3

Psychophysical Theory

The application of the methods described in Chapter 2 yields a quantity, expressed in physical units, called the threshold. The concept of the threshold as an index of absolute and differential sensitivity has been extremely useful in the study of sensory systems. Through the use of this quantity investigators have been able to discover the stimulus conditions to which our sensory systems are the most and the least sensitive. But psychophysicists, not content to work exclusively at this descriptive level, have proposed theories concerning the underlying mechanisms of sensory thresholds. Each theory was proposed to account for empirical data obtained in psychophysical experiments and consists of a description of neurophysiological or psychological processes within the observer which could determine the observer's behavior. The validity of each theory must be evaluated by determining the degree to which precise quantitative deductions from the theory are confirmed by experimental data.

CLASSICAL THRESHOLD THEORY

Early threshold theories were based upon the assumption that the measurements obtained in psychophysical experiments were estimates of a threshold in the observer which could not be measured directly. It was thought that the threshold was a sharp transition point between sensation and no sensation and that a specific, critical amount of neural activity must result from stimulation to exceed threshold. The value of the threshold was assumed to vary with properties of the stimulus (such as duration, area, and wavelength) and also to vary according to the condition of the sensory nervous system (such as state of adaptation and level of background activity). If all the factors affecting the threshold level could be maintained exactly the same from measurement to measurement and if the applica-
tion of the stimulus energy to the receptors could be exactly replicated, a particular stimulus would be expected either always or never to produce a sensation, depending upon whether the stimulus produced enough neural activity to exceed threshold.

Since the level of neural activity increases with the intensity of the stimulus, the predicted psychometric function using the method of constant stimuli would jump sharply from 0 to 100% detection when the stimulus intensity was set at a level that produced the threshold amount of neural activity (Figure 3.1). Empirical psychometric functions do not follow this form but are typically ogive curves, where the proportion of "yes" responses gradually increases with stimulus intensity. The step-like function of Figure 3.1, however, is a theoretical curve representing predicted results in a hypothetical situation where perfect control is maintained over all of the stimulus and biological variables affecting the level of neural activity in the sensory system. Since perfect control cannot be achieved, the function remains a theoretical formulation and must be considered the idealized outcome if the assumptions of the classical threshold theory are correct.

Inasmuch as it has been impossible to test the theory by examining the relationship between proportion of detections and stimulus intensity when all stimulus and biological factors are perfectly controlled, proponents of the theory are obliged to account for results obtained under imperfect conditions. The basic notion in classical threshold theory is that the threshold varies over time. Although an observer's threshold may be a sharp boundary at a particular instant in time, in an experiment it acts as though it were always changing. Factors affecting the threshold fluctuate randomly from moment to moment, and therefore repeated applications of a stimulus of a particular intensity should result in a detection response only on those trials when the momentary threshold is exceeded. The proportion of trials where the momentary threshold is exceeded should increase as an ogival function of stimulus intensity if the variation of momentary thresholds is normally distributed. Curve A of Figure 3.2 illustrates a hypothetical frequency distribution of momentary thresholds expressed in units of stimulus intensity. A detection response should be produced only when stimulus intensity is equal to or greater than the momentary threshold value. The proportion of time that a stimulus will exceed threshold is equal to the proportion of the area under curve A below the value of the stimulus. For example, a weak stimulus of 2 units would exceed momentary thresholds of 2 units or less only 2% of the time because on only 2% the trials will the momentary threshold be lower than 2. The first point on the psychometric function, curve B of Figure 3.2, would be .02 for a stimulus of 2 units. A stimulus of 3 units should be detected on .16 of the trials, sinc
thresholds are 3 units or less on 16% of the trials. The second point on the psychometric function can then be plotted. When stimulus intensity is 4 units, the momentary threshold will be equal to or less than this value on exactly 50% of the trials and this stimulus will be detected on .5 proportion of the trials. Thus, the .5 point on the psychometric function corresponds to the mean of the momentary threshold values. It should now be clear why the early psychophysicists chose 50% detection as the best estimate of the threshold. The proportion of trials on which stimulus intensity is equal to or above momentary threshold is .84 for a stimulus of 5 units and .98 for a stimulus of 6 units.

Classical threshold theory is often identified as the phi–gamma hypothesis, with phi referring to the probability of a response and gamma referring to stimulus intensity. The phi–gamma hypothesis states that the psychometric function in which response probability is plotted against stimulus magnitude should have the ogival form of the cumulative normal distribution. The prediction of an ogive psychometric function follows directly from the assumption that momentary thresholds are normally distributed over time. The unit of measurement of the stimulus is sometimes an important consideration when testing the phi–gamma hypothesis because intensity can sometimes be measured on several different physical scales which are not always linearly related to each other. Sound energy is not proportional to sound pressure, for example, but is instead proportional to sound pressure squared. The cumulative normal distribution cannot possibly describe the psychometric function both when sound intensity is expressed in pressure units and when it is expressed in energy units.

Which unit of stimulus intensity is appropriate for testing the theory? Whenever possible stimulus intensity should be expressed in units that best reflect the operating characteristics of the sensory system. Thurstone was aware of this problem as early as 1928. He noted that for classical threshold theory to predict correctly an ogival psychometric function the stimulus should be expressed in psychological units instead of in physical units. If fluctuations in sensitivity were normally distributed along a psychological dimension of intensity they would not be normally distributed when expressed in units of stimulus intensity unless stimulus values were transformed to reflect psychological intensity. Thurstone made the assumption that Fechner’s law is correct and proposed the phi–log–gamma hypothesis. According to this hypothesis the psychometric function should have the ogival form of the cumulative normal distribution when response probability is plotted as a function of log stimulus magnitude. Since the range of stimulus values is unfortunately so small, the predicted psychometric functions from the phi–gamma hypothesis and the phi–log–gamma hypothesis are so similar that data precise enough to differentiate between the two hypotheses have not yet been obtained.

Because similar ogival psychometric functions have been obtained in detection and discrimination experiments, classical threshold theory has been applied to the measurement of the difference threshold. Investigators assumed that the neural activity for two sensations must differ by some threshold amount to be perceived as different. The size of the stimulus difference required to exceed threshold was assumed to vary randomly from trial to trial, and therefore the probability of detecting a difference should increase as an ogival function of the size of the stimulus difference. A detailed description of the application of classical threshold theory to the problem of sensory discrimination can be found in Boring (1917).

In summary, classical threshold theory was the first attempt to make inferences from psychophysical data about the nature of processes within the observer. Empirically determined psychometric functions have thus been used to make inferences about the underlying nature of sensory mechanisms. The ogival psychometric functions so frequently observed in psychophysical investigations have led to the logical inference that fluctuations in momentary sensory thresholds are normally distributed along some sensory continuum within the observer.

**NEURAL QUANTUM THEORY**

The central assumptions of classical threshold theory are that fluctuations in threshold are random and that the sensory dimension is continuous. The sensory dimension, however, may not be a continuum but instead may be a series of discrete steps. Under these circumstances psychometric functions for difference thresholds could deviate from the ogival form. In fact, not all psychometric functions obtained during measurement of difference thresholds follow the ogival form. Infrequently obtained psychometric functions, where response probability increases from 0 to 1.0 as a linear function of stimulus magnitude, therefore became the basis of a new theory of sensory discrimination. The neural quantum theory, first made explicit by Stevens, Morgan, and Volkman (1941), is an attempt to derive a linear psychometric function from the assumption that discrimination occurs along a sensory dimension within the observer which is made up of small discrete (quantal) steps. Neurophysiology provides no clear evidence as to the nature of the sensory dimension. In the nervous system receptor potentials and postsynaptic potentials vary as a continuous function of stimulus intensity while action potentials constitute all-or-none responses to changes in stimulus intensity. Neither have any final answers been supplied by psychophysics, but it is significant that the psychometric functions of a handful of investigators are remarkably consistent with the neural quantum concept.

In the simplest form of neural quantum theory it is assumed that an observer can detect an increment in a stimulus only when it is large enough to excite one additional neural unit. The size of the necessary stimulus increment will depend on how much the first stimulus is above the threshold of the last excited neural unit. The greater the excess over the threshold of the last unit, the smaller the required stimulus increment to excite the next unit will be.

The first evidence in support of a quantal theory of discrimination was reported by von Békésy (1930). A standard tone lasting .3 sec was presented to an observer and was followed immediately by a .3-sec comparison tone of a different intensity.
first and second stimulus. During a single session sufficient data were obtained to determine the proportion of detections for several sizes of stimulus increment.

Stevens pointed out that there are three features of the data from these experiments which have special importance for neural quantum theory: (a) stimulus increments below a critical size produce no response; (b) above that critical value, the number of increments detected is a linear function of the size of the increment; and (c) increments are always detected when the increment reaches a second critical size, which is twice the size of the largest increment that is never detected. These characteristic results are clearly seen in the data of Stevens, Morgan, and Volkmann (1941) presented in Figure 3.3. The importance of these findings should become clear when the elements of the neural quantum model are understood.

According to the neural quantum model a stimulus of a particular intensity stimulates a certain number of neural quantum units and, as illustrated in Figure 3.4, there may be a stimulus surplus (p) which is insufficient to exceed the threshold of the next unit but is available to combine with a stimulus increment (Δφ). When a stimulus increment occurs it adds to p, and if their sum is large enough to excite one or more additional quantal units the stimulus increment is detected.

