

Psychophysical Methods

WALTER H. EHRENSTEIN and ADDIE EHRENSTEIN

■ Introduction

When Fechner (1860/1966) introduced the new transdisciplinary research program of “*Psychophysik*”, his goal was to present a scientific method of studying the relations between body and mind, or, to put it more precisely, between the physical and phenomenal worlds. The key idea underlying Fechner’s psychophysics was that body and mind are just different reflections of the same reality. From an external, objective viewpoint we speak of processes in the brain (i.e., of bodily processes). Considering the same processes from an internalized, subjective viewpoint, we can speak of processes of the mind. In suggesting that processes of the brain are directly reflected in processes of the mind, Fechner anticipated one of the main goals of modern neuroscience, which is to establish correlations between neuronal (objective) and perceptual (subjective) events.

The goal of this chapter is to present Fechner’s techniques and those extensions and modifications of psychophysical methods that may be helpful to the modern neuroscientist with the time-honored objective of discovering the properties of mind and their relation to the brain.

Inner and Outer Psychophysics

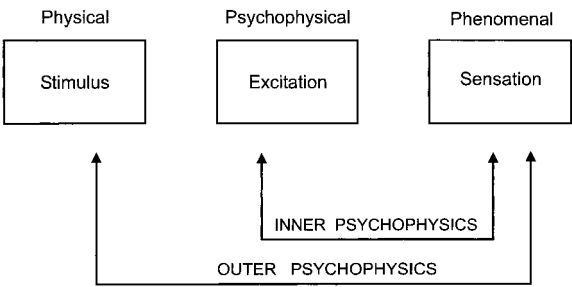
In Fechner’s time there were no physiological methods that enabled the objective recording and study of sensory or neuronal functions. Sensory physiology at that time was essentially “subjective” in that it had to rely on subjective phenomena, that is, on percepts rather than on receptor potentials or neuronal activity. Nonetheless, Fechner referred to neuronal functions in his concept of *inner psychophysics*, or the relation of sensations to the neural activity underlying them (Scheerer 1992). This he distinguished from *outer psychophysics*, which deals with the relation between sensations and the corresponding physical properties and variations of the stimulus itself (see Fig. 1).

For much of the century following Fechner’s publication of *Psychophysik* in 1860, inner psychophysics remained a theoretical concept, whereas the notion of outer psychophysics provided the basis for methods to study sensory and brain processes. The study of subjective phenomena with psychophysical techniques has shaped not only the development of experimental psychology, but also of sensory physiology. Psychophysical methods were used by pioneers in the field of sensory research, such as Aubert, Exner, Helmholtz, Hering, von Kries, Mach, Purkinje and Weber, and provided the basis for

Walter H. Ehrenstein, Institut für Arbeitsphysiologie an der Universität Dortmund, Ardeystr. 67, 44139 Dortmund, Germany (phone: +49-231-1084-274; fax: +49-231-1084-401; e-mail: ehrenst@arb-phys.uni-dortmund.de)

Addie Ehrenstein, Universiteit Utrecht, Department of Psychonomics, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands (phone: +31-30-2534907; fax: +31-30-2534511; e-mail: A.Ehrenstein@fss.uu.nl)

Fig. 1. Fechner's conception of psychophysics. Whereas outer psychophysics was assumed to be based on the methods of physics to describe and control the stimulus, inner psychophysics was a theoretical concept and relied on the methods of outer psychophysics to infer the rules of sensory and neuronal stimulus processing and transformation.



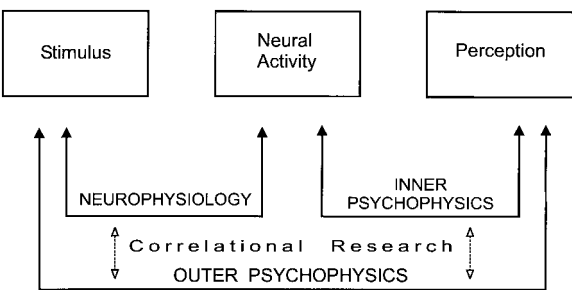
many fundamental insights into and understanding of sensory mechanisms. This psychophysical approach to sensory physiology has come to be referred to as *subjective sensory physiology* (see Jung 1984).

Correlational Research

With the development of various objective methods, such as electrophysiology (e.g., electroencephalography: EEG, Chapter 35 ; visually evoked potentials: VEP, Chapter 36; and single-unit recordings, Chapter 5), magnetoencephalography (MEG, Chapter 37), positron emission tomography (PET, Chapter 39) and functional magnetic resonance imaging (fMRI, Chapter 38), it has become possible to study sensory and brain processes and their locations directly. The relative ease of use and non-invasiveness of most of these techniques has made possible a new interplay between classic psychophysics and modern neuroscience (see Fig. 2). Psychophysical methods have, however, maintained their importance and are used in conjunction with the various objective methods to confirm and complement neurophysiological findings. The complementary research approach that concerns itself with subjective and objective correlates of sensory and neural processes has come to be called *correlational research* (Jung 1961a 1972). This approach, which compares psychophysical and neuronal data on a quantitative, descriptive level (neutral with respect to the question of a material or causal relationship between mental and brain processes), was first established in the study of vision by Jung and colleagues (Jung 1961b, Jung & Kornhuber 1961; Jung & Spillmann 1970; Grüsser and Grüsser-Cornehls 1973). The correlational research approach was soon followed in other sensory areas (see, e.g., Keidel and Spreng 1965; Werner and Mountcastle 1965; Ehrenberger et al. 1966; Borg et al. 1967; Hensel 1976) and has by now become an established venue of research in modern neuroscience (e.g., Spillmann and Werner 1990; Gazzaniga 1995; Spillmann and Ehrenstein 1996;).

As indicated in Fig. 2, the goals of inner psychophysics can be achieved now that the means to directly correlate phenomenal, subjective findings with objective evidence of

Fig. 2. Modern conception of psychophysics. Because of advanced neurophysiological methods, neural activity can be measured objectively, thus allowing for quantitative correlations between psychophysical and neural correlates of perception



sensory and neuronal activity are available. Thus, Fechner's conception of inner psychophysics is no longer dependent on the methodology of outer psychophysics alone. With further progress in correlational research, greater steps in inferring subjective events and perceptual performance by objective techniques are sure to come. For example, perceptual performance losses due to a brain lesion of a given size and location can be examined in great detail with psychophysical tasks. Moreover, in the context of the immensely increased knowledge of sensory and brain functions, inner psychophysics can be addressed much more specifically by choosing stimuli to selectively tap a given mechanism at a certain location. In turn, the hypothesized perceptual (behavioral) significance of a given mechanism or brain area can be determined by means of psychophysical testing (e.g., Wist et al. 1998).

How to Measure Perceptual Experience

Psychophysics starts out with a seeming paradox: It requires the objectification of subjective experience. No apparatus is necessary to obtain percepts; they are immediately present and available to each of us. Thus, the problem is not how to obtain perceptual experience, but how to describe and investigate individual percepts so that they can be communicated and shared by others.

Psychophysics tries to solve this problem by closely linking perceptual experience to physical stimuli. The basic principle is to use the physical stimuli as a reference system. Stimulus characteristics are carefully and systematically manipulated and observers are asked to report their perception of the stimuli. The art of psychophysics is to formulate a question that is precise and simple enough to obtain a convincing answer. An investigation might begin with a simple question such as, "Can you hear the tone?" That is, the task may be one of *detection*.

Sometimes we are not only interested in whether detection has occurred, but in determining which characteristics of the stimulus the observer can identify, e.g., sound characteristics or spatial location. Thus, the problem of sensing something, that of detection, may be followed by that of *identification*.

Detection and identification problems are solved quickly and almost simultaneously when they concern stimuli which are strong and clear. However, under conditions of weak and noisy signals we often experience a stage at which we first detect only that something is there, but fail to identify exactly what or where it is. In such a situation we try to filter out the consistent signal attributes, for instance, the sound of an approaching car, from inconsistent background noise. In such a case, the task is one of *discrimination* of the stimulus, or signal, from a noisy background, and the task is performed under uncertainty. As the car approaches and its sound becomes stronger, the probability of correct discrimination between signal and noise is enhanced. Even if we clearly perceive and identify an object, we may still be faced with the further problem of perceptual judgment, such as, "Is this car dangerously close?" or "Is the rattle under the hood louder than normal?" Questions such as these, concerning "How much x is there?", are part of another fundamental perceptual problem, that of *scaling*, or interpreting, the magnitude of the stimulus on a psychophysical scale.

Outline

In the following sections we will describe the principles of psychophysical methods and give three examples to illustrate their application. First, we present methods that are based on *threshold psychophysics*, starting with the classical procedures along with modern modifications of the classical procedures that allow for adaptive testing. Tech-

niques for control of observer criteria and strategies are also discussed. Second, we describe the methods of *suprathreshold psychophysics*, including the use of reaction time, category scaling, magnitude estimation and cross-modality matching. A third section deals with *comparative psychophysics*, that is, with the special conditions and methods of psychophysical testing in animals.

The description of methods is followed by three specific *examples of psychophysical research*. These examples illustrate how to: (1) study basic mechanisms of adaptation in auditory motion perception, (2) assess impairment of visual function in neurological patients, and (3) measure perceptive fields in monkey and man.

Methods and Procedures

In the following, we will describe the psychophysical tasks and methods that have proven to be most useful in sensory research. Most of the principles are classic, with some having already been worked out by Fechner. The methods of stimulus presentation, response recording and data analysis, however, have been modernized, especially with regard to currently available computer-assisted procedures (see also Chapter 45).

Methods Based on Threshold Measurements

The most basic function of any sensory system is to detect energy or changes of energy in the environment. This energy can consist of chemical (as in taste or smell), electromagnetic (in vision), mechanical (in audition, proprioception and touch) or thermal stimulation. In order to be noticed, the stimulus has to contain a certain level of energy. This minimal or liminal amount of energy is called the *absolute threshold*, and is the stimulus intensity that, according to Fechner, “lifts its sensation over the threshold of consciousness.” The absolute threshold is thus the intensity that an observer can just barely detect. Another threshold, known as the *difference threshold*, is based on stimulus intensities above the absolute threshold. It refers to the minimum intensity by which a variable *comparison* stimulus must deviate from a constant *standard* stimulus to produce a noticeable perceptual difference.

Method of Adjustment

The simplest and quickest way to determine absolute and difference thresholds is to let a subject adjust the stimulus intensity until it is just noticed or until it becomes just unnoticeable (in the case of measurements of the absolute threshold) or appears to be just noticeably different from, or to just match, some other standard stimulus (to measure a difference threshold). The observer is typically provided with a control of some sort that can be used to adjust the intensity, say of a sound, until it just becomes audible (or louder than a standard sound), and then the stimulus intensity is recorded to provide an estimate of the observer’s threshold. Alternatively, the observer can adjust the sound from clearly audible to just barely inaudible (or to match the standard sound), providing another estimate of the threshold. Typically, the two kinds of measurement, that is, series in which the signal strength is increased (ascending series) and series of decreasing signal strengths (descending series) are alternated several times and the results are averaged to obtain the threshold estimate. For example, if a 500-Hz tone is first *heard* at 5 dB on one ascending trial and at 5.5 dB on another, and the tone is first *not* heard at 4 dB on one descending trial and at 4.5 dB on another, the resulting threshold estimate is 4.75 dB.

