Signal Detection and Animal Communication

R. Haven Wiley

DEPARTMENT OF BIOLOGY, UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NORTH CAROLINA 27599, USA

I. Introduction

Although communication consists of associations between signals from one individual and responses by another, in reality these associations are often weak. In recent decades there has been a tendency to explain these weak associations as the result of attempts by signalers to manipulate or exploit receivers and of receivers to resist these attempts. This chapter takes a different approach, although the underlying question remains the same—how can signalers and receivers optimize their behavior? The present approach develops an earlier suggestion that it is the inevitability of errors by receivers that limits optimal behavior by both parties in communication (Wiley, 1994). Signal detection theory provides the basic theory for this approach. The previous applications of this theory, however, have been in psychophysics. To justify its application to the evolution of communication is the purpose of the present chapter.

The problems of signal detection arise especially for signals in their natural contexts. The properties of signals perceived by a receiver inevitably differ from those emitted by the signaler. For instance, acoustic signals like bird songs are altered by attenuation and degradation during propagation through the environment (Naguib, 2003; Naguib and Wiley, 2001; Wiley, 1991; Wiley and Richards, 1982). Although in any one situation, on average, some features of attenuation and degradation are predictable, much remains unpredictable in detail.

Furthermore, a receiver perceives this attenuated and degraded signal against a background of irrelevant energy that shares some features with the signal. An acoustic signal, for instance, is often perceived against a background of sounds with more or less similar frequencies, intervals, or other patterns. These sounds come from nearby individuals of the same or different species and from physical features of the environment such as

wind and water. Finally, any receiver's sensory, associative, and motor neurons always include some unpredictability.

As a result of all of these processes, it is not surprising that signals usually have weak associations with responses. Sometimes when a stimulus occurs, the intended receiver fails to respond; sometimes the receiver responds when there is no stimulus. In the first case, the receiver seems to mistake a stimulus for the background; in the latter it seems to mistake the background for a stimulus. Because of the pervasiveness of these mistakes, receivers fail to achieve maximal performance and signals fail to reach maximal efficiency.

At first sight, these weak associations of signals and responses seem to be just noise in the system without fundamental implications for communication. This chapter, however, develops the view that these mistakes are a result of inescapable constraints on the performance of receivers and that these constraints in turn influence the evolution of both producing and responding to signals (Wiley, 1994). Many current issues in the study of communication, such as honesty and exploitation and the multiplicity and exaggeration of signals, become clearer once we understand the constraints on the performance of receivers.

These constraints on receivers are addressed by signal detection theory (Green and Swets, 1966; Macmillan and Creelman, 1991; McNicol, 1972). Developed originally by electronics engineers, in recent decades this theory has provided a rationale for the psychophysical study of sensory thresholds and perception. Despite its success in these studies, its application to the evolution of communication is still rudimentary.

The initial sections of this chapter provide an introduction to signal detection theory and its applications in psychophysics. The objective of these sections is to identify general principles for the study of adaptations in animal communication. These principles can clarify the properties of signals that affect a receiver's performance. They also suggest ways to extend the theory to the classification as well as detection of signals. These steps lead to hypotheses about the evolution of both signaling and receiving. In particular, signal detection theory leads to natural explanations for the evolution of deception and exaggeration in communication. The final sections take up the design and interpretation of experiments for studying communication in natural situations. The objective of these sections is to suggest practical ways to study the performance of receivers under conditions like those in which communication evolved.

II. Essential Features of Signal Detection

To apprehend the essential features of signal detection theory, it helps to consider a simple situation. Suppose an individual listens for a conspecific vocalization characterized by some feature such as a particular frequency.

In this case, the signal has a single feature, a particular frequency, which varies along a single dimension, its intensity.

Even in this simple case, a receiver in natural situations faces a formidable problem. By the time the signal reaches the receiver, its intensity varies irregularly, as a result of variable attenuation and degradation of the signal during propagation. With some ingenuity and proper instruments, we can measure the intensity of the arriving signal in appropriate time intervals (for instance, the temporal resolution of the receiver's hearing). From this information, we can determine the probabilities of different intensities of the characteristic frequency as the signal reaches the receiver. These probabilities constitute the probability density function (PDF) for the intensity of that frequency during a signal.