The present form of the theory requires the assumption that the observer’s random sensitivity fluctuations are large compared to the size (Q) of the neural units and slow compared with the time needed for the stimulus increment to be made. The relatively large fluctuations in sensitivity cause considerable variability

FIGURE 3.3 Psychometric functions for the detection of an intensity increment added to a 1000-Hz tone of five different intensity levels. Von Békésy’s data are also presented for the detection of increments (circles) and decrements (half circles). The theoretical curves were drawn with the restrictions that they be straight lines with intercepts that stand in a 2-to-1 relation. (From Stevens, Morgan, & Volkmann, 1941.)

On each trial the observer reported whether or not he detected a loudness difference between the two tones. The linear psychometric functions in this experiment were interpreted as support for the quantum nature of loudness discrimination. Similar results were obtained by Stevens, Morgan, and Volkmann (1941) for both loudness and pitch. Some of the linear psychometric functions found in these experiments are presented in Figure 3.3.

Stevens (1972a) reviewed the data from about a dozen investigations carried out over a span of 40 years. Some 140 steplike functions for auditory loudness and pitch and for three types of visual patterns were reproduced in Stevens’ paper as support for neural quantum theory. In the experiments that were reported, observers were presented with a standard stimulus followed immediately by a comparison stimulus which produced an incremental change in stimulation. Observers were required to respond when they detected the incremental change between the

FIGURE 3.4 The diagram on the left illustrates the basic concepts of the neural quantum model. The stimulus activates a certain number of neural units. The stimulus has a surplus p that is insufficient to exceed the threshold of the next unit unless the stimulus increment Δφ is added. (From von Békésy, 1950.) The diagram on the right is the psychometric function predicted for a 2-q quantum criterion. (From Stevens, 1972a. Copyright 1972 by the American Association for the Advancement of Science.)
in the total number of excited neural units. As a result of this large random variability over a range of many neural quanta units one value for the stimulus surplus \( p \) is as likely as any other for the presentation of a particular stimulus and consequently the frequency distribution \( p \) for a particular standard stimulus is rectangular. Thus, if \( Q \) is the stimulus increment that will always succeed in exciting an additional neural quantum, the value of \( \Delta \phi \) that is just sufficient to excite one additional neural quantum is given by

\[
\Delta \phi = Q - p. \tag{3.1}
\]

A given \( \Delta \phi \) will excite an additional neural unit whenever \( \Delta \phi = Q - p \). Since \( p \) is uniformly distributed over the neural quantum and since the greater the value of \( \Delta \phi \) the more likely \( \Delta \phi \) plus \( p \) is to exceed threshold, the proportion of times that \( \Delta \phi \) will excite one additional neural quantum is given by

\[
r_1 = \frac{\Delta \phi}{Q}. \tag{3.2}
\]

Equation (3.2) indicates that for a particular \( Q \) size the proportion of increment detections \( r_1 \) increases as a linear function of \( \Delta \phi \). The function starts at the origin and reaches 1.0 when \( \Delta \phi \) is equal to \( Q \), as seen in Figure 3.4 for the one-quantum criterion. When \( \Delta \phi/Q \) is .8 on 80% of the trials, \( p \) will be equal to or greater than the value necessary to exceed threshold when combined with \( \Delta \phi \). When \( \Delta \phi \) is further reduced so that the value of \( \Delta \phi/Q \) is only .5 on only 50% of the experimental trials, the value of \( p \) should be equal to or greater than the value necessary to exceed threshold when combined with \( \Delta \phi \).

The use of Equation (3.2) predicts the behavior of the observer as he adopts a judgment criterion for reporting an increment in stimulation whenever one additional neural unit is excited. Psychophysical data, however, are more consistent with the hypothesis that the observer reports a stimulus increment when two additional neural quanta are excited. According to Stevens (1972a) the two-quantum criterion is necessary because the random fluctuations in sensitivity produce randomly occurring one-quantum increments and decrements in neural activity. The observer needs to adopt a two-quantum criterion in order to distinguish the presence of the stimulus from the background activity of the nervous system. The prediction from the neural quantum theory for a two-unit threshold is given by

\[
r_2 = \frac{\Delta \phi - Q}{Q}. \tag{3.3}
\]

In Equation (3.3) the proportion of detections first exceeds zero when \( \Delta \phi \) is equal to \( Q \) and becomes 1.0 when \( \Delta \phi \) is equal to \( 2Q \) (see Figure 3.4). Thus the value of \( \Delta \phi \) when the proportion of detections first becomes 1.0 should be exactly twice the value of \( \Delta \phi \) when the proportion of detections first becomes greater than zero. The data reviewed by Stevens (1972a) are remarkably consistent with this prediction.

It is significant that data in support of neural quantum theory have been obtained under a variety of conditions using both auditory and visual stimuli. Stevens argued that the generality of this finding is consistent with the hypothesis that the operation of the neural quantum may be a central neural mechanism of sufficient generality to process information from all sensory modalities under a great variety of stimulus conditions. According to this hypothesis sensory inputs would eventually converge on the same neural center in the brain. The reticular formation of the brain stem is a possible candidate for such a center in that it receives inputs from all sensory modalities. A quantal jump in neural activity level would occur when \( \Delta \phi \) is large enough to exceed the threshold of some switching mechanism in the center.

The difficulty of obtaining linear psychometric functions has been attributed to a lack of precise experimental control over such randomly fluctuating factors as the observer's motivation, attention, and fatigue. Since neural quanta are very small, their operation in the detection or discrimination experiment is often masked by the much larger effects of the fluctuation of these uncontrolled factors on the observer's judgments. Fluctuation of the uncontrolled factors is likely to be normally distributed, and the ogive psychophysical function is therefore obtained. Because it is not precisely clear in this instance what constitutes an acceptable experiment, neural quantum theory is difficult to reject. When the data do not fit the theory one can argue that something was wrong with the experiment. Stevens, however, identified some conditions that seem to be necessary for the production of linear psychometric functions:

1. The stimulus must be carefully controlled. When the standard and the comparison stimuli were bursts of white noise and thus varied randomly over time, a normal ogive rather than a linear function was obtained (Miller, 1947). A jittering stimulus, such as white noise which is constantly changing in amplitude and frequency spectrum, should obscure the stepwise quantal function since large random variations between presentation of standard and comparison stimuli would greatly influence the observer's judgments.

2. If the observer is unable to maintain a constant criterion during an experimental session the psychometric function will tend to be an ogive rather than a straight line. It is no easy task to maintain a fixed criterion during a one- or two-hour session in which a thousand or more stimuli are presented. According to Stevens, some observers simply cannot concentrate well enough to produce linear functions. Best results are often obtained when highly motivated investigators serve as observers.

3. If the size of the neural quantum changes within a session the function will become an ogive. In Figure 3.5 are data obtained by Miller and Garner (1944) for the discrimination of loudness increments of a 1000-Hz tone. Plots A and B clearly support the hypothesis of quantal discrimination with a two-quantum criterion. Plot C shows two functions obtained from one observer in two different sessions. The difference between the functions shows that the size of the neural quantum was larger during the second session. As seen in Plot D, averaging the data from the two sessions produces the ogive function. Plot E contains functions for an increment duration of 200 msec (filled points) and 100 msec (unfilled points).
neural quanta for the two conditions appear to be different in size; averaging the data obscures the step functions and produces the ogive seen in Plot F.

4. The transition from the standard to the comparison stimulus must be rapid. If the delay between the two stimuli is long, the observer’s sensitivity may not be the same for each stimulus presentation. Under these conditions random changes in

sensitivity will be reflected in an ogive psychometric function. For an adequate test of the neural quantum theory the transition from the standard to the comparison stimulus must be nearly instantaneous.

It is therefore under only the most stringent conditions that we can expect neural quanta, if they exist, to manifest themselves in an observer’s performance of a discrimination task. Thus it is not surprising that most psychometric functions have the ogival rather than the linear form. Although Stevens makes a good case for the hypothesis that a well-trained, highly motivated observer under strictly controlled stimulus conditions can make discrimination judgments based on a precise two-quantum criterion, the position of the theory would be made more secure by supporting data from sensory modalities other than vision and audition and by a further delineation of the conditions under which linear psychometric functions are obtained.

Neural quantum theory has been criticized on both methodological and theoretical grounds. Corso (1956, 1973) claims that it is difficult to determine whether the linear psychometric function of neural quantum theory or the classical ogive function is the best-fitting function for a particular set of psychophysical data. The predicted functions do not greatly differ in form; consequently they may fit the same set of data equally well. Furthermore, an adequate statistical test for discrimination between the two hypotheses is lacking. Corso has also been critical of proponents of neural quantum theory for being unable to specify precisely the essential conditions under which the linear psychometric functions predicted from the theory are obtained.

