

# Development of clinical tests of vision: Initial data on two hyperacuity paradigms

JAY M. ENOCH and RICK A. WILLIAMS  
*School of Optometry, University of California, Berkeley, California*

Many basic psychophysical functions offer promise as clinical tests of vision. Here, we discuss problems that one encounters in the clinical setting, how one identifies a psychophysical test for potential clinical development, and an orderly approach to development of suitable test paradigms. Parameters are selected which are relatively insensitive to variables encountered in the field (clinic) in a normal population, but which are sensitive to changes in the response system being studied. Initial data on two hyperacuity tests are presented. These tests are adaptations of hyperacuity paradigms (Westheimer, 1979) to a clinical environment. This particular test set offers promise because it exhibits a unique threshold which is dependent upon neural data processing and is relatively independent of retinal image degradation.

There is a monumental gap between the psychophysicist seeking to develop a test of vision or other modality and the clinician seeking to assess the capabilities and characteristics of the patient. It is crucial that those of us seeking to apply modern science to the clinical environment convey to other scientists with a common interest some of the essentials as to how to proceed in the orderly development of tests having clinical applicability. It is our purpose here to describe several of our recent experiences in this area.

Psychophysical tests have a special advantage in clinical testing in that they provide information about the individual in a noninvasive manner. However, it is a far step from the laboratory in which one has the privilege of 100 or more repetitions of a two-alternative, forced-choice decision to the busy clinic where the patient load is often large, where there are competing demands for the time of the professional, the technicians, or even automated devices, where attention spans of patients (as opposed to paid subjects) may be quite short, and where there is a variance in educational and social background among both dispensers of the test and patients receiving the test. There can be little ambiguity in the result; tests must be valid, rapidly applied, highly reliable, and provide unique knowledge to the clinician. It is often difficult to achieve such a goal. Yet, surely this is a task worthy of the effort, for it is the general uplifting of the human

condition which is one of the major motives that serve as the driving force behind science.

In the clinical environment, it is often necessary to have more than one test available to sample a specific response system, since a given test may not be usable on an individual patient. The clinician has to be versatile enough to know when to abandon a nonproductive routine and to seek alternatives when answers must be obtained. He or she must also be able to recognize the need for further improvement in available or evolving techniques.

## DEVELOPING A NEW TEST

The development of a new clinical test represents a major commitment of time, resources, and personnel. The selection of the specific test requires careful and thoughtful research into the background literature covering the response function of interest. The procedure for developing the selected test for the clinic must be orderly and somewhat standardized so that the final product provides reliable and consistent assessment of the desired function.

Through development, standardization, and extensive clinical trials, tests that were once highly specialized experimental techniques familiar only to a select few vision scientists are now routinely applied for clinical assessment of vision in patients. Consider the transformation of increment threshold experiments into measurements of static perimetric fields on the Goldmann perimeter (e.g., Aulhorn & Harms, 1972; Enoch, 1963, 1980).

The increment threshold paradigms per se existed before the developmental trials were undertaken. The available vision research literature provided a logical case for the selection of this test for development. The experimenters first had to verify

This research was supported in part by National Eye Institute Grants EY 03669 and EY 03674, NIH, Bethesda, Maryland. This manuscript was presented as an invited paper at the International Brain Research Organization meeting in Lausanne, Switzerland, on April 1, 1982. The authors' mailing address is: Visual Studies Center, School of Optometry, University of California, Berkeley, California 94720.

critical experiments reported in the literature and select appropriate parameters.

The selection of parameters is quite possibly the most crucial step in test development, and a practical example is given here in the testing of perimetric visual fields. Parameters are chosen *not* for maximum visual sensitivity, but for maximum reliability in field applications in which poor control conditions often exist. Assume we want to test the just noticeable luminous brightness increment or the static perimetric threshold for a small test object presented for longer than the critical duration at several different points in the monocular visual field. The fixation point is sharply focused and centrally fixated by the observer.

From the days of the mid-19th century, we know that when measuring  $\Delta B$  and varying  $B$  (where  $\Delta B$  is the luminance of the just detectable test target at a given test point and  $B$  is the luminance of the large background plane) two modes of response (ignoring saturation) are recorded (Aguilar & Stiles, 1954). At low background luminance levels,  $B$ , the threshold  $\Delta B$  approaches a constant, and at higher background or adaptation levels, the ratio  $\Delta B/B$  (the Weber fraction) approaches a constant (Figures 1b and 1a).

