/S$) for the separate depth planes showed

no significant trend with increasing depth (decreasing angular size).

In Fig. 5, the perceived mean sizes (PSEs), estimated from a large number of trials, are plotted for two separations, 5 arc min (or 0.02 cm) and 30 arc min (or 1.2 cm), in a common framework based on angular units. If the subjects could scale their judgments appropriately for depth, the objective means in *angular* units should increase systematically with decreasing depth, falling on the diagonal line drawn in each graph. The angular means should, of course, remain constant and fall along the horizontal line also drawn in each graph. All of the angular means follow the predicted horizontal line, although there is a hint of interference from "size constancy" in the 30 min data of subject SM; the mean angular size looks larger to her at the farthest depth plane than at the nearest plane. Both subjects do a fair job of responding to objective size at the larger scale; the means fall close to the oblique line in agreement with prediction. Neither subject could judge objective size appropriately for the target subtending the small angle. The means for the objective and angular size judgments overlap at all except the most extreme disparities, where the objective means separate from the angular means, perhaps reflecting a half-hearted attempt by the subjects to respond in a manner consistent with the feedback reinforcing objective size judgments.* Thus size constancy scaling does not function efficiently for targets subtending a small angular subtense.

*An ANOVA showed that the PSE's for the small objective size were significantly different ($P = 0.02$) from the PSEs for the small angular size provided that the data from the extreme disparities were included in the analysis. When the data from ± 40 min disparities were removed from the analysis, the difference between the angular and objective PSEs for the 5 min size was no longer significant ($P = 0.3$). At the larger 30 min size, the PSE's for objective size were significantly different from the PSE's for angular size with or without the extreme disparities ($P = 0.0001$).

