standing of fracture phenomena, recently reviewed (*6*), goes back to investigations by Griffith in 1920–21 (*7*). Our understanding of deformation characteristics is based on our knowledge of imperfections in crystals, called dislocations. Phenomena relating to dislocations have been intensively studied in recent years, but the essential idea was introduced by Orowan, Taylor, and others in the 1930's. On the basis of data on the mechanical behavior of pure ice and our current level of understanding of the mechanical properties of materials, it has been possible to predict the general behavior of potential ice-alloy systems and to concentrate research activity in areas suggested by our understanding of phenomena rather than by intuitive inventiveness.

Application of materials science and technology to a new system such as ice emphasizes the disparity between concepts limited by discipline (glaciology, metallurgy, ceramics, and so on) and those in which there are no artificial, disciplinary barriers. This is a case where exchange of information between different scientific disciplines has proved to be useful and effective (*9*).

References and Notes

1. E. Moser (Naval Civil Engineering Laboratory), private communication.
2. W. D. Kingery, *J. Glaciol.* **3**, 577 (1960).
3. M. F. Perutz, *ibid.* **1**, 95 (1948).
4. A. H. Cottrell, *Dislocations and Plastic Flow in Crystals* (Clarendon, Oxford, England, 1953).
5. W. T. Read, Jr., *Dislocations in Crystals* (McGraw-Hill, New York, 1953).
6. B. L. Averbach, D. K. Felbeck, G. T. Hahn, D. A. Thomas, Eds., *Fracture* (Technology Press of Massachusetts, Cambridge; Wiley, New York; 1960).
7. A. A. Griffith, *Phil. Trans. Roy. Soc. London* **A221**, 163 (1920–21).
8. M. L. Kronberg, *Acta Met.* **5**, 507 (1957).

9. This study was carried out as part of a program at the Massachusetts Institute of Technology sponsored by the Terrestrial Sciences Laboratory, Geophysics Research Directorate, Air Force Cambridge Research Laboratory, under contract No. AF 19(604)-5994.

Is There a Sensory Threshold?

When the effects of the observer's response criterion are isolated, a sensory limitation is not evident.

John A. Swets

One hundred years ago, at the inception of an experimental psychology of the senses, G. T. Fechner focused attention on the concept of a sensory threshold, a limit on sensitivity. His *Elemente der Psychophysik* described three methods—the methods of adjustment, of limits, and of constants—for estimating the threshold value of a stimulus (*1*). The concept and the methods have been in active service since. Students of sensory processes have continued to measure the energy required for a stimulus to be just detectable, or the difference between two stimuli necessary for the two to be just noticeably different. Very recently there has arisen reasonable doubt that sensory thresholds exist.

The threshold thought to be characteristic of sensory systems has been regarded in the root sense of that word as a barrier that must be overcome. It is analogous to the threshold discovered by physiologists in single neurons. Just as a nervous impulse either occurs or does not occur, so it has been thought that when a weak stimulus is presented we either detect it or we do not, with no shades in between. The analogy with the neuron's all-or-none action, of course, was never meant to be complete; it was plain that at some point above the threshold sensations come in various sizes.

From the start the triggering mechanism of the sensory systems was regarded as inherently unstable. The first experiments disclosed that a given stimulus did not produce a consistent "yes" ("I detect it") response or a consistent "no" ("I do not detect it") response. Plots of the "psychometric function"—the proportion of "yes" responses as a function of the stimulus energy—were in the form of ogives, which suggested an underlying bell-shaped distribution of threshold levels. Abundant evidence for continuous physiological change in large numbers of receptive and nervous elements in the various sensory systems made this picture eminently reasonable. Thus, the threshold value of a stimulus had to be specified in statistical terms. Fechner's experimental methods were designed to obtain good estimates of the mean and the variance of the threshold distribution.

It was also assumed from the beginning that the observer's attitude affects the threshold estimate. The use of ascending and descending series of stimulus energies in the method of limits, to take one example, is intended to counterbalance the errors of "habituation" and "anticipation"—errors to which the observer is subject for extrasensory reasons. Typically, investigators have not been satisfied with experimental observers who were merely well motivated; they have felt the need for elite observers. They have attempted, by selection or training, to obtain observers who could maintain a reasonably constant criterion for a "yes" response.

The classical methods for measuring the threshold, however, do not provide a measure of the observer's response criterion that is independent of the threshold measure. As an example, we may note that a difference between two threshold estimates obtained with the method of limits can be attributed to a criterion change only if it is assumed that sensitivity has remained constant, or to a sensitivity change only if it is assumed that the criterion has remained constant. So, although the observer's response criterion affects the estimate of the threshold, the classical procedures do not permit calibration of the observer with respect to his response criterion.

Within the past ten years methods

The author is associate professor of psychology and a staff member of the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge.

This article is adapted from an address delivered at a centennial symposium honoring Fechner, sponsored by the American Psychological Association and the Psychometric Society, held in Chicago in September 1960.

have become available that provide a reliable, quantitative specification of the response criterion. These methods permit isolation of the effects of the criterion, so that a relatively pure measure of sensitivity remains. Interestingly, the data collected with these methods give us good reason to question the existence of sensory thresholds, to wonder whether anything more than a response criterion is involved in the dichotomy of "yes" and "no" responses. There is now reason to believe that sensory excitation varies continuously and that an apparent threshold cut in the continuum results simply from restricting the observer to two categories of response.

The methods that permit separating the criterion and sensitivity measures, and a psychophysical theory that incorporates the results obtained with these methods, stem directly from the modern approach taken by engineers to the general problem of signal detection. The psychophysical "detection theory," like the more general theory, has two parts. One part is a literal translation of the theory of testing statistical hypotheses, or statistical decision theory. It is this part of the theory that provides a solution to the criterion estimation problem and deals with sensitivity as a continuous variable. The second part is a theory of ideal observers. It specifies the mathematically ideal detection performance—the upper limit on detection performance that is imposed by the environment—in terms of measurable parameters of the signal and of the masking noise (2).

We shall turn in a moment to a description of the theory and to samples of the supporting data. Before proceeding any further, however, we must note that, although Fechner started the study of sensory functions along lines we are now questioning, he also anticipated the present line of attack in both of its major aspects. For one thing, he regarded Bernoulli's ideas on statistical decision as highly relevant to psychophysical theory (3). More important, while advancing the concept of a threshold, he spoke also of what he called "negative sensations"—that is, of a grading of sensory excitation below the threshold. That subsequent workers in the field of psychophysics have shown little interest in negative sensations is apparent from the fact that, 75 years after Fechner's work, Boring could write: "So also a sensation either occurs from stimulation or it does not. If it does not, it has no demonstrable intensity. Fechner talked about negative (subliminal) degrees of intensity, but that is not good psychology today. Above the limen we can sense degrees of intensity, but introspection cannot directly measure these degrees. We are forced to comparison, and there again we meet an all-or-none principle. Either we can observe a difference or we cannot. Introspection as to the amount of difference is not quantitatively reliable" (4).