At the same time, the receiver usually experiences background stimulation that can also include this characteristic frequency. For instance, this frequency might occur in other species' or individuals' vocalizations or in other environmental sources of sound, all irrelevant to the listener. Again, with some care we can determine the PDF for the intensity of this frequency in the background stimulation reaching the receiver.

If the distributions of intensities during the signal and background stimulation overlap, then the receiver (a listener) cannot avoid mistakes. Errors are inevitable whenever a receiver cannot completely separate signal and background. Only an observer with independent access to the source of the signal and the background can measure their properties separately. A receiver has no independent access to the signal. It must instead decide whether or not a particular intensity of the characteristic frequency merits response or not. Past experience with different intensities might lead to different expectations for the frequency of the signal and thus different levels of confidence in its decision to respond or not, but an isolated perception itself provides no basis for certainty.

A simple graph can introduce the issues that arise in this situation. We can plot the overlapping PDFs for intensity during the signal and background stimulation along the same axis (Fig. 1). The subject's criterion for a decision is then represented by a threshold for response. In Fig. 1, the PDFs are represented by normal distributions, with equal variances but different means. This simplified situation applies when background stimulation has a normal distribution of intensities and the signal has a fixed intensity, which is added to the background. Complications are addressed later, but they do not change the basic issues.

Once a threshold for response is chosen, then the total probability of a *correct response* (responding when a signal has occurred) is the integral of the PDF for signals from the threshold to infinity. The probability of a *missed detection* (failing to respond when a signal has occurred) is the integral of the same PDF from the threshold to negative infinity.
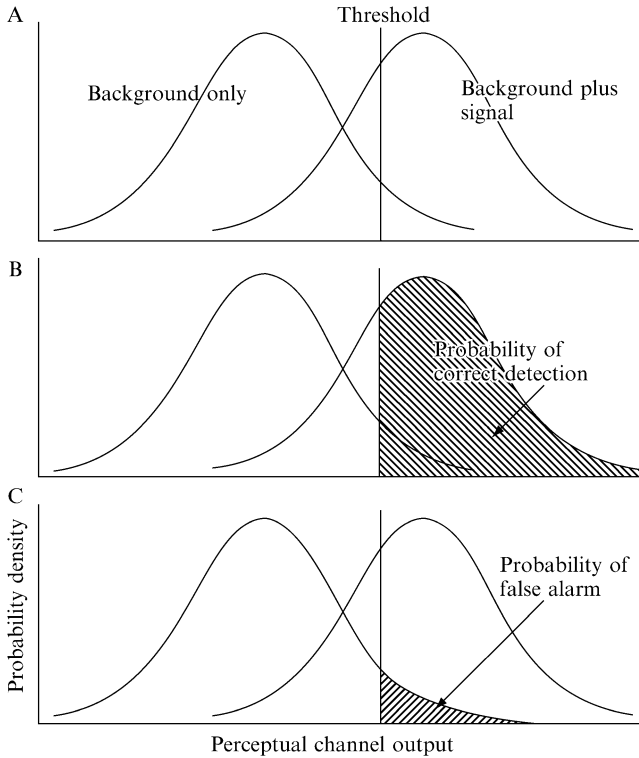
Fig. 1. The basic situation described by signal detection theory. (A) The levels of background stimulation with and without a signal are represented by the outputs from a perceptual channel. The probability of an output as a function of the level of the output is a probability density function, PDF, for the output. A decision to respond involves selecting a criterion (in this case, a threshold in the output of the channel above which a response occurs). (B) Any such threshold results in a probability of correct detections, P_{CD}, the area under the PDF for background plus signal to the right of the threshold. (C) Any threshold also results in a probability of false alarms, P_{FA}, the corresponding area under the PDF for background alone.

Thus when a signal occurs, the probability of correct detection by the receiver equals one minus the probability of missed detection, $P_{CD} = 1 - P_{MD}$. Similarly, we can find the probabilities of *false alarm* (responding to background stimulation), P_{FA}, and *correct rejection* (no response to background stimulation), P_{CR}, from integrals of the PDF for background stimulation. When only background stimulation occurs, $P_{FA} = 1 - P_{CR}$.

