APPENDIX

D

The Signal Detection Approach

In Chapter 1 and Appendix A, we saw that by randomly presenting stimuli of different intensities, we can use the method of constant stimuli to determine a person's threshold—the intensity to which the person reports "I see the light" or "I hear the tone" 50 percent of the time (pages 14, 384). What determines this threshold intensity? Certainly, the physiological workings of the person's eye and visual system are important. But some researchers have pointed out that perhaps other characteristics of the person may also influence the determination of threshold intensity.

To illustrate this idea, let's consider a hypothetical experiment in which we use the method of constant stimuli to measure Lucy's and Cathy's thresholds for seeing a light. We pick five different light intensities, present them in random order, and ask Lucy and Cathy to say "yes" if they see the light and "no" if they don't see it. Lucy thinks about these instructions and decides that she wants to be sure she doesn't miss any presentations of the light. Because Lucy decides to say "yes" if there is even the slightest possibility that she sees the light, we could call her a liberal responder. Cathy, however, is a conservative responder. She wants to be totally sure that she sees the light before saying "yes" and so reports that she sees the light only if she is definitely sure she saw it.

The results of this hypothetical experiment are shown in **Figure D.1**. Lucy gives many more "yes" responses than Cathy does and therefore ends up with a lower threshold. But given what we know about Lucy and Cathy, should we conclude that Lucy's visual system is more sensitive to the lights than Cathy's? It could be that their actual sensitivity to the lights is exactly same, but Lucy's apparently lower threshold occurs because she is more willing than Cathy to report that she sees a light. A way to describe this difference between these two people is that each has a different **response criterion**. Lucy's response criterion is low (she says "yes" if there is the slightest chance a light is present), whereas Cathy's response criterion is high (she says "yes" only when she is sure that she sees the light).

What are the implications of the fact that people may have different response criteria? If we are interested in how one person responds to different stimuli (for example, measuring how a person's threshold varies for different colors of light), then we don't need to take response criterion into account because we are comparing responses within the same person. Response criterion is also not very important if we are testing many people and averaging their responses. However, if we wish to compare two people's responses, their differing response criteria could influence the results. Luckily, an approach called the **signal detection approach** can be used to take differing response criteria into account. We will first describe a signal detection experiment and then describe the theory underlying the experiment.

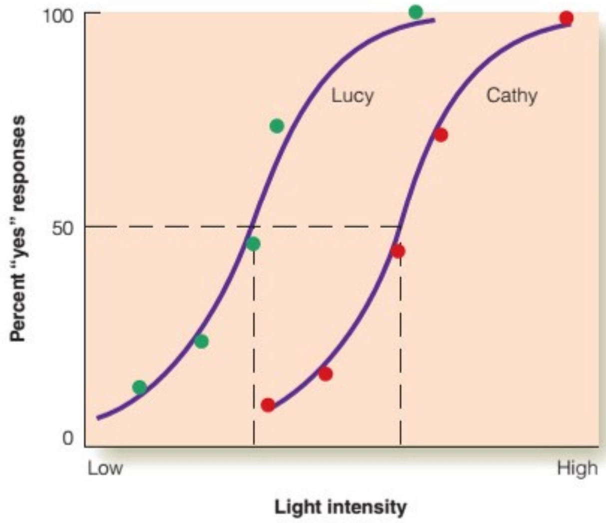

Figure D.1 Data from experiments in which the threshold for seeing a light is determined for Lucy (green points) and Cathy (red points) by means of the method of constant stimuli. These data indicate that Lucy's threshold is lower than Cathy's. But is Lucy really more sensitive to the light than Cathy, or does she just appear to be more sensitive because she is a more liberal responder?

A Signal Detection Experiment

Remember that in a psychophysical procedure such as the method of constant stimuli (see Appendix A), at least five different stimulus intensities are presented and a stimulus is presented on every trial. In a signal detection experiment studying the detection of tones, we use only a single low-intensity tone that is difficult to hear, and we present this tone on some of the trials and present no tone at all on the rest of the trials.

The Basic Experiment

A signal detection experiment differs from a classical psychophysical experiment in two ways: (1) only one stimulus intensity is presented, and (2) on some of the trials, no stimulus is presented. Let's consider the results of such an experiment, using Lucy as our subject. We present the tone for 100 trials and no tone for 100 trials, mixing the tone and no-tone trials at random. Lucy's results are as follows.