Decision Aspects of Signal Detection

How detection theory succeeds in estimating the response criterion may be described in terms of "the fundamental detection problem." The experimenter defines an interval of time for the observer, and the observer must decide whether or not a signal is present during the interval. It is assumed that every interval contains some random interference, or noise—noise that is inherent in the environment, or is produced inadvertently by the experimenter's equipment for generating signals, or is deliberately introduced by the experimenter, or is simply a property of the sensory system. Some intervals contain a specified signal in addition to the background of noise. The observer's report is limited to these two classes of stimulus events—he says either "yes" (a signal was present) or "no" (only noise was present). Note that he does not say whether or not he *saw* (or *heard*) the signal; he says whether, under the particular circumstances, he prefers the decision that it was present or the decision that it was absent.

There is presumably, coinciding with the observation interval, some neural activity in the relevant sensory system. This activity forms the sensory basis—a part of the total basis—for the observer's report. This "sensory excitation," as we shall call it, may be in fact either simple or complex; it may have many dimensions or few; it may be qualitative or quantitative; it may be anything. The exact, or even the general, nature of the actual sensory excitation is of no concern to the application of the theory.

Only two assumptions are made about the sensory excitation. One is that it is continually varying; because of the ever-present noise, it varies over time in the absence of any signal, as well as from one presentation to the next of what is nominally the same signal. The other is that the sensory excitation, insofar as it affects the observer's report, may be represented as a unidimensional variable. In theory, the observer is aware of the probability that each possible excitatory state will occur during an observation interval containing noise alone and also during an observation interval containing a signal in addition to the noise, and he bases his report on the ratio of these two quantities, the likelihood ratio. The likelihood ratio derived from any observation interval is a real, nonzero number and hence may be represented along a single dimension.

The likelihood-ratio criterion. The observer's report after an observation interval is supposed to depend upon whether or not the likelihood ratio measured in that interval exceeds some critical value of the likelihood ratio, a response criterion. The criterion is presumed to be established by the observer in accordance with his detection goal and the relevant situational parameters. If he wishes to maximize the number of correct responses, his criterion will depend upon the a priori probability that a signal will occur in a given interval. If he chooses to maximize the total payoff, his criterion will depend on this probability and also on the values and costs associated with the four possible outcomes of a decision. Several other detection goals can be defined; the way in which each of them determines the criterion has been described elsewhere (5). In any case, the criterion employed by the observer can be expressed as a value of the likelihood ratio. Thus, the observer's decision about an interval is based not only on the sensory information he obtains in that interval but also upon advance information of various kinds and upon his motivation.

Next, consider a probability defined on the variable likelihood ratio—in particular, the probability that each value of likelihood ratio will occur with each of the classes of possible stimulus events: noise alone and signal plus noise. There are, then, two probability distributions. The one associated with signal plus noise will have a greater mean (indeed, its mean is assumed to increase monotonically with increases in the signal strength, but for the moment we are considering a particular signal). Now, if the observer follows the procedure we have described—that is, if he reports that the signal is present whenever the likelihood ratio exceeds a certain criterion and that noise alone is present whenever the likelihood

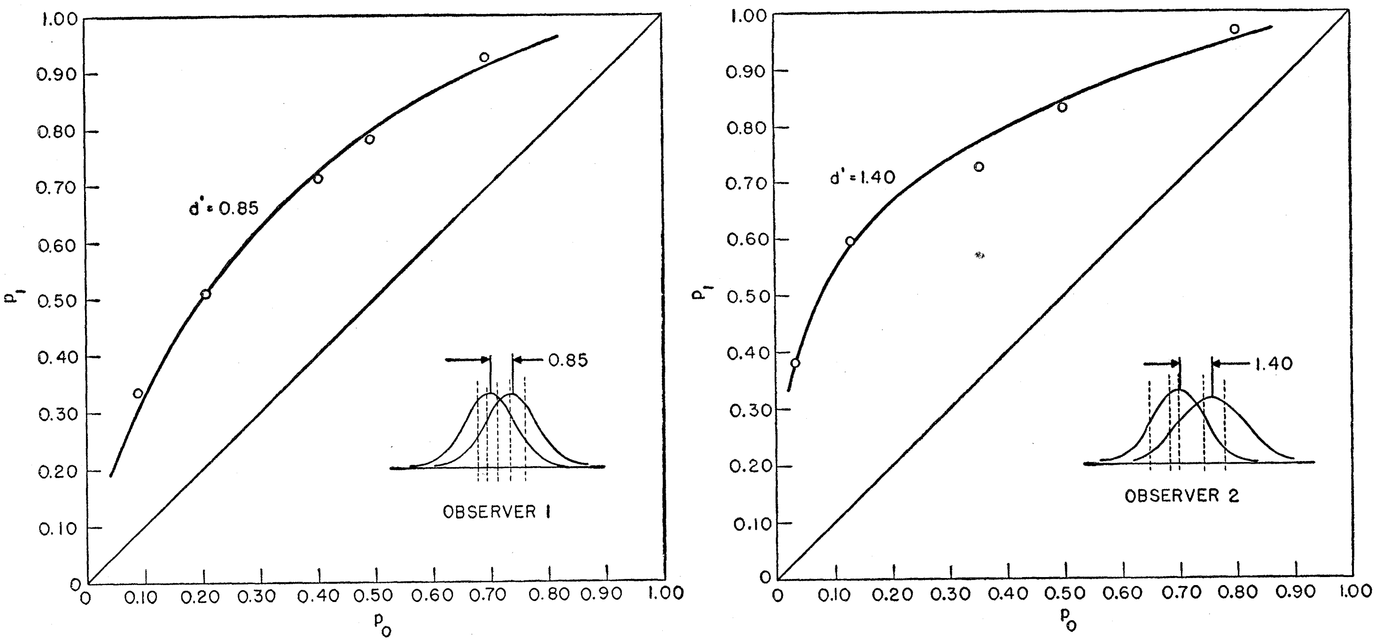

Fig. 1. Two theoretical operating-characteristic curves, with data from a yes-no experiment.

ratio is less than this criterion—then, from the fourfold stimulus-response matrix that results, one can extract two independent measures: a measure of the observer's response criterion and a measure of his sensitivity.

The operating characteristic. The extraction of these two measures depends upon an analysis in terms of the operating characteristic. If we induce the observer to change his criterion from one set of trials to another, and if, for each criterion, we plot the proportion of "yes" reports made when the signal is present (the proportion of hits, or p_1) against the proportion of "yes" reports made when noise alone is present (the proportion of false alarms, or p_0), then, as the criterion varies, a single curve is traced (running from 0 to 1.0 on both coordinates) that shows the proportion of hits to be a nondecreasing function of the proportion of false alarms. This operating-characteristic curve describes completely the successive stimulus-response matrices that are obtained, since the complements of these two proportions are the proportions that belong in the other two cells of the matrix. The particular curve generated in this way depends upon the signal and noise parameters and upon the observer's sensitivity; the point on this curve that corresponds to any given stimulus-response matrix represents the criterion employed by the observer in producing that matrix.

It has been found that, to a good approximation, the operating-characteristic curves produced by human observers correspond to theoretical curves based on normal probability distributions. These curves can be characterized by a single parameter: the difference between the means of the signal-plus-noise and noise-alone distributions divided by the standard deviation of the noise distribution. This parameter has been called d'. Moreover, the slope of the curve at any point is equal to the value of the likelihood-ratio criterion that produces that point.