The first essential feature of signal detection is now apparent. As a receiver changes its threshold for response, *the* P_{FA} *varies with the probability of a correct detection* (P_{CD}). By shifting the threshold for response to the right (toward higher intensities of the characteristic frequency), for example, a receiver can reduce its false alarms (responding when there is no signal present). Simultaneously, however, it increases its missed detections (not responding when a signal occurs). Clearly a receiver in this situation cannot simultaneously both minimize P_{FA} and maximize P_{CD}. This trade-off between correct detections and false alarms has fundamental implications for the evolution of communication (Wiley, 1994).

Another essential feature of signal detection is a *distinction between the receiver's criterion for a response and the detectability of the signal*. In this simple case, the receiver's criterion is represented by a threshold for response; the detectability of the signal is represented by the separation of the PDFs for signal alone and signal plus background (the difference between the means in relation to the standard deviation). A receiver's performance is determined by both of these variables.

When we present signals to animals, such as recordings of calls or songs, we often want to determine the subjects' attitude or responsiveness toward the signal. It is thus the subjects' criterion that interests us. In other cases, such as determination of sensory thresholds, it is the detectability of the signals that interests us. Signal detection theory allows us to separate the criterion for response from the detectability of signals. To see how, we can turn to a well-established application of this theory.

III. APPLICATION OF SIGNAL DETECTION THEORY IN EXPERIMENTAL PSYCHOPHYSICS

The earliest application of signal detection theory to a behavioral problem was the determination of human sensory thresholds. Signal detection theory solved the problem of measuring the detectability of a signal despite differences in subjects' thresholds for responses. Procedures for this purpose are now well established (Green and Swets, 1966; Macmillan and Creelman, 1991; McNicol, 1972).

Before the application of signal detection theory, psychophysicists determined the absolute threshold for hearing sounds of a particular frequency by asking subjects to respond to faint sounds, barely separable from the background. The activity of auditory neurons in response to these sounds would barely differ from their spontaneous activity. These experiments confronted an insurmountable problem, because there was no satisfactory

way to standardize the criteria different subjects used for responding, in other words, their thresholds.

Signal detection theory provides a solution to this problem by a simple modification of the experimental procedure. Subjects listen for a tone during brief intervals indicated by a cuing stimulus, for instance illumination of a light. During half of these intervals, selected at random, there occurs a tone of a particular frequency and intensity; during the remaining intervals there is no tone. The intervals with a tone allow an estimate of P_{CD}; those with no tone allow an estimate of P_{FA}. If the tone is loud enough, subjects detect the tone with high efficiency (high P_{CD} and low P_{FA}). If the tone is faint, this efficiency drops.

The subject's performance in this situation depends on both the *detectability* of the stimulus and the subject's *criterion* for response (in this case, a threshold). The literature in psychophysics often refers to a subject's criterion as a bias. For any constant level of detectability (the distance between the means of the two PDFs relative to the standard deviation), as a subject's threshold increases, P_{CD} increases as a function of $P_{FA.}$ This function, called the *receiver operating characteristic* (ROC), increases monotonically from (0,0) to (1,1) in the unit square (Fig. 2). As an exercise, try generating Fig. 2 from Fig. 1, by varying the threshold for response.

To obtain an ROC, we must measure P_{CD} and P_{FA} at different thresholds for response. Psychophysicists use two basic methods. One involves direct manipulation of the subjects' thresholds, by rewards or instructions that place different weights on correct detections and false alarms. Another method involves asking subjects to rate their certainty for each response (for instance, 0 = absolutely certain no signal occurred, 10 = absolutely certain a signal occurred) (Egan *et al.*, 1959; Macmillan and Creelman, 1991; McNicol, 1972). In the latter case, the experimenter uses different levels of certainty for different thresholds of response. For instance, for a high threshold, take all responses with certainty greater than 9 as positive responses for determining both P_{CD} and $P_{FA.}$ For a lower threshold, take all responses with certainty greater than 8, and so forth. Accuracy in estimating P_{CD} and P_{FA} at each threshold requires repeated tests for each subject.