A more serious criticism of neural quantum theory comes from Wright (1974) who has argued that the special procedures employed by neural quantum theorists bias the observer to be very conservative in reporting weak stimuli. He has shown that the results predicted from neural quantum theory can also be predicted from the theory of signal detection when it is assumed that observers adopt conservative judgment strategies in the neural quantum test situation. The theory of signal detection will be extensively discussed later in this chapter. Furthermore, Wright was able to show that as the observer’s judgment strategy becomes more conservative the results predicted from the theory of signal detection more closely approach those predicted from neural quantum theory. Under these conditions it is not possible to decide between the two alternative explanations of the data without further experimental investigation. It is unlikely, however, that the neural quantum question will be adequately answered in the near future. Little research has been conducted on the problem during the last ten years, and with the death of S. S. Stevens the theory is left without a major defender. In a 1973 paper Norman pointed out that the rise of the theory of signal detection has led to an almost complete lack of attention to problems of how observers respond in noise-free situations (Norman, 1973). It is unlikely that a satisfactory answer to the question will come as a side effect in the study of detection and discrimination in noisy situations, since such experiments would surely obscure neural qua
exist. What is needed, and will come with a revival of interest in the problem, are experiments specifically designed to test neural quantum theory.

While neural quantum theory was developed to account for discrimination data obtained in a relatively noise-free situation, the theory of signal detection applies to situations in which the observer must detect weak stimuli presented against a noisy background. In psychophysics today perhaps the most powerful arguments that sensation changes on a continuum rather than in discrete steps come from the proponents of the theory of signal detection. Before describing this theory in detail, some evidence will be presented in support of the hypothesis that, in most situations, an observer’s judgments of weak stimuli are not determined by the abrupt discontinuity in neural activity implied by the absolute threshold concept, but instead are determined by an adjustable judgment criterion.

EVIDENCE AGAINST THE THRESHOLD CONCEPT

The old concept of the absolute threshold as a boundary or limit below which no sensation can occur is cast in doubt by the results of many recent psychophysical experiments. It is apparent that threshold measurement does not consist of an observer’s simply reporting the presence or absence of sensations. Experimenters have therefore directed their attention to the problem of discovering exactly what it is that observers do when they detect stimuli.

Early psychophysicists assumed a close correspondence between the verbal reports of a well-trained observer and concurrent neurological changes in the sensory system caused by stimulation. They worked to obtain results that were pure sensory functions uncontaminated by factors not directly related to the sensory system (e.g., the observer’s attitudes and expectations concerning the task). Experimenters assumed that in a well-controlled psychophysical experiment the probability of a “yes” response \( p(\text{yes}) \) for a particular stimulus presentation was entirely a function of the stimulus and the biological state of the sensory system. Since interest was mainly in the sensory system, this assumption simplified matters considerably. The results of recent experiments, however, indicate that many nonsensory variables, even when well-trained observers are used, strongly influence performance in the detection situation.

One nonsensory variable consistently found to affect the \( p(\text{yes}) \) is the probability of stimulus occurrence \( p(S) \). In the early psychophysical experiments, \( p(S) \) was always 1.0, for a stimulus was presented on every trial. It seems likely that even for the most conscientious observers the extremely high expectation of stimulus occurrence associated with presenting a stimulus on every trial would itself influence the probability of saying “yes” when a stimulus is presented. In fact, when \( p(S) \) is systematically varied, \( p(\text{yes}) \) is found to increase with \( p(S) \). In such an experiment fairly weak stimuli are typically used, and during a session several hundred trials may be administered. The observer’s task is to report whether or not

<table>
<thead>
<tr>
<th>( p(S) )</th>
<th>Number of stimulus trials</th>
<th>Number of no stimulus trials</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>.90</td>
<td>180</td>
<td>20</td>
<td>Stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Yes .99</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No .01</td>
</tr>
<tr>
<td>.70</td>
<td>140</td>
<td>60</td>
<td>Stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Yes .91</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No .09</td>
</tr>
<tr>
<td>.50</td>
<td>100</td>
<td>100</td>
<td>Stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Yes .69</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No .31</td>
</tr>
<tr>
<td>.30</td>
<td>60</td>
<td>140</td>
<td>Stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Yes .36</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No .64</td>
</tr>
<tr>
<td>.10</td>
<td>20</td>
<td>180</td>
<td>Stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Yes .05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No .95</td>
</tr>
</tbody>
</table>

a stimulus occurred on a particular trial. A value of \( p(S) \) is chosen and the trials on which the stimulus is presented are determined randomly. Several sessions are usually conducted using different values of \( p(S) \).

Consider the possible outcomes of a single trial in the detection situation. When the stimulus is present, the observer may report “yes” (a hit) or he may report “no” (a miss). On trials when the stimulus is absent, a false alarm is made if the observer says “yes” but he makes a correct rejection if he says “no.” A \( 2 \times 2 \) table containing the experimentally obtained proportions for each of the four possible outcomes summarizes the results of a series of trials for a particular set of stimulus conditions. In an experiment on \( p(S) \) we would have a \( 2 \times 2 \) table for each of the values of \( p(S) \). Table 3.1 shows the kind of results the experiment might yield. We need only consider \( p(\text{yes}) \), since \( p(\text{no}) \) is equal to 1.0 minus \( p(\text{yes}) \) and is therefore not an independent measure of performance in a two-choice situation. It is quite clear that the probability of a “yes” response when the stimulus is present increases as the probability of occurrence of the stimulus is made higher. This relationship has been found to exist for weak, moderate, and strong stimuli.

The relationship between the probability of reporting “yes” when the stimulus is present \( p(\text{yes}|\text{stimulus}) \) and the probability of reporting “yes” when it is absent \( p(\text{yes}|\text{no stimulus}) \) can be illustrated by a graph called a receiver-operating characteristic curve or ROC curve (Figure 3.6). Each point on the ROC curve represents the data obtained under a specific \( p(S) \) condition. Each ROC curve represents data obtained for a stimulus of fixed intensity. Thus, the ROC curve is a means of illustrating the often dramatic effects of a nonsensory factor on performance in the detection task. As seen in Figure 3.6, increasing the stimulus intensity influences the function by making it arch higher, and a decrease in
stimulus expectancy is low and therefore the probability of a correct detection should be low. But since the value of a correct detection and the cost of failing to detect a stimulus are very high, the probability of a hit is kept high and the probability of a miss low. Thus, if the payoff for correct detections and the punishment for misses are made great enough, accurate detection can be maintained even in situations where stimulus occurrence is extremely infrequent. But what happens on occasions when the stimulus is not present? Since the payoff for correctly detecting stimuli is high the radar scope observer will frequently say "yes" in the absence of a stimulus (false alarm). Thus, when response consequences are changed, $p(\text{yes}|\text{stimulus})$ and $p(\text{yes}|\text{no stimulus})$ tend to change together. This principle is also evident when false alarms are punished and correct identifications of stimulus absence are rewarded. In this case the probability of "yes" responding when there is no stimulus is kept low, but likewise the probability of detecting a stimulus by saying "yes" when it is present is low.

If stimulus intensity and stimulus probability are held constant in a psychophysical experiment, an ROC curve can be generated by manipulating the costs and values in the detection situation. Typically, at the start of an experimental session the observer is told the stimulus probability and payoff conditions. The payoff conditions are specified by a payoff matrix such as the one shown in Table 3.2. Money is often used to reward correct responses while loss of money is used to

---

**FIGURE 3.6** Receiver-operating characteristic curve. The proportion of "yes" responses when a stimulus is presented is plotted against the proportion of "yes" responses when no stimulus is presented. Each data point corresponds to a different probability of stimulus occurrence. The points on a single curve result from the presentation of a stimulus of a particular intensity.

**FIGURE 3.7** Psychometric functions for the detection of 60-Hz vibration on the fingertip when $p(S)$ was .3 and when $p(S)$ was .7. It is apparent that $p(S)$ has a large effect on the measured threshold. (From Gescheider, Wright, Weber, & Barton, 1971.)
punish incorrect responses. It is the various combinations of costs and values which, by changing the likelihood of reporting a stimulus on a particular trial, result in different data points on an ROC curve. If the value of reporting a stimulus when it is presented is high and the cost of incorrectly reporting it when it is absent is low, the observer will exhibit a high probability of reporting stimuli. However, if the cost is high for incorrectly reporting a stimulus and the value is low for correctly reporting it, the probability of reporting stimuli will be low. It is significant that ROC curves that have been generated by changing payoff conditions have exactly the same form as those generated by changing stimulus probability. It seems that the two variables affect the same psychological process. Some investigators have identified this process as response bias.

Thus at least two nonsensory factors, stimulus probability and response consequences, have been found to have large effects on detection. The evidence for the effects of nonsensory variables on detection performance strongly suggests that the absolute threshold concept does not apply to stimulus detection behavior. If thresholds do in fact exist, it is almost impossible to measure them using classical psychophysical techniques. Human judgments about sensory information are plainly biased by the prevailing conditions of the detection situation.

The early psychophysicists were well aware that response biasing factors such as expectancies and payoff contingencies could contaminate their experimental results. Attempts to control for the effects of biasing factors were made in several ways. Extensive training prior to the experiment which teaches the observer to maintain a consistent approach in making his judgments was one way of obtaining stable data where response bias was held constant, if not eliminated.