When the tone is presented, Lucy

- Says "yes" on 90 trials. This correct response—saying "yes" when a stimulus is present—is called a hit in signal detection terminology.
- Says "no" on 10 trials. This incorrect response—saying "no" when a stimulus is present—is called a miss.

When no tone is presented, Lucy

- Says "yes" on 40 trials. This incorrect response—saying "yes" when there is no stimulus—is called a false alarm.
- Says "no" on 60 trials. This correct response—saying "no" when there is no stimulus—is called a correct rejection.

These results are not very surprising, given that we know Lucy has a low criterion and likes to say "yes" a lot. This gives her a high hit rate of 90 percent but also causes her to say "yes" on many trials when no tone is present, so her 90 percent hit rate is accompanied by a 40 percent false-alarm rate. If we do a similar experiment on Cathy, who has a higher criterion and therefore says "yes" much less often, we find that she has a lower hit rate (say, 60 percent) but also a lower false-alarm rate (say, 10 percent). Note that although Lucy and Cathy say "yes" on numerous trials on which no stimulus is presented, that result would not be predicted by classical threshold theory. Classical theory would say "no stimulus, no response," but that is clearly not the case here. By adding the following new wrinkle to our signal detection experiment, we can obtain another result that would not be predicted by classical threshold theory.

Payoffs

Without changing the tone's intensity at all, we can cause Lucy and Cathy to change their percentages of hits and false alarms. We do this by manipulating each person's motivation by means of payoffs. Let's look at how payoffs might influence Cathy's responding. Remember that Cathy is a conservative responder who is hesitant to say "yes." But being clever experimenters, we can make Cathy say "yes"" more frequently by adding some financial inducements to the experiment. We tell Cathy that we are going to reward her for making correct responses and are going to penalize her for making incorrect responses by using the following payoffs.

Hit:	Win $100
Correct rejection:	Win $10
False alarm:	Lose $10
Miss:	Lose $10

What would you do if you were in Cathy's position? You realize that the way to make money is to say "yes" more. You can lose $10 if a "yes" response results in a false alarm, but this small loss is more than counterbalanced by the $100 you can win for a hit. Although you decide not to say "yes" on every trial—after all, you want to be honest with the experimenter about whether you heard the tone—you decide to stop being so conservative. You decide to change your criterion for saying "yes." The results of this experiment are interesting. Cathy becomes a more liberal responder and says "yes" a lot more, responding with 98 percent hits and 90 percent false alarms.

This result is plotted as data point L (for "liberal" response) in **Figure D.2**, a plot of the percentage of hits versus the percentage of false alarms. The solid curve going through point L is called a receiver operating characteristic (ROC) curve. We will see why the ROC curve is important in a moment, but first let's see how we determine the other points on the curve. Doing this is simple: all we have to do is to change the payoffs. We can make Cathy raise her criterion and therefore respond more conservatively by means of the following payoffs.

Hit:	Win $10
Correct rejection:	Win $100
False alarm:	Lose $10
Miss:	Lose $10

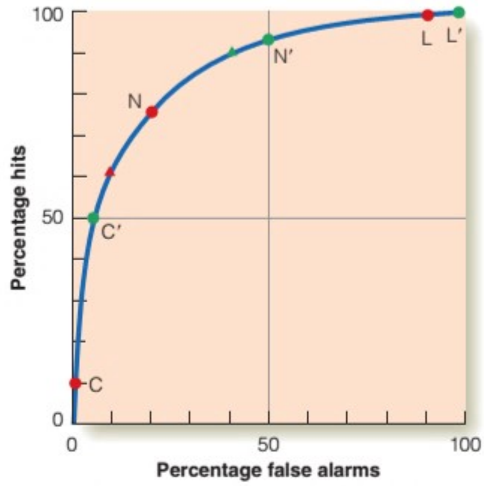

Figure D.2 A receiver operating characteristic (ROC) curve determined by testing Lucy (green data points) and Cathy (red data points) under three different criteria: liberal (L and L'), neutral (N and N'), and conservative (C and C'). The fact that Cathy's and Lucy's data points all fall on this curve means that they have the same sensitivity to the tone. The triangles indicate the results for Lucy and Cathy for an experiment that did not use payoffs.