In a routine clinical test situation, we seek to define visual sensitivity at specific points in the visual field. In the field (clinic), one may inquire as to the effect of pupil size variations [as a population variable, or due to age, refractive error, and preadaptation conditions, or due to alterations due to the addition of miotic (pupil contracting) or mydriatic (pupil dilating) pharmaceutical agents, etc.], uniformity of illumination of the background field, fluctuation of line voltage at peak load times, aging effects on the light source, etc. If response is governed by  $\Delta B = \text{constant}$  (i.e., low background levels), every fluctuation of the just listed factors will alter measured  $\Delta B$  over and above intrinsic sensitivity variations in the visual system. On the other hand, if both  $\Delta B$  and  $B$  are provided by the same light source and the luminance or adaptation level,  $B$ , is set in the range in which  $\Delta B/B$  is a constant, these several factors will have little effect on response as long as parametric alterations are modest. For example, if the pupil is smaller by 1 or 2 mm, retinal illuminance, the visual stimulus, is reduced *for both*  $\Delta B$  and  $B$ , but the ratio is not altered! On the other hand, if the visual system response is anomalous, measured  $\Delta B/B$  will probably be altered. Thus, setting the level of  $B$  in the range in which  $\Delta B/B$  is a constant is advantageous. However, if background luminance is too high, problems are encountered, particularly in the elderly, whose small pupil size and incipient cataract formation tend to reduce contrast and blur boundaries. In perimetry we are moving towards an adaptation/background luminance level

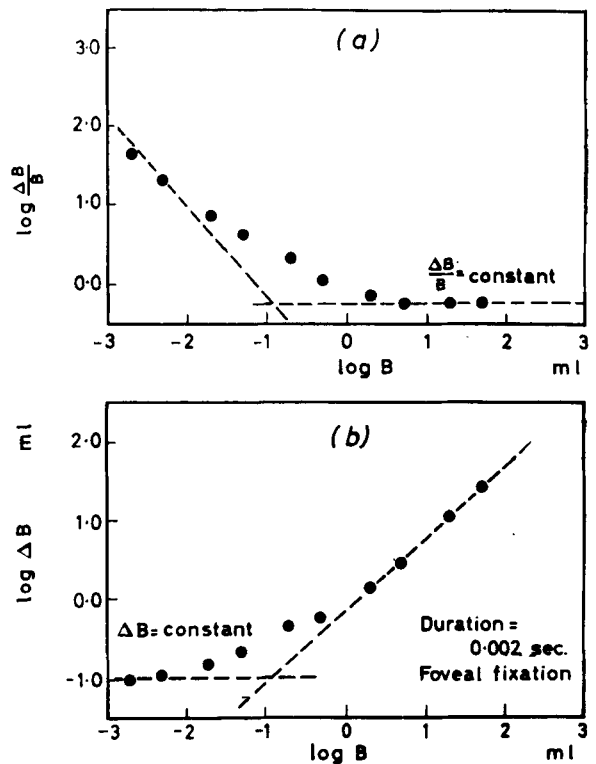


Figure 1. The same set of foveal response data is plotted on this pair of graphs. The ordinate on each graph is different. In (a),  $\log \Delta B/B$  is plotted against the logarithm of the luminance. From this graph, it is evident that at higher luminance levels  $\Delta B/B$  is a constant; that is, the curve appears to asymptote on a line parallel to the abscissa. Actually, at very high luminance levels, this relation no longer holds and the value of  $\Delta B/B$  starts to rise again. In (b), one is concerned with events taking place at low luminance levels. It is evident that the hypothetical curve passing through these data is approaching another asymptote. In this case, it is  $\Delta B$  which is equal to a constant. The value of the constant is the absolute threshold. (Enoch, 1963, Fig. 115, p. 239; reproduced courtesy of Butterworths, Ltd.)

of approximately 31.83 cd/m<sup>2</sup> (100 apostilbs, international, or 10 mL; see Enoch, 1979).

Thus, parameter selection is crucial. Again, the prime determinant is not maximum sensitivity per se, but, rather, reliable realization of the goal of the test with minimal false positives and negatives. In essence, one expects the psychophysical test to provide information in several areas:

First, is the measured function normal or abnormal, compared with a normal age-matched population?

Second, if it is not normal, how abnormal is the measured function? If a routine scaling (e.g., Snellen acuity) cannot be achieved in clinical situations, a rather classic categorization scheme is used: 0 (normal), +1 (mild/early error), +2 (modest error), +3 (advanced error), +4 (extreme error, anomalous appearance, or behavior).

Third, are there alternative measures that one

may use if the primary technique fails, cannot be applied, or is contraindicated?

Fourth, how reliable is the test? A reasonable number of replications of a determination within a single visit should give consistent results.

Fifth, how relevant is the test to the response system one seeks to evaluate? *Information is gleaned from the total experience*, not just the test in question. The clinician-examiner is considering correlated behavior, that is, specific responses elicited by the test stimulus, and these results are placed in a larger, broader context, and not taken in isolation.

Finally, assurance is needed that the therapeutic measures taken are beneficial to the patient. The techniques at hand should provide fine measurements of visual (or other) function that are useful in properly modulating a therapeutic regimen.