The ROC then allows us to determine the detectability of a signal in a way that is independent of subjects' thresholds for response. As the detectability of a signal increases (the PDFs for signal alone and signal plus background move apart), the ROC moves away from the positive diagonal toward the upper left corner of the unit square, the point where performance is ideal ($P_{CD} = 1$, $P_{FA} = 0$). The closer the ROC approaches the upper left corner, the greater the detectability of the signal.

As the subject's threshold changes, on the other hand, its performance moves one way or the other along the ROC. As its threshold increases,
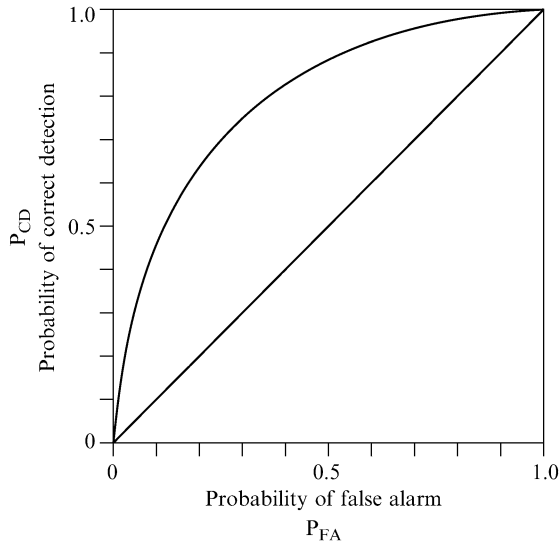
F<small>IG</small>. 2. A receiver operating characteristic (ROC) results from plotting P_{CD} as a function of $P_{FA,}$ as the threshold for response varies. The ROC is symmetrical about the negative diagonal of the unit square provided the two PDFs have normal distributions and equal variances. The separation of the means of the PDFs determines how far the ROC lies from the positive diagonal and thus how nearly it approaches the point of ideal performance, the upper left-hand corner. This illustration shows the ROC when the means are separated by one standard deviation ($d' = 1$).

a subject's performance approaches the origin ($P_{FA} = 0$, $P_{CD} = 0$). As its threshold decreases, its performance approaches the upper right corner ($P_{CD} = 1$, $P_{FA} = 1$). Thus changes in detectability of a signal shift the ROC away from or toward the diagonal, while changes in the subject's threshold shift its performance upward or downward along the ROC. Some study of Figs. 1 and 2 can clarify these relationships between the detectability of a signal, the threshold for response, and a subject's performance (its P_{CD} and P_{FA}).

  Measurement of detectability is straightforward when the PDFs for background alone and for signal plus background are normally distributed with equal variance. The ROC in this case is symmetrical about the negative diagonal. If we plot the normal deviates or z-transforms of P_{CD} and $P_{FA,}$ then the ROC is a straight line with unit slope (Fig. 3), and the difference in z-scores, $z(P_{FA}) - z(P_{CD})$, is the same for all points on this line. This difference, usually represented by d', represents the detectability of the signal. It equals the separation of the PDFs for background alone
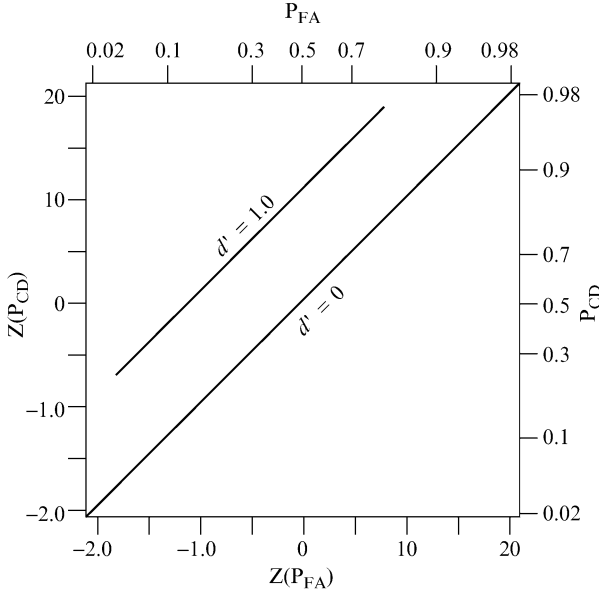
FIG. 3. An ROC plotted on probability (z-transformed) axes is a straight line with slope = 1 in the case of normally distributed PDFs with equal variance. This illustration shows the same ROC as Fig. 2.

and signal plus background divided by their standard deviation. For alternative measures of detectability, all highly correlated with d', see discussions by Green and Swets (1966), McNicol (1972), or Macmillan and Creelman (1991).