This schedule of payoffs offers a great inducement to respond conservatively because there is a big reward for saying "no" when no tone is presented. Cathy's criterion is therefore shifted to a much higher level, so Cathy now returns to her conservative ways and says "yes" only when she is quite certain that a tone is presented; otherwise she says "no." The result of this newfound conservatism is a hit rate of only 10 percent and a minuscule false-alarm rate of 1 percent, indicated by point C (for "conservative" response) on the ROC curve. We should note that although Cathy hits on only 10 percent of the trials in which a tone is presented, she scores a phenomenal 99 percent correct rejections on trials in which a tone is not presented. (If there are 100 trials in which no tone is presented, then correct rejections + false alarms = 100. Because there was 1 false alarm, there must be 99 correct rejections.)

Cathy, by this time, is rich and decides to put a down payment on the Miata she's been dreaming about. (So far she's won $8,980 in the first experiment and $9,090 in the second experiment, for a total of $18,070! To be sure you understand how the payoff system works, check this calculation yourself. Remember that the signal was presented on 100 trials and was not presented on 100 trials.) However, we point out that she may need a little extra cash to have a satellite audio system installed in her car, so she agrees to stick around for one more experiment. We now use the following neutral schedule of payoffs.

Hit:	Win $10
Correct rejection:	Win $10
False alarm:	Lose $10
Miss:	Lose $10

With this schedule, we obtain point N (for "neutral") on the ROC curve: 75 percent hits and 20 percent false alarms. Cathy wins $1,100 more and becomes the proud owner of a Miata with a satellite radio system, and we are the proud owners of the world's most expensive ROC curve. (Do not, at this point, go to the psychology department in search of the nearest signal detection experiment. In real life, the payoffs are quite a bit less than in our hypothetical example.)

What Does the ROC Curve Tell Us?

Cathy's ROC curve shows that factors other than sensitivity to the stimulus determine a person's response. Remember that in all of our experiments the intensity of the tone has remained constant. Even though we changed only the person's criterion, we succeeded in drastically changing the person's responses.

Other than demonstrating that people will change how they respond to an unchanging stimulus, what does the ROC curve tell us? Remember, at the beginning of this discussion, we said that a signal detection experiment can tell us whether Cathy and Lucy are equally sensitive to the tone. The beauty of signal detection theory is that the person's sensitivity is indicated by the shape of the ROC curve, so if experiments on two people result in identical ROC curves, their sensitivities must be equal. (This conclusion is not obvious from our discussion so far. We will explain below why the shape of the ROC curve is related to the person's sensitivity.) If we repeat the above experiments on Lucy, we get the following results (data points L′, N′, and C′ in Figure D.2):

Liberal Payoff
Hits = 99 percent
False alarms = 95 percent

Neutral Payoff
Hits = 92 percent
False alarms = 50 percent

Conservative Payoff
Hits = 50 percent
False alarms = 6 percent

The data points for Lucy's results are shown by the green circles in Figure D.1. Note that although these points are different from Cathy's, they fall on the same ROC curve as do Cathy's. We have also plotted the data points for the first experiments we did on Lucy (open triangle) and Cathy (filled triangle) before we introduced payoffs. These points also fall on the ROC curve.

That Cathy's and Lucy's data both fall on the same ROC curve indicates their equal sensitivity to the tones. This confirms our suspicion that the method of constant stimuli misled us into thinking that Lucy is more sensitive, when the real reason for her apparently greater sensitivity is her lower criterion for saying "yes."

Before we leave our signal detection experiment, it is important to note that signal detection procedures can be used without the elaborate payoffs that we described for Cathy and Lucy. Much briefer procedures, which we will describe shortly, can be used to determine whether differences in the responses of different persons are due to differences in threshold or to differences in response criteria.

What does signal detection theory tell us about functions such as the spectral sensitivity curve (Figure 2.15, page 31) and the audibility curve (Figure 11.8, page 265), which are usually determined using one of the classical psychophysical methods? When the classical methods are used to determine these functions, it is usually assumed that the person's criterion remains constant throughout the experiment, so that the function measured is due not to changes in response criterion but to changes in the wavelength or some other physical property of the stimulus. This is a good assumption because changing the wavelength of the stimulus probably has little or no effect on factors such as motivation, which would shift the person's criterion. Furthermore, experiments such as the one for determining the spectral sensitivity curve usually use highly experienced people who are trained to give stable results. Thus, even though the idea of an "absolute threshold" may not be strictly correct, classical psychophysical experiments run under well-controlled conditions have remained an important tool for measuring the relationship between stimuli and perception.