The following methods of threshold determination differ from the adjustment method in that they do not allow the observer to control the stimulus intensity directly. As they rely on the experimenter's rather than on the subject's control, they provide a more standardized method of measurement.

Method of Limits

In the method of limits, a single stimulus, say a single light, is changed in intensity in successive, discrete steps and the observer's response to each stimulus presentation is recorded. As in the previous method, the stimulus should initially be too weak to be detected, so that the answer is "not seen"; intensity is then increased in steps until the stimulus becomes visible (ascending series), or it is changed from a clearly visible intensity until it becomes invisible (descending series). The average of the intensity of the last "seen" and the first "not seen" stimuli in the ascending trials, or vice versa in the descending trials, is recorded as an estimate of the absolute threshold (for an example, see Table 1). Ascending and descending series often yield slight but systematic differences in thresholds. Therefore, the two types of series are usually used in alternation and the results are averaged to obtain the threshold estimate.

The determination of the difference threshold requires stimuli, such as two flashes of light, which may be presented simultaneously, one next to the other, or successively, one after the other. While the intensity of the standard stimulus is kept constant, the intensity of the comparison stimulus is changed in a series of steps. The comparison stimulus is either initially weaker (ascending series) or initially stronger (descending series) than the standard. A series terminates when the observer's response changes from "weaker" to "stronger" or vice versa. The difference threshold is then the intensity difference between the stimuli of the first trial on which the response differs from the previous one. As before, ascending and descending series are alternated and the results averaged to obtain the threshold estimate.

Method of Constant Stimuli

In the method of constant stimuli the experimenter chooses a number of stimulus values (usually from five to nine) which, on the basis of previous exploration (e.g., using the Method of Adjustment) are likely to encompass the threshold value. This fixed set of

Table 1. Method of Limits. Determination of Absolute Threshold.
Response (Stimulus Perceived): yes (Y), no (N).

Stimulus Intensity	Alternating Ascending and Descending Series					
0	N		N		N	
1	N		N		N	
2	N		N	N	N	
3	N	N	N	Y	N	N
4	N	Y	N	Y	N	Y
5	N	Y	Y	Y	Y	Y
6	Y	Y		Y		Y
7		Y		Y		Y
Transition Points	5.5	3.5	4.5	2.5	4.5	3.5

$$\text{Threshold} = \text{Average Transition Points} = (5.5 + 3.5 + 4.5 + 2.5 + 4.5 + 3.5) / 6 = 24 / 6 = 4$$

Fig. 3. Psychometric function which shows the relationship between the percentage of times that a stimulus is perceived and the corresponding stimulus intensity. The threshold is defined as the intensity at which the stimulus is detected 50 percent of the time.

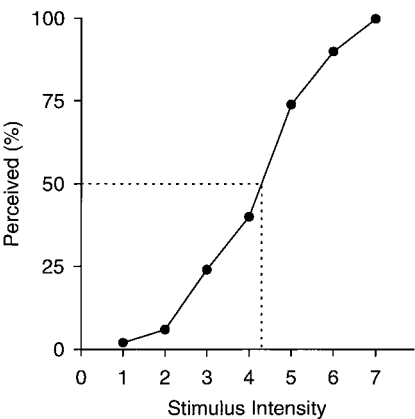


Table 2. Method of Constant Stimuli (50 Presentations for each Stimulus Intensity)

Stimulus Intensity (arbitrary units)	1	2	3	4	5	6	7
Frequency of Perceived Stimuli	1	3	12	20	37	45	50
Percentage of Perceived Stimuli	2	6	24	40	74	90	100

stimuli is presented multiple times in a quasi-random order that ensures each will occur equally often. After each stimulus presentation, the observer reports whether or not the stimulus was detected (for the absolute threshold) or whether its intensity was stronger or weaker than that of a standard (for computing a difference threshold). Once each stimulus intensity has been presented multiple times (usually not less than 20), the proportion of “detected” and “not detected” (or, “stronger” and “weaker”) responses is calculated for each stimulus level (for an example, see Table 2). The data are then plotted with stimulus intensity along the abscissa and percentage of perceived stimuli along the ordinate. The resulting graph represents the so-called *psychometric function* (see Fig. 3).

If there were a fixed threshold for detection, the psychometric function should show an abrupt transition from “not perceived” to “perceived.” However, psychometric functions seldom conform to this all-or-none rule. What we usually obtain is a sigmoid (S-shaped) curve that reflects that lower stimulus intensities are detected occasionally and higher values more often, with intensities in the intermediate region being detected on some trials but not on others. There are various reasons why the psychometric function obeys an S-shaped rather than a sharp step function. A major source of variability are the continual fluctuations in sensitivity that are present in any biological sensory system (due to spontaneous activity or internal noise). Those inherent fluctuations mean that an observer must detect activity elicited by external stimulation against a background level of activity.

In any case, the threshold thus occurs with a certain *probability* and its intensity value must be defined statistically. By convention, the absolute threshold measured with the method of constant stimuli is defined as the intensity value that elicits “perceived” responses on 50% of the trials. Notice that in the example shown in Table 2 and Fig. 3, no stimulus level was detected on exactly 50% of the trials. However, level 4 was detected 40% of the time and level 5, 74% of the time. Consequently, the threshold value of 50% lies between these two points. If we assume that the percentage of trials in which the stimulus is detected increases linearly between these intensities (which is justified given that sigmoid functions are approximately linear in the middle range), we can determine

the threshold intensity by linear interpolation as follows:

$$T = a + (b - a) \cdot \frac{50 - p_a}{p_b - p_a}$$

where T is the threshold, a and b are the intensity levels of the stimuli that bracket 50% detection (with a being the lower intensity stimulus), and p_a and p_b the respective percentages of detection. For the present case we obtain the following result:

$$T = 4 + (5 - 4) \cdot \frac{50 - 40}{74 - 40} = 4 + \frac{10}{34} = 4.29$$

Although the method of constant stimuli is assumed to provide the most reliable threshold estimates, its major drawback is that it is rather time-consuming and requires a patient, attentive observer because of the many trials required.

Adaptive Testing

Adaptive testing procedures are used to keep the test stimuli close to the threshold by adapting the sequence of stimulus presentations according to the observer's response. Since a smaller range of stimuli need be presented, adaptive methods are relatively efficient. An example of such an adaptive procedure is the *staircase method* first introduced by von Békésy (1947), who applied it to audiometry.

The staircase method is a modification of the Method of Limits. A typical application of this method is shown in Fig. 4, where the stimulus series starts with a descending set of stimuli. Each time the observer says "yes" (I can detect the stimulus), the stimulus intensity is decreased by one step. This continues until the stimulus becomes too weak to be detected. At this point we do not, as in the method of limits, end the series, but rather reverse its direction by increasing the stimulus intensity by one step. This procedure continues with increasing the intensity if the observer's response is "no" and decreasing the intensity if it is "yes." In this way, the stimulus intensity flips back and forth around the threshold value. Usually six to nine such reversals in intensity are taken to estimate the threshold, which is defined as the average of all the stimulus intensities at which the observer's responses changed, i.e., the transition points as defined in the Methods of Limits (see Table 1).

Staircase Method

In the staircase method, most of the stimulus values are concentrated in the threshold region, making it a more efficient method than the method of limits. A problem with this *simple staircase* procedure is that an observer may easily become aware of the

Fig. 4. Adaptive testing technique using a single staircase procedure. This example shows a descending staircase for which stimulus intensity is decreased when the stimulus is perceived and increased when it is not perceived.

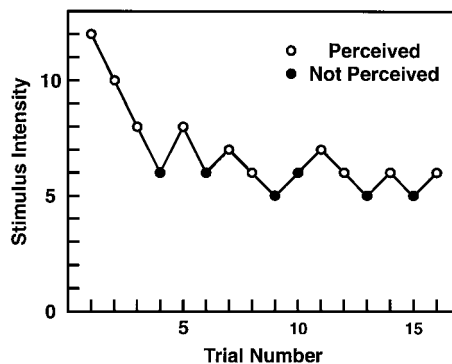
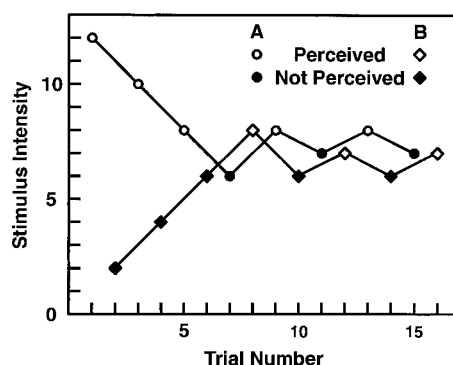


Fig. 5. Example of using two interleaved staircases. Stimuli from the respective descending and ascending staircases are presented on alternate trials.



scheme that governs stimulus presentation, which could lead him or her to anticipate the approach of threshold and change his or her response before the threshold is actually reached.

To overcome this problem, one may use two (or more) *interleaved staircases*, as shown in Figure 5. On trial 1, staircase A begins with a well above threshold intensity. On trial 2, staircase B starts with a well below threshold stimulus. On trial 3, the next stimulus of staircase A is presented, on trial 4, the next stimulus from staircase B, and so on. Over the course of the trials, both staircases converge at the threshold intensity. The two staircases may also be interleaved in a random rather than regular sequence to prevent the observer from figuring out which staircase to expect from trial to trial (Cornsweet 1962).

Best PEST Procedure

An even more efficient adaptive testing method is that of *Parameter Estimation by Sequential Testing* (PEST), which uses maximum-likelihood estimation to select the most efficient ("best") stimulus intensity for a given trial (Best PEST; Lieberman and Pentland 1982). This method, which can easily be implemented on a personal computer, is similar to, but faster and more accurate than, conventional staircase procedures. Basically, it adjusts the amount of change in the stimulus on the basis of information already gathered according to the observer's performance, thereby allowing the collection of more precise information about the threshold on subsequent trials. The Best PEST procedure usually assumes that the underlying psychometric function has a standard sigmoid form; it can, however, also accommodate differently shaped functions as specified by the researcher (Lieberman and Pentland 1982).

Adaptive Variants of the Method of Constant Stimuli

A problem with using constant stimuli is that, because only stimuli near the threshold provide relevant information, many of the stimuli presented are too far away from the threshold to be of use. This inefficiency can be avoided by pre-testing (e.g., using the Method of Adjustment) in order to determine the exact range of critical stimuli to use, tailored to the sensitivity of the observer. Alternatively, one may adapt the stimulus set while the experiment itself is in progress (Farell and Pelli 1999). This approach of *sequential estimation* is preferable over pre-testing since a stimulus set optimized during a short pre-testing period might not be optimal for the entire experiment due to fluctuations in the observer's sensitivity. The most sophisticated and efficient versions of these adaptive strategies combine the experimenter's prior knowledge of the appropriate stimulus range and the observer's response on past trials to choose the signal strength for the next trial using a Bayesian adaptive psychometric method (Watson and Pelli 1983; King-Smith et al. 1994). In the Bayesian adaptive method the threshold is treated as a normally distributed random variable. After each response, the threshold's Gaussian probability density function is updated using Bayes's rule to integrate the pri-

or probability-of-detection information. Each trial is placed at the current maximum-likelihood estimate of threshold, i.e., the mode. The final threshold estimate is also the mode (Farell and Pelli 1999).