Detectability is a measure of a receiver's ability to separate a signal from background stimulation; the analogous measure of ability to separate two signals is *discriminability*. The methods just described for measurement of the detectability of a signal also permit measurement of the discriminability of two signals. Instead of comparing responses to a signal and background stimulation, we compare responses to two signals in the presence of constant background stimulation.

IV. GENERAL ASSUMPTIONS OF SIGNAL DETECTION THEORY

The theory of signal detection derives from assumptions about the nature of signals and their processing by receivers. This section considers these assumptions in order to establish the wide application of this theory.

General assumptions are separated from some specific ones so that we do not discard the general theory entirely on the basis of questions about specifics. This section addresses general assumptions; the next considers specifics.

The theory accommodates a broad definition of a signal. Elsewhere, I have proposed that a signal is any pattern of energy produced by one individual (the signaler) and evoking a response from another individual (the receiver) without providing all of the power necessary to effect the response (Wiley, 1994). Some power is necessary to produce an alteration in the receiver's sensors, but the receiver itself provides essential power for the response. It is the necessary role of the receiver in producing a response that creates the essential elements of signal detection and, ultimately, all communication. Although a signal is similar to any stimulus that evokes a response, the term ''signal'' serves to emphasize the crucial importance of the limited contribution of power for the response.

The restriction of the sources and receivers of signals to living individuals (or their components) serves to include just those cases in which signalers and receivers might coevolve. This restriction is not essential, however, as signal detection theory addresses the optimization of a receiver's performance regardless of the source of the signals. Nevertheless, when both source and receiver are living organisms or their components, the possibility of coevolution raises particularly interesting issues, a topic we discuss later.

Signal detection theory also accommodates a broad scope for receivers. The two essential components of a receiver are a sensor and a mechanism for decisions. Each sensor is a perceptual *channel* tuned to a particular feature or dimension of stimulation (such as a particular band of frequencies of sound, a particular direction of a visual object, or a particular spectrotemporal pattern of sound). A decision to respond then depends on the output from one or more of these perceptual channels (Fig. 4). Any channel is specified by its characteristic feature (for instance, the frequency of sound for maximal response from an auditory neuron) and its selectivity (often presented as its tuning curve or pass band).

Each channel produces an output that depends on the energy in its pass band within the broader range of energy impinging on the organism. This stimulation can include background energy of no interest to the organism (including irrelevant signals produced by other species or individuals and energy from the physical environment). The physiological mechanisms of channels often also produce spontaneous output. Consequently, a decision to respond based on the output of a channel often includes the possibility of false alarms and missed detections.

This model has broad generality (Green and Swets, 1966, Chapter 1). It is perhaps the most general model for an organism's responses to
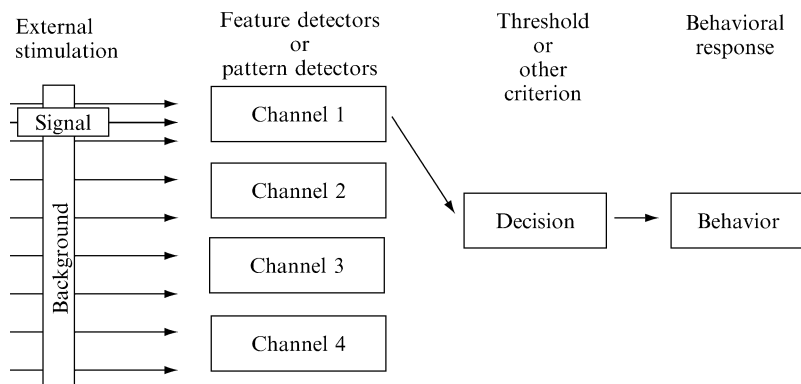
FIG. 4. The general model for signal detection involves perceptual channels that analyze features or patterns in stimulation impinging on the receiver. The output of one or more channels forms the basis for a decision to respond (in the form of a multidimensional criterion for response). Channels and decisions might represent distinct neurons or populations of neurons, or a single neuron might combine these two properties.