Another technique frequently employed to control response bias was the application of statistical procedures to the data. Early experimenters assumed that detection responses would occur on a certain proportion of trials where the stimulus did not exceed threshold. These detection responses were considered false and were attributed to guessing due to response bias. Such false detection responses were often observed when catch trials containing no stimulus were presented. It was assumed that sensory events on catch trials never exceeded threshold and, therefore, that the proportion of false alarms on catch trials would give a good estimation of the guessing rate. The proportion of hits observed in an experiment was thought to be the sum of the proportion of trials where threshold was exceeded by the stimulus and the proportion of trials where the stimulus did not exceed threshold but where the observer guessed anyway and made the detection response. The following equation describes this relationship:

\[ p(\text{hits}) = p^*(\text{hits}) + p(\text{false alarms})[1 - p^*(\text{hits})]. \]  

(3.4)

The empirically obtained proportion of hits, \( p(\text{hits}) \), is equal to the proportion of hits when threshold is exceeded, \( p^*(\text{hits}) \), plus the proportion of hits when threshold was not exceeded, \( p(\text{false alarms})[1 - p^*(\text{hits})] \). The proportion of hits when threshold was not exceeded is the proportion of stimulus trials when threshold was not exceeded, \( 1 - p^*(\text{hits}) \), multiplied by the guessing rate as estimated by the proportion of false alarms on catch trials, \( p(\text{false alarms}) \). Rearrangement of Equation (3.4) yields a correction for guessing that can be applied to the proportion of hits obtained in an experiment:

\[ p^*(\text{hits}) = \frac{p(\text{hits}) - p(\text{false alarms})}{1 - p(\text{false alarms})}. \]  

(3.5)

Implicit in the use of this equation to determine the proportion of hits corrected for guessing, \( p^*(\text{hits}) \), is the assumption of classical psychophysics that a threshold exists. Thus the equation was not merely a statistical tool used by early psychophysicists but was a theoretical statement as well. Fortunately, this theoretical statement can easily be tested, since Equation (3.4) states that the

![Figure 3.8](image-url)  
**Figure 3.8** A family of ROC curves predicted from threshold theory. Each line represents changes in \( p(\text{yes}|\text{stimulus}) \) and \( p(\text{yes}|\text{no stimulus}) \) caused by changes in the guessing rate for the detection of a stimulus of a particular intensity.
relation between the empirically determined proportion of hits and false alarms should be linear. The straight lines start from various values of $p(yes|stimulus)$ when $p(\text{yes}|\text{no stimulus})$ is zero. The higher the signal strength, the higher the value of $p(yes|stimulus)$ will be. As the observer increases his guessing rate, $p(yes|stimulus)$ and $p(\text{yes}|\text{no stimulus})$ should increase linearly until both become 1.0. A family of ROC curves predicted from threshold theory is seen in Figure 3.8.

It is important to note that if Equation (3.5) is applied to the values of $p(yes|stimulus)$ on the ordinate, all of the curves become horizontal lines illustrating the assumed independence of $p(yes|\text{hits})$ and $p(\text{false alarms})$. It is unfortunate for the proponents of threshold theory that ROC curves have been found to deviate consistently from linearity. The curved shape of empirically obtained ROC curves is thought to constitute powerful evidence for rejection of the threshold concept in favor of a new conception of the observer’s behavior in the detection situation (Swets, 1961). We must discard the theory of a threshold that is exceeded only when a stimulus of sufficient strength is presented. But the concept of a lower threshold that is frequently exceeded by spontaneous activity in the nervous system is not in disagreement with the shape of empirical ROC curves. Later in the chapter low threshold theory will be discussed.

### THE THEORY OF SIGNAL DETECTION

The discovery that expectancy and payoff have such a dramatic influence upon detection behavior has been incorporated into a new theoretical conception of the detection situation. Tanner and Swets (1954) proposed that statistical decision theory and certain ideas about electronic signal-detecting devices might be used to build a model closely approximating how people actually behave in detection situations. The model is called the theory of signal detection (TSD) and is described in detail by Green and Swets (1966).

Signals (stimuli) are always detected—whether by electronic devices or by humans—against a background level of activity. The level of this background activity, called noise, is assumed to vary randomly and may be either external to the detecting device or caused by the device itself (e.g., physiological noise caused by spontaneous activity of the nervous system). In the detection situation the observer must therefore first make an observation ($x$) and then make a decision about the observation. On each trial the observer must decide whether $x$ is due to a signal added to the noise background or to the noise alone. When a weak signal is applied, the decision becomes difficult and errors are frequent. One factor contributing to the difficulty of the problem is the random variation of background noise. On some trials the noise level may be so high as to be mistaken for a signal and on other trials it may be so low that the addition of a weak signal is mistaken for noise. This state of affairs can be represented graphically by two probability distributions describing the random variation of noise ($N$) and the signal plus noise ($SN$) (Figure 3.9). Since the signal is added to the noise, the average sensory observation magnitude will always be greater for the signal-plus-noise distribution, $f_{SN}(x)$, than for the noise distribution, $f_{N}(x)$. However, the difference between the means becomes smaller and smaller as the signal strength is decreased, until the distributions are essentially the same. It is when the two distributions greatly overlap that decision-making becomes difficult.

On a specific trial the observer makes a sensory observation $x$ which consists of a sample from one or the other of the distributions and is required to decide on the correct distribution. The ordinate of $f_{N}(x)$ gives the probability density, or likelihood, of $x$ occurring when only noise is presented. Similarly, the ordinate of $f_{SN}(x)$ gives the likelihood of $x$ occurring when a signal is presented. Each value of $x$ can now be expressed in terms of these two likelihoods or probability densities. For each value of $x$ there exists a particular likelihood ratio, $l(x)$, defined as

$$l(x) = \frac{f_{SN}(x)}{f_{N}(x)}.$$  (3.6)

The likelihood ratio provides the observer with a basis for making a decision since it expresses the likelihood of $x$ in the SN situation relative to the likelihood of $x$ in

---

$^1$The term probability density is used because $x$ is continuous rather than discrete. With a limited number of discrete values each could be described as having a particular probability of occurrence. In either case the ordinate gives the relative likelihood of a particular value of $x$. 

---

**Figure 3.9** Theoretical frequency distributions of noise $f_{N}(x)$ and signal plus noise $f_{SN}(x)$ for three different values of signal strength.
3. PSYCHOPHYSICAL THEORY

Criterion

\[ f_N(x) \] \hspace{1cm} \[ f_{SN}(x) \]

"No" region \hspace{1cm} "Yes" region

**Magnitude of sensory observation (x)**

FIGURE 3.10 Theoretical frequency distributions of noise, \( f_N(x) \), and signal plus noise, \( f_{SN}(x) \). The location of the observer's criterion determines whether a particular sensory observation results in a "no" or a "yes" response.

The **N** situation. Even though \( x \) may vary on several dimensions (e.g., hue, saturation, brightness, shape), each \( x \) can be located on a single dimension of likelihood ratio since for each \( x \) there exists single values of \( f_N(x) \) and \( f_{SN}(x) \). Thus, the observer’s final decision of whether \( x \) is due to \( N \) or \( SN \) can be based on a single quantity.

One of the assumptions of TSD is that an observer establishes a particular value of \( l(x) \) as a cutoff point, or **criterion** (\( \beta \)), and that his decision will be determined by whether or not a particular observation \( x \) is above or below the criterion. Proponents of the theory assume that the observer operates by a **decision rule**: when \( l(x) \) is equal to or greater than \( \beta \) he should choose \( SN \), and when \( l(x) \) is below \( \beta \) he should choose \( N \) (Figure 3.10). If the observer properly sets his criterion he will perform optimally in a long series of observations.

Swets, Tanner, and Birdsall (1961) consider the detection situation to be analogous to a game of chance in which three dice are thrown. Two of the dice are ordinary, but the third is a special die with three spots on each of three sides and no spots on the other three sides. When the dice are thrown, the player is told only the total number of spots on all three dice. This information is analogous to the information given for each observation in a detection situation. On the basis of the total number of spots showing, the player must decide whether the usual die showed a zero or a three. Similarly, in the detection situation the observer must decide whether his observation was a product of noise alone or of signal plus noise.

To come out ahead in the long run the player of the dice game would compute the probability of occurrence of each of the possible totals (2 to 12) when the unusual die shows zero and, likewise, the probabilities of each of the totals (5 to 15) when the unusual die shows three. The results could be plotted as two probability distributions and should be thought of as the analogs of the noise and signal-plus-noise distributions (Figure 3.11). Furthermore, as in the detection situation a criterion should be set so that if the total number of spots were greater than some number the player would say "three," and if the total were less than the number he would say "zero." In our example, where the probabilities of a three and a zero are both .30 and the costs and values are the same for the various decision outcomes, the optimal criterion is the point where the two curves cross. In a detection situation, where the stimulus probability is .50 and the costs and values are equal for the various decision outcomes, the optimal criterion is also the point on the observation magnitude dimension where the two distributions cross.

It can be demonstrated mathematically that in the dice game the optimal cutoff point changes when the conditions of the game are changed. For example, if the unusual die were changed to one having three spots on five of the six sides and no spots on only one side, the probability of obtaining three spots would be .83 instead of .50 and the optimal criterion would be lowered. A good player would lower his cutoff point for saying "three." Likewise the optimal criterion would be raised if only one side of the die had three spots and the other five sides had none. In the detection situation, changing the stimulus probability is assumed to have a similar effect on the observer's criterion. In order to perform optimally in a situation where stimulus occurrence is highly probable, an observer will report a signal after a less intense sensory observation than when the stimulus is relatively improbable.

The location of the optimal cutoff point in the dice game is also influenced by changes in the payoff conditions. For example, if the reward is great for correctly saying "three" and the punishment slight for saying "three" when "zero" is correct, the optimal criterion will, of course, be relatively low for saying "three." Rewards and punishments in the detection situation are assumed to have a similar effect on the observer's criterion. A radar scope observer, for instance, maintains a low criterion for reporting signals because of the extreme importance of detecting enemy aircraft and the possible disastrous consequences of failing to do so.