Forced-Choice Methods

All the psychophysical methods discussed so far rely on the observer's subjective report of what he or she has perceived. These methods may be termed "subjective," because the experimenter cannot control whether the observer's report is correct or not. In such subjective experiments, results may depend on the criterion that the observer uses for judging whether or not a stimulus was perceived. The *forced-choice method* provides a more objective approach. In this method, the observer is required to make a positive response on every trial – regardless of whether he or she "saw" (or heard, etc.) the stimulus. For example, the stimulus (e.g., a light) may be presented above or below a fixation point, and the observer may be required to indicate – on every trial – which position was occupied by the stimulus. The forced-choice method was first devised by Bergmann (1858; see also Fechner 1860/1966, p. 242). In order to measure visual acuity, Bergmann varied the orientation of a test grating, and instead of asking whether a particular grating was visible, he forced the observer to identify the grating's orientation. A century later, this approach was "re-invented" and achieved an established position in psychophysics (Blackwell 1952).

The use of forced-choice methods reveals that many observers can discern lights so dim or sounds so weak that they claim they cannot see or hear them. For example, if one first measures the absolute threshold for a light by the method of adjustment and then uses this threshold intensity in a forced-choice experiment in which the dim light is flashed either above or below a fixation point and the task is to indicate its location, the performance is often found to be correct throughout. This accomplishment stands in contrast to the observer's prevailing impression that her or his responses were mere guesses and that nothing was actually visible. If the light is then presented again, now at an intensity somewhat below the previously determined threshold, one still can obtain 70 to 75 percent correct choices, which is well above chance level. Typically, forced-choice testing confirms that stimulus intensities can be discerned below the absolute thresholds defined by an unforced, more subjective procedure (Sekuler and Blake 1994).

It appears that the amount of stimulus information necessary to support a decision is greater in an unforced-choice than in a forced-choice situation. A comparison between unforced and forced-choice testing is also useful to factor out possible *criterion* differences among observers. The criterion can be defined as an implicit rule that an observer obeys in converting sensory information into overt responses. It has received much consideration within the theory of signal detection, to be discussed in the next section.

Signal Detection Approach

This psychophysical approach that concentrates on sensory decision processes was derived from *Signal Detection Theory* (SDT; Green and Swets 1966). Its precursors are found in Fechner's *Theory of Discrimination* (Fechner 1860/1966, pp. 85–89; see also Link 1992) and in Thurstone's (1927) *Law of Comparative Judgment*. SDT provides the basis for a set of methods used to measure both the *sensitivity* of the observer in performing some perceptual task and any *response bias* that the observer might have. According to SDT, the sensory evidence that indicates the presence of a stimulus (the "sig-

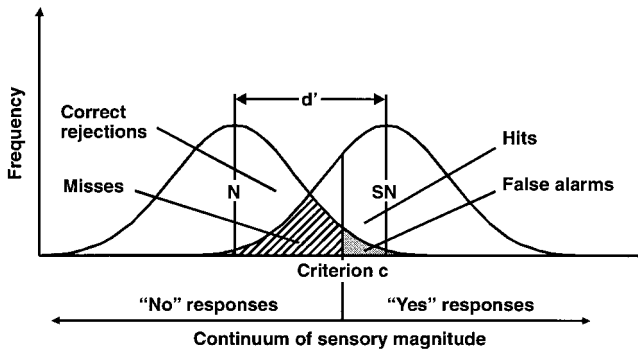


Fig. 6. Assumed psychophysical distributions in a signal detection task. Sensory magnitude (from weak to strong) is plotted from left to right along the horizontal axis and sensory excitation is plotted on the vertical axis. The distribution of noise alone (N) is shown on the left, that of signal plus noise (SN) on the right. The presence of signals is assumed to shift the sensory magnitude, but leaves the shape of distribution unchanged. The index d' refers to observer *sensitivity* and c to the observer's *criterion*.

nal") can be represented on a continuum (the continuum of sensory evidence). The strength of the signal, or, to put it in SDT terms, the evidence that a signal is indeed present, is assumed to vary from trial to trial. That is, the signal is assumed to be characterized best as a distribution of values on a continuum of sensory evidence rather than as a single value. Also, on any given trial there is some "noise" present in addition to the signal. Therefore, trials on which a signal is present are typically called signal-plus-noise trials.

Even on trials where no stimulus is present, there is assumed to be some evidence suggesting that a stimulus might be present. For instance, there may be some background noise or variability in the sensory registration process that is interpreted as evidence of the to-be-detected stimulus. Thus, a distribution of values of "noise" strength is also assumed.

Signal detection methods can be applied whenever there is some overlap of the signal-plus-noise and noise distributions. That is, whenever there is some range of values on the sensory evidence continuum for which the observer is unsure whether a signal was presented or not. If the overlap of the distributions is minimal, the signal-plus-noise and noise trials are relatively easy to tell apart and the observer will appear to be very sensitive. If the distributions overlap more, so that the means of the distributions are relatively close together, observer sensitivity in detecting the signal will be relatively low. Thus, the distance between the signal-plus-noise and noise distributions can be taken as a measure of sensitivity (see Figure 6).

Another assumption of SDT is that, for a given session, the observer sets some response "criterion." If the stimulus energy on a given trial exceeds the criterion, the observer responds, "Yes, a signal was present"; if the criterion is not exceeded, the response is "No, a signal was not present". Because the observer does not know whether the signal was presented or not on a given trial, the criterion is assumed to be the same for both signal-plus-noise and noise trials.

Signal Detection Procedures

Observer sensitivity and response bias are measured by examining performance on both signal-plus-noise and noise trials. First, the "hit" and "false alarm" rates are computed. The hit rate is the probability that the observer said "Yes" when a signal was in fact present. The false alarm rate is the probability that the observer responded "Yes" when the signal was not present. Miss (saying "No" when the signal was present) and

correct rejection (saying “No” when no signal was present) rates can easily be computed from the hit and false alarm rates, respectively (e.g., $1 - \text{Hit Rate} = \text{Miss Rate}$), but these measures are not needed for further computations.

As already mentioned, the hit rate is the probability of correctly identifying a signal. As such, it can be defined as the area of the signal-plus-noise distribution that lies to the right of the response criterion (see Figure 6). Similarly, the false alarm rate is the area of the noise distribution to the right of the criterion. It is often assumed that both the signal-plus-noise and noise distributions are normal and have equal variance. Under this assumption, a table of areas under the normal curve can be used to convert the hit and false alarm rates to measures of distance along the sensory evidence continuum, and, thus, to compute a common *measure of sensitivity*, d' . The formula for d' and an example are given later in the chapter.

An observer's performance will depend not only on how much overlap there is between the signal and noise distributions, but also on where the criterion is located. If it is set at a relatively low level of sensory evidence, the observer will detect most of the signals (i.e., the signal energy will most often exceed the criterion value), but many false alarms will be made as well. Such a lenient observer is called “liberal.” If the criterion is set to a higher level of evidence, fewer false alarms will be made, but the hit rate will also decrease. High criterion settings characterize a stringent or “conservative” observer. High hit rates combined with low false alarm rates are representative of good performance (i.e., high sensitivity).

Receiver Operating Characteristic (ROC) curves can be created for an observer by plotting the hit rate against false alarm rate. All of the points along a single ROC curve reflect the same sensitivity or d' . In other words, ROC curves are *isosensitivity functions* (see Fig. 7). Since the degree to which the ROC curve approaches the upper left-hand corner of the ROC plot depends on the sensitivity of the observer, different curves are generated by manipulating the discriminability of the stimulus or the sensitivity of the observer.

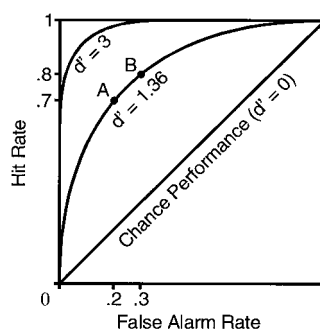
Different points along a single ROC curve represent different levels of bias. Technically, bias is defined as the slope of the ROC curve at a given point. Observers may be naturally biased to respond either liberally or conservatively, but bias can also be manipulated by the experimenter. For instance, the probability of a signal trial may be changed (a lower percentage of signal trials will tend to lead to more conservative performance) or observers may be given different incentives to make relatively few false alarms or, alternatively, relatively few misses.

As mentioned above, a popular measure of observer sensitivity is d' . It is defined as the distance between the means of the signal-plus-noise and noise distributions (see Fig. 6) and is calculated from the observer's hit and false alarm rates. The z-score is a measure

Isosensitivity Functions

Computing Observer Sensitivity

Fig. 7. Receiver operating characteristic curves for three sensitivities (d'). Note that the same sensitivity can result from different hit and false alarm rates: For a given d' , a more conservative criterion is indicated by both lower hit and false alarm rates (A). A more liberal criterion is reflected in higher hit and false alarm rates (B).



of the distance of a score from the mean of a distribution in standard deviation units. Thus, z-scores can be used to measure sensitivity. The z-transformation (z is the inverse of the normal distribution function) converts the hit and false alarm rates to z-scores. The difference between $z(\text{hit rate})$ and $z(\text{false alarm rate})$ tells us how far apart in standardized units the means of the signal-plus-noise and noise distributions are. The formula for d' thus is:

$$d' = z(\text{hit rate}) - z(\text{false alarm rate})$$

A d' of 0 describes chance-level discrimination, i.e., a complete overlap of the signal-plus-noise and signal distributions and a complete lack of discrimination. A d' of 1 is considered moderate performance, a d' of 4.65 (when the hit rate reaches 0.99 with a false alarm rate of 0.01) is considered a ceiling value or optimal performance (see Fig. 7).

Example As an example of applying signal detection methods, consider the case of evaluating the effects of alcohol on performance in a visual detection task. Suppose that 100 trials were presented to an observer, 50 noise trials and 50 signal plus noise trials, on two different occasions. One set of 100 trials was performed without ingesting alcohol, and the other after, say, two drinks.

For each set of trials, we compute the hit and false alarm rates. Suppose that without alcohol, 35 of the signals were correctly detected, giving a hit rate of $35/50 = 0.70$, and suppose that 10 of the noise trials were erroneously classified as signal trials, giving a false alarm rate of $10/50 = 0.20$. Suppose that the hit rate and false alarm rate after ingesting alcohol were 0.80 and 0.30, respectively. Looking at the hit rates, it appears that performance actually improved after ingesting alcohol. To test whether this is really the case, we need to compute d' . To compute d' , we simply look up the hit and false alarm rates in the body of a table of areas under the normal curve, and take the corresponding z-scores (e.g., Macmillan and Creelman 1991). We compute $d' = z(\text{hit rate}) - z(\text{false alarm rate})$ for the two sets of observations. Without alcohol, this result is $0.542 - (-0.842) = 1.366$. With alcohol, the result is $0.842 - (-0.542) = 1.366$.

Naturally, such perfect agreement is seldom seen in the real world (and a statistical test of differences in d' may be required), but the example is clear: *Sensitivity*, as measured by d' , was not affected by alcohol. What did change is the *bias* of the observer. One measure of bias, c , can be computed as $c = -0.5[z(\text{hit rate}) + z(\text{false alarm rate})]$. Accordingly, we find a bias of 0.159 for the alcohol-free performance and a bias of -0.159 for performance with alcohol. Positive values of c reflect more conservative performance, and negative values of c more liberal performance. Thus we can see, in our example, that alcohol had the effect of making our observer more liberal. These data are illustrated in Fig. 7; point A on the ROC curve refers to alcohol-free performance, point B to performance with alcohol.