Signal Detection Theory

We will now discuss the theoretical basis for the signal detection experiments we have just described. Our purpose is to explain the theoretical bases underlying two ideas: (1) the percentage of hits and false alarms depends on a person's criterion, and (2) a person's sensitivity to a stimulus is indicated by the shape of the person's ROC curve. We will begin by describing two key concepts of signal detection theory (SDT): signal and noise. (See Swets, 1964.)

Signal and Noise

The **signal** is the stimulus presented to the person. Thus, in the signal detection experiment we just described, the signal is the tone. The **noise** is all the other stimuli in the environment, and because the signal is usually very faint, noise can sometimes be mistaken for the signal. Seeing what appears to be a flicker of light in a completely dark room is an example of visual noise. Seeing light where there is none is what we have been calling a false alarm, according to signal detection theory. False alarms are caused by the noise. In the experiment we just described, hearing a tone on a trial in which no tone was presented is an example of auditory noise.

Let's now consider a typical signal detection experiment, in which a signal is presented on some trials and no signal is presented on the other trials. Signal detection theory describes this procedure not in terms of presenting a signal or no signal, but in terms of presenting signal plus noise (S + N) or noise (N). That is, the noise is always present, and on some trials, we add a signal. Either condition can result in the perceptual effect of hearing a tone. A false alarm occurs when the person says "yes" on a noise trial, and a hit occurs when the person says "yes" on a signal-plus-noise trial. Now that we have defined signal and noise, we introduce the idea of probability distributions for noise and signal plus noise.

Probability Distributions

Figure D.3 shows two probability distributions. The one on the left represents the probability that a given perceptual effect will be caused by noise (N), and the one on the right represents the probability that a given perceptual effect will be caused by signal plus noise (S + N). The key to understanding these distributions is to realize that the value labeled "Perceptual effect (loudness)" on the horizontal axis is what the person experiences on each trial. Thus, in an experiment in which the person is asked to indicate whether a tone is present, the perceptual effect is the perceived loudness of the tone. Remember that in an SDT experiment the tone always has the same *intensity*. The *loudness* of the tone, however, can vary from trial to trial. The person perceives different loudnesses on different trials, because of either trial-to-trial changes in attention or changes in the state of the person's auditory system.

The probability distributions tell us what the chances are that a given loudness of tone is due to (N) or to (S + N). For

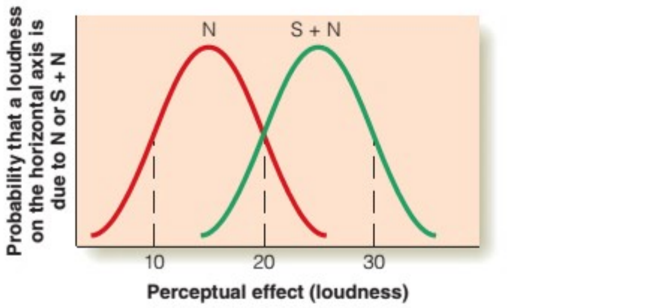

Figure D.3 Probability distributions for noise alone (N, red curve), and for signal plus noise (S + N, green curve). The probability that any given perceptual effect is caused by the noise (no signal is presented) or by the signal plus noise (signal is presented) can be determined by finding the value of the perceptual effect on the horizontal axis and extending a vertical line up from that value. The place where that line intersects the (N) and (S + N) distributions indicates the probability that the perceptual effect was caused by (N) or by (S + N).

example, let's assume that a person hears a tone with a loudness of 10 on one of the trials of a signal detection experiment. By extending a vertical dashed line up from 10 on the "Perceptual effect" axis in Figure D.3, we see that the probability that a loudness of 10 is due to (S + N) is extremely low, because the distribution for (S + N) is essentially zero at this loudness. There is, however, a fairly high probability that a loudness of 10 is due to (N), because the (N) distribution is fairly high at this point.

Let's now assume that, on another trial, the person perceives a loudness of 20. The probability distributions indicate that when the tone's loudness is 20, it is equally probable that this loudness is due to (N) or to (S + N). We can also see from Figure D.3 that a tone with a perceived loudness of 30 would have a high probability of being caused by (S + N) and only a small probability of being caused by (N).