**FIGURE 3.11** Probability distributions for the dice game. (From Swets, Tanner, & Birdsall, 1961. Copyright 1961 by the American Psychological Association. Reprinted by permission.)
The criterion set in the dice game will also be influenced by the degree of overlap of the two distributions. If the number of spots on the unusual die is made zero or four instead of zero or three, the distributions will overlap less because the distribution when four is correct will be shifted further up the scale, and the game becomes easier. Since the point where the two distributions cross is also shifted up the scale, the optimal criterion will be higher. This manipulation in the dice game can be translated into an increase in stimulus intensity in the detection situation. When the signal-plus-noise distribution is shifted to a higher point on the observation magnitude dimension, detection becomes easier and the optimal criterion is higher.

In summary, the detection of energy changes in our environment involves, according to TSD, the establishing of a decision rule in the same way as the efficient playing of a game of chance does. The decision rule is the setting of a criterion determining which hypothesis about a given piece of information will be accepted and which rejected. The location of the optimal criterion is a function of (a) the probabilities of the N and SN presentations, and (b) the costs and values for the various decision outcomes.

In the detection situation, where the costs and values of the various decision outcomes and the probability of signal presentation are precisely known, the optimum value of the criterion, $\beta_{\text{opt}}$, can be calculated by

$$\beta_{\text{opt}} = \frac{p(N)}{p(SN)} \times \frac{\text{value(correct rejection) - cost(false alarm)}}{\text{value(hit) - cost(miss)}}. \quad (3.7)$$

$\beta_{\text{opt}}$ is the value of the likelihood ratio, $l(x)$, which, when used as the criterion, will result in the largest possible winnings in the long run; $p(N)$ is the probability of a noise trial; $p(SN)$ is the probability of a signal-plus-noise trial; value is the amount given to the observer for each correct decision; and cost is the amount taken away from the observer for each incorrect observation.

When the value of $\beta$, as calculated from the judgments of observers, is compared with $\beta_{\text{opt}}$, it is generally found that observers do fairly well at optimizing their winnings. An exception to this rule occurs, however, when $\beta_{\text{opt}}$ is very small or very large, in which case $\beta$ will not be as extreme as $\beta_{\text{opt}}$ and the observer will fail to optimize his winnings. Observers refuse to set extremely low or extremely high criteria even though the conditions of the situation clearly demand such a strategy for optimal performance.

**THEORETICAL SIGNIFICANCE OF THE RECEIVER-OPERATING CHARACTERISTIC CURVE**

One of the main sources of evidence supporting TSD is the experimental manipulation of variables resulting in data plotted as ROC curves. In fact, shapes of ROC curves for various stimulus intensities can be generated from the postulates of the theory and checked against empirical data. It should be recalled that the high threshold theory was rejected because of its failure to predict empirical ROC curves.

The manner in which ROC curves are predicted from TSD is illustrated in Figure 3.12. The upper ROC curve represents a situation where the signal strength is sufficient to result in only a slight overlap of the N and SN probability distributions. The vertical lines represent the locations of the criterion that might be associated with specific conditions of stimulus probability and payoff. Each point on an ROC curve, according to the theory, is determined by the location of the observer's criterion on the $x$ dimension. If an observation is to the right of the criterion the observer will say "yes." The proportion of the area under the curve to
the right of the criterion gives the proportion of "yes" decisions. Therefore, the values of \( p(\text{yes}|\text{N}) \) and \( p(\text{yes}|\text{SN}) \) can be determined by finding the areas under the N and SN distribution curves, respectively, which are located to the right of the criterion. As the criterion is changed from high to low, the values of \( p(\text{yes}|\text{N}) \) and \( p(\text{yes}|\text{SN}) \) change and, when plotted, form an ROC curve. The illustration in the lower half of Figure 3.12 is a predicted ROC curve when the signal strength is so weak as to result in considerable overlap of the N and SN distributions. The finding that the theoretical ROC curves generated in this manner are very similar to those obtained experimentally is strong support for the theory.

In the classical psychophysical experiment the results expressed as thresholds were a function of both stimulus detectability and the location of the observer's criterion. Thus, as a measure of sensitivity to stimuli the threshold may be hopelessly contaminated by changes in the observer's criterion. Such contamination can lead to faulty conclusions about the results of an experiment. For example, there have been cases in which investigators incorrectly attributed large changes in thresholds to changes in the sensitivity of sensory processes. In fact, as revealed by subsequent experimentation, the only thing that had changed was the criterion. TSD and its associated methodology afford a means of independently measuring each of these factors. The theory proposes that \( d' \), a measure of detectability, is equal to the difference between the means of the SN and N distributions \( (M_{\text{SN}} - M_{\text{N}}) \) expressed in standard deviation units of the N distribution \( (\sigma_{N}) \):

\[
d' = \frac{M_{\text{SN}} - M_{\text{N}}}{\sigma_{N}}. \tag{3.8}
\]

Because the location of the SN distribution with respect to that of the N distribution is entirely a function of stimulus intensity and properties of the sensory system, \( d' \) is a pure index of stimulus detectability which is independent of the location of the observer's criterion.

But how can this theoretical concept of signal detectability be measured? Since different \( d' \) values predict different ROC curves, the value of \( d' \) in a particular situation can be ascertained by determining on which member of the family of ROC curves an observer's response probabilities fall. A family of ROC curves corresponding to \( d' \) values ranging from 0 to 3.0 is seen in Figure 3.13. Because only a limited number of curves are generally presented in such a graph, it is best to use them when only approximate values of \( d' \) are needed. Fortunately, simple methods are available for the determination of exact values of \( d' \). The value of \( d' \) can be quickly derived from the empirical values of \( p(\text{yes}|\text{SN}) \) and \( p(\text{yes}|\text{N}) \). By the use of a table of the normal distribution found in any textbook on statistics, these proportions are simply converted into z scores. To obtain \( d' \), the z score for false alarms is subtracted from the z score for hits. The value of \( d' \) can also be determined directly for any combination of false alarm and hit proportions by using a table provided by Elliot (1964).

By reference to a table of the normal distribution, the \( p(\text{yes}|\text{SN}) \) and \( p(\text{yes}|\text{N}) \) values of Figure 3.12 can be converted to z scores. If this is done it will be seen that the \( d' \) value for each pair of z scores will be 2.0 for the upper graph and 1.0 for the lower graph. Furthermore, it should be clear that for a particular separation of the N and SN distributions the value of \( d' \) will remain constant for all possible criterion positions. Thus, an ROC curve is a description of performance changes which are accounted for by a constant \( d' \) and a continuously variable criterion.

It has been experimentally demonstrated for both visual stimuli (Swets, Tanner, & Birdsall, 1955) and auditory stimuli (Tanner, Swets, & Green, 1956) that \( d' \), as a measure of sensitivity, is not contaminated by the effects of variables which shift an observer's response criterion. Furthermore, \( d' \) values, unlike the different threshold values obtained through the use of the various classical psychophysical methods, remain relatively invariant when measured by different experimental procedures. When observers were required to say "yes" or "no" in response to a designated time interval that sometimes contained a signal, \( d' \) estimates were found to approximate those obtained when the observer had to choose one of two designated intervals, one of which always contained a signal (Swets, 1959; Tanner & Swets, 1954).

Once the correct ROC curve has been determined, the location of the observer's criterion, \( \beta \), can be determined by observing exactly where on the ROC curve the point is located. If the point is near the bottom of the ROC curve where the slope is great, the criterion is high; if the point is near the top of the curve where the slope is
slight, the criterion is low. The exact value of $\beta$ is equal to the slope of the ROC curve at a particular point.

To reiterate, $\beta$ is a value of the likelihood ratio. It is the ratio of the ordinate of the SN distribution at the criterion to the ordinate of the N distribution at the criterion, as follows:

$$\beta = \frac{f_{SN}(x) \text{ at criterion}}{f_{N}(x) \text{ at criterion}}$$

(3.9)

Figure 3.10 illustrates that moving the criterion to the right increases the value of $\beta$ and moving it to the left decreases $\beta$. A low value of $\beta$ represents a lax criterion where the observer will be liberal about reporting signals, while a high value of $\beta$ represents a strict criterion where the observer will be conservative about reporting signals.

The value of $\beta$ can be calculated from a pair of hit and false alarm rates. The ordinate of the N distribution at criterion can be estimated as the ordinate value given in the table of the normal distribution that corresponds to the false alarm rate. Likewise the ordinate of the SN distribution at criterion is obtained by converting the hit rate into the ordinate value on the normal distribution curve. For example, ordinate values for a false alarm rate of .20 and a hit rate of .85 are .2801 and .2333, respectively:

$$\beta = \frac{.2333}{.2801} = .83.$$  

If the investigator wishes to study the effects of a particular variable, he is equipped with a technique for finding out whether the effects of the variable are on detectability or on the location of the criterion. He has only to observe whether systematic changes in the variable result in different points along a single ROC curve or points located on different ROC curves. Also the values of $d'$ and $\beta$ can be calculated for various experimental conditions. In some experiments manipulation of an independent variable has led to changes in both $\beta$ and $d'$.