Suprathreshold Methods

So far we have dealt with methods to determine the observer's threshold or ability to detect ambiguous signals. Since threshold stimuli are per definition difficult to detect or discriminate, these methods cannot be used in situations in which all stimuli are easily perceived. Even when all stimuli are above threshold, not all stimuli are equally easy to detect. For example, some stimuli are more conspicuous than others; they "pop out" and are rapidly discerned. For instance, a red object stands out more against a green background than does a blue one of similar intensity. Methods for comparing performance with above-threshold stimuli are discussed in this section.

Chronometric Methods

To measure suprathreshold stimulus differences, Münsterberg (1894) advocated the use of response or reaction times as the proper tool in psychophysics. Reaction time (RT) is defined as the time between the onset of a stimulus and the overt motor response elicited by the stimulus. In the case where two stimuli are presented on each trial and the observer is to indicate whether they are the same or different, RT can be seen as a measure of the ease with which we can discriminate between the two stimuli, or, according to Münsterberg, as “a measure of their subjective difference” (see also Petrusic 1993).

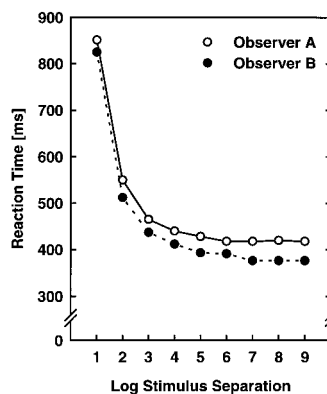
In a typical experiment, on any given trial the observer presses a key to indicate whether the two stimuli were the same or different on a certain dimension, e.g., color. Typically, RT increases as the two stimuli become more similar, or in other words, it increases as the difference between the stimuli decreases. An example with results from two observers in a brightness discrimination task is given in Fig. 8. The stimuli were pairs of squares with luminances ranging from 0.1 to 10 candela/m² in ten logarithmic steps, resulting in 9 perceptually equal stimulus separations (Ehrenstein et al. 1992).

We can distinguish between two varieties of RT: simple and choice. In *simple RT* tasks, a response is made upon detection of any stimulus event. In *choice RT* tasks, the response to be made depends on the identity of the stimulus. Thus, simple RT can be said to require only the detection of a stimulus event, whereas choice RT requires sensory discrimination and further processes, such as stimulus identification and response selection (Sanders 1998).

Generally, simple RTs are used to study sensory performance. When the stimulus intensity is low and near threshold (although still strong enough to be readily distinguishable), the response is slower than when the stimulus is well above threshold. Typically, both simple and choice RT decrease as stimulus intensity and separation increase up to a certain stimulus strength, with little change in RT thereafter (see Fig. 8). Thus, RT differentiates better between sensory performance in the lower than upper intensity range of stimuli.

Simple RT reflects not only variations of stimulus intensity; it can also be affected by other stimulus attributes such as spatial frequency. For instance, with sine-wave gratings of equal mean luminance (i.e., without differences in average intensity) one finds shorter RTs for low spatial frequencies (coarse gratings) than for high spatial frequencies (fine gratings). Breitmeyer (1975), who first showed this spatial-frequency effect on simple RT, concluded that low spatial frequencies preferentially stimulate *transient* visual mechanisms whereas the high spatial frequencies trigger *sustained* visual processing. Breitmeyer’s study is a good example of using an RT task to test sensory mecha-

Fig. 8. Reaction times for same-different judgments as a function of perceptually equal steps of brightness contrast (Ehrenstein et al., 1992). In both observers (A and B), reaction time decreases with increasing contrast in a hyperbolic fashion.



nisms (and their preferred sensitivities) rather than, as done in conventional psychophysics, relating changes in stimulus intensity to that of RT. The distinction between transient and sustained mechanisms has meanwhile received further support from functional anatomic studies (Livingstone and Hubel 1988), resulting in the distinction of two visual pathways, the magnocellular system with good temporal but poor spatial resolution and the parvocellular system with good spatial but poor temporal resolution.

Choice RT is often used to study more complex sensorimotor or cognitive behavior as it involves stages, in addition to detection, such as discrimination, identification and response selection (Proctor and van Zandt 1994; Sanders 1998). In sensory research, however, one has to be careful in ruling out or accounting for these additional, potentially complicating factors. Typically, the observer is provided with two response keys, one for “same” and one for “different” responses. Pairs of stimuli that vary in similarity from identical to very different are presented, and RTs are collected. The observer has to respond as rapidly as possible, but to keep incorrect responses to a minimum (incorrect responses are often identified to the observer, e.g., by sounding a buzzer following an incorrect response, and excluded from the final analysis). Identical stimuli are shown on some trials so that the observer has to decide whether the stimuli are in fact different. However, the subsequent data analysis usually uses only responses to dissimilar pairs of stimuli. RTs for multiple trials (typically 15 to 30) with a given stimulus pair are averaged, and the resulting means or medians are plotted as a function of stimulus separation (as in Fig. 8).

Rank-orderings of RTs can also be further analyzed by using *Multidimensional Scaling* (MDS) methods. MDS refers to various computational procedures by which the degree of perceived similarity between stimuli is mapped to spatial distances. For example, by means of MDS, it is possible to construct the color space of a color-normal observer in comparison to that of a color-deficient observer (Müller et al. 1992).

Chronometric methods are also employed to study behavior that involves cognitive (e.g., attention, short-term memory) processes in addition to perceptual functions, e.g., in order to infer the amount of information processed in the working memory or the complexity of stimulus-response relations (see Ehrenstein et al. 1997; Sanders 1998). Although these studies involve various aspects of perception, they clearly transcend the reach of sensory psychophysics.

Scaling Methods

In addition to the use of RTs, there are other psychophysical approaches that apply to suprathreshold stimulus intensities. Most notable is that of *scaling*. A scale is a rule by which we assign numbers to objects or events. Psychophysical scaling has the goal of assigning numbers to perceptual events.

Category Scaling

First attempts to use scaling as a psychophysical method go back to Sanford in 1898 (Coren et al. 1994, p. 49). Sanford had observers judge 109 envelopes, each of which contained different weights (ranging between 5 and 100 grams), by sorting the envelopes into 5 categories (with category 1 for the lightest weights, category 5 for the heaviest, and the remaining weights distributed in the other categories). The average of all the weights placed in each category was plotted on a logarithmic ordinate, against linear units (category numbers) on the abscissa to obtain a measure of subjective magnitude. This method is called *category scaling*, and since category scaling uses judgment of single stimuli rather than relative comparison between stimuli, it is also referred to as “absolute judgment” (e.g., Haubensak 1992; Sokolov and Ehrenstein 1996).

Category scaling requires more stimulus levels than categories and, as we cannot know in advance what a correct category assignment might be, there are no right or

wrong responses. Sensation scales based on category judgments turn out to be relatively stable over variations such as the number of categories used and the labels applied to the categories (number versus words). They depend, however, on the stimulus spacing, that is, on whether the stimuli vary in steps that are equal in absolute or relative (logarithmic) terms (Marks 1974). Linear spacing is useful if a smaller stimulus range (1–2 log units) is to be scaled; logarithmic spacing is more appropriate when scaling larger stimulus ranges. Category scaling has the advantage of allowing direct responses of the observer to variations in the stimulus. It is, however, limited in that responses are necessarily restricted to relatively few categories. Thus, similar stimuli that may give rise to different sensations may nonetheless be grouped into the same category. Another scaling technique, magnitude estimation, avoids these problems.

In its basic formulation, magnitude estimation is relatively simple. The observer is presented with a series of stimuli in irregular order and instructed to tell how intense they appear by assigning numbers to them. For example, the observer may be asked to assign to the first stimulus, say a sound signal, any number that seems appropriate and then to assign successive numbers in such a way that they reflect subjective differences in perceived intensity. In order to make each number match the perceived intensity, one may use whole numbers or decimals. That is, if one sound appears slightly louder than another, the louder sound may be assigned 1.1, 11 or 110 with the softer sound being 1.0, 10 or 100, respectively. Usually there is no limit to the range of numbers that an observer may use, although sometimes the experimenter may arbitrarily assign a given number to one particular stimulus level. In this case the subject might be told to let a rating of 100 correspond to the loudness of a 65-dB sound, but from there on the ratings would be entirely up to the subject. Given this latitude in assigning numbers to perceived intensities, one might expect the outcome to be a jumble of random numbers, but in fact this does not happen. Whereas the particular numbers assigned to particular stimulus intensities will vary from observer to observer, the order and spacing between numbers show a high degree of regularity among individuals.

Magnitude Estimation

When applying this *direct ratio-scaling* method, the experimenter should use as large a range of stimulus values as possible or practical. This is to assure that the sensory magnitudes will be well distributed over the perceptual range. However, the sessions should be made a reasonable length so as not to overtax the subjects. It is preferable to use multiple sessions rather than one extremely long session.

Instructions in a magnitude estimation task should reflect that magnitude estimation is an estimating rather than a matching approach. That is, it should be made clear that there is no fixed standard stimulus and that there is no modulus. Early experiments typically used both a standard and a modulus. For example, a stimulus might be selected from the middle of the range, assigned the number “10”, and then be presented at a given interval or per subject request. However, the choice of standard has been shown to influence the shape of the obtained psychophysical function and has not proved to have any clear benefits (see Marks 1974).

If the underlying relations are expected to be similar for all individuals tested, the data may be averaged. To average magnitude estimates, the *geometric mean*:

$$M_g = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} = \left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}}$$

is preferred over the *arithmetic mean*:

$$M_a = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

since it is less susceptible to the extreme values that are likely to occur in experiments where subjects pick their own numbers. The distribution of magnitude estimates given to a single stimulus level often approximates log-normality (i.e., is approximately normal after a log transformation – this is another way of saying the distribution has a long tail at the high end of the distribution), and the geometric mean gives an unbiased estimate of the expected value of the logarithms of the magnitude estimates. However, since the geometric mean will be zero when any of the observations are zero, the median (which is generally not preferred because it “wastes” information about the distribution of values) may be preferred when there are zero values (as will occur if some subjects cannot detect some of the stimuli). As Marks (1974) recommends: “Use the geometric mean whenever possible (even if it means dropping a few stimuli that were given zero values), but use the median in a pinch.”

Once an average is decided on, plot the data. Magnitude estimates often plot as lines on double-logarithmic coordinates. Note that a straight line on a double log plot indicates a simple power function between stimulus and sensation.

Determining Psychophysical Functions

It is often the case that the lines representing the relation of a sensation to stimulus values are not quite straight when plotted on log-log paper. Thus, simply fitting a power function to the data may not adequately describe the data. Simply connecting the data points with line segments can serve to show the basic trends in the data, and drawing free-hand curves can sometimes serve to display the trends more clearly, but at the cost of precision. It can be of advantage to fit some simple equation to the data to characterize the psychophysical functions. This can serve to summarize a good deal of data with just a few parameters.