### HYPERACUITY: AN EXAMPLE OF THE DEVELOPMENT OF A NEW CLINICAL TEST

In order to illustrate the orderly process of clinical test development, we now present data on two hyperacuity tests that we are currently developing in our laboratory for clinical application. One of these tests will be utilized in the test battery we are developing, termed "quantitative layer-by-layer perimetry" (Enoch, Fitzgerald, & Campos, 1981), and the second is expected to prove useful in the assessment of central retinal and visual pathway integrity in the presence of irregular, translucent, or partially opaque ocular media.

#### Selection of a Test for Development

Several questions must be addressed as a possible test is selected for development in order to determine if the task is worthy of pursuit.

(1) Is the new test unique in providing the clinician some valuable piece of information not currently available? If not, how does it overlap or partially replicate other tests?

Previous evidence suggests that vernier-acuity tasks probably reflect central visual processing. A class of visual responses, including vernier judgments, stereoscopic judgments, the detection of the slope or tilt of a line, etc., all having a nearly common and extremely fine threshold, has been termed "hyperacuity" by Westheimer and his co-workers (Westheimer, 1975, 1979; Westheimer & McKee, 1977, 1978). These responses all have just detection thresholds on the order of 3 to 8 sec of arc. These judgments are dimensionally far finer than the retinal receptor mosaic. Since the stimuli can be presented dichoptically (Westheimer & Hauske, 1975) with little or no change in threshold and comparable magnitude responses exist in stereoscopic judgments, it is assumed that there must be

some centrally located data-processing stage for such capability. The notion of a central locus within the visual pathway attracted us to this stimulus set for purposes of development of a test suitable for clinical application. Assuming that such a locus can be isolated, this clinical research can be of great value in defining the organization of the visual process, per se.

The parallel remarkable insensitivity of this set of stimuli to image degradation (Stigmar, 1971; Westheimer, 1979; and see Figure 4, below) also offers great clinical promise. Simply stated, there are few events encountered that are quite as disappointing to a patient and clinician than to have a good surgical result following removal of ocular media opacities only to reveal subsequent poor visual function. The ophthalmologist seeks assurance that meaningful visual fields and central retinal function exist prior to the decision to operate. Generally, doubt increases in proportion to the magnitude of translucency/opacification of the media. Numerous tests have been developed, most of which fail or give ambiguous results when one needs them most. In preliminary studies, we have been greatly impressed by the hyperacuity paradigms in assessment of function through models of various types of media alterations.

(2) Is the parametric function which forms the basis of the test robust; that is, are large changes recorded when the stimulus is altered? Clearly, a 0.1 log unit effect is *not* appropriate, although careful parametric studies may reveal an enhanced response that is practical for clinical application. Data, which we discuss below (Figure 2), show that hyperacuity thresholds are highly dependent upon the stimulus spatial configuration (over 1.0 log unit changes in threshold can be observed).

(3) Do standards for the test response exist, or must they be formulated? A psychophysical test that has already seen considerable experimentation among normal subjects is often preferable to a new, less well-seasoned paradigm. The characteristics of the hyperacuity responses have recently received considerable attention (e.g., Westheimer, 1979), and related visual phenomena were reported in the late 19th and early 20th centuries (Anderson & Weymouth, 1923; Berry, 1948; Best, 1900).

(4) Must the clinician perform a statistical analysis, or is the test result clear from the recorded data? Most experimental methods of measuring a hyperacuity threshold involve a statistical analysis (probit) of multiple trial results, although these and other statistical algorithms are readily implemented on small microcomputers. The preferred approach, however, is to seek alternative testing paradigms that are fast and efficient and that yield a clear result with minimal data analysis (see below).

(5) What is the cost in terms of time and equip-

ment for development of the test? Can it be implemented on available clinical devices? Hyperacuity tests as presently conceived employ computer-generated line stimuli presented on a TV (CRT) screen. These tests could readily be incorporated into a small microprocessor system that might already be serving in today's clinics as an automated filing system.

(6) Is the test readily calibrated and maintained in the field? The test should be relatively insensitive to critical and difficult calibration of stimuli. Hyperacuity responses are surprisingly insensitive to stimulus alterations such as contrast, luminance, and blur. With solid state electronics used for generating stimulus displays, calibration of target displacements should be highly reliable and stable over time.

(7) Finally, and importantly, how much training is needed by the examiner and who will provide the training? If the procedure is complex, it will be difficult to establish standards in the field, and technical assistance can be a major problem. Many of the hyperacuity configurations involve only a simple judgment on the patient's part as to the relative location of neighboring stimulus components; that is, is the upper line of a vernier pair to the right or the left of the lower line? Training of both patient and clinician will be minimal.

In answering these questions, one must distinguish between a test to be applied in a primary-, secondary-, or tertiary-care facility, that is, a generalist's office, a specialist's office, or a regional care/referral center. The more restricted the applicable population and the more delicate and difficult the technique, the more likely it is that the test will be limited to secondary and tertiary centers. Often the decision is dependent on population size; that is, a test applicable to diabetics or the aged clearly has broader application than one for relatively rare, blue-cone monochromacy.