**TESTING THE ASSUMPTIONS OF THE THEORY OF SIGNAL DETECTION**

The form of the ROC curve predicted from TSD can be more easily subjected to experimental tests if the values of $p(\text{yes}|\text{SN})$ and $p(\text{yes}|\text{N})$ obtained in an experiment are plotted on the ROC curve as $z$ scores. If the N and SN distributions are normal in form and also have equal variances ($\sigma^2$), the ROC curves should be linear with a slope of 1.0 when $z$ scores for hits are plotted against $z$ scores for false alarms. In the normal distribution and equal variance situation, when the criterion is shifted by a particular $z$ score distance on the N distribution, it is also shifted by exactly the same distance on the SN distribution. The linearity prediction follows from the assumption that the N and SN distributions are normal in form. The prediction of a slope of 1.0 follows from the assumption of equal N and SN variances. The prediction from TSD is that ROC curves plotted as $z$ scores should be linear with a slope of 1.0, as in Figure 3.14. The points on the two functions were obtained by converting to $z$ scores the $p$ values plotted in the two ROC curves of Figure 3.12. It is not difficult to determine whether or not experimental results confirm the theory. The standard procedure is to determine the best-fitting straight line for the data plotted as $z$ scores. The use of the method of least squares will provide the best estimate of the intercept and slope of the function. If the data points do not significantly deviate from the function, the assumption of normal distribution is supported. If the slope of the function does not significantly deviate from 1.0, the equal variance assumption is supported.

Empirical ROC curves plotted as $z$ scores are almost always linear and therefore there is rather general acceptance of the hypothesis that the N and SN distributions are normal distributions. This kind of analysis of detection data, however, has also revealed that certain details of the original version of TSD were not correct. For example, the assumption that the N and SN distributions have equal variances is
not supported in most experiments. The slope of the ROC curve is frequently found to be less than 1.0. This result is usually explained by assuming that the variance is greater for the SN than for the N distribution. A more general statement of this assumption is that the variance of the SN distribution increases as the mean of the distribution increases.

Figure 3.15 is an ROC curve for an observer in an experiment on the detection of auditory signals (Tanner, Swets, & Green, 1956). Because the ROC curve was asymmetrical, the assumption was made that the standard deviation of the SN distribution (σ_{SN}) was greater than the standard deviation of the N distribution (σ_{N}). The deviation of the results of this experiment from the equal variance of N and SN distributions can better be seen when the data are plotted as z scores (Figure 3.16). If it is assumed that both N and SN distributions are of normal form, then the reciprocal of the slope of the ROC curve is equal to \( \frac{\sigma_{SN}}{\sigma_{N}} \). The slope of the ROC curve for observer 1 is 1.0, and therefore the N and SN distributions have the same variance since \( \sigma_{SN}/\sigma_{N} = 1.0 \). The slope of the ROC curve for observer 2 is .75 and \( \sigma_{SN} \) is greater than \( \sigma_{N} \) since the reciprocal of .75, the value of \( \sigma_{SN}/\sigma_{N} \), is 1.33.

In cases when the variance of the SN distribution is greater than that of the N distribution the symbol \( \Delta m \), rather than \( d' \), is sometimes used to denote the difference between the means of normal N and SN distributions. Thus, the quantities \( d' \) and \( \Delta m \) are symbols for the same measures of signal detectability applied to the cases of equal and unequal variances, respectively.

Note, however, that for observer 2 the value of \( d' \), conventionally determined as the value of the difference between \( z(\text{yes}|N) \) and \( z(\text{yes}|SN) \), is not constant along the ROC curve. In cases where the ROC curve slope is less than 1.0, \( \Delta m \) may be used as the measure of detectability. The value of \( \Delta m \) is equal to the absolute difference between \( z(\text{yes}|N) \) and \( z(\text{yes}|SN) \) at a point where \( z(\text{yes}|SN) \) is equal to 0. Since we start with the mean of the SN distribution \( z(\text{yes}|SN) = 0 \) and \( p(\text{yes}|SN) = .5 \) and determine the corresponding z score for \( \text{yes}|N \), the value of \( \Delta m \) is expressed in the standard deviation units of the N distribution (σ_{N}). In the Tanner et al. (1956) experiment, \( \Delta m \) for observer 2 was 1.35. Notice that for observer 1 the value of \( d' \) is .85 at all points on the ROC curve.

A measure of signal detectability that is sometimes used instead of \( \Delta m \) is \( d' \). The value of \( d' \) is the absolute difference between \( z(\text{yes}|N) \) and \( z(\text{yes}|SN) \) at a point on the ROC curve where it crosses the negative diagonal. The primary benefit of using this measure is that it gives equal weight to \( \sigma_{N} \) and \( \sigma_{SN} \). In the present example the value of \( d' \) for observer 2 is 1.23.

FIGURE 3.15 Receiver-operating characteristic curve for an observer in an experiment on the detection of auditory signals. The asymmetrical ROC curve is consistent with the hypothesis that the noise and signal plus noise distributions are normal in form but have unequal variances. (From Tanner, Swets, & Green, 1956.)

FIGURE 3.16 Receiver-operating characteristic curves plotted in z score units from an experiment on the detection of auditory signals. The open points are for the observer whose data are plotted in Figure 3.15. The filled points are for another observer. (From Tanner, Swets, & Green, 1956.)
Theodor (1972) has pointed out that the correct equation for calculating $d'$ from the proportion of hits and false alarms is

$$d' = \left[ \frac{\sigma_{SN}}{\sigma_N} \right] z(hits) - z(false\ alarms).$$  \hspace{1cm} (3.10)

In cases where the equal variance assumption is made the $z$ score for the proportion of false alarms is simply subtracted from the $z$ score for the proportion of hits. It should be recalled that, as indicated in Equation (3.8), $d'$ is the difference between the means of the $N$ and $SN$ distributions expressed in terms of the standard deviation units of the $N$ distribution ($\sigma_N$). In Equation (3.10), $\sigma_{SN}/\sigma_N$ represents the scaling factor necessary to convert $z(hits)$ into $\sigma_N$ units. The value of $\sigma_{SN}/\sigma_N$ is obtained by calculating the reciprocal of the slope of the ROC curve plotted in $z$-score units. The practical significance of Equation (3.10) is that $d'$ can be calculated for any pair of hit and false alarm proportions as readily when $\sigma_{SN}>\sigma_N$ as when $\sigma_{SN}=\sigma_N$. Since the slope of the ROC curve plotted in $z$ units must be known it is necessary to determine the ROC curve during each experiment. This requirement would be very time consuming if an entire experiment on the effects of signal probability or payoff had to be conducted in order to obtain data points for the ROC curve. Fortunately, the confidence rating procedure discussed later in this chapter provides an exceptionally economical method for obtaining the data needed to determine the ROC curve.

One virtue of TSD made apparent by the above discussion of the ROC curve is that the theory is experimentally testable. Precise, quantitative predictions of what should happen in the detection situation under a variety of conditions can be made from the theory. Experimental results that do not correspond to predicted results have served as a basis for modifying the quantitative statements of the theory. Furthermore, the range of predictions from TSD is comprehensive and extends far beyond the limited confines of earlier psychophysical theories which dealt exclusively with sensory processes. Experimental data continue to accumulate rapidly in support of TSD. The use of detectability measures such as $d'$, $\Delta_m$, and $d_e'$, and of criterion measures such as $\beta$ to separate the observer's sensitivity from the location of his decision criterion therefore becomes increasingly justifiable.

LOW THRESHOLD THEORY

We have seen that high threshold theory predicts ROC curves which are not consistent with experimental data. The theory of signal detection, an alternative theory consistent with empirically determined ROC curves, was then discussed. In TSD the concept of sensory threshold, so central to classical psychophysics, is rejected in favor of an adjustable decision criterion. In fairness to threshold theory, however, it should be pointed out that one version of the threshold concept predicts ROC curves which fit the data about as well as the predictions from TSD. In this theory the threshold is assumed to be much lower, located slightly above the mean of the $N$ distribution, than that of classical threshold theory. The theory was originally described by Swets, Tanner, and Birdsell (1955, 1961). In low threshold theory some of the false alarms on catch trials are due to an observation being above threshold while others are due to guessing when the observation is below threshold. It should be recalled that all false alarms on catch trials were attributed to guessing in classical theory. The shape of the ROC curve predicted from low threshold theory is seen in Figure 3.17. The curved shape of the ROC curve up to $p(yes|N)$ of about .50 is assumed to be due, as it would be in TSD, to a progressive lowering of the observer's decision criterion and would result in an increase in the proportion of hits and false alarms. Beyond this point the function is linear because the criterion is lower than the threshold, which at this point would begin to determine the observer's decisions. For observations below threshold response bias changes the guessing rate and, therefore, this upper branch of the ROC curve must be linear. Thus, the linear segment of the curve and the linear curve of classical threshold theory are both attributed to change in guessing rate.

Luce (1963) proposed another version of low threshold theory. In Luce's theory the threshold is assumed to exist somewhere between the middle and the upper end of the noise distribution. During a sensory observation an observer is in a detect state if the observation exceeds threshold and in the nondetect state if the observation is below threshold. As in other threshold theories, observations below threshold are assumed to be indiscriminable. But observations above threshold are also assumed to be indiscriminable from one another. Thus for the purpose of detecting signals the observer can discriminate between two kind

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{roc_curve.png}
\caption{Receiver-operating characteristic curve predicted from low threshold theory.}
\end{figure}
that both observer and energy thresholds exist, that neither type of threshold exists, or that only one of the two types of threshold exists. We have seen that the evidence for observer thresholds is equivocal. What can be said about energy thresholds?

