barely visible". You then take that same "threshold" light intensity and present it to the same observer in a forced-choice experiment. The dim light is flashed either to the left or to the right of a fixation point while the observer tries to identify its location. Although such comments are not part of the actual data-collection procedure, the observer may frequently volunteer that her responses are mere guesses and that nothing was actually visible. After many trials, you tally up the number of times the observer was correct. Despite the observer's claims, you find that she was correct 100 percent of the time. Next you decrease the light's intensity to a level *below* the threshold value previously determined by the method of adjustment. When you repeat the forced-choice testing with this new, even dimmer light, the observer protests even more strongly, insisting after every trial that she hadn't seen the light and that she was only guessing about its location. Surprisingly, though, the observer continues to do well—perhaps getting 70 to 75 percent of the choices right. Note that this is well above chance level; if the observer were really just guessing, on average only 50 percent of the responses would be correct.

Similar results have been obtained in studies of the other senses. Typically, forced-choice testing confirms that stimuli can be discerned whose intensities are well below the absolute thresholds defined by Fechner's subjective methods. Don't misconstrue the import of this observation. People aren't lying when they make subjective threshold settings; they are relying on information necessary to generate a conscious experience of the stimulus, and the amount of information needed to support that decision appears to be greater than that required for forced-choice performance (where consciousness experience is not a necessary component of the judgment).

Forced-choice testing is also useful for eliminating extraneous (nonsensory) differences among observers. In subjective (nonforced-choice) tests, results can be strongly influenced by the criterion that the observer uses for saying whether or not a light was visible. The **criterion** is the implicit rule that the observer uses to translate sensory information into overt responses. For example, one observer may have a strict criterion; she will not report "seeing" a stimulus unless the sensory evidence is quite strong. Another observer may have a more lax criterion; she is satisfied with weaker sensory evidence. The first observer's responses might lead you to conclude that the observer's threshold was considerably higher than the second observer's when in fact their apparent differences could have been caused by criterion differences alone.

Forced-choice methods, then, should be employed whenever you want to factor out possible criterion differences among observers. The same holds true if you are dealing with groups of observers whose criteria are likely to differ. Since many elderly people tend to be more reluctant about saying "Yes, I detect it," forced-choice methods are useful in comparing the sensory capacities of older and younger observers. To take another example, hospitalized schizophrenics may be reluctant to admit that they see anything that they are not absolutely sure about. So forced-choice methods are important in comparing the vision of schizophrenic and normal observers. In these comparisons and others, if *criterion* differences cannot be ruled out, it is impossible to be conclusive with respect to the source of the differences in *sensory* capacities. More commonly, researchers have to worry about the constancy of a single observer's criterion from one test to another. This would be important if one is interested in whether some treatment—such as perceptual training—changes an observer's ability to see, or whether the training had affected only the observer's willingness to *say* he sees something.

Before concluding this outline of forced-choice methods, we should note that some of Fechner's original methods can be converted into forced-choice versions. You can devise a forced-choice method of limits or a forced-choice method of constant stimuli. In fact, forced-choice staircases have become particularly popular, with the correctness or incorrectness of a person's responses determining whether the stimulus is increased or decreased from trial to trial.

## SENSORY DECISION THEORY

There's another important topic that must be included in any survey of psychophysical methods. That topic is **sensory decision theory (SDT),** a term that covers a set of procedures and a sophisticated psychophysical theory (Green and Swets, 1966). SDT, sometimes also called **signal detection theory** in recognition of its origins in electrical engineering, offers psychophysics two distinct but complementary benefits. One benefit comes from SDT's procedures for expressing precisely and quantitatively what information is contained in some stimulus. In a visual stimulus, for example, the information includes the spatial distribution of light from different parts of the stimulus and a description of how that distribution changes with time. Characterizing the stimulus in this way allows you to determine the efficiency with which a human observer uses the potential information. Defining the potential information contained in a stimulus specifies an upper theoretical limit to observer performance. Knowing that theoretical limit makes it possible to compare the performance of a human observer against the performance of an ideal or perfect observer, that is, one who uses all the stimulus's information. If, as usually happens, information were lost by the eye or by other parts of the visual system, performance would be less than ideal. Take one example. Some light incident on the eye's cornea is either absorbed or reflected rather than transmitted into the eye itself. This loss of light, which is governed by the cornea's structure, reduces the amount of information potentially available in the stimulus. The eye's detection performance is limited, then, because the visual system never gets access to the information lost to absorption or reflection. With appropriate calculations, one can define precisely how much reduction in performance is mandated by this or any other information loss. Geisler (1989) puts this aspect of SDT to excellent use in a psychophysical analysis of the stages in visual processing.

SDT's other main strength lies in its explicit recognition that perceptual measurements are influenced not only by an observer's sensory capacities, but also by various *non*sensory factors. Such nonsensory influences on performance include the observer's criterion, which in turn is determined by motivational variables (Green, 1964). SDT offers a way to distinguish sensory from nonsensory factors, providing separate measures of both sorts of influences (MacMillan and Creelman, 1991). At the completion of a typical SDT experiment, one has two measures of an observer's performance. One measure, $d'$, reflects the observer's sensory capacity; the other measure, $\beta$, (beta), reflects the observer's criterion for acting on the information provided by the senses.*

As we've said, SDT recognizes that in a vision experiment the response "Yes, I see it" depends on two factors—sensory capacity and criterion. To distinguish the two, SDT compares the frequency with which the observer says "Yes" when some dim light *has* been presented and the frequency with which the observer says "Yes" when *no* light has been presented. Take an example. Suppose some observer says "Yes" every single time a very dim light is presented. You might think his eyes are very sensitive. However, you discover that the same observer also says "Yes" when no light whatever is presented. Clearly, you should not take every *yes* at face value.