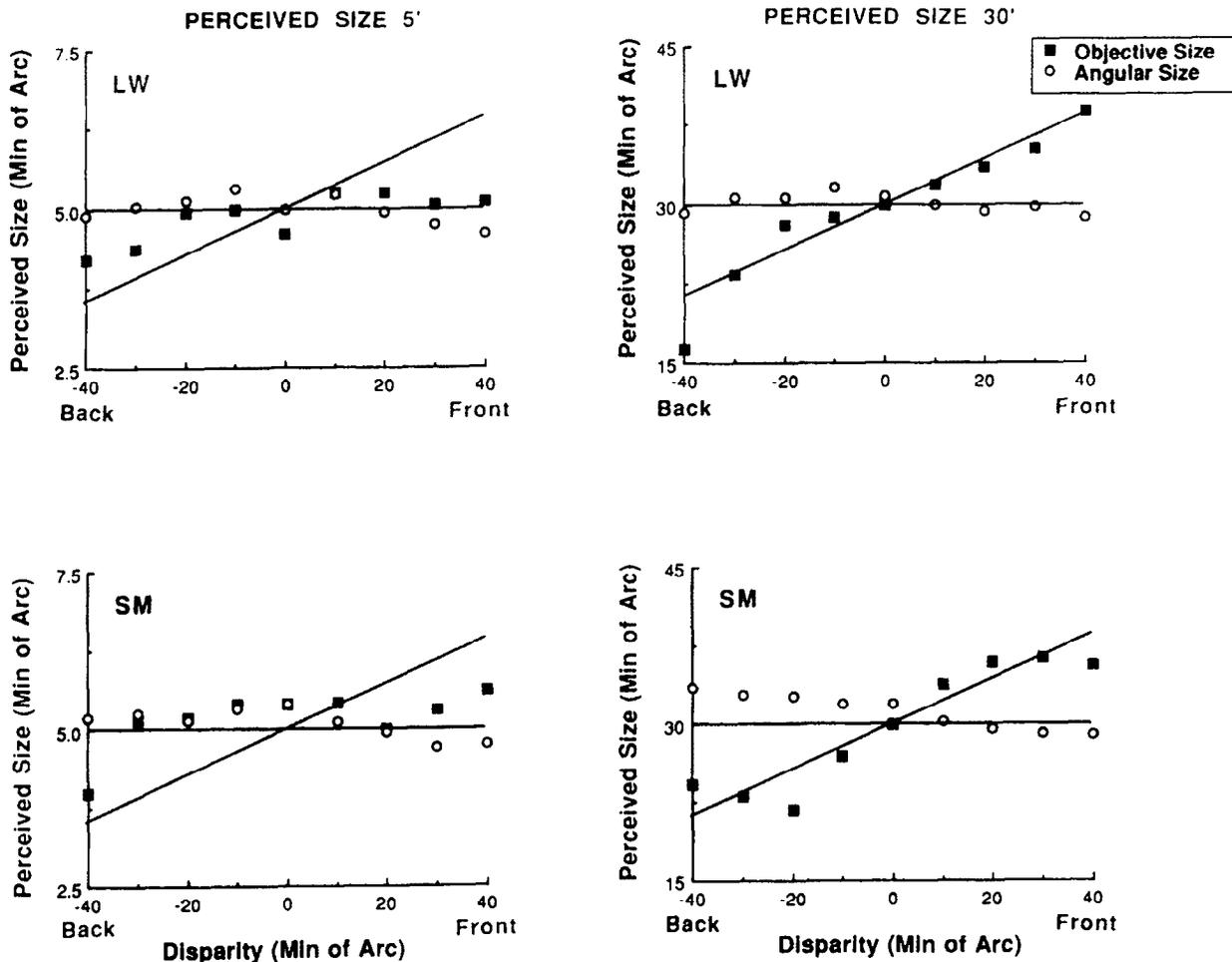


FIGURE 5. The PSEs (stimulus value corresponding to 50% point on psychometric function) plotted as a function of target disparity for both angular and objective size judgments at two different sizes (5 and 30 arc min, or 0.22 and 1.3 cm). The oblique line in each graph shows the predicted change in angular subtense for objective size, the horizontal lines being the prediction for the angular size. Subjects are unable to compensate for depth at small scales. Target duration 150 msec.

If our observers were attempting to use a different angular reference size for every depth plane, as scaling models of size constancy imply (Andrews, 1964; Gregory, 1966, 1987; Morgan, 1992), then the major source of noise in their objective thresholds is reference uncertainty. White *et al.* (1992) have argued that reference uncertainty has a multiplicative effect on increment thresholds for separation, even for targets presented in a single plane (see also Klein & Levi, 1987; Morgan, 1991). It is easy to see why this noise is multiplicative. Subjects make these judgments using an implicit reference probably estimated by taking a running average of the preceding three or four trials. Even if they are perfect at this averaging process, the average would often be either too small or too large with the result that the stimulus on the current trial could be "perfectly" encoded, but still judged incorrectly. Typically, the size of the incremental steps are increased proportionally with the size of the mean separation, since the threshold also grows proportionally, i.e. Weber's law holds over a large range of separations. Thus, the errors introduced by reference uncertainty are necessarily magnified by the step size, increasing the increment thresholds by a constant factor at all scales.

Reference uncertainty, *per se*, would affect the angular and objective thresholds equally, but the objective thresholds could suffer from increased uncertainty due to constancy scaling. Observers might misjudge the depth plane of the target and apply the wrong reference, or incorrectly estimate the implicit reference by averaging trials from adjacent depth planes. Since, in the objective size condition, the disparity-dependent variations in angular subtense are also proportional to the mean size, errors in constancy scaling would amplify the multiplicative effect of reference error. The problem with this argument is that there is no particular reason why reference uncertainty should have more effect at small scales than large. In the upper half of Fig. 6, we have plotted the ratio of objective to angular thresholds as a function of target separation. If constancy scaling has a multiplicative effect, due to increased uncertainty about the reference size, this ratio should remain constant across mean sizes. Instead, the ratio declines steadily, reaching a value close to 1 at mean separations between 10 and 20 arc min.

The noise in the objective judgments appears to be additive, rather than the multiplicative. The graph of the differences between the objective and angular thresholds, shown in the lower half of Fig. 6, suggests

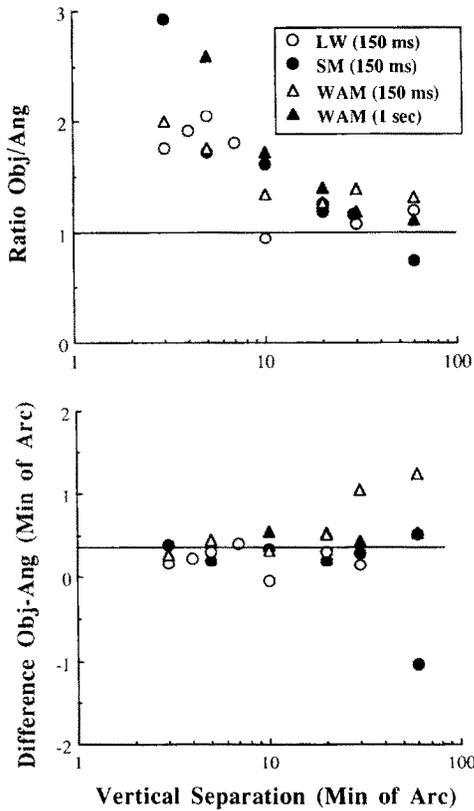


FIGURE 6. Upper graph shows ratio of objective to angular size thresholds as a function of mean angular size. Lower graph shows the difference between objective and angular size thresholds as a function of mean angular size. These data based on results shown in Fig. 4.

that some source of noise adds approx. 20" to the objective thresholds. Perhaps the constancy-scaling calculation adds a tiny amount of noise to these judgments—about one synapse worth? Or is there really some special source of noise at the small separations that diminishes at larger separations? We speculate that, at small scales, the angular thresholds are superior to the objective thresholds because, in the angular condition, observers can use local shifts in luminance or contrast to judge separation. In the objective size condition, these luminance or contrast cues are obscured by the scaled changes in angular subtense with disparity. We shall consider this second explanation at greater length in the Discussion section.