Even if observer thresholds do not exist, the psychophysical threshold as an experimentally determined energy value necessary for some specified detection performance level can be used to measure sensitivity. In fact, our present understanding of sensory systems would be impossible if it were not for the thousands of experimental studies on energy thresholds conducted since Fechner outlined the psychophysical methods in 1860. The chief limitation of psychophysical energy thresholds as measures of sensitivity is that they may be affected by the observer's decision criterion. If care is taken to control for the effects of criterion shifting, the use of energy thresholds as measures of sensitivity is highly recommended.

The best evidence for energy thresholds comes from studies which illustrate that an observer's performance is the same when a weak stimulus is presented as when no stimulus is presented, unless the weak stimulus is above some critical value. Data from an experiment by Gescheider, Wright, Weber, and Barton (1971) illustrate an energy threshold for the detection of a 60-Hz vibration on the fingertip. Data plotted in Figure 3.19 illustrate that $d'$ was near zero for all stimulus strengths less than 1 μm peak-to-peak displacement of the vibrator contactor. These data, though indicative of an energy threshold, are not inconsistent with TSD. They do not necessarily imply the existence of an observer threshold. Unlike the observer threshold concept, the energy threshold is not tied to the assumption that there is a boundary on the continuum of sensory magnitude below which events cannot be discriminated and above which events can be discriminated. In

![Figure 3.18](image1)

**Figure 3.18** Receiver-operating characteristic curves predicted from two-state theory.

![Figure 3.19](image2)

**Figure 3.19** The relation between $d'$ and signal amplitude for detecting 60-Hz vibration on the fingertip. The data indicate an energy threshold corresponding to a signal amplitude of 1.0-μm peak-to-peak displacement of the stimulator. (From Gescheider, Wright, Weber, & Barton, 1971.)
the context of TSD, an energy threshold simply implies that as stimulus intensity is increased above zero a critical intensity value must be exceeded before the mean of the SN distribution becomes greater than the mean of the N distribution. The use of \( d' \) as a performance measure for determining energy thresholds is an advisable procedure. Under such conditions the value of the threshold will not be contaminated by variations in the observer's judgment criterion. The threshold is defined as the point on the stimulus scale where the observer's \( d' \) first becomes greater than zero.

Vendrik and Eijkman (1968) found that the probability of detecting mechanical and electrical stimuli on the skin does not change until a certain stimulus strength is exceeded. The perception of warmth and cold, on the other hand, showed a linear increasing function throughout the entire range of stimulus intensities. Vendrik and Eijkman concluded that the temperature sensory system does not have a measurable threshold while both the tactile system and the system stimulated by electrical current do have energy thresholds.

In conclusion, a few definite statements can be made about the present status of the threshold concept. It can be said with confidence that if there is an observer threshold it is not the high threshold of classical theory but a much lower threshold located somewhere near the mean of the noise distribution. Since the observer's criterion would usually be higher than this threshold his judgments of signals would be based on whether a sensory observation is above or below criterion rather than above or below threshold. The results of hundreds of experiments have indicated that TSD is basically correct as a model of detection behavior. Further experimentation is needed, however, to determine whether TSD with a low threshold will be needed to account for all the data of detection experiments. Regardless of the outcome of this particular research problem, it can be said that the usefulness of the energy threshold concept is not in question. Energy thresholds can be measured if care is taken to control for the effects of changes in the observer's criterion. Criterion problems can be dealt with by using well-trained observers who can maintain a constant criterion for all experimental conditions or by using TSD measures which are insensitive to changes in the observer's criterion. When properly measured, energy thresholds provide very useful indices of the sensitivity of sensory systems.

THE THREE BASIC PSYCHOPHYSICAL PROCEDURES OF THE THEORY OF SIGNAL DETECTION

The procedures of TSD are designed to provide the psychophysicist with data in the form of response proportions that can be readily converted into the theoretical constructs of sensitivity, criterion, distribution variance, and distribution shape. TSD can be tested by comparing the values of the constructs predicted from the

theory with those that are derived from response proportion data. In those circumstances where the data support the applicability of TSD, the theory can be used to solve many empirical problems. The situation in which a variable is found to have a large effect on response proportions is illustrative. By converting response proportion data into theoretical terms such as \( d' \) and \( \beta \), an investigator can determine whether the effect was due to changes in the observer's sensitivity, his criterion, or both his sensitivity and his criterion. Today there are three basic procedures of TSD used to solve such problems in psychophysics.

The Yes-No Procedure

With the yes-no procedure observers are given a long series of trials, usually more than 300 in a session, in which they must judge the presence or absence of a signal. Some proportion of the trials is SN and the remaining proportion is N. At the start of a session the observer is usually told what the proportion of SN trials will be and what the costs and values associated with the various decision outcomes will be. In many experiments an observation interval is designated on each trial by the presentation of a light, a sound, or some other cue during which SN or N alone is presented in the sensory modality under consideration. In a study on auditory detection, for example, a light one second in duration might be presented every five seconds. The observer must judge as quickly as possible whether or not a tone was presented during the period of time when the light was on. Knowledge of results is usually given after each judgment.

An ROC curve for a single signal strength can be plotted if the proportions of hits and false alarms are obtained for several criterion locations. Generally, payoff contingencies and signal probability are kept constant for an experimental session so that the observer's criterion will remain stable for the session. Data for different criterion levels are often obtained by changing signal probability or payoff contingencies for different sessions. The ROC curve in Figure 3.20 was obtained by Tanner, Swets, and Green (1956) in an experiment on the detection of tones against a background of white noise. Signal probability was either .10, .30, .50, .70, or .90. Each data point on the ROC curve was obtained by using one of these values for a block of 600 trials. As expected the data points ordered themselves on the ROC curve according to signal probability. The ROC curve fitted to the data is the theoretical curve for normal N and SN distributions of equal variance. The theoretical N and SN distributions are shown as an insert in the figure. The value of \( d' \) was .55 and dashed lines in the insert indicate the location of the five criteria that the observer employed for the five signal probabilities.

In the second part of the experiment the same observer was again induced by variation of payoff conditions to vary his criterion for detecting the same stimulus. Signal probability was .50 and the payoff varied from being relatively high for correct responses on SN
The procedure outlined above was used in an experiment to determine the effects of an auditory stimulus on the detection of a tactile signal applied to the fingertip (Gescheider, Barton, Bruce, Goldberg, & Greenspan, 1969). An attempt was made to measure both the detectability of a tactile stimulus and the location of the observer's criterion when the auditory stimulus was set at various intensity levels. The observer was required to decide whether or not a stimulus had been presented in an observation interval. Over a series of such trials a random half of the observation intervals contained a tactile signal while the other half contained no signal. Both the value of $d'$ and the value of $\beta$ were estimated from the proportions of "yes" responses made on signal and on the nonsignal trials. At two different tactile signal intensities $d'$ was found to decrease slightly while $\beta$ increased as a function of the auditory stimulus intensity (Figure 3.22). Thus, the disruptive effect of intense auditory stimulation on tactile signal detection performance is primarily due to the observer setting a relatively high criterion.

When ROC curves are not available to check the validity of the normal distribution and equal variance assumptions, measures of sensitivity not requiring these assumptions should be used whenever possible. Such a nonparametric
assumption is made that in the absence of response bias toward one or more of the observation intervals the observer chooses the observation interval containing the largest sensory observation. Since the observer’s criterion is not a factor in such a judgment the proportion of correct responses, p(c), can be used as a measure of sensitivity. The value of p(c) will be underestimated when response bias toward one of the observation intervals exists. Procedures for correcting the p(c) obtained when response bias exists are found in Green and Swets (1966).

Confidence Rating Procedure

Often it is desirable to obtain an ROC curve from data in a single session within which signal probability and payoff contingencies are fixed. The confidence rating method is very economical since data for several points on an ROC curve can be obtained for a single experimental condition by having the observer make a confidence rating for each of his yes–no judgments. For example, the observer might be instructed to say “five” if he is sure a signal was presented, “four” if he is fairly sure a signal was presented, “three” if he is not sure, “two” if he is fairly sure a signal was not presented, and “one” if he is sure a signal was not presented. It is assumed that to make his ratings the observer sets up n minus 1 criteria along the sensory continuum to delineate his rating categories (Figure 3.23). The number of criteria in Figure 3.23 is one less than the number of categories. In this particular example, “five” is given to observations that are equal to or greater than C5, “four” to observations that are equal to or greater than C4 but less than C5, “three” to observations equal to or greater than C3 but less than C4, “two” to observations equal to or greater than C2 but less than C3, and “one” to observations that are less than C1.