To achieve its goals, SDT must always compare the observer's responses in two different circumstances. When vision is being assessed, SDT determines the observer's responses to a weak light (called *signal*) as well as to no light (called *noise*). Typically, these signal and noise trials are randomly intermixed. After each trial, the observer responds "Yes" or "No." A "Yes" response to a signal (such as during a trial involving presentation of light) is termed a "hit" (because the observer was correct, or made a hit); a "Yes" response to a noise trial (i.e., one *without* a light) is termed a "false alarm" (because the observer erred, saying he saw something when nothing had been presented). After presenting many signal trials randomly interspersed among noise trials, the experimenter tallies the proportion of signal trials on which the observer responded "Yes" and the proportion of noise trials on which the observer responded "Yes."

The two proportions, the hit rate and false alarm rate, can be plugged into equations to get the sought-after values of $d'$ and $\beta$. It's not necessary to work

*One sensory area to which SDT has been applied is the question of whether acupuncture truly reduces the sensation of pain or merely makes a patient less willing to report the presence of pain. Experimental results with SDT are mixed; some data suggest a change only in the patient's criterion for reporting pain (Clark and Yang, 1974), other data suggest a genuine sensory change (Chapman, Chen and Bonica, 1977).

SDT has also been applied to problems in a wide variety of nonsensory areas. These include the study of memory (Bujey et al., 2000; Dobbins et al., 2000; Slotnick et al., 2000), anxiety (Grossberg and Grant, 1978), medical diagnosis (Emmerich and Levine, 1970; Swets, 1979), identification in police lineups (Wells, Lindsay and Ferguson, 1979), perception of hazards in mine shafts (Blignaut, 1979), and decision making in many other areas (Swets, Dawes and Monahan, 2000). In each, SDT has been useful because of its ability to separate informational and motivational influences on judgments.

through the details of these calculations here; in fact, researchers generally rely on published tables to convert their hit and false alarm rates into $d'$.

To make sure that you've got the idea of SDT, consider a hypothetical experiment. Suppose that two observers are tested with the same dim light and that both achieve the same hit rate. However, one observer produces a higher false alarm rate than the other. Which observer has the higher sensitivity to the dim light? The answer is, the observer whose false alarm rate is lower—her responses indicate that she was superior in discriminating the presence of light from the absence of light. Sensory decision theory treats all tests of detection as tests of an observer's ability to discriminate the presence of a stimulus from its absence (recall Fechner's "inner source of light sensation"). Good discrimination is shown by the combination of a high hit rate and a low false alarm rate. A large difference between an observer's responses when there is a light and when there isn't a light signifies that she can distinguish between the two. Poor discrimination is evidenced when the hit and false alarm rates are equal or nearly equal. In the extreme, when the two rates are equal, you know that the observer has completely failed to distinguish between the absence and presence of a light—performance is at chance level, no matter how high the hit rate might be.

Our discussion of the observer's criterion for reporting sensory information has emphasized differences among observers. But you should recognize that any single observer's criterion varies, depending on a number of factors. SDT specifies how an observer's criterion is likely to change along with changes in the relative importance, to the observer, of hits and false alarms. If an observer is in a situation where it is vital to detect *all* the stimuli and the cost of making a few false alarms is trivial, then SDT predicts the observer will adopt a liberal criterion. This is exemplified in the situation of a radar operator who must monitor the radar for any sign of approach by an enemy. The operator must not miss any possible enemy intrusions whatever. However, if false alarms are costly—in monetary terms or in psychological ones—the observer should adopt a stricter criterion. This is exemplified in the situation of a person who on several successive nights has imagined the sound of a burglar and therefore called the police. He must be very cautious about sounding yet another false alarm.

SDT also specifies how the criterion might change along with changes in the probability that a stimulus will occur. If the observer knows ahead of time that a stimulus is very likely to occur, then optimally one should adopt a liberal criterion for reporting the presence of the stimulus. However, if the observer knows that a stimulus is very unlikely, the optimal strategy is to adopt a strict criterion (that is, to require more powerful evidence before reporting that the very unlikely event has actually occurred).

## PSYCHOPHYSICAL FUNCTIONS FROM PSYCHOMETRIC DATA

Each technique described in the previous sections produces a measure of the threshold, either an absolute threshold or a difference threshold. Any of these technique generates data that specify how performance changes with some stimulus variable that is scaled along an intensive dimension. For instance, Appendix Figure 1 plots the percentage of "seen" responses against light intensity. Curves of this form, as mentioned earlier, are called psychometric functions, and from such a curve we may derive a single intensity value as an estimate of the threshold.

In studying perception we are often interested in how the threshold changes with some stimulus variable. This requires that we measure the threshold repeatedly, over a range of values along some stimulus dimension. For instance, the visual threshold for detecting a spot of light varies depending on the wavelength of that light (recall Figure 3.30); the contrast threshold for detecting a grating pattern varies with the spatial frequency of the pattern (recall Figure 5.13). Likewise, in hearing the threshold intensity for detecting a tone varies with the tone's frequency (recall Figure 11.2 ). These represent measurements of absolute thresholds as the function of different values along a stimulus dimension of interest (for instance, spatial frequency). These kinds of curves—plots of threshold as the function of some stimulus variable—are called **psychophysical functions.** A psychophysical function represents a family of thresholds, not just a single threshold value. Thus, a psychophysical function more completely summarizes the operation of a sensory system than does a single threshold point from that function.

The examples given above are psychophysical functions for detection thresholds. It is also possible to generate a psychophysical function showing how a difference threshold varies with some stimulus variable.