#### **Procedure for Developing a Test**

The procedure for developing a new test must be orderly and somewhat standardized so that the final product provides reliable and consistent assessment of the desired function. In a general procedure for test development, the psychophysicist might:

(1) Replicate previously established results in order to verify the techniques in the present environment.

(2) Choose a selected number of test parameters, which together define the functional response under study. Parametric values for the test are chosen for maximum reliability and information transfer.

(3) Define the responses to the test format within a normal subject population that is directly comparable to the target patient population.

(4) Establish a simple calibration technique that will be field applicable and standardize the data-reporting format.

(5) Enhance the speed and efficiency of the test with minimal loss of reliability and validity.

(6) Establish written instructions for both examiner and patient. The instruction format must often be altered in the transition from trained psychophysical subject to patient.

(7) Run initial trials within a small sample of patients to point up problems and advantages of the selected test.

(8) Correlate initial results using the new test with the results of established tests that might be applied under similar circumstances.

(9) The final stage involves clinical trials, usually in a double-blind situation, to establish the general utility of the new technique.

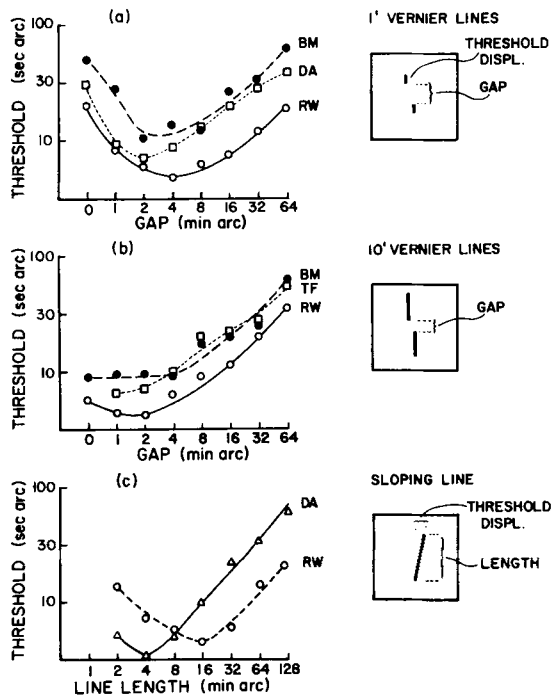
To illustrate this procedure, we will use two hyperacuity paradigms as examples. We are developing the two test sets in parallel: one as part of a quantitative layer-by-layer perimetric test battery, and the second as a test of visual function in the presence of ocular opacities.

Initially, testing is performed with *all* suitable controls exercised. The high trial rates, the precise specification of all stimuli, the randomness of trials, etc., are all employed. It is only later, after the baseline parametric searches have been made and normal populations have been tested, that the test format is simplified for clinical application. Stated another way, most of the trials and psychophysics are performed *before* one ever sees the patient!

#### **Hyperacuity as Part of a Layer-by-Layer Perimetric Test Battery**

As discussed above, the evidence suggests that hyperacuity responses are probably based on visual processing at a central locus within the visual pathway. This is our prime motivation for selecting this test for development and incorporation into our layer-by-layer test battery.

**Step 1. Verification.** Sample data obtained using three of many possible test paradigms are presented in Figure 2. These stimulus patterns are viewed on a CRT screen and are produced by a laboratory computer that is programmed to time stimulus presentation, to deliver feedback to subjects, and to perform simple statistical analyses of the summed responses. The data were collected using a two-alternative, forced-choice procedure with stimulus levels presented in a random order based on the method of constant stimuli. Each data point represents the threshold displacement, in seconds arc of visual angle, estimated from a probit analysis (Finney, 1971) of 200 to 300 trials distributed over 7 to 10 stimulus displacements (i.e., 1, 2, 3, . . . sub-



**Figure 2.** Comparison of hyperacuity responses between subjects to three stimulus configurations: (a) Two spots, each 1 min arc  $\times$  1 min arc, separated by a variable-sized gap. Threshold displacement (i.e., the displacement of the upper spot to the right or to the left of the lower spot which could be detected on 75% of the presentations) in seconds of arc vs. gap size (minutes of arc) are plotted on log-log coordinates. Threshold values differ among the subjects, but the shape of the curves is consistent. (b) Two lines, each 10 min arc long and 0.25 min arc wide, separated by a variable-sized vertical gap. (c) A single vertical line, 0.25 min arc wide and of variable length, is presented tilting slightly either to the left or right of vertical. Threshold displacement is measured as the just detectable linear displacement of the upper end of the line relative to the lower end, in seconds of arc of visual angle.

units to the right or left of alignment). The observer's task was to report on any given trial whether the upper line (Figures 2a and 2b) was displaced to the right or left of the lower line, or whether the slope of the single line (Figure 2c) was angled to the right or left of vertical. (In a clinical environment, alternative phrasing might be used, such as "toward or away from the examiner" relative to the reference, as left/right is a surprisingly confusing notion to some.) These results verify and extend those of others (e.g., Westheimer & McKee, 1977).