Random changes in disparity degrade angular judgments

At larger sizes, angular and objective thresholds are almost equally precise, but both are significantly higher than the thresholds for targets presented only in the fixation plane. As Fig. 7 shows, the ratio of the random disparity angular thresholds to the fixation plane thresholds is nearly a factor of 2 at all separations, indicating that the noise elevating the random disparity angular thresholds is multiplicative. An observer can judge angular size better if he or she keeps one eye closed, obscuring the random shifts in disparity. For example, we measured a "random disparity" angular threshold with one screen covered so that the target (one stereo half-image) shifted left or right from trial-to-trial over a 40 min range; the threshold measured with

these random lateral displacements was identical to the threshold for the fixation plane condition where the target was presented in one position. Thus, changes in depth, and not simply changes in retinal position, are responsible for the elevation in the thresholds.

Disparity could, in itself, increase the noise, because the neural units that code large disparities may be much coarser (larger receptive fields) than the units that code small disparities (Tyler, 1975; Marr & Poggio, 1979). Large receptive fields are spatially extended and have shallow weighting functions, so their signals provide less precise information about location than smaller receptive fields. If spatial localization of the horizontal target lines is mediated by these coarser disparity units, the random disparity thresholds would necessarily be less precise than the fixation plane thresholds. Why would the brain use the signals from these coarse disparity units to encode position when more precise smaller units are available? Here we must assume that position and disparity are jointly encoded, that there is an obligatory coupling between depth and location. To register the disparity of the target, an individual with normal stereopsis must use the coarser disparity units at a cost in the precision of spatial location. This coupling between disparity and location has been observed in another context; McKee *et al.* (1990a) found that the precise signal associated with a monocular vernier target was obscured when the vernier target was paired stereoscopically with a disparate target in the other eye.

In natural conditions, observers can change their convergence to examine objects of interest, thereby removing the potentially degrading effects of disparity. To simulate this natural state of affairs, we increased target duration to 1500 msec for one test separation (1.3 cm and 30 arc min) and asked subjects to converge to the plane of the target on every trial. The 1500 msec duration is certainly long enough to permit accurate convergence movements (Riggs & Niehl, 1960), but making them repeatedly is very demanding—a bit like doing oculomotor push-ups—and subject LW was unable to perform the task because it induced severe headaches. Increasing target duration did improve performance for the other two subjects. The angular and

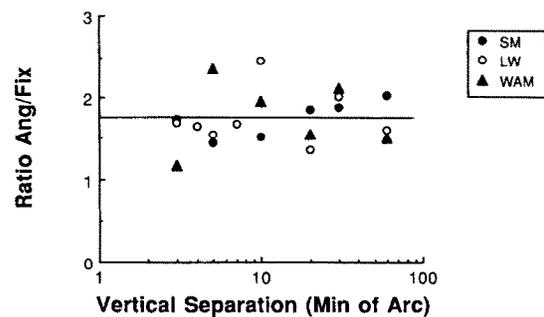


FIGURE 7. Ratio of angular size thresholds, measured with random changes in disparity, to size thresholds measured only in the fixation plane. Target duration 150 msec. The horizontal line shows mean ratio for these measurements. The random disparity thresholds for angular size are consistently higher than size thresholds measured in a single plane.

objective thresholds of subject WAM decreased by about 40%; his fixation plane showed a somewhat smaller decrease. Still, the pattern of his results at 1500 msec mimicked his results for 150 msec (see Fig. 4); his objective size threshold ($\Delta S/S = 0.08$) remained higher than his angular size threshold ($\Delta S/S = 0.06$), and both were significantly higher than his fixation plane threshold ($\Delta S/S = 0.03$). For subject SM, the objective and angular thresholds were exactly equal ($\Delta S/S = 0.05$) at the longer duration, but again both were stubbornly higher than the fixation plane threshold ($\Delta S/S = 0.03$) despite the reduction in target disparity produced by converging to the plane of the target.

What accounts for this persistent decrement in the precision of the random disparity *angular* judgments? In the long duration condition, convergence may not have been perfect, leaving the target with a small uncorrected disparity. Undoubtedly, vergence induced changes in accommodation, thereby introducing a slight amount of blur in the target (maximally $\frac{1}{3}$ D). The repetitive vergence movements could also have increased fixation instability, thus increasing motion blur. However, the thresholds for a 30 min separation are not likely to be degraded by a minor amount of disparity, motion blur or defocus (Klein & Levi, 1990b). There is another curious possibility—the well-learned scaling of angular subtense with apparent depth, may interfere with the precise judgment of *angular* size. Size constancy scaling produces a powerful illusion in most individuals—the apparent change in the size of an after-image with a change of apparent depth, i.e. Emmert's Law. Even if the perceptual aspects of size constancy are weak, as in subject WAM, the constant connection between angular subtense and depth may produce some uncertainty in these angular judgments—multiplicative noise of the type associated with uncertainty about the size of the reference.