The Forced Choice Procedure

An excellent technique for obtaining a measure of the observer’s sensitivity which is uncontaminated by fluctuations in his criterion is the forced choice procedure. On a particular trial two or more observation intervals are presented and it is the observer’s task to report which observation interval contained a signal.

\[
A' = \frac{1/2 + \frac{4p(hits) - p(false alarms)}{[4p(hits)][1 - p(false alarms)]}}{[1 + p(hits) - p(false alarms)]} \quad (3.11)
\]

In the example above where p(hits) was .84 and p(false alarms) was .50

\[
A' = \frac{1/2 + \frac{.84 - .50}{4 \times .84} \left(1 - .84 - .50\right)}{1 - .84 - .50} = .77.
\]
TABLE 3.3
Determination of an ROC Curve by the Confidence Rating Procedure

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>SN</td>
<td>.01</td>
<td>.08</td>
<td>.28</td>
<td>.36</td>
<td>.27</td>
</tr>
<tr>
<td>N</td>
<td>.10</td>
<td>.23</td>
<td>.47</td>
<td>.17</td>
<td>.03</td>
</tr>
<tr>
<td>C1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

During the experiment the proportion of responses for each of the rating categories for the SN trials and for the N trials are determined. Sample data from an experiment on the detection of a 2.78-μm amplitude vibration on the fingertip (Gescheider, Wright, & Polak, 1971) are shown in Table 3.3. The bottom part of the table lists the calculated hit and false alarm rates that would occur if the observer were induced to set his yes–no decision criterion at each of the four criterion points defined by the five rating categories. For the C4 criterion the estimated hit rate of .27 corresponds to the proportion of "five" responses given on SN trials and the estimated false alarm rate of .03 corresponds to the proportion of "five" responses on N trials. For the C3 criterion the estimated hit rate of .63 is the proportion of "four" responses plus the proportion of "five" responses for the SN trials since this would include all of the SN observations above C3. For the same reason the estimated false alarm rate of .20 for C3 is the proportion of "four" responses plus the proportion of "five" responses for the N trials. The estimated hit rate of .91 for C2 is the proportion of "three" responses plus the proportion of "four" responses plus the proportion of "five" responses on SN trials while the estimated false alarm rate of .67 for C2 is the proportion of "three" responses plus the proportion of "four" responses plus the proportion of "five" responses on N trials. Finally, the estimated hit rate of .99 for C1 is the summation of proportions of "five," "four," "three," and "two" responses for SN trials and the estimated false alarm rate of .90 for C1 is the summation of proportions of "five," "four," "three," and "two" responses for N trials. Each of the four pairs of hit and false alarm proportions that result for this procedure provides a point for an ROC curve.

In Figure 3.24 the open points of the ROC curve for a 2.78-μm signal amplitude are the values found in Table 3.3. The open points were obtained when the observer was expecting weak signals, while the filled points were obtained when he was expecting strong signals. Signal detectability was apparently not affected by changes in signal strength expectancy since the data for both conditions could be fitted by a single ROC curve. However, it is also evident that when the observer is expecting weak signals he sets a lower criterion than when he is expecting strong ones. Notice that this finding held for all values of stimulus amplitude. Changing stimulus amplitude did, however, have a large effect on d', as can be seen from the family of ROC curves.

Because sufficient data can be quickly obtained for constructing an ROC curve by the confidence rating procedure, its use can provide a convenient means of testing the hypotheses of normality of N and SN distributions and equal variance of N and SN distributions. When z(hits) is plotted against z(false alarms), a linear ROC curve is consistent with the normality of distributions assumption and a slope of 1.0 is consistent with the equal variance assumption. When the ROC curve is linear but the slope is not 1.0, the value of \( \sigma_{SN}/\sigma_N \) can be obtained by calculating the reciprocal of the slope of the function.

The validity of confidence rating data is supplied by the finding that the yes–no procedure and the rating procedure generally yield very similar values of signal detectability (Green & Swets, 1966; Markowitz & Swets, 1967). The values of d' obtained from the yes–no procedure and the rating procedure in the study of vibrotactile sensitivity by Gescheider et al. (1971) were plotted against signal amplitude (Figure 3.25). The open and filled points represent d' values obtained from the yes–no procedure when the observer's criterion was low and high, respectively. The squares are d' values from the rating procedure. The correspondence of values obtained with the two methods is remarkable. An important
The continuity–noncontinuity issue has a long history in several areas of psychology besides the study of sensory processes. The essential form of the issue, however, is always the same: do psychological processes change on a continuum or do they change in discrete steps? For years psychologists have been concerned with the problem of whether learning is an all-or-none or continuous process. Does the learning curve represent a number of small discrete increments in learning or does it represent a gradual and continuous change in the amount learned? A closely related problem arises in the study of recognition memory. To be recognized must a stimulus exceed some threshold below which memory strength is zero or must the stimulus exceed a criterion on a memory–strength continuum?

In a recognition–memory test an observer is first exposed to a series of stimulus items such as words, nonsense syllables, or visual nonsense forms which he attempts to remember. Subsequently the observer is exposed to a series of stimulus items, some of which are old items from the earlier series and some of which are new items. The observer is required to report "old" or "new" for each presentation of an item. Reporting "old" for an old stimulus is a hit, while a false alarm is reporting "old" for a new stimulus.

In the TSD model of recognition memory, each item, whether old or new, is assumed to be located on a continuum of memory strength. Variability in memory strength for different items is assumed to form two overlapping normal distributions on the memory-strength continuum (Figure 3.26). The distributions for new and old items are analogous to the N and SN distributions of the detection situation. According to the TSD analysis the observer will report "new" if the memory strength of the item is below a criterion but will report "old" if the memory strength is above the criterion. The experiment might be repeated several times with different sets of items. By inducing the observer to change the location of his criterion for each new experiment several pairs of hit and false alarm proportions can be obtained and plotted as an ROC curve. If the TSD analysis is correct and memory strength varies on a continuum, the ROC curve, called a memory operating characteristic (MOC) curve, should be curvilinear when plotted as proportions and linear when plotted as z scores.

In the noncontinuity model of recognition memory it is postulated that an item either has a suprathreshold memory strength which always results in a recognition response or is below the recognition threshold and results in a recognition response only when the observer guesses. This model of recognition memory is exactly
3. PSYCHOPHYSICAL THEORY

analogous to the high threshold theory of detection. The threshold is assumed to be unaffected by changes in the observer’s expectations and motivation. From the threshold model the MOC curve is predicted to be linear when plotted as proportions. The prediction that the proportion of hits will be a linear function of the proportion of false alarms follows from the assumption that all old items above threshold will be recognized while the new items will never be above threshold and therefore will never be recognized. The false alarm proportion is therefore the rate at which the observer will guess “old” when an item is below threshold and he is in the non-recognition state. The hit rate is equal to the proportion of old items of threshold plus the proportion of old items below threshold multiplied by the false alarm rate. When something is done to the observer to induce him to change his rate of guessing old items, the proportion of false alarms and hits will increase. A review of the discussion of Equation (3.4) will reveal that the logic behind the prediction of a linear ROC curve for detection and a linear MOC curve for recognition memory is exactly the same.

The general finding has been that MOC curves are curvilinear rather than linear when plotted as proportions (see Banks, 1970; Lockhart & Murdock, 1970, for reviews of the experimental findings). Thus a model with a high recognition threshold located above the mean of the new item distribution appears to be untenable, while the data are consistent with the TSD model. A mathematically precise statement of a TSD model of memory, in which old and new items are assumed to vary in memory strength along a continuum, has been developed by Wickelgren and Norman (1966).

CONCLUDING STATEMENT ON THE THEORY OF SIGNAL DETECTION

The theory of signal detection has been a major advancement in experimental psychology. Along with the work of S. S. Stevens on psychophysical scaling it has been responsible for recent intense interest in psychophysics. The discovery that human observers behave in ways closely paralleling statistical decision theory when detecting signals immersed in noisy backgrounds is not evidence that observer thresholds do not exist, however. Experimental data clearly indicate that when the observer is placed in a situation in which there is an N distribution and an SN distribution, he behaves as if he were testing a statistical hypothesis. When he is presented with a sensory observation he tries to decide from which distribution it came. An observer will tend to choose one distribution over the other depending on such circumstances as costs and values of decision outcomes and signal probability. Behavior in such noisy situations does not appear to be governed by an all-or-none threshold: instead decisions appear to be determined by the location of the adjustable decision criterion.

But what can we say of the observer’s behavior in situations where noise is absent or greatly reduced? Stevens (1961b) suggested that TSD applies to noisy situations where the threshold is obscured. He claimed that when noise is absent the all-or-none step function of the threshold will emerge. Accordingly, Stevens (1961b) suggests that

We should continue to explore the fertile and heuristic domain of detection theory (because signals often do in fact occur in noise), and we should study methods for reducing the noise in our experiments on differential sensitivity in order to see how the nervous system operates on pure signals, unobscured by noise. A complete suppression of noise may not be possible, of course, but a sufficient reduction may be achieved to allow a quantal step function to manifest itself in the action of the sensory system [p. 808].

Stevens compared this basic problem in psychophysics to the problem of the nature of electricity:

When, after years of effort, R. A. Millikan finally succeeded in suppressing enough sources of noise in his oil drop experiment, he was able to show that the charge on the electron is not normally distributed, as some evidence had suggested, but has a fixed, all-or-none value [p. 808].

We have seen that a theory of a high threshold near the upper end of the N distribution is not consistent with experimental data. We have also seen that there are no definite experiments which prove that thresholds do not exist at some point relatively low in the N distribution. Furthermore, there are no clear-cut data that disprove the neural quantum theory of the difference threshold. Perhaps further experimentation will reveal that some combination of the postulates of TSD and low threshold theory will best account for the facts of signal detection behavior. Or perhaps the prediction by Urban (1930) of a new psychophysics without the threshold concept as its cornerstone will be confirmed.