The yes-no experiment. The procedure employed in the fundamental detection problem is often referred to as the "yes-no procedure," and we shall adopt this terminology. Two operating-characteristic curves resulting from this procedure are shown in Fig. 1. The

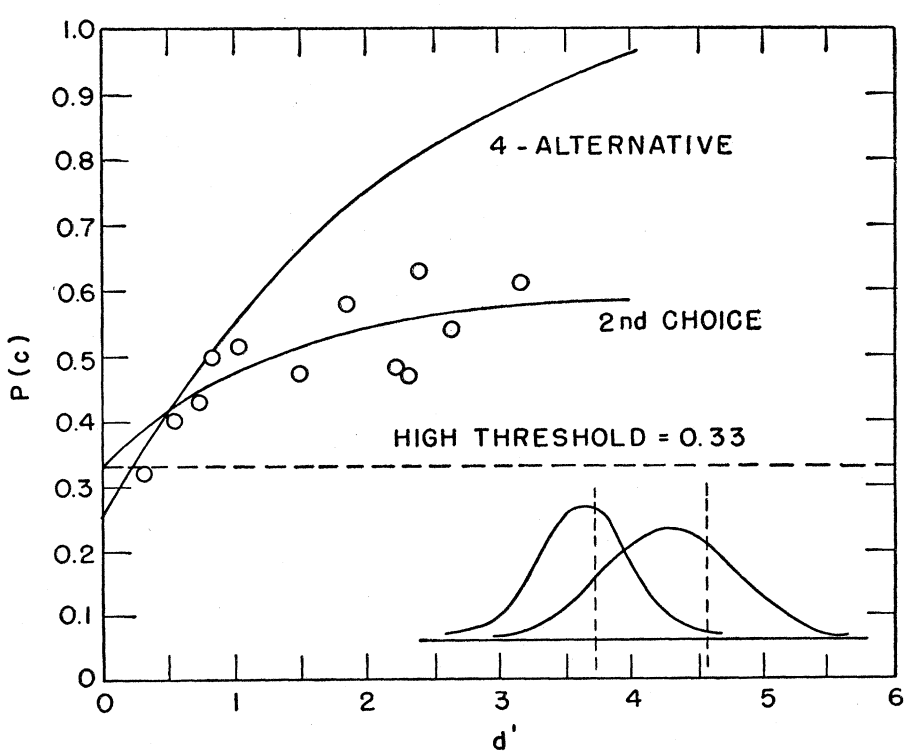

Fig. 2. The results obtained in a second-choice experiment, shown with the prediction from detection theory. [Data from J. A. Swets, W. P. Tanner, Jr., T. G. Birdsall (*5*)]

data points were obtained in an auditory experiment in which the observers attempted to detect a tone burst in a background of white noise. The curves are the theoretical curves that fit the data best. The inserts at lower right in the two graphs show the normal probability distributions underlying the curves, and the five criteria corresponding to the data points. In this particular experiment the observers changed their criteria from one set of trials to another as the experimenter changed the a priori probability of the occurrence of the signal. The distance between the means of the two distributions is shown as 0.85 for observer No. 1 and as 1.40 for observer No. 2; this distance is equal to d' under the convention that the standard deviation of the noise distribution is unity.

We may note that the curve fitted to the data of the first observer is symmetrical about the negative diagonal, and that the curve fitted to the data of the second observer is not. Both types of curves are seen frequently; the second curve is especially characteristic of data collected in visual experiments. Theoretically, the curve shown in the graph at left will result if the observer knows the signal exactly—that is, if he knows its frequency, amplitude, starting time, duration, and phase. A theoretical curve like the one shown in the graph at right results if the observer has inadequate information about frequency and phase, or, as is the case when the signal is a white light, if there is no frequency and phase information. The probability distributions that are shown in the inserts reflect this difference between the operating-characteristic curves.

Both of the curves shown are based on the assumption that sensory excitation is continuous, that the observer can order values of sensory excitation throughout its range. Two other experiments have been employed to test the validity of this assumption: one involves a variant of the forced-choice procedure; the other involves a rating procedure. We shall consider these experiments in turn.

The second-choice experiment. In the forced-choice procedure, four temporal intervals were defined on each trial, exactly one of which contained the signal. The signal was a small spot of light projected briefly on a large, uniformly illuminated background. Ordinarily, the observer simply chooses the interval he believes most likely to have contained the signal. In this experiment the observer made a second choice as well as a first.

The results are shown in Fig. 2. The top curve is the theoretical function relating the proportion of correct first choices to d'; the lower curve is the theoretical relation of the proportion of correct second choices to d'. The points on the graph represent the proportions of correct second choices obtained by experiment. They are plotted at the value of d' corresponding to the observed proportion of correct first choices.

It may be seen that the data points are fitted well by the theoretical curve. The rather considerable variability can be attributed to the fact that each point is based on less than 100 observations. In spite of the variability, it is clear that the points deviate significantly from the horizontal dashed line. The dashed line may be taken as a baseline; it assumes a sensory threshold such that it is exceeded on only a negligible proportion of the trials when noise alone is presented. Should such a threshold exist, the second choice would be correct only by chance. The data indicate that the observer is capable of ordering values of sensory excitation well below this point. Two sensory thresholds are shown in the insert at lower right in Fig. 2. The threshold on the right, at three standard deviations from the mean of the noise distribution, corresponds to the horizontal dashed line in the upper part of the figure. The data indicate that, were a threshold to exist, it would have to be at least as low as the left-hand threshold, at approximately the mean of the noise distribution.

The rating experiment. In the rating procedure, as in the yes-no procedure, a signal is either presented or not presented in a single observation interval. The observer's task is to reflect gradations in the sensory excitation by assigning each observation to one of several categories of likelihood of occurrence of a signal in the interval.

The results of a visual experiment are displayed in Fig. 3. The abscissa represents a six-point scale of certainty concerning the occurrence of a signal. The six categories were also defined in terms of the a posteriori probability of occurrence, but, for our purpose, only the property of order need be assumed. The ordinate shows the proportion, of the observations placed in each category, that resulted from the presentation of the signal.