How should one deal with departures from linearity in log-log plots of psychophysical functions? The typical finding is that the functions become steeper as the stimulus values approach zero. More correctly, one could say that the functions become steeper as the threshold of sensation is approached. Stevens (1975) and others have suggested that a threshold parameter should be subtracted from the value of each stimulus. This suggestion follows from the reasoning that, only after the threshold is passed, can we expect a power relation.

Thus, one equation one might fit to the data is:

$$\Psi = k (\phi - \phi_0)^\beta,$$

where Ψ is estimated sensory magnitude, ϕ is stimulus intensity, ϕ_0 is the threshold parameter as defined above, and k and β are constants.

Cross-Modality Matching

Another popular procedure for scaling sensory experience relies exclusively on sensory magnitudes, avoiding the assignment of numbers altogether. In this method, called *cross-modality matching*, an observer equates perceived intensities arising from the stimulation of two different sensory modalities (Stevens 1975). That is, an observer adjusts the intensity of one stimulus (e.g., a light) until it appears to be as intense as a stimulus presented in another modality (e.g., a tone). Cross-modality matching can be regarded as a combination of the methods of scaling with that of adjustment. The adjusted stimulus value, such as the luminance of a light, is recorded by a photometer and taken as an estimate of perceived sound intensity. These data (averaged across repeated trials) are plotted on log-log axes as for magnitude estimation data. Despite the fact that the observers no longer make numerical estimates, the data typically still obey the power function (see above).

Since the subject needs to adjust one of the sensory variables to give a response, rather than simply report a number or assign a category label, cross-modality matching is often more difficult to use than the other scaling methods. Nevertheless, because of its conceptual simplicity (with respect to measurement theory), in that the observer relies on sensory rather than “cognitive” magnitudes (numbers or verbal categories), cross-modality matching can be regarded as an optimal method of suprathreshold sensory psychophysics.

Comparative Psychophysics

Although psychophysical methods were originally developed to study human perception, they can also be used in the study of sensory performance of non-human species. This so-called *comparative psychophysics* has been applied to a wide range of species, including mammals, birds, fish and insects (e.g., Berkley and Stebbins 1990; Blake 1999). Comparative psychophysics is essentially the same as human psychophysics in that it uses highly restricted sorts of stimuli and even more restricted simple responses. Although it is much easier to work with human subjects, the advantages of animal work are obvious. It affords the comparison of anatomical, physiological and behavioral information within a species as well as comparison across species with differently organized neuronal and sensory systems.

Because verbal instructions are of no use in comparative psychophysics, animals need to be trained non-verbally to perform in a sensorimotor task (see also Chapter 44). This training essentially relies on the two main forms of behavioral control or reinforcement, appetitive or aversive. Appetitive control is accomplished by giving food or water, respectively, to a hungry or thirsty animal for correct performance; aversive control relies on presentation of some undesirable stimulus (such as a loud tone or a brief electric shock) in association with a sensory event or in consequence of an incorrect response. Aversive control should evoke some brief, mild degree of discomfort rather than pain.

Behavioral techniques for establishing stimulus control and assessing sensory performance include reflex responses, stimulus-associated (classical) conditioning and operant, behavior-linked conditioning.

Reflex Methods

Reflex responses, such as optokinetic nystagm, the looming reflex and orienting reactions, require no training because the animal's behavior occurs automatically and stereotypically in response to the adequate (releasing) stimulus. For instance, orienting reactions to light refer to the fact that some animals are naturally attracted to light (phototaxis) while others are repelled by light (photophobia). Both stereotypes of behavior can be used in determining a psychophysical threshold for seeing, indicated by the intensity of an optic stimulus necessary to just elicit an orienting reflex. Although reflex methods are attractive to the investigator in that no training is required, their use is limited to reflexes elicited by just a few stimuli. These methods are used when they happen to be so appropriate that conditioning is unnecessary, or when it is not clear that the animal can be conditioned. If a training method seems called for, one may decide on stimulus-associated or operant conditioning methods.

Stimulus-Associated Conditioning

Procedures that use stimulus-associated (also called “classical”) conditioning have proven particularly popular in comparative psychophysics because initial training is simple

and relatively rapid. With these procedures, a sensory stimulus (the conditioned stimulus) is repeatedly paired with an unconditioned stimulus that itself reliably elicits a reflexive (unconditioned) response. Following relatively few such pairings, the sensory stimulus elicits the response on its own (i.e., the conditioned response). For example, an electric shock delivered across the animal's body produces heart-rate acceleration (unconditioned response) which can be registered by implanted electrodes connected to an electrocardiograph. Initially neutral sensory stimuli are paired with the shock stimulus several times until they elicit heart-rate acceleration in the absence of electric shock.

In contrast to the finding that "classical conditioning does not lend itself to the measurement of discrimination thresholds" (Blake 1999, p. 145) in the visual modality, various examples of successful application of classical conditioning in measuring *discrimination thresholds* exist for the auditory modality (Delius and Emmerton 1978; Grunwald et al. 1986; Lewald 1987a,b; Klump et al. 1995).

For example, studies of sound localization in pigeons (Lewald 1987a,b) have used a conditioned (acoustic) stimulus paired with an unconditioned stimulus (a weak electric shock causing an increase in heart rate). In these studies, each experimental series started with a random number (between 5 to 11) of identical presentations of a *reference tone* from one loudspeaker. Subsequently, the conditioned or *test tone* was presented (with the same frequency, intensity and duration as the reference tone, but from another loudspeaker location) and followed by an electric shock. The pigeon was repeatedly exposed to these stimulus conditions until its heart rate accelerated selectively at the onset of the test tone. Starting with a simple task in which the sound locations were well separated, the pigeon was then presented with increasingly more difficult discriminations (smaller angles between the sound locations) until its heart rate failed to differentiate between the test and the reference stimuli. An easier task was then presented to enable the bird to relearn its response before the limits of discrimination, which define the difference threshold, were tested again.

In visual psychophysics, the response readily generalizes to stimuli other than the conditioned stimulus so that it is difficult to test conditions in which an animal is required to respond to one sensory stimulus but not to another. For this reason, stimulus-associated conditioning in vision works primarily for the measurement of *absolute thresholds* where the animal simply detects the presence of a sensory stimulus.

In addition to the limitation that classical conditioning applies unequally well to auditory and visual modalities (Delius and Emmerton 1978; Blake 1999), a general disadvantage of this method is the use of aversive stimuli which may put the welfare of the animal at risk.

Operant Conditioning

The technique of operant or instrumental conditioning refers to procedures where the animal's behavior determines the outcome of a given trial. For comparative psychophysics, instrumental behavior is controlled by a sensory stimulus. For example, animals may learn that a reward is contingent on the presence of one stimulus, but not on another, i.e., the type of stimulus cues the animal about which response produces reward. If an animal is trained to respond to one of two visual stimuli presented simultaneously, it must select which one of the stimuli is reliably connected with reward delivery.

Absolute thresholds are determined by using one stimulus which is always sufficiently weak so as to be invisible (e.g., a zero-contrast grating) paired with an initially visible test grating whose contrast is varied systematically over trials (e.g., by a staircase procedure), to find the luminance contrast at which it becomes barely distinguishable from the blank comparison stimulus.

Difference thresholds are determined with both stimuli visible, but differing in some dimension (e.g., grating contrast, spatial frequency). The animal is initially trained to respond to one of the two easily discriminable stimulus patterns; then the difference between the stimuli is diminished until the value at which the two are barely discriminable is reached. The trial is terminated when the animal fails to respond correctly. As in human psychophysics, the intensity or value of the stimulus can be varied according to the method of constant stimuli or adaptive staircase procedures.

Animal Studies and Inner Psychophysics

If we argued that we cannot know what an animal actually perceives, we should be aware that this problem generalizes to all other individuals, animal and human. Thus the challenge in both animal and human psychophysics is to design and control stimulus conditions that rule out all possible extraneous cues that might support successful detection or discrimination. The study of animal perception has been much advanced by the development and refinement of psychophysical paradigms that can be used to train animals to perform tasks involving sensory detection and discrimination. Some of these methods have also proved useful for the study of human abilities, especially in infants (Atkinson and Braddick 1999).

Comparative psychophysics can be regarded as coming closest to Fechner's *inner psychophysics* in that it allows for the most direct linking of neural activity to corresponding perceptual performance. For example, by recording single cell activity from alert behaving animals, variations in neural responsiveness can be simultaneously compared with perceptual performance under the same task demands and stimulus conditions within the same animal. Such experiments have shown, e.g., that when the stimulus displays are matched for preferred size, speed and direction of motion, *neurometric* and *psychometric* functions (based on neuronal firing rates and perceptual response rates, respectively) become statistically indistinguishable (see Britten, et al. 1992).

Experimental Examples

A: Auditory Motion Aftereffect

(Source: Ehrenstein 1994, in which further references are given).

Problem

After being exposed to visual motion for some time, and then shifting our gaze to a non-moving stimulus, we perceive apparent motion in the direction opposite to that of the previously observed motion. For example, when we shift our gaze from a running waterfall to adjacent rocks, the rocks appear to be moving upward (this is called the *waterfall illusion*). Motion aftereffects are quite compelling, both when induced experimentally (e.g., by gazing at a rotating spiral) or naturally (e.g., by looking at a waterfall). Whereas motion aftereffects are well known and well investigated in vision, little is known about possible analogs of these effects in the auditory modality. Auditory motion aftereffects may be studied by first allowing observers to adapt to auditory motion in a given direction, and then asking them to use a fast psychophysical method, e.g., that of adjustment to indicate the location of a static sound. Biases in the localization of the sound would indicate that observers were influenced by motion aftereffects.

Methods

Apparatus To protect the observer (and the results) from the influence of ambient noise, the experiment was conducted in an acoustically shielded, sound-attenuating room. The sound signal, a narrow-band (0.5 octave band-pass) signal with 1 kHz mean frequency and 60 dB sound pressure level, was generated by a noise generator (Rohde and Schwarz, SUF) linked to a function generator (WAVETEK VCG/VCA, Model 136). Sound pressure level was calibrated using an artificial ear (Brüel and Kjaer 4152) in combination with a sound-level meter (Brüel and Kjaer 2209). By means of a ramp generator and multiplier the signal was presented via headphones (KOSS, Type PRO14AA) with reciprocally increasing intensity (in one ear) and decreasing intensity (in the other ear) to obtain simulated auditory motion. That is, the tone was perceived as moving within the head along a line between the two ears. For instance, with a ramp timing of 3.11 to 0.11 s and a rising ramp presented to the right ear (with the left ear receiving the inverted signal), the sound is heard to move at a moderate speed from left to right, and then to suddenly (within 0.11 s) return from the right ear to its starting position at the left ear. If the ramp inputs to both ears are inverted, motion from right to left is perceived.

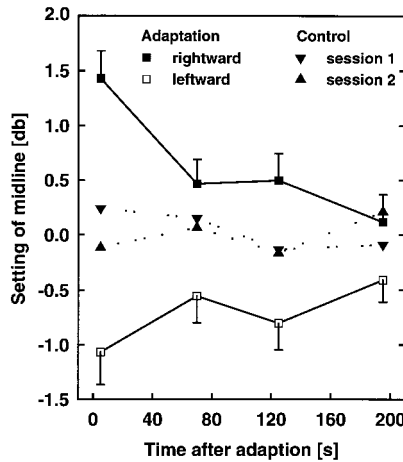
When the signal is not ramped, the same narrow-band stimulus can be presented at various interaural intensities by means of a two-channel logarithmic attenuator. The result of this is that the signal is perceived at different intracranial locations. The perceived location of the sound could be shifted by the observer by adjusting the knob of a ten-turn potentiometer connected to the attenuator, with each rotation resulting in a constant variation of the interaural intensity (ΔI), and with clockwise rotation shifting the perceived location of the sound to the right and counter-clockwise rotation shifting the sound to the left. The actual potentiometer settings were transformed into the respective ΔI -values (dB) and recorded for later analysis.