In a final appraisal of the relative contributions of disparity and reference uncertainty to the noise in the angular thresholds, we compared our standard random disparity condition (9 planes spanning ± 40 arc min) to a condition in which the test target was presented only in 2 planes (the fixation plane and at an uncrossed disparity of 30 arc min). In both conditions, the test target was presented for a brief duration of 150 msec. Reference uncertainty presumably depends on the number of simultaneously-held references (see Morgan, 1992), so, if reference uncertainty constitutes the major source of noise in the random disparity angular thresholds, the thresholds in the two-planes condition should be superior to the 9-planes condition. On the other hand, if disparity is solely responsible for the noise, the results from the 2-planes and 9-planes conditions should be comparable, since, in both conditions, the thresholds are based on data pooled from small and large disparity trials. As Table 1 shows, the thresholds for the 2-planes condition are somewhat higher than in the fixation-plane condition, at least for 2 of the 6 comparisons. Note, however, that the 30 min angular thresholds for subject SM and WAM are actually lower in this two-

plane condition, than in 1500 msec condition described above, where convergence movements should have reduced target disparity below 30 min on most trials. Thus, disparity *per se* accounts for a portion of the noise, but it has a less consistent effect than reference uncertainty, represented by the 9-planes condition. Physically, the angular subtense between the target lines did not change with disparity, so our subjects could have employed a single angular reference for every depth. They seemed unable to make consistent use of a single reference, a result that argues that human observers do not have direct access to information about angular subtense.

"Anti-constancy" vs constancy

We were surprised to discover that subject WAM had difficulty judging objective size, so, for demonstration purposes, we constructed a target consisting of two identical pairs of lines, each pair separated by the same angular separation. Both pairs were presented simultaneously, but one pair was shown with an uncrossed disparity of 10 arc min. For subject SM, the separation between the distant pair looked larger than the separation of the nearer pair, but to subject WAM, the separation between both pairs looked identical, even given time for scrutiny. Apparently, he has no perceptual size constancy based on disparity alone. Thouless (1931) also noted considerable individual variation in reported phenomenal size in his studies of size constancy. Without phenomenal size constancy, how was subject WAM able to make the objective size judgments? He said that he had learned to use different references for different distances to minimize error feedback. For this subject, size constancy was not a natural response induced by disparity, but was instead an arbitrary recalibration.

Could the other two subjects learn to respond on some completely arbitrary basis? To answer this question, we created an experimental condition in which the angular size of the target *increased* with increasing distance, exactly inverting the natural relationship, a condition we

TABLE 1. The Weber fractions for angular size judgments

	Fixation plane only	2 planes 0 and 30'	9 planes -40' to +40'
<i>5 min mean separation</i>			
LW	0.034 \pm 0.002	0.035 \pm 0.002	0.054 \pm 0.009
SPM	0.041 \pm 0.002	0.056 \pm 0.005	0.057 \pm 0.003
WAM	0.054 \pm 0.002	0.055 \pm 0.013	0.093 \pm 0.007
Average	0.043	0.048	0.068
<i>30 min mean separation</i>			
LW	0.036 \pm 0.002	0.046 \pm 0.001	0.050 \pm 0.006
SPM	0.026 \pm 0.004	0.027 \pm 0.004	0.052 \pm 0.003
WAM	0.048 \pm 0.007	0.050 \pm 0.003	0.095 \pm 0.013
Average	0.037	0.041	0.066

In the "fixation plane" condition, the target was presented only in the fixation plane. In the "two planes" condition, the target was presented at random either in the fixation plane or with 30 min uncrossed disparity. In the "9 planes" condition, the target was presented at random in 1 of 9 depth planes, covering ± 40 arc min range. Target disparity accounts for a portion of the noise in random disparity angular judgments, but reference uncertainty also contributes to the loss in precision.

TABLE 2. The Weber fractions for 30 min angular subtense, for the conditions in which target disparity changed from trial-to-trial to 1 of 9 depth planes chosen at random: angular size judgments; objective size judgments ("constancy"), and "anti-constancy" judgments

	Many planes angular	Many planes objective	Many planes "anti-constancy"
SPM 30' separation	0.059 (0.004)	0.065 (0.004)	0.08 (0.006)
LW 30' separation	0.055 (0.005)	0.059 (0.004)	0.10 (0.006)

labeled "anti-constancy". As in the previous experiment on size constancy, subjects were required to judge target size while compensating for an orderly change in angular subtense with disparity, but in this case, their responses were judged according to criteria that had no natural counterparts. Error feedback was given throughout the experiment to assist the subjects in establishing an appropriate set of references for all depth planes. Following a small amount of practice, subjects LW and SM collected enough data (≥ 900 trials) to estimate "anti-