**Step 2. Parameters.** Note that in Figure 2 the threshold values for equivalent stimuli differ among the subjects, while for each stimulus configuration the shapes of the functions are consistent. These results and similar data published by Westheimer and his co-workers (e.g., Westheimer & McKee, 1977) suggest that hyperacuity capabilities are highly

dependent on the spatial configuration of the stimulus (i.e., separation of vernier lines, distance between the line endpoints in the line-orientation task). This dependency produces functions with distinct shapes (some "U"-shaped, others more monotonic in form) that are consistent across observers. Such a response characteristic is one that might be extremely useful in clinical research for evaluation of the integrity of the information processing which seems to be involved in the hyperacuity response.

If the analysis were limited to these data, we might select gap sizes 0, 4, and 16 min arc for the stimulus displays shown in Figures 2a and 2b and lines 2, 4, and 32 min arc long for the display shown in Figure 2c in order to define a common curve shape in the normal population. A host of other parameters (e.g., luminance, contrast, stimulus duration, etc.) would have to be assessed as well.

We have begun to define the hyperacuity response in parafoveal regions in a normal population (also see Westheimer, 1982). For one subject, the vernier threshold (1-min-arc lines, 4-min-arc gap) increases fairly rapidly from about 5 sec arc for foveal fixation to about 10 sec arc at 1-deg eccentricity and to 40-50 sec arc at 4-deg eccentricity. The appropriate hyperacuity configurations must be defined at eccentricities ranging from 1 deg up to 20 deg in the horizontal, vertical, and oblique meridians.

**Step 3. Patient populations.** The tests we have so far described could be applied within a large range of target patient populations, including stroke patients who have recovered significant verbal abilities and occipital lobe head-injury patients. In our study of cataract patients, for example, the appropriate control group would be an age-matched population of normal subjects, since, in this case, age might be a contributing factor to the normal response. In effect, we will try every type of hyperacuity configuration that seems promising in order to define those that elicit the most stable, reliable responses from a population of observers (both normals and patients).

**Step 4. Calibration of stimulus displacements.** Calibration of stimulus displacements is fairly straightforward. One can simply photograph the stimuli displayed on the CRT screen and include a reference reticle in the same plane as the stimulus. Conversion from displacement at the target plane to visual angle is a simple trigonometric calculation.

**Step 5. Test efficiency.** The test procedures that we have adopted from Westheimer's lab require 200 to 300 trials to obtain a single threshold estimate, with a standard error of less than 10% of the threshold value. In a clinical setting, in which the patient's and clinician's time is at a premium and in which a hyperacuity test may be only one of many tests to be run in a diagnostic series, the number of trials

and the time required for each test must be kept to a minimum.

We have therefore begun to develop faster and more efficient testing procedures that are compatible with the basic hyperacuity stimulus configuration. One method that appears to be quite successful is a staircase procedure (Cornsweet, 1962; Wetherill & Levitt, 1965). For a vernier stimulus, the upper line is initially offset to a relatively large distance from the lower line. On each trial, the displacement is either decreased or increased, depending on whether the observer's judgment was correct or incorrect. Several staircase criteria have been tried, each providing an estimate of a different point on the underlying psychometric function. We have found the following scheme to be optimal: during an estimation phase, the vernier offset is decreased by one step every time the observer makes a correct judgment. The first incorrect response terminates the estimation phase, after which two consecutive correct responses are required before the offset is further decreased. A single incorrect response causes the offset to increase by one step on the next trial. Using this protocol, a hyperacuity threshold estimate can be obtained in 15-25 trials, depending on the observer's response variability. The data in Figure 3 compare threshold offset estimates that are obtained with this staircase procedure with those resulting from the method of constant stimuli with a probit analysis of the results. In all but one case, the staircase estimates agree to within one standard error of the probit-estimated values. Such comparisons of the modified technique with the established method are extremely important in this phase of test development, and we are continuing to exercise this staircase procedure in other subjects and for other stimuli.

We have also experimented with adjustment techniques. The observer pushes a left button to move the upper line to the left or a right button to move it to the right. When he/she is satisfied with the alignment, the computer records the resulting offset

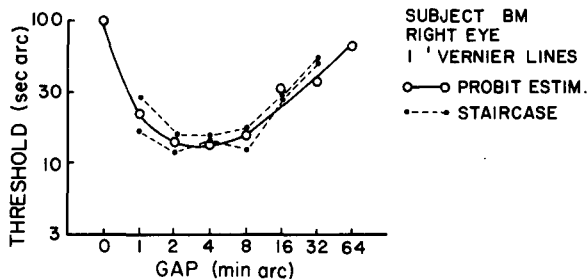


Figure 3. Comparison of results obtained with two procedures. Open circles—method of constant stimuli, 200 trials/point, thresholds obtained from probit analysis. Closed circles—efficient staircase procedure, two repetitions.

and begins another adjustment session. An average of 10 to 20 settings gives a reliable estimate of the observer's vernier acuity threshold.