Selection of Subjects Before testing, subjects were selected according to normal hearing ability, assessed by a standard hearing test (e.g., PHONAK Selector A).

Procedure

1. Subjects were instructed to adjust the ΔI of the sound by turning a knob of the potentiometer so that the sound was perceived as being exactly on the midline between the two ears.
2. Initial practice trials were given to acquaint the subject with the task and, especially, with the psychophysical adjustment procedure. Since the dichotic midpoint has a very salient perceptual quality, most subjects learn the task quite easily.
3. Before the trials on which the subject adapted to simulated auditory motion, each subject made a total of 10 settings of the intracranial auditory midline. Beginning with $\Delta I = 0$, four randomly presented initial starting points ($\Delta I = 3, 6, -3, -6$ dB) were presented and the sequence was repeated once. These settings were used as control measurements.
4. Following the control trials, 10 trials of sound localization after motion adaptation were presented. Simulated motion was presented as an adapting stimulus for approximately 90 s (28 ramp cycles). Immediately following this adaptation, the same narrow-band stimulus was presented without ramping at various interaural intensities (resembling different intracranial locations) and the observer adjusted the location of the sound to correspond to the auditory midline in the same way as in the control measurements. These settings were made over approximately 4 minutes, and the times of the settings were marked for later analysis to plot the time course of any mo-

Fig. 9. Inter-aural midline settings as a function of time after adapting to leftward or rightward dichotic motion (Ehrenstein 1994). Settings are maximally displaced (in opposite direction to that of adapting motion) immediately after adaptation and soon return to control-measurement level.



tion aftereffect. To control for possible effects of different initial starting points, the measurements at each of the respective four non-zero starting points were pooled together.

5. In order to avoid possible confounding effects of motion adaptation in one direction with those caused by adaptation in the other direction, the experiment was performed in two different sessions on different days, and each subject was tested for only one direction of motion within a given session.

Results

Following the exposure to simulated sound motion no counter-motion is heard. That is, the test stimulus appears to be stationary. However, the observer's settings of the inter-aural midline of the sounds were displaced in the direction opposite to the direction of adaptation (on average by 1.2 dB). This displacement effect was reduced as a function of time after adaptation (see Fig. 9).

Conclusion

Having established that an aftereffect of auditory motion occurs and having determined the strength of the effect and its time course by the adjustment method, it would be appropriate to investigate it further with a more sophisticated psychophysical method, for example, with a computer-assisted adaptive staircase procedure or method of constant stimuli.

The observed auditory aftereffect may be analogous to visual motion aftereffects since it is direction-specific; however, because it does not induce apparent motion and occurs with dichotic stimulation, it might also resemble disparity-specific stereoscopic aftereffects.

B: Interocular Latency Differences in Neurological Patients

(Source: Ehrenstein, Manny and Oepen 1985; with further references)

Problem

The optic nerves are among the earliest and most frequently involved sites of demyelinating plaques in multiple sclerosis (with fibers of the foveal ganglion cells being par-

ticularly affected). In the early stages of multiple sclerosis, the conduction time for visual stimuli is often prolonged for one eye. Typically, the change in conduction time is revealed by a prolonged latency in the visually evoked cortical potential (VEP). In this example, we look at a psychophysical procedure for testing foveal differences between conduction times of the two eyes.

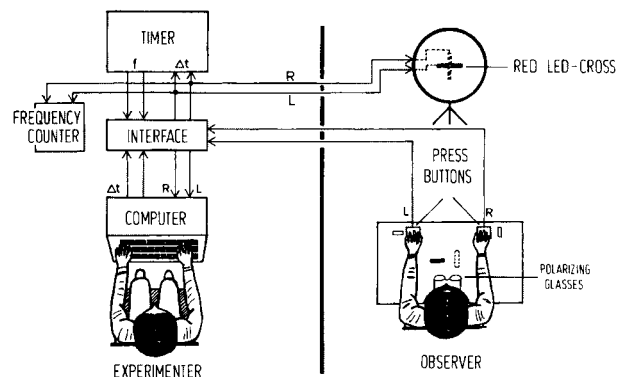
Methods

- Rationale** Assume that due to demyelination the conduction time of the right visual pathway is prolonged by 20 ms. Then a stimulus presented to the right eye has to be delivered 20 ms before a stimulus to the left eye in order for the two to appear simultaneous.
- Subjects** Subjects were patients clinically suspected of multiple sclerosis (based on different groups of definitive, probable, or possible neurological symptoms) and healthy control subjects, matched for age and sex (incidence of multiple sclerosis is much higher in females than in males). All subjects had normal or corrected-to-normal visual acuity.
- Apparatus** The stimulus consisted of a small cross formed by four rectangular light-emitting diodes (LEDs). The LEDs were covered by polarizing filters so that an observer wearing polarizing glasses could see the horizontal bar only with her or his left eye and the vertical bar only with the right eye. A circular, non-polarized LED in the center of the cross served as a fixation cross, visible to both eyes. The horizontal and vertical bars of the LED cross were presented for 80 ms with the onset asynchrony (Δt) of the two bars varied by a 4-channel tachistoscope timer (Scientific Prototype) controlled by a computer in steps of either 30 or 15 ms for the patients, and 15 ms for normal controls (see Fig. 10).

Procedure

A chin-and-forehead rest supported the head of the subject. A warning click preceded each trial by 500 ms. The observer's task was to press one of two keys in order to indicate which bar, horizontal or vertical, appeared first. Observers were allowed to withhold a response whenever both bars appeared to occur simultaneously, or when the order of succession was uncertain.

Fig. 10. Schematic representation of the experimental set-up to measure interocular time thresholds (Ehrenstein et. al., 1985).



The threshold for stimulus onset asynchrony was determined by a modified method of constant stimuli. A series of trials was presented that started with one bar (e.g., the horizontal bar, left-eye stimulus) preceding the other bar by 150ms, with the interval being reduced by 15-ms steps until the vertical bar (right eye stimulus) preceded the horizontal bar by 150 ms; the order of presentation was then reversed, going from vertical bar (right eye) first to horizontal bar (left eye) first. So as not to overtax the patients, only a total of 10 such sequences of trials were given. Thresholds were determined as 50% of the responses left before right, or right before left, respectively. The interocular latency difference (the arithmetic mean of the two time thresholds “right eye before left” and “left eye before right”) was computed for each subject. For example, if an observer needs a delay of -67.5 ms to just see “left before right” and a delay of $+45$ ms to just see “right before left”, the latency difference is $(-67.5 + 45) / 2 = -11.25$ ms.

Initial practice trials were conducted to acquaint the observers with the task, but also to check whether the patients’ time discrimination could be tested within the range of 150ms, or whether a larger Δt (i.e., 300ms with a step size of 30ms) would be needed. Catch trials with zero or maximal temporal delays were inserted at random intervals to check on the observer’s attention and reliability. Non-zero delay trials on which the subject withheld the response were repeated. Patients with motor problems, who found pressing the response key difficult, were allowed to give their responses verbally.

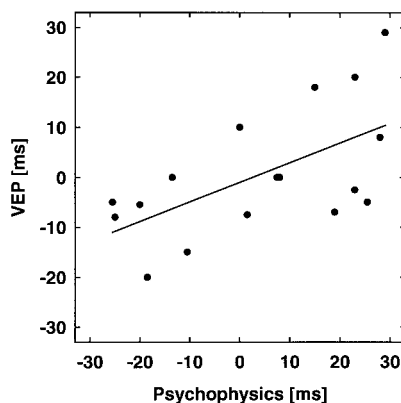
VEPs (P2-latencies) of the patients, recorded under clinical routine conditions in response to foveal stimulation (contrast reversal of a small rectangle), served as an objective reference for comparison with the psychophysical data.

Results

Patients suspected of multiple sclerosis exhibited much higher interocular latency differences (up to 29ms) than normal controls (up to 12ms). This indicates unilateral or asymmetrical impairment of the visual pathways in these patients. VEP and psychophysical latency differences showed a highly significant positive correlation ($r = 0.59$; $p < 0.01$, see correlation diagram in Fig. 11). Psychophysical latencies were better measures of performance than VEP latencies, since they extended over a slightly larger time range.

The psychophysical data also yielded a higher diagnostic validity than the VEP data (higher hit rate in indicating a neurologically confirmed history of optic neuritis). Thus, the diagnostic rate can be significantly raised: from approximately 60% (based on VEP evidence alone) to 90% (VEP and psychophysical testing combined).

Fig. 11. Correlation diagram of the interocular latency differences obtained by VEP recordings and psychophysically in 17 patients suspected of multiple sclerosis (redrawn from Ehrenstein et al., 1985).



Conclusion

It seems to be useful to combine VEP and psychophysical latency measures to diagnose multiple sclerosis. As a clinical routine, the established VEP examination might be preferred, since the recording process is simple and less dependent on the patient's cooperation than the psychophysical procedure. However, psychophysical examination of latency differences is indicated in cases where the VEP provides unclear or conflicting evidence (e.g., in patients with a confirmed history of optic neuritis, but normal VEP latencies).

C: Perceptive Fields in Monkey and Man

(Source: Oehler 1985; see also Spillmann, Ransom-Hogg and Oehler 1987 with further references).

Problem

A key concept in sensory physiology is that of the receptive field. A receptive field is defined as the area on the receptor surface (e.g., retina) within which a change in stimulation leads to a change in the discharge rate of a corresponding single neuron. Jung and Spillmann (1970) introduced the term *perceptive field* to denote the psychophysical correlate of a receptive field, as in the perception of illusory changes in brightness in Hermann grids. Westheimer (1965) introduced a psychophysical paradigm that affords precise measurement of perceptive fields. In the present study perceptive fields were measured by using the same Westheimer paradigm in monkey and human observers.

Methods

Rationale The Westheimer paradigm requires measuring the threshold for a small test spot as a function of background size (Fig. 12). To reduce the effects of stray light, both the test spot and background are centered on a large ambient field. The test spot is varied in luminance according to one of the methods to determine the absolute threshold, such as the method of constant stimuli.

Typical results obtained with this method are shown schematically in Fig. 13. The threshold for detecting the test spot first increases with the diameter of the background, reaching a peak and then decreases and finally levels off with further increases in the background area.

Fig. 12. Stimulus configuration used to measure visual perceptive fields (Westheimer paradigm); redrawn from Spillmann et al. (1987).