constancy" PSEs for single mean size (1.3 cm equal to 30 arc min at 1.5 m) at each of the 9 depth planes. These PSEs are plotted in Fig. 8, along with the PSEs for normal constancy, redrawn from Fig. 5. Although both authors were fairly good at making these "anti-constancy" judgments, neither perceived the target as being of a constant size. On the contrary, the most distant target appeared enormous, jointly magnified by the increased angular size programmed into our display, and by size constancy—our well-learned tendency to perceive the size of more distant objects as larger than nearer objects. We both adopted a similar strategy for making these judgments. We imagined a three-dimensional wedge with the small end pointed towards us, and, as each target was briefly displayed, we judged the target against this imagined reference frame—a visualization of the rescaling needed to reduce the number of errors signaled by the computer. As Table 2 shows, the thresholds for "anti-constancy" are somewhat less precise (roughly a factor of 1.5) than for constancy, but they are based on a couple of days of practice. The ease with which we learned this unnatural trick suggests that natural size constancy may also be learned response (Helmholtz, 1868; Morgan, 1992).

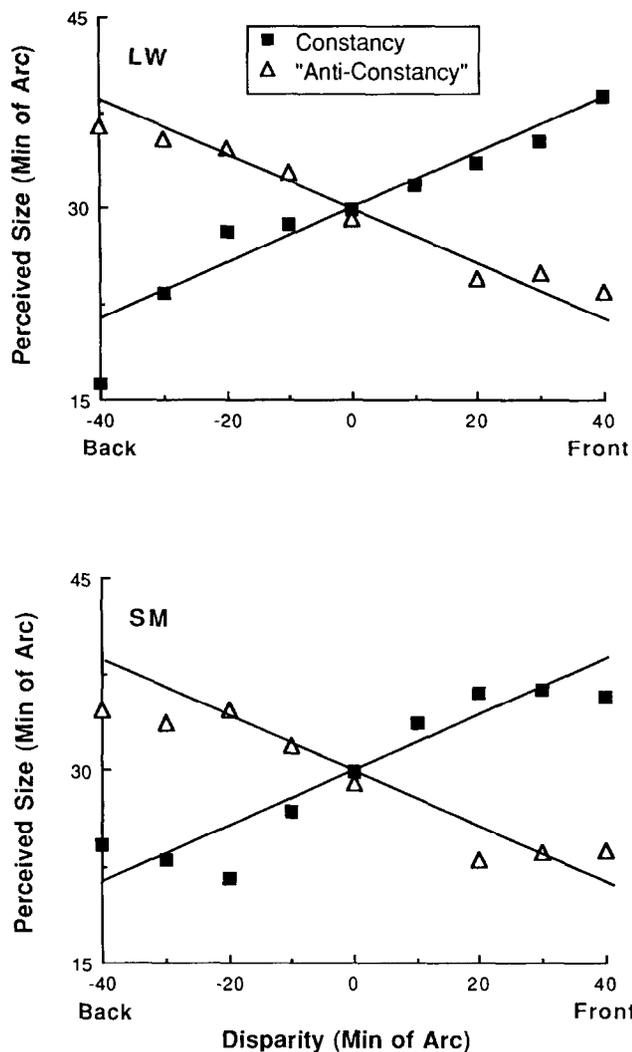


FIGURE 8. The PSEs for constancy and "anti-constancy" as a function of target disparity for one size (1.3 cm subtending 30 arc min at 1.5 m). In the constancy condition, the angular subtense decreased with increasing depth consistent with the natural geometry associated with increasing physical distance. In the "anti-constancy" condition, the angular subtense increased with increasing depth completely inverting the natural relationship. Target duration 150 msec.

DISCUSSION

In our artificial viewing conditions, the best subjects could judge objective size with a precision of 5–6%. These low thresholds may nonetheless underestimate the precision of size constancy in natural surroundings, where the rich array of visual information could substantially reduce uncertainty about depth and size. We were therefore pleased to discover that the most precise Weber fractions derived from Gilinsky's data were between 6–7%, since Gilinsky had performed her experiments in an open field and had allowed her subjects unlimited viewing time. Our experiments lead to two unexpected conclusions. First, observers cannot judge objective size (cm) with the same precision as they judge angular size (deg) for targets subtending < 10 arc min, but, beyond this range, some observers can judge objective and angular size with about equal proficiency. Second, observers do not have direct access to information about angular subtense, at least not for targets viewed binocularly.

Why should there be a lower limit on size constancy? As described above, the small differences between the angular and objective thresholds are consistent with additive noise. If target disparity accounted for most of the noise in both the angular and objective thresholds,

the small additive component associated with the objective thresholds might be attributed to the neural calculation required for size constancy. However, the case for additive noise is not compelling for a number of reasons. First, disparity explains a portion of the noise in our thresholds, but not all of it. Second, scaling or reference models of constancy (Andrews, 1964; Morgan, 1992) generally predict multiplicative noise in the objective thresholds, rather than additive noise. Finally, some studies have reported that objective judgments at large scales are actually more precise than angular judgments (Leibowitz & Harvey, 1969; Burbeck, 1987a; and subject SM for the 60 arc min separation in the present study). The perceived size data suggest a different kind of explanation. At 5 arc min, the PSEs for objective size are thoroughly intermixed with the PSEs for angular size (see Fig. 5) as though the subjects could not develop a strategy for making the objective size judgments. The "anti-constancy" judgments are just as imprecise as the 5 arc min objective judgments, but the "anti-constancy" PSEs show the predicted relationship with disparity (see Fig. 8), because the subjects had a clear strategy for scaling these unnatural judgments. Apparently, our subjects could not manipulate or scale the information they normally use to judge changes in small separations.