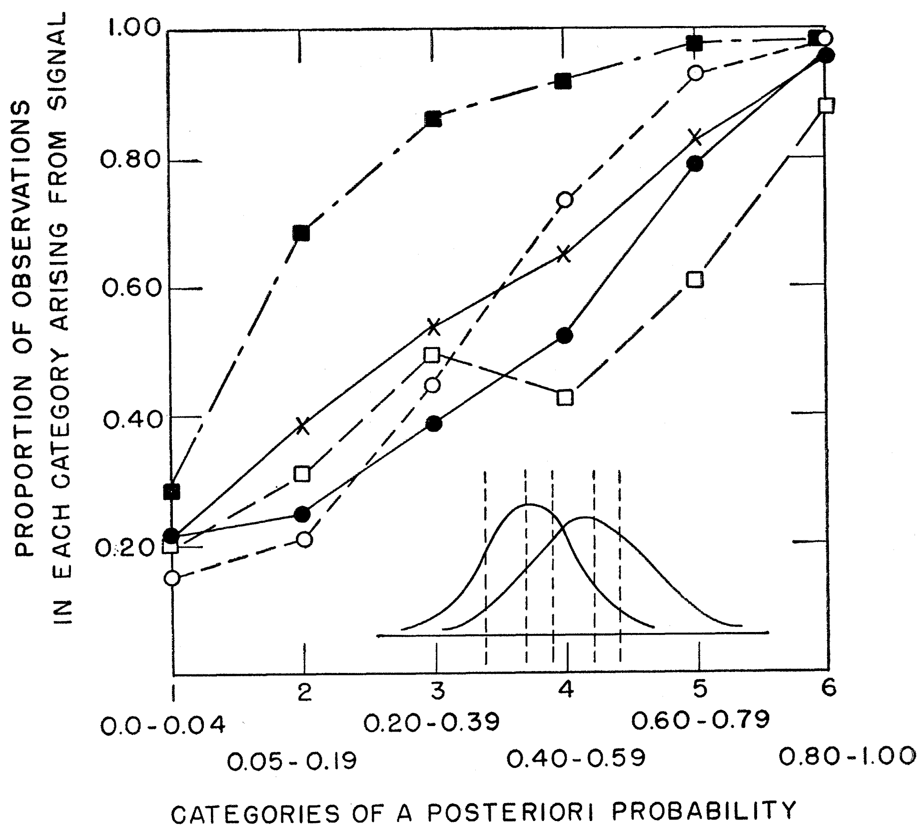

Fig. 3. The results of a rating experiment. [Data from J. A. Swets, W. P. Tanner, Jr., T. G. Birdsall (5)]

Five curves are shown in Fig. 3. Four of them correspond to the four observers; the fifth, marked by ×'s, represents the average. It may be seen that the curves for three of the four observers increase monotonically, while that for the fourth has a single reversal. The implication is that the human observer can distinguish at least six categories of sensory excitation.

It is possible to compute operating-characteristic curves from these data, by regarding the category boundaries successively as criteria. The curves (not shown here) are very similar in appearance to those obtained with the yes-no procedure (*5*). By way of illustration, the five criteria used by one of the observers (the one represented by solid circles) are shown in the insert at lower right in Fig. 3.

The experimental invariance of d'. It has been found experimentally, in vision (*5*) and in audition (*6*), that the sensitivity measure d' remains relatively constant with changes in the response criterion. Thus, detection theory provides a measure of sensitivity that is practically uncontaminated by the factors that might be expected to affect the observer's attitude.

It has also been found that the measure d' remains relatively invariant with different experimental procedures. For vision (*7*) and audition (*8*) the estimates of d' from the yes-no procedure and from the four-interval, forced-choice procedure are very nearly the same. Again, consistent estimates are obtained from forced-choice procedures with 2, 3, 4, 6, and 8 intervals (*8*). Finally, the rating procedure yields estimates of d' indistinguishable from those obtained with the yes-no procedure (*9*).

Thus, the psychophysical detection theory has passed some rather severe tests—the quantity that is supposed to remain invariant does remain invariant. This finding may be contrasted with the well-known fact that estimates of the threshold depend heavily on the particular procedure used.

Theory of Ideal Observers

Detection theory states, for several types of signal and noise, the maximum possible detectability as a function of the parameters of the signal and the noise. Given certain assumptions, this relationship can be stated very precisely. The case of the "signal specified exactly" (in which everything about the signal is known, including its frequency, phase, starting time, duration, and amplitude) appears to be a useful standard in audition experiments. In this case, the maximum d' is equal to the quantity $(2E/N_0)^{\frac{1}{2}}$, in which E is the signal energy and N_0 is the noise power in a one-cycle band. An ideal observer for visual signals has also been defined (*10*).

It can be argued that a theory of ideal performance is a good starting point in working toward a descriptive theory. Ideal theories involve few variables, and these are simply described. Experiments can be used to uncover whatever additional variables may be needed to describe the performance of real observers. Alternatively, experiments can be used to indicate how the ideal theory may be degraded—that is, to identify those functions of which the ideal detection device must be deprived—in order to accurately describe real behavior.

Given a normative theory, it is possible to describe the real observer's efficiency. In the present instance, the efficiency measure η has been defined as the ratio of the observed to the ideal $(d')^2$. It seems likely that substantive problems will be illuminated by the computation of η for different types of signals and for different parameters of a given type of signal. The observed variation of this measure should be helpful in determining the range over which the human observer can adjust the parameters of his sensory system to match different signal parameters (he is, after all, quite proficient in detecting a surprisingly large number of different signals), and in determining which parameters of a signal the observer is not using, or not using precisely, in his detection process (*11*).

The human observer, of course, performs less well than does the ideal observer in the great majority of detection tasks, if not in all. The interesting question concerns not the amount but the nature of the discrepancy that is observed.

The human observer performs less well than the ideal observer defined for the case of the "signal specified exactly." That is to say, the human observer's psychometric function is shifted to the right. More important, the slope of the human observer's function is greater than that of the ideal function for this particular case—a result sometimes referred to as "low-signal suppression." Let us consider three possible reasons for these discrepancies.

First, the human observer may well have a noisy decision process, whereas the ideal decision process is noiseless. For example, the human observer's response criterion may be unstable. If he vacillates between two criteria, the resulting point on his operating-characteristic curve will be on a straight line connecting the points corresponding to the two criteria; this average point falls below the curve (a curve with smoothly decreasing slope) on which the two criteria are located. Again, the observer's decision axis may not be continuous. It may be, as far as we know, divided into a relatively small number of categories—say, into seven.

A second likely cause of deviation from the ideal is the noise inherent in the human sensory systems. Consistent results are obtained from estimating the amount of "internal noise" (that is, noise in the decision process and noise in the sensory system) in two ways: by examining the decisions of an observer over several presentations of the same signal and noise (on tape) and by examining the correlation among the responses of several observers to a single presentation (*12*).

A third, and favored, possibility is faulty memory. This explanation is favored because it accounts not only for the shift of the human observer's psychometric function but also for the greater slope of his function. The reasoning proceeds as follows: If the detection process involves some sort of tuning of the receptive apparatus, and if the observer's memory of the characteristics of the incoming signal is faulty, then the observer is essentially confronted with a signal not specified exactly but specified only statistically. He has some uncertainty about the incoming signal.

If uncertainty is introduced into the calculations of the psychometric function of the ideal detector, it is found that performance falls off as uncertainty increases, and that this decline in performance is greater for weak signals than for strong ones (*13*). That is, a family of theoretical uncertainty curves shows progressively steeper slopes coinciding with progressive shifts to the right. This is what one would expect; the accuracy of knowledge about signal characteristics is less critical for strong signals, since strong signals carry with them more information about these characteristics.

It has been observed that visual data (*10*) and auditory data (*14*) are fitted well, with respect to slope, by the theoretical curve that corresponds to uncertainty among approximately 100 orthogonal signal alternatives. It is not difficult to imagine that the product of the uncertainties about the time, location, and frequency of the signals used in these experiments could be as high as 100.

It is possible to obtain empirical corroboration of this theoretical analysis of uncertainty in terms of faulty memory. This is achieved by providing various aids to memory within the experimental procedure. In such experiments, memory for frequency is made unnecessary by introducing a continuous tone or light (a "carrier") of the same frequency as the signal, so that the signal to be detected is an increment in the carrier. This procedure also eliminates the need for phase memory in audition and location memory in vision. In further experiments a pulsed carrier is used in order to make unnecessary memory for starting time and for duration. In all of these experiments a forced-choice procedure is used, so that memory for amplitude beyond a single trial can also be considered irrelevant. In this way, all of the information thought to be relevant may be contained in the immediate situation. Experimentally, we find that the human observer's psychometric functions show progressively flatter slopes as more and more memory aids are introduced. In fact, when all of the aids mentioned above are used, the observer's slope parallels that for the ideal observer without uncertainty, and it deviates as little as 3 decibels from the ideal curve in absolute value (*14*).