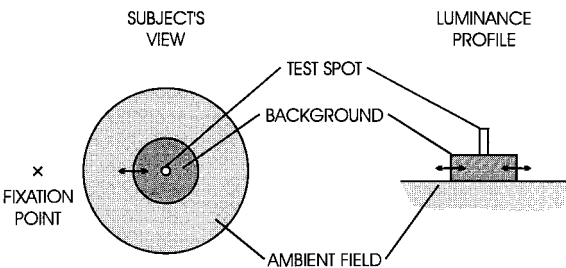
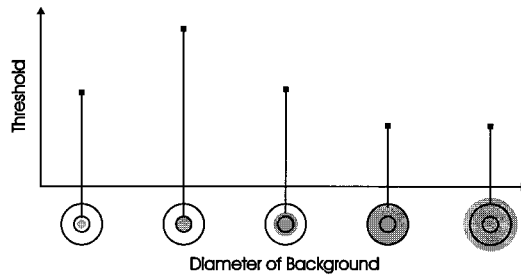


Fig. 13. Typical threshold data as a function of background size and their presumed relation to receptive field organization of an on-center neuron. The threshold increases until the background reaches center size. The threshold decreases once the background encroaches on the inhibitory surround region until the total field size is reached and finally levels off; a further increase in background size does not change the adapting level (redrawn from Spillmann et al., 1987).



The neurophysiological interpretation of this function, relating change in background area to threshold for detection, is depicted (for an on-center neuron) below the abscissa in Fig. 13. As the size of the background increases, the neuronal adapting level is increased, which causes desensitization due to spatial summation of receptor input. As a consequence, the threshold value continues to increase until the background covers the entire receptive field center. As the background encroaches upon the antagonistic surround, the adaptation level is lowered, which leads to sensitization caused by lateral inhibition, hence the threshold decreases. This decrease continues until the background covers the entire receptive field. A further increase in background size does not change the adapting level; hence the threshold remains constant.

According to this interpretation, two measures of perceptive fields can be derived: the center size, indicated by the diameter at which the threshold curve peaks, and the total field size, corresponding to the background diameter at which the threshold curve levels off.

Two rhesus monkeys (*Macaca mulatta*) were tested; one was an adult female of 3.5 kg body weight, the other an adolescent 6 kg male; both were emmetropic as determined by skiascopy (assessing refraction by light reflection through the pupil of the eye). Threshold measurements were repeated with two emmetropic human observers, a 30-year-old woman and a 21-year-old man.

Subjects

The animal, a rhesus monkey, was placed on a "primate chair" with her or his head fixed to it by means of a permanently implanted metal headpiece. Attached to the chair were a response lever and an infrared eye movement monitor. One eye of the monkey was occluded; the monkey used the other eye to fixate the center of a semi-cylindrical screen (see Fig. 14). The stimulus configuration (as shown in Fig. 12) was back-projected onto a translucent screen.

Apparatus

The optical system for presenting the stimuli consisted of two channels, one producing the test spot, the other the background stimulus. The test spot's luminance (produced by a halogen bulb) was varied by a neutral density wedge, and its duration (100 ms) was controlled by an electromagnetic shutter. The background diameter was varied by a series of apertures. The background was centered relative to the test spot by means of an adjusting mirror (*M* in Fig. 14). Background and test spot beams were united by means of a pellicle (*P* in Fig. 14). The ambient field was superimposed onto the luminance configuration using another projector, which projected onto the front side of the screen.

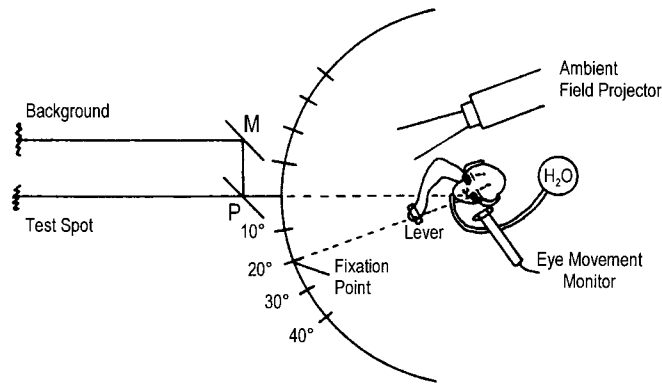


Fig. 14. Schematic apparatus used to measure perceptive fields in monkey and man (redrawn from Oehler 1985).

Training

A reaction-time paradigm was used to train the animal in two steps. First, the animal was trained to fixate on the stimulus location using a dimming task in which the monkey had to pull a lever at the onset of a red fixation light, hold the lever for a random duration (0.4 to 4.9 s), and release it within an allowed RT of 0.8 s following the dimming of the fixation point. Correct responses were rewarded by a drop of water, whereas responses occurring too early or too late were indicated by a high or low tone and punished by a prolonged "time-out." In the second step, the dimming of the fixation point (which was now held at a constant intensity) was replaced by the onset of the test spot as signal for the monkey as when to release the lever.

As the threshold measurements were to be made at defined eccentricities up to 40 degrees in the periphery, the monkey was required to dissociate between the locus of fixation and the focus of visual attention. To control for accurate fixation, eye movements were continuously recorded and trials automatically interrupted when the animal departed from proper fixation.

At the beginning of the training sessions, every correct response was rewarded. After a few days of training, subthreshold stimuli were interspersed with clearly visible stimuli (in an intermittent reinforcement schedule). Before each experimental session, the animal was water-deprived for 22 hours; during each session the monkey obtained 120–200 ml of water.

Threshold Measurement

The method of constant stimuli was used. Releasing the lever within the given time was regarded as a correct response ("stimulus seen"), whereas releasing too late was regarded as "not seen." (Releasing the lever too early occurred rarely, presumably because it was punished by a long time-out.) Each stimulus intensity was presented 10 times in quasi-random order and the rate of correct responses was plotted as a function of test spot luminance with "50% seen" as the threshold value.

Although human observers do not require a special training session to learn the task, threshold measurements were made for comparison under essentially the same conditions as for the monkey, i.e., using the reaction-time paradigm and using a lever to make responses.

Results

With increasing retinal eccentricity, the threshold curves peak at increasingly larger background sizes, indicating an increase of perceptive center size from 0.25° (at 5° eccentricity) to 1.5° (at 40° eccentricity). Total perceptive field size, indicated by the points

where the threshold curves level off, increases from about 1° to 3° (from 0.5° to 40° eccentricity). Rhesus monkey and human data are fairly similar with respect to center size, whereas total perceptive fields are larger for humans.

Conclusion

Psychophysical methods can be used to study neuronal organization underlying visual perception in human and non-human primates. Human perceptive field sizes can be compared to those obtained in trained monkeys; these in turn can be related to receptive field sizes and their morphological substrate, dendritic fields, in monkeys. This allows for correlations among psychophysical, neurophysiological and histological measurements.

Concluding Remarks

The objective of this chapter was to outline some of the main methodological issues of psychophysics in their relation to neuroscientific approaches. As illustrated by the three experimental examples, psychophysical methods can be applied to various aspects of sensory and perceptual problems, from basic and comparative research to clinical diagnostics. Due to the current progress in the cognitive and brain sciences, along with the cumulative progress of psychophysics itself, psychophysics is beginning to emerge as a discipline in its own right. Its methods apply not only to questions concerning the sensory environment and perceptual performance afforded by it (outer psychophysics), but more and more to the functional states and processes of sensory and neural systems (inner psychophysics).

The limitations in the use of psychophysical methods reside in that they require a conscious and cooperating subject, an understanding of the task and a way to reliably report the sensed events. In clinical applications, there is cause for some concern that observers may purposely “cheat” in order to simulate or exaggerate a performance loss. To some extent, methods such as those of the signal detection approach allow the experimenter to check on the observer’s response bias and sensitivity, but these techniques cannot be used to correct for purposely biased reports. Here we reach a principal limitation of subjective testing: Cooperation on the part of the subject is required. This limitation of psychophysical methods can be overcome by combining subjective with objective methods in the examination of sensory function.

We have made a distinction between methods that rely on threshold stimuli (i.e., methods that force the subject to reach a performance limit, i.e., to fail) and supra-threshold conditions which more realistically cover the daily-life range of stimulus intensities. Both approaches of measurement have advantages. Threshold measures allow for an accurate analysis of sensitivity of a detailed sensory function, whereas supra-threshold measures afford the study of performance and integrative functioning of sensory systems over their entire operational range. It is difficult, however, to prescribe in advance the appropriate technique for a particular application (some useful sources of information about current methods and software developments as well as suppliers of experimental equipment and computer software are given in the Chapter Appendix).

As a general rule, one should look for the easiest and most convincing method to answer the experimental question. More sophisticated procedures are not *per se* better than simpler methods, particularly if they introduce difficult technical problems. For example, if accurate timing over very short time intervals is necessary, the use of easily controllable (and inexpensive) LEDs might be preferable to using a computer monitor

that relies on a cathode-ray tube display with complex problems of timing accuracy (depending, e.g., on the sampling rate and on rise and decay functions of the phosphor; see Robson 1999). Sometimes, new methods evolve out of special experimental questions. The problem of alignment of auditory and visual spatial coordinates, for example, led to the development of a new psychophysical technique, i.e., that of using laser pointing to acoustic targets (Lewald and Ehrenstein 1998).

The methods presented here are recipes for psychophysical research, but they need to be complemented by detailed information from the respective field of research application and should be followed flexibly, to stimulate rather than hinder creative modification and development of new experimental paradigms.

■ References

- Atkinson J, Braddick O (1999) Research methods in infant vision. In: Carpenter RHS, Robson JG, eds., *Vision research. A practical guide to laboratory methods*. Oxford University Press, Oxford, pp. 161–186
- Békésy G von (1947) A new audiometer. *Acta Otolaryngol* 35: 411–422
- Bergmann C (1858) Anatomisches und Physiologisches über die Netzhaut des Auges. In: Henle J, Pfeufer C von (eds) *Zeitschrift für rationelle Medicin, Dritte Reihe, II. Band*. Winter, Leipzig & Heidelberg, pp 83–108
- Berkley MA, Stebbins WC, eds. (1990) *Comparative perception*. Wiley, New York.
- Blackwell HR (1952) Studies of psychophysical methods for measuring visual thresholds. *J Opt Soc Amer* 42: 624–643
- Blake R (1999) The behavioural analysis of animal vision. In: Carpenter RHS, Robson JG, eds., *Vision research. A practical guide to laboratory methods*. Oxford Univ. Press, Oxford, pp. 137–160.
- Borg G, Diamant H, Ström L, Zotterman Y (1967) The relation between neural and perceptual intensity: A comparative study on the neural and psychophysical response to taste stimuli. *J Physiol* 192: 13–20
- Breitmeyer BG (1975) Simple reaction time as a measure of the temporal response properties of transient and sustained channels. *Vision Res* 15: 1411–1412
- Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: A comparison of neuronal and psychophysical performance. *J Neurosci* 12: 4745–4765
- Coren S, Ward LM, Enns JT (1994) *Sensation and perception*. Harcourt Brace & Co, Forth Worth TX, 4th ed
- Cornsweet TN (1962) The staircase-method in psychophysics. *Amer J Psychol* 75: 485–491.
- Delius JD, Emmerton J (1978) Stimulus dependent asymmetry in classical and instrumental discrimination learning by pigeons. *Psychol Rec* 28: 425–434
- Ehrenberger K, Finkenzeller P, Keidel WD, Plattig KH (1966) Elektrophysiologische Korrelation der Stevensschen Potenzfunktion und objektive Schwellenmessung am Vibrationssinn des Menschen. *Pflügers Arch* 290: 114–123
- Ehrenstein A, Schweickert R, Choi S, Proctor RW (1997) Scheduling processes in working memory: Instructions control the order of memory search and mental arithmetic. *Q J Exp Psychol* 50A: 766–802
- Ehrenstein WH (1994) Auditory aftereffects following simulated motion produced by varying interaural intensity or time. *Perception* 23: 1249–1255
- Ehrenstein WH, Manny K, Oepen G (1985) Foveal interocular time thresholds and latency differences in multiple sclerosis. *J Neurol* 231: 313–318
- Ehrenstein WH, Hamada J, Müller M, Cavanaugh CR (1992) Psychophysics of suprathreshold brightness differences: a comparison of reaction time and rating methods. *Perception* 21, suppl 2: 82
- Farell B, Pelli DG (1999) Psychophysical methods, or how to measure a threshold, and why. Carpenter RHS, Robson JG, eds., *Vision research. A practical guide to laboratory methods*. Oxford University Press, Oxford, pp. 129–136
- Fechner GT (1860/1966) *Elemente der Psychophysik*. Breitkopf & Härtel, Leipzig (reprinted in 1964 by Bonset, Amsterdam); English translation by HE Adler (1966): *Elements of psychophysics*. Holt, Rinehart & Winston, New York
- Gazzaniga MS, ed (1995) *The cognitive neurosciences*. MIT Press, Cambridge, MA