There is increasing evidence that different cues are used for judging small distances than for judging large ones (Burbeck, 1987b; Klein & Levi, 1985, 1987; Levi, Klein & Yap, 1988; Levi & Klein, 1990a; Wilson, 1991). The cues at small distances are more akin to contrast judgments than to distance judgments, in that observers use subtle changes in the light distribution to detect changes in position. To take an extreme example, consider a pair of bright lines separated by a barely discernible gap of 2 arc min. If the separation is decreased, the lines fuse into a single bright bar; if the separation is increased, the gap becomes darker and more distinct. At slightly larger distances, e.g. 5 arc min, the apparent width of the central dark gap can be compared to the apparent width of either of the bright target lines. While these particular cues are specific to our task, thresholds for judging the dimensions of features that subtend small angles (< than 10 arc min) are probably limited by noise in mechanisms that code local changes in luminance or contrast. It is difficult to use these subtle contrast-dependent cues in the angular size condition because of the random variations in disparity, but, in the objective size condition, these cues are completely obscured by the changes in angular subtense accompanying the variations in disparity. In order to use these contrast-dependent cues to judge objective size, the observer would need a separate template of the light distribution of the reference size for every tested disparity. An ideal observer could readily store these multiple templates, but the real observer has many opportunities to make mistakes—to misjudge the disparity or to use the wrong template—with a resulting decrement in precision.

Doesn't the real observer have this problem with objective judgments at all sizes? Why should this problem disappear at larger scales? The observer uses a

different strategy to judge the distance between widely-separated features. When there is no overlap between the retinal images of the target lines, the light distribution provides little information about separation, so instead each target line is separately localized according to its spatial co-ordinates ("its local sign"). The evidence for this premise is that, once target separation exceeds 5–10 arc min, thresholds are not affected by target contrast, by target spatial frequency content, or even by whether the targets are of opposite contrast sign (Morgan & Regan, 1987; Burbeck, 1987b; Levi & Westheimer, 1987). At small scales, the contrast-dependent changes in the light distribution provide more precise information about separation than the spatial co-ordinates of the target lines. At large scales, this contrast-dependent information is no longer available so separation is economically coded as the distance between the spatial co-ordinates corresponding to the "local signs". We speculate that size constancy operates only on the information supplied by the "local signs", perhaps by re-scaling the distance between the co-ordinates of each individual contour with changes in depth (Andrews, 1964), or perhaps by maintaining a separate reference size for every disparity (Morgan, 1992). Certainly it would require less information (fewer "bits") to store a neural representation of the distance between the spatial co-ordinates than to store a complete template of the luminance distribution of the reference size for every tested disparity.

On which spatial map do these co-ordinates lie? Traditionally, "local signs" are thought to refer to retinal co-ordinates, but our data indicate that an observer with normal stereopsis viewing a target binocularly does not have access to retinal or monocular information (see also McKee, Levi & Bowne, 1990a), else there would be no discrepancy between the fixation plane thresholds and the angular thresholds measured with random changes in disparity. The physical information available at the eye is angular subtense, but it is unlikely that information about angular subtense or retinal location is preserved beyond the binocular confluence occurring at striate cortex. The first primitive map of location may be generated at the striate level, but it already represents a transformation of the retinal co-ordinates (Levi *et al.*, 1985). For off-horopter targets, perceived location necessarily reflects signals from both retinae, and is not congruent with either (Nelson, 1977; Sheedy & Fry, 1979; Rose & Blake, 1988). Thus, there must be multiple representations of spatial location which depend on target disparity. How does the brain co-ordinate these multiple maps of location so that a feature on one depth map can be compared to a feature on a different map? In making *angular* judgments under natural conditions, the observer can change convergence so that all comparisons are made in the fixation plane, where there should be no question about angular size. We were therefore puzzled when the discrepancy between the fixation plane angular thresholds and the random disparity thresholds persisted, even when our subjects were given adequate time (and encouragement) to change their convergence. Convergence errors probably account for

part of the discrepancy, but a portion may be reasonably attributed to reference uncertainty, the confusion introduced by the parallel calculation of objective size.