stimulation: a decision to respond or not depends on the output of a channel that receives combined signal and background. Green and Swets (1966) showed that the best rule for a decision to respond is a likelihood ratio that takes into account the expected frequencies of occurrence of signals. These basic ideas have a long history in psychology (Broadbent, 1958) and are familiar to ethologists and neuroethologists studying releasing mechanisms, stimulus filtering, and feature detectors.

The literature of psychophysics often contrasts ''signal detection theory'' with ''threshold theory'' (Green and Swets, 1966; Luce, 1963; Luce and Green, 1974; Macmillan and Creelman, 1991; Yonelinas, 2002). The distinction, however, is not fundamental. Threshold theory assumes some threshold above which a signal is always detected without error. Below this threshold, signals are detected with some fixed P_{FA} and P_{CD} (or some fixed ratio of these values). This theory thus requires at least two channels for the analysis of any one feature of a signal, one error-free for signals above the threshold, the other error-prone for signals below the threshold. These two channels, however, are equivalent to a single channel without normally distributed PDFs for background and for signal plus background (in this case the PDFs are rectangular; for full discussion, see Green and Swets, 1966; McNicol, 1972; Macmillan and Creelman, 1991). Only if we restrict the term ''signal detection theory'' to normally distributed PDFs with equal variance, are we forced to draw a distinction between this theory and ''threshold theory.'' If we relax these restrictions, threshold theory

becomes a special case of a general signal detection theory, based on a model of signal detection without restrictions on the distributions of outputs from perceptual channels.

Debate about these alternatives complicates much of the psychological literature on signal detection. In many cases, signal detection theory with additional assumptions of normality and equal variance can explain the properties of experimentally determined ROCs. The assumptions of normality and equal variance are best approached by examining the procedures for measuring detectability.

V. Specific Assumptions of Signal Detection Theory: Measuring Detectability

Signal detection theory, as applied routinely in psychophysical determinations of sensory thresholds, involves calculation of d' from measurements of P_{CD} as a function of P_{FA}. As shown in Section III earlier, this calculation is made simple by assuming normal PDFs with equal variances. In this special case, a single pair of measurements of P_{CD} and P_{FA} determines the ROC and thus d', as calculated from the standardized deviates, or z-scores, of P_{CD} and P_{FA}.

Calculation of d' from a single pair of measurements and determinations of absolute sensory thresholds requires some specific conditions: (1) normally distributed PDFs with equal variance; (2) fixed criteria for responses; and (3) cuing of responses. This section considers each of these requirements. Although each is critical in special cases, none is necessary for measurements of detectability in general.

A. Normal Distributions with Equal Variance

For sensory discriminations under laboratory conditions, the relevant PDFs are often nearly normal with nearly equal variances. The clearest evidence is an ROC symmetrical around the negative diagonal in the unit square and linear in probability space (with z-transformed axes for P_{CD} and P_{FA}) with slope equal to 1 (Green and Swets, 1966; McNicol, 1972). In this case, $d' = z(P_{FA}) - z(P_{CD})$ provides an unambiguous measure of detectability.

If the PDFs are not normally distributed or have unequal variances, then the picture changes. If variances are not equal, the ROC lacks symmetry around the negative diagonal. When plotted in z-transform space, the ROC has a slope equal to the ratio of variances. If the PDFs are not normally distributed, the ROC changes shape and is no longer linear in z-transform space.

Consequently, when either normality or equal variance is violated, $d' = z(P_{FA}) - z(P_{CD})$ makes little sense as a measure of detectability.