- Green DM, Swets JA (1966) Signal detection theory and psychophysics. Wiley, New York
- Grüsser OJ, Grüsser-Cornehls U (1973) Neuronal mechanisms of visual movement perception and some psychophysical and behavioral correlations. In: Jung R (ed) Handbook of sensory physiology, vol. VII/3A, Springer, Berlin, pp. 333–429
- Grunwald E, Bräucker R, Schwartzkopff J (1986) Auditory intensity discrimination in the pigeon (*Columba livia*) as measured by heart-rate conditioning. *Naturwiss* 73: 41
- Haubensak G (1992) The consistency model: A process model for absolute judgments. *J Exp Psychol: Hum Perc Perf* 18: 303–309
- Hensel H (1976) Correlations of neural activity and thermal sensation in man. In: Zotterman Y (ed) Sensory functions of the skin in primates. Pergamon Press, Oxford, pp 331–353
- Jung R (1961 a) Korrelationen von Neuronentätigkeit und Sehen. In: Jung R, Kornhuber HH (eds.) *Neurophysiologie und Psychophysik des visuellen Systems*. Springer, Berlin, pp. 410–435
- Jung R (1961 b) Neuronal integration in the visual cortex and its significance for visual information. In: Rosenblith W (ed.) *Sensory communication*. M.I.T. Press, New York, pp. 629–674.
- Jung R (1972) Neurophysiological and psychophysical correlates in vision research. In: Karczmar AG, Eccles JC (eds.) *Brain and Human Behavior*. Springer, Berlin, pp 209–258
- Jung R (1984) Sensory research in historical perspective: Some philosophical foundations of perception. In: Brookhart JM, Mountcastle VB (eds) *Handbook of Physiology*, vol. III. American Physiological Society. Washington DC, pp 1–74
- Jung R, Kornhuber H, eds (1961) *Neurophysiologie und Psychophysik des visuellen Systems*. Springer, Berlin
- Jung R, Spillmann L (1970) Receptive-field estimation and perceptual integration in human vision. In: Young FA, Lindsley DB (eds), *Early experience and visual information processing in perceptual and reading disorders*. National Academy Press, Washington, DC, pp. 181–197
- Keidel WD, Spreng M (1965) Neurophysiological evidence for Stevens' power function in man. *J Acoust Soc Amer* 38: 191–195
- King-Smith PE, Grigsby SS, Vingrys AJ, Benes SC, Supowit A (1994) Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation. *Vision Res* 34: 885–912
- Klump GM, Dooling RJ, Fay RR, Stebbins WC, eds. (1995) *Methods in comparative psychoacoustics*. Birkhäuser, Basel
- Lewald J (1987a) The acuity of sound localization in the pigeon (*Columba livia*). *Naturwiss* 74: 296–297
- Lewald J (1987b) Interaural time and intensity difference thresholds of the pigeon (*Columba livia*). *Naturwiss* 74: 449–451
- Lewald J, Ehrenstein WH (1998) Auditory-visual spatial integration: A new psychophysical approach using laser pointing to acoustic targets. *J Acoust Soc Am* 104: 1586–1597.
- Lieberman HR, Pentland AP (1982) Microcomputer-based estimation of psychophysical thresholds: The Best PEST. *Beh Res Meth Instr* 14: 21–25
- Livingstone M, Hubel D (1988) Segregation of form, color, movement and depth: Anatomy, physiology, and perception. *Science* 240: 740–750.
- Link SW (1992) *The wave theory of difference and similarity*. Erlbaum, Hillsdale, NJ
- Macmillan NA, Creelman CD (1991) *Detection theory. A user's guide*. Cambridge Univ. Press, Cambridge
- Marks LE (1974) *Sensory processes. The new psychophysics*. Academic Press, New York
- Müller M, Cavanus CR, Mollon JD (1992) Constructing the color space of the deuteranomalous observer. In: Drum B., Moreland JD, Serra J (eds.) *Colour vision deficiencies X*. Kluwer, Dordrecht, pp. 377–387
- Münsterberg H (1894) Studies from the Harvard psychological laboratory: A psychometric investigation of the psycho-physic law. *Psychol Rev* 1: 45–51
- Oehler R (1985) Spatial interactions in the rhesus monkey retina: a behavioural study using the Westheimer paradigm. *Exp Brain Res* 59: 217–225
- Petrusic WM (1993) Response time based psychophysics. *Beh. Brain Sci* 16: 158–159
- Proctor RW, Van Zandt T (1994) *Human factors in simple and complex systems*. Allyn & Bacon, Boston
- Robson T (1999) Topics in computerized visual-stimulus generation. In: Carpenter RHS, Robson JG (eds) *Vision research. A practical guide to laboratory methods*. Oxford University Press, Oxford, pp 81–105
- Sanders AF (1998) *Elements of human performance: Reaction processes and attention in human skill*. Erlbaum, Mahwah, NJ

- Scheerer E (1992) Fechner's inner psychophysics: Its historical fate and present status. In: Geissler HG, Link SW, Townsend JT (eds) *Cognition, information processing, and psychophysics*. Erlbaum, Hillsdale NJ, pp 3–21
- Sekuler R, Blake R (1994) *Perception*, 3rd ed. McGraw-Hill, New York
- Sokolov AN, Ehrenstein WH (1996) Absolute judgments of visual velocity. In: Masin S (ed) *Fechner Day 96*, CLEUP, Padua, pp. 57–62.
- Spillmann L, Ehrenstein WH (1996) From neuron to Gestalt: Mechanisms of visual perception. In: Greger R, Windhorst U (eds) *Comprehensive human physiology*, vol. 1. Springer, Berlin, pp 861–893
- Spillmann L, Ransom-Hogg A, Oehler R (1987) A comparison of perceptive and receptive fields in man and monkey. *Hum Neurobiol* 6: 51–62
- Spillmann L, Werner JS, eds (1990) *Visual perception: The neurophysiological foundations*. Academic Press, San Diego.
- Stevens SS (1975) *Psychophysics: Introduction to its perceptual, neural, and social prospects*. Wiley, New York
- Thurstone LL (1927) A law of comparative judgment. *Psychol Rev* 34:273–286.
- Watson AB, Pelli DG (1983) QUEST: a Bayesian adaptive psychometric method. *Perc Psychophys* 33: 113–120.
- Werner G, Mountcastle VB (1965) Neural activity in mechanoreceptive cutaneous afferents: Stimulus-response relations, Weber functions, and information transmission. *J Neurophysiol* 28: 359–397
- Westheimer G (1965) Spatial interaction in the human retina during scotopic vision. *J Physiol* 181: 881–894
- Wist ER, Ehrenstein WH, Schrauf M (1998) A computer-assisted test for the electrophysiological and psychophysical measurement of dynamic visual function based on motion contrast. *J Neurosci Meth* 80: 41–47.

Further Useful Sources of Information

There are two special issues on the use of cathode-ray tube displays in visual psychophysics; they contain papers both on techniques and a series of hardware and software notes: *Spatial Vision*, vol. 10/ 4 (1997) and vol. 11/1 (1997) VSB International Science Publishers, Zeist, NL, <http://www.vspub.com>.

Current information about psychophysical methods and computer software can also be found in the following journals:

- **Behavior Research Methods, Instruments & Computers** Psychonomic Society, <http://www.sig.net/~psysoc/home.htm>.
- **Journal of the Acoustical Society of America** ASA, <http://asa.aip.org>.
- **Journal of Neuroscience Methods** Elsevier, <http://elsevier.nl>.
- **Journal of the Optical Society of America** OSA, <http://www.osa.org>.
- **Perception** Pion Ltd, <http://www.perceptionweb.com>.
- **Perception & Psychophysics** Psychonomic Society, <http://www.sig.net/~psysoc/home.htm>.
- **Vision Research** Pergamon/ Elsevier, <http://elsevier.nl>.

Suppliers Most up-to-date information is now available on the internet. A list of all the interesting sites would soon be obsolete, but some starting points are <http://asa.aip.org> (for auditory psychophysics) and <http://www.visionscience.com> (for visual psychophysics).

Some Suppliers of Experimental Equipment

- Brüel & Kjaer Sound & Vibration, <http://bk.dk>
- Displaytech, <http://www.displaytech.com>
- Hewlett-Packard, <http://www.hp.com>

- ISCAN Inc., <http://www.iscaninc.com>
- Ledtronics, <http://www.ledtronics.com>
- Oriel Instruments, <http://www.oriel.com>
- Permobil Meditech, <http://www.tiac.net/users/permobil>
- SensoMotoric Instruments (SMI), <http://www.smi.de>
- Skalar Medical, <http://www.wirehub.ul/~skalar/eye.htm>
- Texas Instruments, <http://www.ti.com>

Some Suppliers of Computer Software

Eye Lines 2.5 (1997) by W. K. Beagley, <http://www.alma.edu/EL.html>

This software is for generating and running experiments in visual perception and sensorimotor coordination (on any Macintosh or IBM PC). Paradigms include the Method of Adjustment with a set of open-ended tools allowing for real-time adjustments of the size, orientation, position or brightness of stimulus elements on multiple stimulus fields. Data can be sorted by subject, trial or condition.

Reference: Beagley WK (1993) Eye Lines: Generating data through image manipulation, issues in interface design, and teaching of experimental thinking. *Behavior Research Methods, Instruments, & Computers* 25: 333–336.

VideoToolbox (1995) by D. Pelli, <ftp.stolaf.edu/pub/macpsych>

This is a collection of 200 subroutines written in C and several demonstration and utility programs for visual psychophysics with Macintosh computers. Some programs, like the threshold estimation program, are in standard C and can also be used on other computers.

Reference: Pelli DG, Zhand L (1991) Accurate control of contrast on microcomputer displays. *Vision Research* 31: 1337–1350.

Auditory Perception (1994) by Cool Spring Software, <http://users.aol.com/Coolspring>

This program for the Macintosh (with stereo sound output) provides tools to examine auditory acuity (for amplitude and pitch using pure tones and speech), neglect (auditory agnosia, using environmental sounds), and cerebral dominance (dichotic listening). Users may also enter their own auditory stimuli by means of an inexpensive sound digitizer.

Reference: Psychology Software News 5: 95 (April 1995)