In practice, we believe that some combination of these techniques will be required in the clinical atmosphere in order to meet the needs of a variety of patients and their individual preferences.

**Step 6. Written instructions.** The instruction set included with a vernier test might be presented as follows: "When we begin, two short, bright lines, one above the other, will appear on the screen in front of you. The upper line is located either slightly towards me (your right) or away from me (your left). You must indicate to which side you think the upper line is located by pushing one of the two buttons you hold in your hand. If you cannot tell for sure which side to choose, just make your best guess. The computer will indicate whether your judgment was correct by presenting an appropriate message on the screen. After you make your choice and press one of the buttons, the screen will go blank for a few seconds. Then a new display will appear and you must make another choice. This sequence will be repeated 15 to 25 times."

**Step 7. Initial clinical trials.** Our progress on this hyperacuity paradigm is currently at the stage of initiating clinical trials. We must select patients with early pathology or anomaly of known type which has a known site of action, in order to assess the effects of this pathology or anomaly on measured hyperacuity functions. At this step, we must evaluate our initial choices of test parameters and modify these choices if the clinical results so dictate.

**Step 8. Correlation with established tests.** For the case of hyperacuity as part of a layer-by-layer test battery, there are no established analogous tests. We are essentially breaking new ground. What we must do, however, is to relate changes in hyperacuity functions to losses determined in other visual functions that are localized to lower levels within the visual pathway. Only in this way can lower level anomalies be distinguished from centrally localized anomalous functions using the hyperacuity test. This is often termed upstream-downstream analysis; that is, how do alterations on distal responses alter more proximal (CNS) responses, and vice versa?

**Step 9. Clinical trials.** Of course, this stage depends on the success of all previous stages. Relevant parameters must be completely defined, clinical testing techniques must be efficient and reliable, and the format for presentation of results must have been standardized.

### Hyperacuity as an Indicator of Visual System Integrity in the Presence of Ocular Opacities

A second clinical application of the hyperacuity paradigm is indicated by the observation that cer-

tain stimulus configurations can be significantly blurred with only minimal effect on the observer's performance (Stigmar, 1971; Westheimer, 1979). In light of the difficulty of predicting postsurgical visual function in cataract and other ocular opacity cases, we are proceeding in parallel with the above work on this second clinical application of hyperacuity tests. Our results suggest that the detrimental effects of reducing the high spatial frequency information in the retinal image are highly dependent on the spatial configuration of the stimulus.

The development of this second application of hyperacuity tests has involved preliminary work that is different from that described above. However, the same orderly development procedure can be followed.

Steps 1, 2, and 3 were essentially combined in order to quantify the dependence of hyperacuity responses on stimulus integrity. We chose ground-glass blur as a means to mimic the possible effects of media opacities on the retinal image. We were able to systematically reduce the high spatial-frequency information in the stimuli by placing a ground-glass plate between the CRT screen and the observer. When the ground glass is placed close to the CRT, mainly high frequencies are affected. As the CRT-to-glass distance is increased, the stimulus becomes progressively more blurred and the reduction of high spatial-frequency information extends also to lower frequencies (Stigmar, 1971; Westheimer & McKee, 1980).

The effects of this ground-glass blur on the hyperacuity responses of one normal subject are shown in Figure 4. Results that were obtained with 1-min-arc vernier lines (Figure 4a), 10-min-arc vernier lines (Figure 4b), and the sloping line stimulus (Figure 4c) are shown. The detrimental effect of this type of stimulus degradation is highly dependent on the distance separating the two features whose relative localization is to be judged. Short 1-min-arc vernier lines, separated by a small vertical gap (less than 8 min arc) and short sloping lines (less than 32 min arc) are most seriously affected by ground-glass blur. However, selected stimulus configurations are very moderately, if at all, affected. For example, thresholds for vernier lines (either 1 or 10 min arc in length) separated by a 16- to 32-min-arc vertical gap merely double (from about 12 to 24 sec arc) when the stimulus line half-width is blurred from 1 to over 20 min arc. It is remarkable that when the line profile of a vernier target is over 20 min arc wide at midheight, displacements of only 22 sec arc can still be reliably detected by the normal observer.