When normality or equal variance does not apply, we must use an alternative measure of detectability. A simple one is the area between the ROC and the positive diagonal of the unit square. This area measures the displacement of the ROC away from the positive diagonal and toward the point of maximal performance at the upper left-hand corner; d' provides a measure of this displacement only for a symmetrical ROC.

B. OPTIMAL CRITERIA

Accurate measurement of absolute sensory thresholds requires that subjects use an optimal criterion or rating scale for any set of experimental conditions. Variation among subjects, or variation among trials for any one subject, results in an underestimate of d' for maximal performance and also an underestimate of any difference in variances between signal and background. In carefully conducted psychophysical experiments, these possible errors turn out to be slight (Macmillan and Kaplan, 1985; McNicol, 1972, pp. 202–204).

This assumption that subjects use an optimal criterion is less critical for an investigation of communication, when an organism's actual performance has greater interest than its maximally possible performance. In this case, we can combine observations from different subjects by averaging z-scores to obtain a composite value of d' (Macmillan and Kaplan, 1985; McNicol, 1972, p. 112). If subjects' criteria or ratings vary, these composite measurements of detectability reflect expected average performance. Alternatively, we could study each individual's ability to detect or to discriminate signals.

C. CUING OF RESPONSES

Any measurement of the detectability of a stimulus requires null (background only) presentations, which permit measurement of P_{FA}, the probability of response without the signal present. In laboratory experiments, a cuing stimulus identifies intervals in which the subject must make a decision. This procedure assures equal decisions with and without the signal present. In field experiments this device is not possible. However, we can still include null presentations with no stimulus; even better, white noise or prerecorded background sounds might serve as a null stimulus. Alternatively, one could abandon attempts to measure the absolute detectability of any one stimulus and consider only the discriminability of two signals. In this case, a balanced experimental design could include equal numbers of presentations of the two signals.

The absence of null presentations confounds interpretation of a large body of research on human vigilance (Davies and Parasuraman, 1982; Mackie, 1977). Studies of vigilance and field studies of responses to playback have some similarities. In both cases, subjects experience long intervals between infrequent occurrences of a stimulus. The long periods without signals inevitably make P_{FA} very small during any brief interval when the signal is absent. Consistently small P_{FA} makes a meaningful ROC difficult to construct. Despite some suggestions for ways to circumvent this problem (Egan *et al.*, 1961a; Watson and Nichols, 1976), there seems to be no convincingly satisfactory solution. When we cannot measure false alarms, by means of cuing, null presentations, or comparisons of two signals, determination of an ROC is problematic. Measurement of P_{FA} is essential for a full understanding of a receiver's performance. A later section discusses some practical possibilities for solving this problem in field studies of animal communication by means of playbacks.

The two general results of signal detection theory—the interdependence of P_{CD} and P_{FA} and the distinction between the receiver's criterion and the detectability of the signal—do not depend on the specific assumptions of normality and equal variance and are not affected by the practical difficulties of measuring detectability or discriminability. These two general features of signal detection are alone sufficient to clarify the determinants of a receiver's performance.

VI. Properties of Signals That Affect a Receiver's Performance

Signal detection theory makes it clear that any receiver's performance in detecting or discriminating signals has limits. Furthermore, these limits are in part determined by properties of the signals. Predictions about these determinants of a receiver's performance have in many cases been repeatedly confirmed by psychophysical studies of humans, but the results of these studies have broad application to signal detection in general and thus to all forms of communication. Consider three properties of a signal that influence a receiver's performance: (1) contrast, (2) redundancy, and (3) uncertainty. We shall see that the inevitable effects of these three properties of signals explain a lot of "receiver psychology."

A. Contrast

Contrast and detectability are so closely related that it requires care to distinguish them carefully. As explained earlier, detectability is the difference between the means, in relation to the standard deviations, of background

alone and background plus signal in the output of some perceptual channel (for instance, in the responses of an experimental subject). Contrast is an analogous difference in the stimulation at the input to a channel (in the stimulation impinging on the subject). Unlike detectability, contrast depends only on the properties of the external stimulation reaching an organism and not on the properties of the organism's perceptual channels.