Relationship of the Data to Various Threshold Theories

Although there is a limit on detection performance, even ideally, and although the human observer falls short of the limit, these facts do not imply a sensory threshold. We have just seen that the human observer's performance can be analyzed in terms of memory, and, conceivably, additional memory aids could bring his performance closer to the ideal. Moreover, consideration of ideal observers concerns an upper rather than a lower limit. The human observer, while falling short of the ideal, can still detect signals at a high rate. Ideally, any displacement of the signal-plus-noise distribution from the noise-alone distribution will lead to a detection rate greater than chance. Although it is difficult to obtain data near the chance point, the theoretical curves that fit the plots of d' against signal energy for human observers go through zero on the energy scale.

This last-mentioned result, of course, based as it is on extrapolation, cannot stand by itself as conclusive argument against the existence of a threshold. The result also depends on a measure of performance that is specific to detection theory. So we shall not be concerned with it further. It is possible, however, to relate the various threshold theories that have been proposed to the experimental results discussed earlier—results obtained with the yes-no, second-choice, and rating procedures, as shown in Figs. 1, 2, and 3. We shall examine these results in relation to threshold theories proposed by Blackwell (*15*), Luce (*16*), Green (*17*), Swets, Tanner, and Birdsall (*5*), and Stevens (*18*).

Blackwell's high-threshold theory. Blackwell's theory assumes that, whereas the observer may be led to say "yes" when noise alone is presented, only very infrequently is his threshold exceeded by the sensory excitation arising from noise—so infrequently, in fact, that these instances can be ignored. There is a "true" value of p_0—call it p_0'—that for all practical purposes is equal to zero. Corresponding to p_0', there is some true p_1', the value of which depends on the signal strength. Since the observer is unable to order values of sensory excitation below $p_0' \approx 0$, if he says "yes" in response to such a value he is merely guessing and will be correct on a chance basis. The operating-characteristic curve (for a given signal strength) that results from this theory is that of Fig. 4. It is a straight line from (p_0', p_1') through ($p_0 = 1.00$, $p_1 = 1.00$). The insert at lower right shows the location of the threshold. The data of observer 1 shown in Fig. 1 are reproduced for comparison.

This theoretical curve is described by the equation

$$p_1 = p_1' + p_0(1 - p_1') \qquad (1)$$

The observed proportion of "yes" responses to a signal (p_1) equals the proportion of true "yes" responses (p_1') plus a guessing factor (p_0) modified by the opportunity for guessing ($1 - p_1'$). The beauty of this high-threshold theory is that, if it is correct, the influence of spurious "yes" responses can be eliminated, the proportion of true "yes" responses being left. The familiar correction for chance success

$$p_1' = \frac{p_1 - p_0}{1 - p_0} \qquad (2)$$

is a rearrangement of Eq. 1. The correction serves to normalize the psychometric function so that, whatever the observer's tendency to guess, the stimulus threshold can be taken as the signal energy corresponding to $p_1' = 0.50$.

However, the theory does not agree with the data. The empirical curve shown in Fig. 4, like the great majority of operating-characteristic curves that have been obtained, is not adequately fitted by a straight line. The horizontal line in Fig. 2, which follows from this theory, does not fit the second-choice data shown there. The rating data of Fig. 3 also indicate ordering of values of sensory excitation below a p_0 of approximately zero. Further, yes-no and forced-choice thresholds calculated from this theory are not consistent with each other (*15*).

Luce's low-threshold theory. Luce has suggested that a sensory threshold may exist at a somewhat lower level relative to the distribution of noise—that is, that p_0' may be substantial. Apart from this, the low-threshold theory is like the high-threshold theory, only twice so. Whereas Blackwell's theory permits the observer to say "yes" without discrimination when the sensory excitation fails to exceed the threshold, Luce's theory also permits the observer to say "no" without discrimination when the sensory excitation does exceed the threshold. Thus the operating-characteristic curve of this theory contains two linear segments, as shown in Fig. 5. Again, the data for observer 1 in Fig. 1 are shown for comparison. The location of the threshold indicated by these data is shown in the insert at lower right.

It may be seen that the two-line curve fits the yes-no data reasonably well, perhaps as well as the nonlinear curve of detection theory. Although the calculations have not been performed, it seems probable that this theory will also be in fairly good agreement with the second-choice data of Fig. 2. It provides for two categories of sensory excitation, and two categories would seem sufficient to produce a proportion of correct second choices significantly

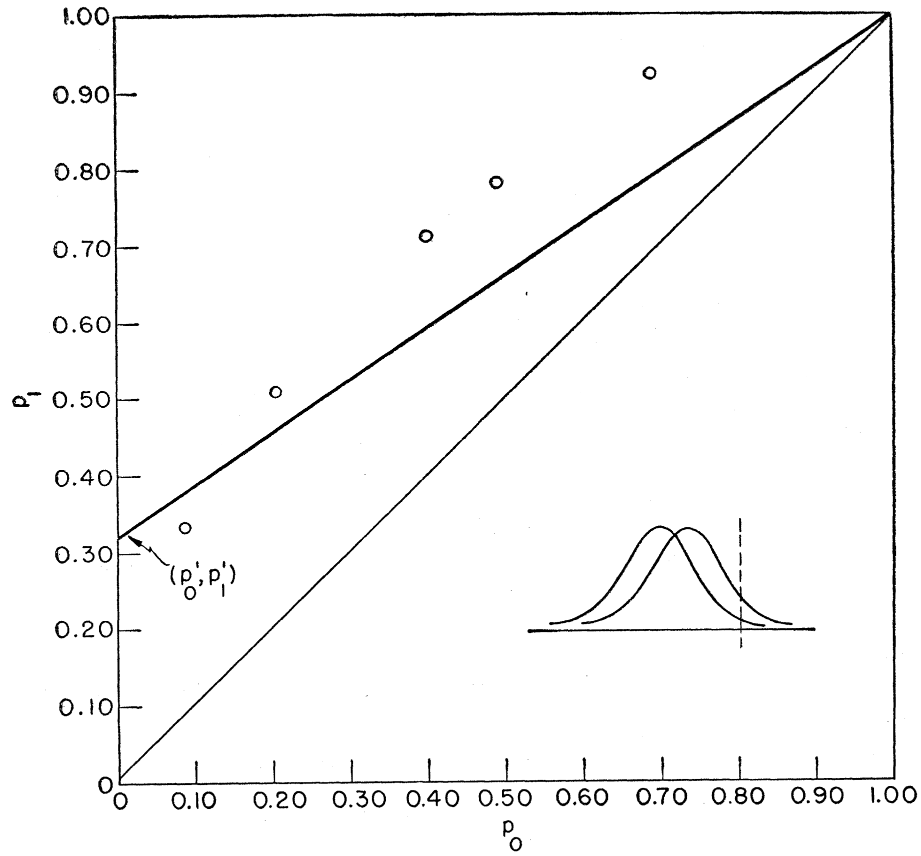

Fig. 4. The results of a yes-no experiment, and a theoretical function from Blackwell's high-threshold theory.