These results suggest that specific hyperacuity tasks are resistant to substantial degradation of high spatial-frequency information in the stimulus. One might ask whether displacement thresholds on the

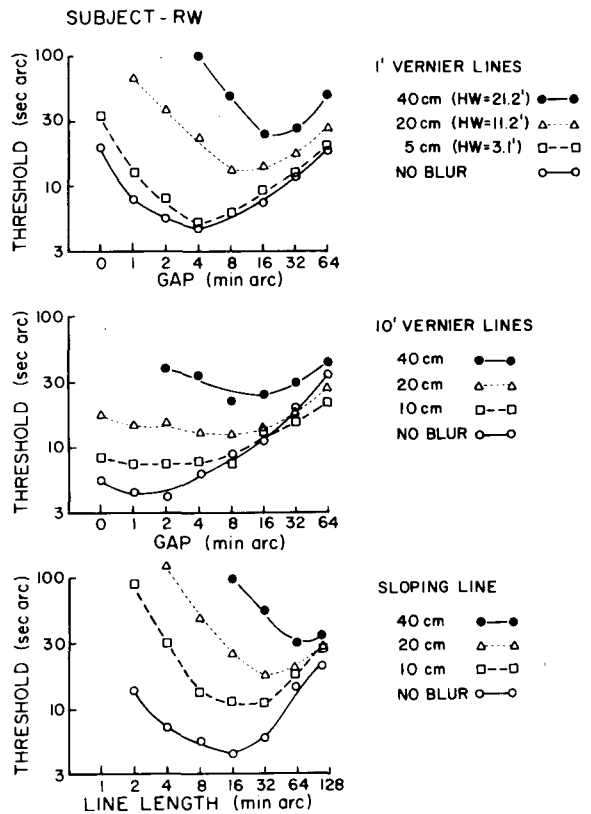


Figure 4. Hyperacuity responses from a single subject with blurred retinal images. A ground-glass screen placed at variable distances from the CRT uniformly spreads the luminance distribution of the stimulus lines. In each figure, results were obtained with no blur and with the ground glass placed at 5, 10, 20, or 40 cm from the CRT. At these distances, the half-widths (HW) of each stimulus line ranged from less than 1 min arc (no blur) to 3.1 min arc (5-cm CRT-to-glass distance), to 11.2 min arc (20-cm CRT-to-glass distance), to 21.2 min arc (40-cm CRT-to-glass distance). Data are shown for 1-min arc lines, 10-min-arc lines, and the tilting line stimulus.

order of 30 sec arc (0.5 min arc) can still be defined as hyperacuity responses. The task with blurred stimuli still involves relative localization of two light distributions as opposed to resolution, and, in this sense, a hyperacuity mechanism is being tapped.

Results obtained from a second subject using ground-glass blur are consistent with those shown in Figure 4. We have also obtained preliminary results using lens blur. Ground-glass blur mainly reduces the high spatial-frequency content of the image. The effects of lens blur are more complicated, producing both amplitude and phase changes in the high-frequency region of the stimulus Fourier spectrum (Westheimer & McKee, 1980). In these experiments, a 3-mm artificial pupil is placed in front of the subject's eye while the other eye is occluded. Defocusing the well-corrected eye by 1.5 diopters has about the same effect on vernier thresh-

olds with 10-min-arc-long lines as the ground-glass blur with the glass placed 20 cm from the CRT. Although lens blur (simulating uncorrected refractive error) can be qualitatively compared with ground-glass blur, their effects on the retinal image are very different. These different physical effects, however, seem to produce similar effects on hyperacuity performance. If spectacle (lens) blur produces high-frequency noise (striae, etc.), a superior result may be obtained by using a lower frequency display, that is, using some ground-glass blur!

If parameter selection were limited to the data sets in Figure 4, for the display used in Figure 4a, gaps of 32 or 64 min arc would be selected. For the stimulus display in Figure 4b, gaps of 16-64 min arc might be used; for the display in Figure 4c, 128 min arc would be best. In each case, threshold is minimally altered for these values.

Calibration techniques (Step 4), test efficiency (Step 5), and patient instructions (Step 6) would be equivalent to those discussed in relation to the layer-by-layer test battery. Initial trials (Step 7) will be run on selected cataract patients both before and after surgery, so that initial results with this new test can be correlated (Step 8) with results of established tests (VEP, interference acuity).

In summary, we seek a specific subset of hyperacuity configurations which are relatively resistant to our imposed image degradations. It is likely that this subset will also be resistant to real ocular opacities. This statement will be tested in extensive clinical trials. We have already shown that lens blur and ground-glass blur, although having different retinal image effects, have analogous effects on hyperacuity responses. The data in Figure 4 are suggestive of at least one class of hyperacuity configuration that possesses this resistance to image degradation. It seems that as long as the features of the stimulus whose relative localization is to be judged are separated by distances approximately the same as the line-profile half-widths of the blurred stimuli, the hyperacuity performance is only slightly affected. A tentative conclusion might be that the appropriate stimulus configuration should be chosen with regard to the density or severity of the ocular opacity. In general, short 1-min-arc vernier lines separated by 32 min arc or more, or a sloping line that is 64 min arc or more long, provide remarkably precise hyperacuity responses in the presence of severe image degradation. Of course, if points or lines are separated by larger distances, or the length of the single line is further increased, non-central foveal responses might be sampled.

One final point: In some individuals who exhibit modest refractive error there are superimposed distracting high-frequency, spuriously resolved patterns of stria or other forms of noise. In these in-

dividuals, performance is greatly increased by providing modest ground-glass blur which effectively removes high-frequency spurious details from the stimulus to be evaluated.