Unlike the results of Burbeck's study (1987a), our objective thresholds were never equal to thresholds for targets presented in a single plane. Her observers only had to cope with 2 planes, whereas ours were dealing with 9. Morgan (1992) has argued persuasively that size constancy reflects the human ability to use multiple references in judging size. In the case of size constancy, target disparity acts as the cue to the appropriate reference. Our "anti-constancy" results show that a subject can quickly learn an orderly, but arbitrary, reference system in judging size. Morgan (1992) asked subjects to judge the width separating a pair of lines presented at one of four different orientations on interspersed trials; there was a different implicit reference width for each orientation, so subjects were forced to switch their size criterion with changes in target orientation. He found that the precision of size judgments was unaffected when the size criterion changed systematically with changes in orientation or position, for four different reference sizes. A slight, but perceptible, loss in precision is apparent in his data as the number of references increased from four to eight; Weber fractions for his best subject increased from about 3% to about 5%. This small increase in thresholds is comparable, in our best subjects, to the difference between their single plane, single criterion thresholds, and their multi-plane, multi-criteria thresholds for both objective and *angular* size.

We began this study with the simplistic picture drawn in Fig. 1; either the human observer had access to information about angular subtense, or the observer lost this information in the calculation of objective size and had to estimate angular subtense indirectly. Regrettably, the complicated diagram shown in Fig. 9 is probably a better description of the operational flow chart in human size processing. All retinal signals are affected by the

noise in cortical binocular units, and there is undoubtedly more noise or uncertainty associated with the off-horopter units than with units serving the fixation plane. For targets subtending very small angles, the best information about changes in size or distance is supplied by delicate changes in the light distribution; it is difficult or impossible to scale these contrast-dependent changes with the depth-dependent changes in angular subtense, so size constancy is very inefficient at small scales. At larger sizes, we think that estimates of angular and objective size depend on two separate parallel processes and that neither of these perceptual dimensions is derived directly from the other. Far from being only the "raw data" of experience, angular size is the basis of a powerful cue to depth—size or texture gradients (Gibson, 1950)—so there is a good computational reason for preserving this useful information. If the diagram in Fig. 9 is correct, the normal observer with good stereopsis always has two conflicting estimates of size.

Fifty years ago, Thouless (1931) concluded that perceived size is a compromise between angular subtense and perfect size constancy. He also noted considerable individual variation in the balance between these extremes, with artists tending to match their size percepts closer to the angle subtended at the retina. If, as we suppose, estimates of angular and objective size are both learned manipulations of the spatial co-ordinates on the different depth maps, then it is hardly surprising that these parallel estimates have different weights in different individuals. Most individuals can supply fairly precise information about either angular or objective size as required, but, in our demanding experimental conditions, Thouless' compromise—the parallel calculation of angular and objective size—increases reference uncertainty.

REFERENCES

- Andrews, D. P. (1964). Error-correcting perceptual mechanisms. *Quarterly Journal of Experimental Psychology*, 16, 105–115.
- Boring, E. G. (1943). The moon illusion. *American Journal of Physics*, 11, 55–60.
- Burbeck, C. A. (1987a). On the locus of spatial frequency discrimination. *Journal of the Optical Society of America A*, 4, 1807–1813.
- Burbeck, C. A. (1987b). Position and spatial frequency in large-scale localization judgments. *Vision Research*, 27, 417–427.
- Carlson, V. R. (1960). Overestimation in size constancy. *American Journal of Psychology*, 73, 199–213.
- Carlson, V. R. (1977). Instructions and perceptual constancy judgments. In Epstein, W. (Ed.), *Stability and constancy in visual perception: Mechanisms and processes* (pp. 217–254). New York: Wiley.
- Foley, J. M. (1976). Successive stereo and vernier position discrimination as a function of dark interval duration. *Vision Research*, 16, 1269–1274.
- Gibson, J. J. (1950). *The perception of the visual world*. Boston, Mass.: Houghton-Mifflin.
- Gilinsky, A. S. (1955). The effect of attitude upon the perception of size. *American Journal of Psychology*, 68, 173–192.
- Gogel, W. C. (1969). The sensing of retinal size. *Vision Research*, 9, 1079–1094.
- Gregory, R. L. (1966). *Eye and brain* (pp. 151–154). London: Weidenfeld & Nicholson.
- Gregory, R. L. (1987). *The Oxford companion to the mind* (p. 218–219). New York: Oxford University Press.

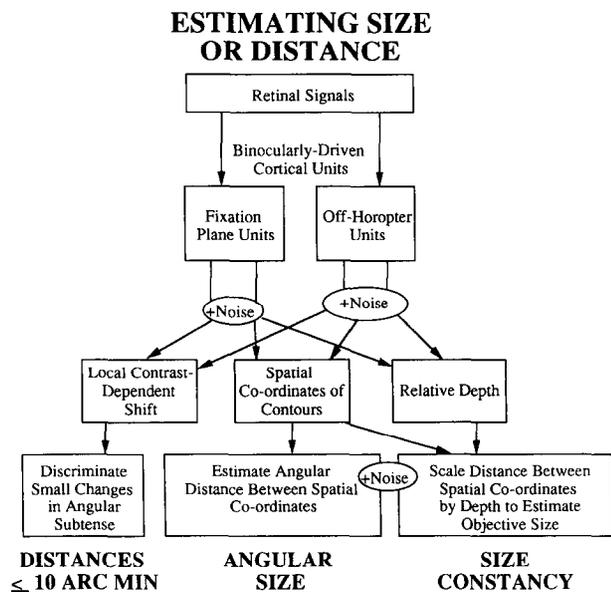


FIGURE 9. Flow chart showing the operations involved in the human estimate of angular and objective sizes. See text.