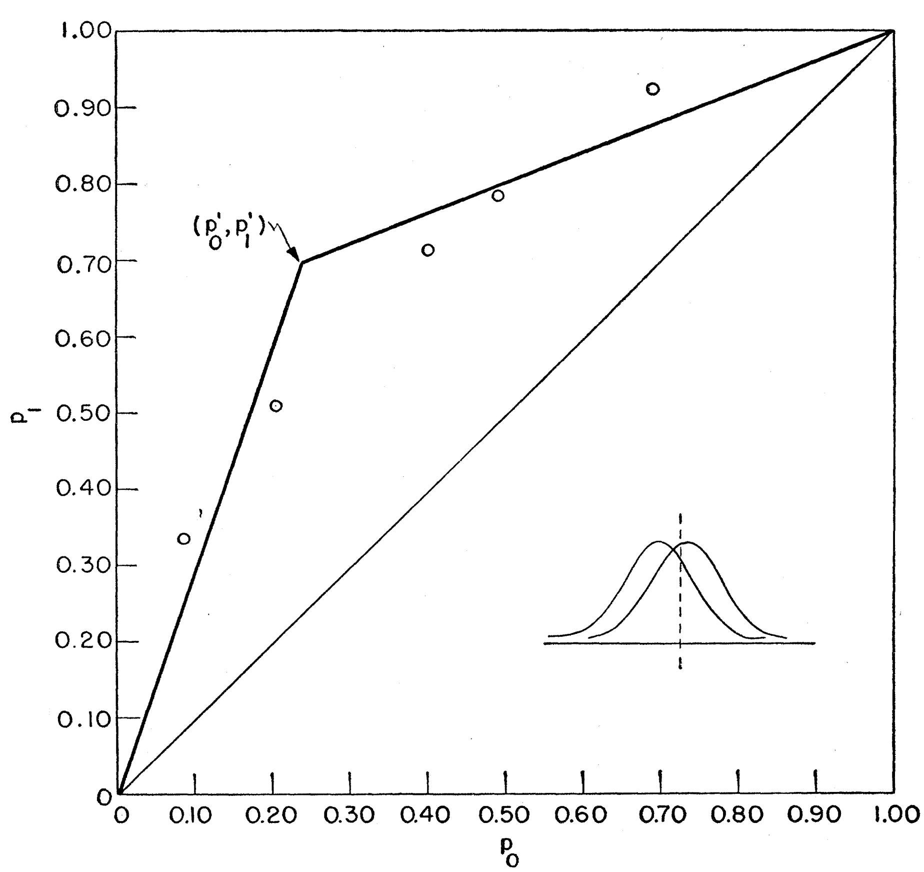

Fig. 5. The results of a yes-no experiment, and a theoretical function from Luce's low-threshold theory.

above the chance proportion. However, on the face of it, a two-category theory is inconsistent with the six categories of sensory excitation indicated by the rating data of Fig. 3. (We may note in passing that the theory raises the interesting question of how another threshold, the one above which a more complete ordering exists, might be measured.)

Green's two-threshold theory. Green has observed that operating-characteristic data, perhaps adequately fitted by Luce's curve of two segments, are certainly better fitted by a curve with three linear segments. This curve, shown in Fig. 6, corresponds to a theory that includes a range of uncertainty between a lower threshold, below which lies true rejection, and an upper threshold, above which lies true detection. The insert at lower right shows the location of the two thresholds.

As is evident from Fig. 6, the curve of three line segments fits the yes-no data at least as well as the nonlinear curve of detection theory. Again, the calculations have not been performed, but it seems very likely that a three-category theory can account for the second-choice data. Even a three-category theory, however, is inconsistent with the six categories of sensory excitation indicated by the rating data.

There is, of course, no need to stop at two thresholds and three categories. A five-threshold theory, with a curve of six line segments, would fit any operating-characteristic data very well indeed and would also be entirely consistent with the second-choice and rating results. However, such a theory is irrelevant to the question under consideration. It is hardly a threshold theory in any important sense. It may be recalled that we considered it earlier as a variant of detection theory.

Swets, Tanner, and Birdsall's low-threshold theory. Tanner, Birdsall, and I proposed a threshold theory that may be described as combining some of the features of Blackwell's and Luce's theories. This theory permits ordering of values of sensory excitation above the threshold but locates the threshold well within the noise distribution. The corresponding operating-characteristic curve is composed of a linear segment above some substantial value of p_0 (say, 0.30 to 0.50) and a curvilinear segment below this value. Inspection of Fig. 1 shows that such a curve fits yes-no data rather well. It is evident that the second-choice data, and rating data

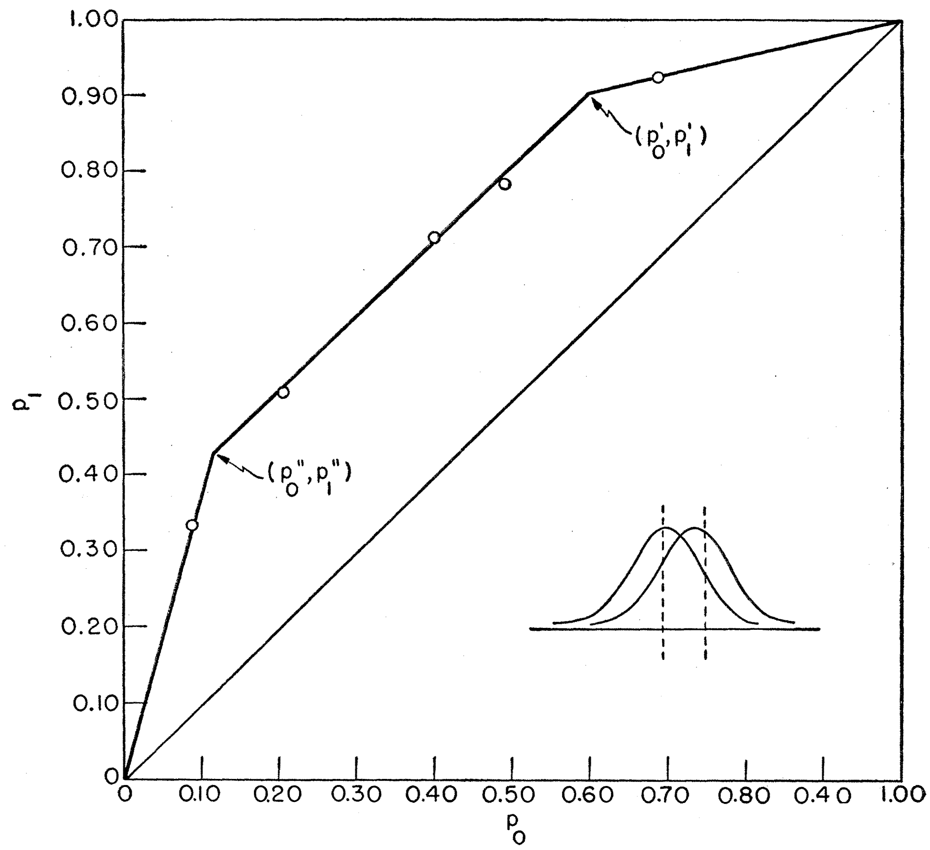

Fig. 6. The results of a yes-no experiment, and a theoretical function from Green's two-threshold theory.

exhibiting six categories, could also be obtained without ordering below this threshold.