Now that we understand the curves of Figure D.3, we can appreciate the problem confronting the person. On each trial, she has to decide whether no tone (N) was present or whether a tone (S + N) was present. However, the overlap in the probability distributions for (N) and (S + N) means that for some perceptual effects this judgment will be difficult. As we saw before, it is equally probable that a tone with a loudness of 20 is due to (N) or to (S + N). So, on a trial in which the person hears a tone with a loudness of 20, how does she decide whether the signal was presented? According to signal detection theory, the person's decision depends on the location of her criterion.

The Criterion

We can see how the criterion affects the person's response by looking at **Figure D.4**. In this figure, we have labeled three different criteria: liberal (L), neutral (N), and conservative (C). Remember that we can cause people to adopt these different criteria by means of different payoffs. According to signal detection theory, once the person adopts a criterion, he or she

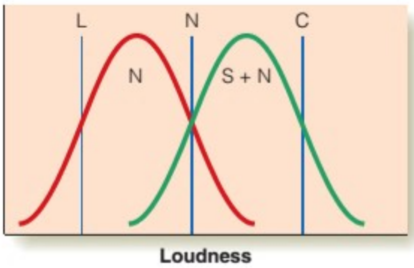

Figure D.4 The same probability distributions from Figure D.3, showing three criteria: liberal (L), neutral (N), and conservative (C). When a person adopts a criterion, he or she uses the following decision rule: Respond "yes" ("I detect the stimulus") when the perceptual effect is greater than the criterion, and respond "no" ("I do not detect the stimulus") when the perceptual effect is less than the criterion.

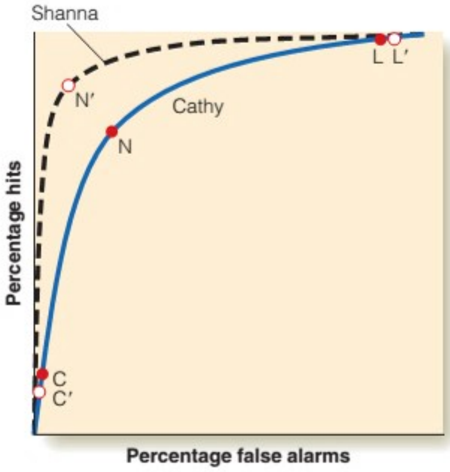

Figure D.5 ROC curves for Cathy (solid curve) and Shanna (dashed curve) determined using liberal (L, L′), neutral (N, N′), and conservative (C, C′) criteria.

uses the following rule to decide how to respond on a given trial: If the perceptual effect is greater than (to the right of) the criterion, say "Yes, the tone was present"; if the perceptual effect is less than (to the left of) the criterion, say "No, the tone was not present." Let's consider how different criteria influence the person's hits and false alarms.

To determine how the criterion affects the person's hits and false alarms, we will consider what happens when we present (N) and when we present (S + N) under three different criteria.

Liberal Criterion

1. Present (N): Because most of the probability distribution for (N) falls to the right of the criterion, the chances are good that presenting (N) will result in a loudness to the right of the criterion. This means that the probability of saying "yes" when (N) is presented is high; therefore, the probability of a false alarm is high.
2. Present (S + N): Because the entire probability distribution for (S + N) falls to the right of the criterion, the chances are excellent that presenting (S + N) will result in a loudness to the right of the criterion. Thus, the probability of saying "yes" when the signal is presented is high; therefore, the probability of a hit is high. Because criterion L results in high false alarms and high hits, adopting that criterion will result in point L on the ROC curve in **Figure D.5**.

Neutral Criterion

1. Present (N): The person will answer "yes" only rarely when (N) is presented because only a small portion of the (N) distribution falls to the right of the criterion. The false-alarm rate, therefore, will be fairly low.
2. Present (S + N): The person will answer "yes" frequently when (S + N) is presented because most of the (S + N) distribution falls to the right of the criterion. The hit rate, therefore, will be fairly high (but not as high as for the L criterion). Criterion N results in point N on the ROC curve in Figure D.5.

Conservative Criterion

1. Present (N): False alarms will be very low because none of the (N) curve falls to the right of the criterion.
2. Present (S + N): Hits will also be low because only a small portion of the (S + N) curve falls to the right of the criterion. Criterion C results in point C on the ROC curve in Figure D.5.