- Helmholtz H. V. (1868). The recent progress of the theory of vision. Translated by Pye-Smith, P. H. In Kline, M. (Ed.), *Popular scientific lectures*, New York: Dover.
- Holway, A. H. & Boring, E. G. (1941). Determinants of apparent visual size with distance variant. *American Journal of Psychology*, *54*, 21–37.
- Klein, S. A. & Levi, D. M. (1985). Hyperacuity thresholds of 1 sec: Theoretical predictions and empirical validation. *Journal of the Optical Society of America A*, *2*, 1170–1190.
- Klein, S. A. & Levi, D. M. (1987). Position sense of the peripheral retina. *Journal of the Optical Society of America A*, *4*, 1543–1553.
- Leibowitz, H. W. & Harvey, L. W. (1969). Effect of instructions, environment, and type of test object on matched size. *Journal of Experimental Psychology*, *81*, 36–43.
- Levi, D. M. & Klein, S. A. (1990a). The role of separation and eccentricity in encoding position. *Vision Research*, *30*, 557–586.
- Levi, D. M. & Klein, S. A. (1990b). Equivalent intrinsic blur in spatial vision. *Vision Research*, *30*, 1971–1993.
- 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.
- Levi, D. M., Klein, S. A. & Aitsebaomo, P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, *25*, 963–977.
- Levi, D. M., Klein, S. A. & Yap, Y. L. (1988). Weber's Law for position: Unconfounding the role of separation and eccentricity. *Vision Research*, *28*, 597–603.
- Marr, D. & Poggio, T. (1979). A computational theory of human stereoscopic vision. *Proceedings of the Royal Society, B204*, 301–328.
- McKee, S. P. & Levi, D. M. (1987). Dichoptic hyperacuity: The precision of nonius alignment. *Journal of the Optical Society of America A*, *4*, 1104–1108.
- McKee, S. P. & Welch, L. (1989). Is there a constancy for velocity? *Vision Research*, *29*, 553–561.
- McKee, S. P., Levi, D. M. & Bowne, S. F. (1990a). The imprecision of stereopsis. *Vision Research*, *30*, 1763–1779.
- McKee, S. P., Welch, L., Taylor, D. G. & Bowne, S. F. (1990b). Finding the common bond: Stereoacuity and the other hyperacuties. *Vision Research*, *30*, 879–891.
- Morgan, M. J. (1989). Vision of solid objects. *Nature*, *339*, 101–103.
- Morgan, M. J. (1991). Hyperacuity. In Regan, D. (Ed.), *Vision and visual dysfunction: Spatial vision*. London: Macmillan.
- Morgan, M. J. (1992). On the scaling of size judgments by orientational cues. *Vision Research*, *32*, 1433–1445.
- Morgan, M. J. & Regan, D. (1987). Opponent model for line interval discrimination: Interval and vernier performance compared. *Vision Research*, *27*, 107–118.
- Nelson, J. I. (1975). Globality and stereoscopic fusion in binocular vision. *Journal of Theoretical Biology*, *49*, 1–88.
- Riggs, L. A. & Niehl, E. W. (1960). Eye movements recorded during convergence and divergence. *Journal of the Optical Society of America*, *50*, 913–920.
- Rose, D. & Blake, R. (1988). Mislocalization of diplopic images. *Journal of the Optical Society of America A*, *5*, 1512–1521.
- Ross, J., Jenkins, B. & Johnstone, J. R. (1980). Size constancy fails below half a degree. *Nature*, *283*, 473–474.
- Sheedy, J. E. & Fry, G. (1979). The perceived direction of the binocular image. *Vision Research*, *19*, 201–211.
- Thouless, R. H. (1931). Phenomenal regression to the real object. *British Journal of Psychology*, *21*, 338–359.
- Tyler, C. W. (1975). Spatial organization of binocular disparity sensitivity. *Vision Research*, *15*, 583–590.
- Westheimer, G. & McKee, S. P. (1977). Spatial configurations for visual hyperacuity. *Vision Research*, *17*, 941–947.
- White, J. M., Levi, D. M. & Aitsebaomo, A. P. (1992). Spatial localization without visual references. *Vision Research*, *32*, 513–526.
- Wilson, H. R. (1991). Psychophysical models of spatial vision and hyperacuity. In Regan, D. (Ed.), *Vision and visual dysfunction: Spatial vision*. London: Macmillan.

Acknowledgements—This research was supported by the Air Force Office of Scientific Research Grant #89-0035 and NIH Core Grant 5 P-30-EY-01186. We thank Professor Michael J. Morgan for his thoughtful critique of the manuscript, Dr Samuel Bowne for helpful discussions and Doug Taylor for his carefully conceived programs and technical assistance.