Stevens' quantal-threshold theory. The quantal-threshold theory advocated by Stevens cannot be treated on the same terms as the other threshold theories. The data of Figs. 1, 2, and 3 are not directly relevant to it. The reason is that, whereas the other threshold theories give a prominent place to noise, collection of data in accordance with the quantal theory requires a serious attempt to eliminate all noise, or at least enough of it to allow the discontinuities of neural action to manifest themselves.

We may doubt, a priori, that noise can in fact be reduced sufficiently to reveal the "grain" of the action of a sensory system. Although the other theories we have examined apply to experiments in which the noise is considerable and, as a matter of fact, are typically applied to experiments in which noise (a background of some kind) is added deliberately, they are not generally viewed as restricted to such experiments. In adding noise we acknowledge its universality. The assumption is that the irreducible minimum of ambient noise, equipment noise, and noise inside the observer is enough to obscure the all-or-none quality of individual nervous elements in a psychological experiment. Noise is added in order to bring the total, or at least that part of it external to the observer, to a relatively constant level, and to a level at which it can be measured.

A recent article reviewing the experiments that have sought to demonstrate a quantal threshold has questioned whether any of the experiments suffices as a demonstration (*19*). Even if we ignore some technical questions concerning curve-fitting procedures and grant that some experiments have produced data in agreement with the quantal-threshold theory, we must observe that obtaining such data evidently depends upon the circumstance of having elite experimenters as well as elite observers (*18*). A relatively large amount of negative evidence exists; several other experimenters have attempted to reproduce the conditions of the successful experiments without success (*19*).

A striking feature of the quantal-theory experiments, in the present context, is the stimulus-presentation procedure employed. Although not contingent upon anything in the theory, the recommended procedure is to present signals of the same magnitude on all the trials of a series and to make known to the observer that this is the case. This procedure provides an unfortunate protection for the theory; if the observer is likely to make noise-determined "yes" responses, the fact will not be disclosed by the experiment. Licklider has expressed aptly the growing discomfiture over this procedure: "More and more, workers in the field are growing dissatisfied with the classical psychophysical techniques, particularly with the [methods that ask the observer] to report 'present' or 'absent' when he already knows 'present.' It is widely felt that the 'thresholds' yielded by these procedures are on such an insecure semantic basis that they cannot serve as good building blocks for a quantitative science" (*20*). Although the original intent behind the use of this procedure in the quantal-theory experiments was to make the task as easy as possible for the observer, from the point of view of detection theory the procedure presents a very difficult task—it requires that the observer try to establish the response criterion that he would establish if he did not know that the signal was present on every trial.

Thus the advocates of the quantal theory specify a procedure that makes detection theory inapplicable. The result is that, as things stand, the conflict between the two theories cannot be resolved to the satisfaction of all concerned, as it conceivably could be if both theories could be confronted with the same set of data. However, there is reason to hope—since the quantal-theory procedure is not intrinsic to the theory but rests rather on a sense of experimental propriety, which is a relatively labile matter—that such a confrontation will some day be possible.

Is There a Sensory Threshold?

We have considered the data of three experiments—the yes-no, second-choice, and rating experiments—in relation to five competing theories concerning the processes underlying these data. The three sets of data are in agreement with detection theory, a theory that denies the existence of a sensory threshold, and also with the version of a low-threshold theory proposed by Tanner, Birdsall, and me. Blackwell's high-threshold theory is inconsistent with all three sets of results. Luce's low-threshold theory is consistent with the first, perhaps consistent with the second, and inconsistent with the third. Green's two-threshold theory fits the first two sets of results but not the last. We also considered the only other explicit threshold theory available—the quantal theory, to which the three experiments are not directly relevant.

The outcome is that, as far as we know, there may be a sensory threshold. The possibility of a quantal threshold cannot be discounted, and certainly not on the basis of data at hand. On another level of analysis, there may be what we have termed a low threshold, somewhere in the vicinity of the mean of the noise distribution. The low-threshold theory proposed by Tanner, Birdsall, and me fits all of the data we examined. If the rating experiment can be dismissed (there is now no apparent reason for giving it less than full status), then Luce's and Green's theories, which involve a low threshold, fit the remaining data.

On the other hand, the existence of a sensory threshold has not been demonstrated. Data consistent with the quantal theory are, at best, here today and gone tomorrow, and the theory has yet to be tested through an objective procedure. With respect to a low threshold, we may ask whether demonstration of such a threshold is even conceivable.

It is apparent that it will be difficult to measure a low threshold. Consider the low-threshold theory that permits complete ordering above the threshold in connection with the forced-choice experiment. The observer conveys less information about his ordering than he is capable of conveying if only a first choice is required. We saw in the preceding discussion that the second choice conveys a significant amount of information. Another experiment, in which the observer tried to be incorrect, indicated that he can order four choices (*6*). Thus it is difficult to determine when enough information has been extracted to yield a valid estimate of a low threshold.

Again, it is difficult to imagine how one might determine the signal energy corresponding to the thresholds of Luce's and Green's theories. The determination is made especially difficult by the fact that, in general, empirical operating-characteristic curves for various signal energies are fitted well by the theoretical curves of detection theory. Consequently, the line-segment curves that best fit the data have lines intersecting at a value of p_0 that depends upon the signal energy. The implication is that the location of the threshold depends on the signal energy that is being presented.

Implications for Practice

We have, then, the possibility of a threshold, but it is no more than a possibility, and we must observe that since it is practically unmeasurable it will not be a very useful concept in experimental practice. Moreover, even if the low threshold proposed by Tanner, Birdsall, and me did exist, and were measurable, it would not restrict the application of detection theory. We may note that yes-no data resulting from a suprathreshold criterion depend upon the criterion but are completely independent of the threshold value. The same limitation applies to the quantal threshold. It appears that a compelling demonstration of this concept will be difficult to achieve, so that in practice a theory and a method that deal with noise will be required.

Accordingly, with any attempt to measure sensitivity by means of "yes" and "no" responses, a measure of the observer's response criterion should be obtained. The only way known to obtain this measure is to use catch trials—randomly chosen trials that do not contain a signal. The methods of adjustment, limits, and constants in their usual forms, in which the observer knows that the signal is present on every trial, are inappropriate.

A large number of catch trials should be presented. It is not sufficient to employ a few catch trials, enough to monitor the observer, and then to remind him to avoid "false-positive" responses each time he makes one. This procedure merely forces the criterion up to a point where it cannot be measured, and it can be shown that the calculated threshold varies by as much as 6 decibels as the criterion varies in this unmeasurable range (*5*). Precision is also sacrificed when, because highly trained observers are employed, the untestable assumption is made that they do maintain a constant high criterion. Even if all laboratories should be fortunate enough to have such observers, we would have to expect a range of variation of 6 decibels among "constant criterion" observers in different laboratories. To be sure, for some problems, this amount of variability is not bothersome; for others it is.

The presentation of a large number of catch trials—enough to provide a good estimate of the probability of a "yes" response on such a trial—is still inadequate if this estimate is then used to correct the proportion of "yes" responses to the signal for chance success. The validity of the correction for chance depends upon the existence of a high threshold that is inconsistent with all of the data that we examined. It should be noted that the common procedure of taking the proportion of correct responses that is halfway between chance and perfect performance as corresponding to the threshold value of the signal is entirely equivalent to using the chance correction.