You can see that applying different criteria to the probability distributions generates the solid ROC curve in Figure D.5. But why are these probability distributions necessary? After all, when we described the experiment with Cathy and Lucy, we determined the ROC curve simply by plotting the results of the experiment. The reason the (N) and (S + N) distributions are important is that, according to signal detection theory, the person's sensitivity to a stimulus is indicated by the distance (d') between the peaks of the (N) and (S + N) distributions, and this distance affects the shape of the ROC curve. We will now consider how the person's sensitivity to a stimulus affects the shape of the ROC curve.

The Effect of Sensitivity on the ROC Curve

We can understand how the person's sensitivity to a stimulus affects the shape of the ROC curve by considering what the probability distributions would look like for Shanna, a person with supersensitive hearing. Shanna's hearing is so good that a tone barely audible to Cathy sounds very loud to Shanna. If presenting (S + N) causes Shanna to hear a loud tone, this means that her (S + N) distribution should be far to the right, as shown in **Figure D.6**. In signal detection terms, we would say that Shanna's high sensitivity is indicated by the large separation (d') between the (N) and the (S + N) probability distributions. To see how this greater separation between the probability distributions will affect her ROC curve, let's see how she would respond when adopting liberal, neutral, and conservative criteria.

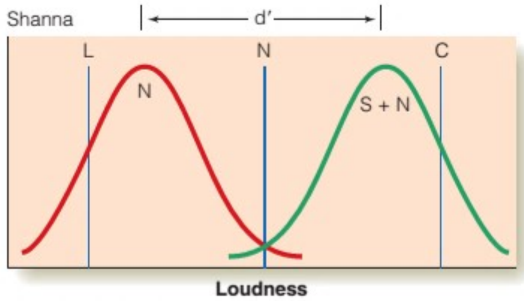

Figure D.6 Probability distributions for Shanna, a person who is extremely sensitive to the signal. The noise distribution (red) remains the same, but the (S + N) distribution (green) is shifted to the right compared to the curves in Figure D.4. Liberal (L), neutral (N), and conservative (C) criteria are shown.

Liberal Criterion

1. Present (N): high false alarms.
2. Present (S + N): high hits.

The liberal criterion, therefore, results in point L′ on the ROC curve of Figure D.5.

Neutral Criterion

1. Present (N): low false alarms. It is important to note that Shanna's false alarms for the neutral criterion will be lower than Cathy's false alarms for the neutral criterion because only a very small portion of Shanna's (N) distribution falls to the right of the criterion, whereas more of Cathy's (N) distribution falls to the right of the neutral criterion (Figure D.4).
2. Present (S + N): high hits.

In this case, Shanna's hits will be higher than Cathy's because almost all of Shanna's (S + N) distribution falls to the right of the neutral criterion, whereas less of Cathy's does (Figure D.4). The neutral criterion, therefore, results in point N′ on the ROC curve in Figure D.5.

Conservative Criterion

1. Present (N): low false alarms.
2. Present (S + N): low hits.

The conservative criterion, therefore, results in point C′ on the ROC curve.

The difference between the two ROC curves in Figure D.5 is obvious because Shanna's curve is more "bowed." But before you conclude that the difference between these two ROC curves has anything to do with where we positioned Shanna's L, N, and C criteria, see whether you can get an ROC curve like Shanna's from the two probability distributions of Figure D.4. You will find that, no matter where you position the criteria, there is no way that you can get a point like point N′ (with very high hits and very low false alarms) from the curves of Figure D.4. In order to achieve very high hits and very low false alarms, the two probability distributions must be spaced far apart, as in Figure D.6.

Thus, increasing the distance (d′) between the (N) and the (S + N) probability distributions changes the shape of the ROC curve. When the person's sensitivity (d′) is high, the ROC curve is more bowed. In practice, d′ can be determined by comparing the experimentally determined ROC curve to standard ROC curves (see Gescheider, 1976), or d′ can be calculated from the proportions of hits and false alarms that occur in an experiment by means of a mathematical procedure we will not discuss here. This mathematical procedure for calculating d′ enables us to determine a person's sensitivity by determining only one data point on an ROC curve, thus using the signal detection procedure without running a large number of trials.