## CONCLUSIONS

We seek to encourage those psychophysicists who are considering entrance into the "arena" of clinical testing. Here, using a series of three different examples, we have tried to introduce the reader to some of the complexities of the test development task and to a rational approach to the analysis of a problem. The task is neither simple nor overwhelming, but, rather, quite orderly. There are aspects of this sort of research which are not easy and not quickly rewarding. Specific pathological conditions at the right stage of development are sometimes difficult to obtain. Often one has to wait for just the right case to be able to ask the right questions, and the equipment must be ready to serve the needs of the investigation at the critical time. In essence, the task is taken in small steps, each sometimes separated by substantial time periods. It is therefore the psychophysicist's responsibility to bring the research together into a logical and orderly package.

Here we have used two evolving hyperacuity tests to demonstrate how a psychophysical test may be developed for use in a clinical environment. This is a challenging area for the psychophysicist—one which is opening up rapidly and one which can be serviced by the psychologist/clinician. The clinicians do need, indeed are asking for, the help. This has been declared an area of high program relevance by the National Advisory Eye Council of the National Eye Institute, NIH. When one sees the relative paucity of information provided by routine testing, the virtual lack of application of modern developments in making clinical judgments, the inadequate training in many pertinent areas among clinicians, and the vast opportunities available to the scientist, one cannot fail to be impressed. We encourage the establishment of such efforts within the scientific community, and are willing to help those who are encountering difficulties.

## REFERENCES

- AGUILAR, M., & STILES, W. S. Saturation of the rod mechanism of the retina at high levels of stimulation. *Optica Acta*, 1954, 1, 59-65.
- ANDERSEN, E. E., & WEYMOUTH, F. W. Visual perception and the retinal mosaic. I. Retinal mean local sign—An explanation of the fineness of binocular perception of distance. *American Journal of Physiology*, 1923, 64, 561-594.
- AULHORN, E., & HARMS, H. Visual perimetry. In D. Jameson & L. M. Hurvich (Eds.), *Handbook of sensory physiology*



- (Vol. VII/4): *Visual psychophysics*. New York: Springer-Verlag, 1972.
- BERRY, R. N. Quantitative relations among vernier, real depth, and stereoscopic depth acuities. *Journal of Experimental Psychology*, 1948, **38**, 708-721.
- BEST, F. Über die Grenze der Erkennbarkeit von Lagenunterscheiden. *A. von Graefes Archiv für Ophthalmologie*, 1900, **51**, 453-460.
- CORNSWEET, T. The staircase method in psychophysics. *American Journal of Psychology*, 1962, **75**, 485-491.
- ENOCH, J. M. Physiology. In A. Sorsby (Ed.), *Modern ophthalmology* (Vol. 1) *Basic aspects* (1st ed.). London: Butterworths, 1963.
- ENOCH, J. M. (Ed.). *Perimetric standards and perimetric glossary of the International Council of Ophthalmology*. The Hague: Junk, 1979.
- ENOCH, J. M. First interprofessional standard for visual field testing: Report of Working Group 39, *Advances in Ophthalmology*, 1980, **40**, 173-224.
- ENOCH, J. M., FITZGERALD, C. R., & CAMPOS, E. C. *Quantitative layer-by-layer perimetry: An extended analysis*. New York: Grune & Stratton, 1981.
- FINNEY, D. J. *Probit analysis* (3rd ed.). Cambridge: Cambridge University Press, 1971.
- STIGMAR, G. Blurred visual stimuli. II. The effect of blurred visual stimuli on vernier and stereo acuity. *Acta Ophthalmologica*, 1971, **49**, 364-379.
- WESTHEIMER, G. Visual acuity and hyperacuity. *Investigative Ophthalmology*, 1975, **14**, 570.
- WESTHEIMER, G. The spatial sense of the eye. *Investigative Ophthalmology and Visual Science*, 1979, **18**, 893-912.
- WESTHEIMER, G. The spatial grain of the parafoveal visual field. *Vision Research*, 1982, **22**, 157.
- WESTHEIMER, G., & HAUSKE, G. Temporal and spatial interference with vernier acuity. *Vision Research*, 1975, **15**, 1137-1141.
- WESTHEIMER, G., & MCKEE, S. P. Spatial configurations for visual hyperacuity. *Vision Research*, 1977, **17**, 941-947.
- WESTHEIMER, G., & MCKEE, S. P. Stereoscopic acuity for moving retinal images. *Journal of the Optical Society of America*, 1978, **68**, 450.
- WESTHEIMER, G., & MCKEE, S. P. Stereoscopic acuity with defocused and spatially filtered retinal images. *Journal of the Optical Society of America*, 1980, **70**, 772.
- WETHERILL, G. B., & LEVITT, H. Sequential estimation of points on a psychometric function. *British Journal of Mathematical and Statistical Psychology*, 1965, **18**, 1-10.

(Manuscript received June 21, 1982;  
revision accepted for publication September 9, 1982.)