In summary, in measuring sensitivity it is desirable to manipulate the response criterion so that it lies in a range where it can be measured, to include enough catch trials to obtain a good estimate of this response criterion, and to use a method of analysis that yields independent measures of sensitivity and the response criterion. One qualification should be added: We can forego estimating the response criterion in a forced-choice experiment. Under the forced-choice procedure, few observers show a bias in their responses large enough to affect the sensitivity index d' appreciably. Those who do show such a bias initially can overcome it with little difficulty. As a result, the observer can be viewed as choosing the interval most likely to contain a signal, without regard to any criterion. For this reason, the forced-choice procedure may be used to advantage in studies having an emphasis on sensory, rather than on motivational or response, processes (*21*).

References and Notes

1. G. T. Fechner, *Elemente der Psychophysik* (1860).
2. The general theory of signal detectability is presented in W. W. Peterson, T. G. Birdsall, W. C. Fox, *IRE Trans. Profess. Group on Information Theory* **PGIT-4**, 171 (1954). Psychophysical theories similar to it have been suggested by M. Smith and Edna A. Wilson, *Psychol. Monographs* **67**, No. 9, Whole No. 359 (1953), and W. N. Munson and J. E. Karlin, *J. Acoust. Soc. Am.* **26**, 542 (1956). The first application of detection theory in psychophysics is described in W. P. Tanner, Jr., and J. A. Swets, *Psychol. Rev.* **61**, 401 (1954).
3. E. G. Boring, *A History of Experimental Psychology* (Appleton-Century-Crofts, New York, ed. 2, 1950), p. 284.
4. E. G. Boring, *The Physical Dimensions of Consciousness* (Century, New York, 1933).

5. J. A. Swets, W. P. Tanner, Jr., T. G. Birdsall, "The evidence for a decision-making theory of visual detection," *Electronic Defense Group, Univ. Michigan, Tech. Rept. No. 40* (1955). This material will appear as "Decision processes in perception," *Psychol. Rev.*
6. W. P. Tanner, Jr., J. A. Swets, D. M. Green, "Some general properties of the hearing mechanism," *Electronic Defense Group, Univ. Michigan, Tech. Rept. No. 30* (1956).
7. W. P. Tanner, Jr., and J. A. Swets, *Psychol. Rev.* **61**, 401 (1954).
8. J. A. Swets, *J. Acoust. Soc. Am.* **31**, 511 (1959).
9. J. P. Egan, A. I. Schulman, G. Z. Greenberg, *ibid.* **31**, 768 (1959).
10. W. P. Tanner, Jr., and R. C. Jones, "The ideal sensor system as approached through statistical decision theory and the theory of signal detectability," proceedings of the Armed Forces–NRC Vision Committee meeting, Washington, D.C., Nov. 1959.
11. W. P. Tanner, Jr., and T. G. Birdsall, *J. Acoust. Soc. Am.* **30**, 922 (1958).
12. J. A. Swets, E. F. Shipley, M. J. McKey, D. M. Green, *ibid.* **31**, 514 (1959).
13. W W. Peterson, T. G. Birdsall, W. C. Fox, *IRE Trans. Profess. Group on Information Theory* **PGIT-4**, 171 (1954).
14. D. M. Green, *J. Acoust. Soc. Am.* **32**, 1189 (1960).
15. H. R. Blackwell, "Threshold psychophysical measurements," unpublished. Also, *Univ. Mich. Eng. Research Inst. Bull. No. 36* (1953).
16. R. D. Luce, *Science* **132**, 1495 (1960).
17. D. M. Green, personal communication.
18. S. S. Stevens, in *Sensory Communication*, W. A. Rosenblith, Ed. (Technology Press and Wiley, New York, in press); ———, *Science* **133**, 80 (1961).
19. J. F. Corso, *Psychol. Bull.* **53**, 371 (1956).
20. J. C. R. Licklider, in *Psychology: A Study of Science*, S. Koch, Ed. (McGraw-Hill, New York, 1959), vol. 1.
21. The preparation of this paper (technical report No. AFCCDD TR 61–10) was supported by the U.S. Army Signal Corps, the Air Force (Office of Scientific Research and Operational Applications Laboratory), and the Office of Naval Research.

Science and the News

Hanford and Stanford: The Issue Is Clear but the Politics Are Complex

The complicated politicking that has linked a $95 million proposal to add power generating facilities to the new Hanford, Washington, plutonium reactor and the $114 million proposal to build a giant electron accelerator at Stanford University grew even more complicated last week when the House of Representatives knocked the Hanford proposal out of the Atomic Energy Commission authorization bill. The House defeat set in motion an elaborate stratagem by supporters of Hanford to save the project, and the stratagem, until nearly the last minute, involved a threat to kill the Stanford accelerator, although this move was finally abandoned.

The Hanford proposal involves building a generating plant to use the steam produced by the cooling system of the plutonium reactor. If built, it would produce 700,000 kilowatts of power, and would be the largest atomic power plant in the world. Heavy opposition developed from the private power industry, which was immensely displeased at the idea of the government's going into the production of atomic power, and from the coal industry and coal-producing areas generally, which felt that if the Hanford plant were not converted to power production, new coal-fired generating plants would be built to provide for the Northwest's power needs. One West Virginia Democrat from a coal-mining area who had a nearly perfect record of support for the Administration took the floor to dissociate himself from the arguments of the opponents of public power. He said his vote would simply reflect the fact that he was representing West Virginia, not the whole United States, and that he could not vote for a proposal that would, he feared, just put more West Virginians on the dole.

In the Pacific Northwest, on the other hand, conservative Republicans joined the Democrats in supporting the proposal, and the strongly conservative Portland *Oregonian*, after the House vote, published a bitter editorial railing at the "incredible piece of Congressional stupidity" based on "arguments as phony as a lead wedding ring."

The debate, then, was essentially over the issue of an expansion of public power, but with many departures from normal voting patterns, based on sectional interests.

Democrats on the Atomic Energy Committee hoped to save the proposal after the House defeat by restoring the Hanford authorization in conference, with the chance that the conference report might be pushed through the House. This was the tactic that enabled the Administration to get through its minimum wage bill after a preliminary defeat in the House.

A conference report is supposed to represent a compromise between rival House and Senate bills. To create at least an illusion of something that could be compromised between the House and Senate, the Democrats talked of knocking some provision out of the bill in the Senate. The conference committee, controlled by supporters of Hanford, could then arrange a "compromise" in which the House would give in on Hanford and the Senate would graciously restore whatever it had knocked out. Since the Stanford accelerator was the only project in the bill even remotely comparable in importance to Hanford, the Democrats planned to try to kill Stanford in the Senate.

This was a peculiarly transparent scheme since the same Democratic Senators who as members of the Joint Atomic Energy Committee had unanimously voted in favor of Stanford would now have to take the floor of the Senate to argue that it should be killed. But it was seriously talked about, privately of course, by leading members of the Joint Committee up until the day before the Senate vote last Tuesday, and apparently finally abandoned only when it became clear that there was no way to get a majority of the Senate to go along with it.

Last Tuesday, opponents of Hanford in the Senate, led by Hickenlooper of Iowa, fought hard to kill the Hanford project then and there by knocking it out of the Senate version, thus leaving no chance for it to be restored in conference. But after 3 hours of debate supporters of the project carried the vote by 54 to 36. This left Hanford in in the Senate version, out in the House, and with a majority in the conference