Contrast usually increases detectability. The influence of contrast on a subject's performance is so clear that it has received little explicit study by psychophysicists. One such study, included in one of the first applications of signal detection theory to perception, showed that log d' increased linearly with log intensity for a signal in the presence of constant background stimulation (Tanner and Swets, 1954).

Because we define contrast by the properties of a signal in relation to the background stimulation impinging on an organism, detectability of the signal depends on both its contrast and the selectivity of the perceptual channel. This dual determination of detectability is the basis for a procedure in psychophysics for determining bandwidths of sensory channels. In the case of hearing, the intensity of broad-spectrum background sound (white noise) that can mask a signal of a particular frequency depends on the bandwidth of the auditory channel. In fact, the signal-to-noise ratio (a measure of contrast) for complete masking of a single frequency with broad-spectrum noise equals the effective bandwidth of the auditory channel for that frequency.

The dual determination of detectability implies that the intensities of signals and background stimulation impinging on an organism do not alone allow us to predict an organism's performance. For instance, the intensity of a particular frequency of sound, or hue of light in a signal, and in the background are not enough to allow us to predict the detectability of that sound or light for a particular organism. To determine the influence of contrast on detectability, we must study the organism's responses, at either the neural or behavioral levels.

Study of contrast and detectability in natural situations is still rudimentary (Klump, 1996). For instance, despite many studies of sound propagation in natural environments and its influence on the evolution of bird songs (reviewed by Naguib and Wiley, 2001; Wiley, 1991), we know little about the properties of background sound in relation to acoustic signals in natural situations. Such studies of acoustic contrast would require recordings of signalers with omnidirectional microphones at typical positions for conspecific listeners. To extend these studies to detectability would require adjustments for the directionality and selectivity of the listeners' hearing.

Only one study has shown how background noise affects the detectability of acoustic signals in natural situations. Measurements of auditory

thresholds in great tits *Parus major*, in the absence of noise, reveal greatest sensitivity to frequencies between 2 and 4 kHz, lower than most of this species' vocalizations. However, critical bandwidths remain nearly constant over a wide range of frequencies up to 8 kHz. Consequently, in the presence of wind in a forest, which produces noise decreasing exponentially in intensity with increasing frequency, the frequency for greatest detectability shifts to 8 kHz (Langemann *et al.*, 1998). It is also clear that birds and mammals can increase the intensity of vocal signals in the presence of background sound, presumably to improve the contrast of their signals with the background (Brumm, 2004; Brumm and Todt, 2002; Brumm *et al.*, 2004; Cynx *et al.*, 1998; Leonard and Horn, 2005). Shifts in frequency to increase contrast with background noise are not so well documented. The clearest case is again the great tit, which uses higher dominant frequencies in its songs in urban environments with predominantly low-frequency noise (Slabbekoorn and Peet, 2003).

Contrast and detectability of visual signals is more complex. Unlike acoustic signals, for which the signaler generates the power to produce the signal, visual signals usually rely on reflectance or scattering of light from other sources. As Endler (1990, 1993) explains, the spectrum of light arriving at a receiver's eyes from an object depends on the product of the irradiance spectrum, the reflectance spectrum of the object, and the transmission spectrum (the spectra of the incident, reflected, and transmitted light, Q, R, and T). The contrast between a visual signal and its background thus depends on the contrast between QRT for the signal and the background. Q, which depends on the photic properties of the environment, can vary substantially with microhabitat (Endler, 1993; Gomez and Théry, 2004). These principles apply to male manakins, small birds that use bright colors in their plumage to produce visual displays at leks in the understory of tropical forests. Both the reflectance spectra of patches in their plumage and the placement of their leks in the forest serve to increase the contrast of their displays with the visual background (Endler and Théry, 1996; Heindl and Winkler, 2003). Furthermore, Uy and Endler (2004) have shown that, in one species, males increase the contrast of their plumage with the background by clearing fallen leaves from their display sites. Contrast between different parts of a signal is also affected by choice of location (Endler, 1993; Heindl and Winkler, 2003), but this within-signal contrast is a form of structural redundancy, discussed in the next Section VI.B.

One consequence of the dependence of visual signals on environmental irradiance is that changes in habitats can drastically alter contrast of signals with background. A case in point are the numerous endemic species of cichlids in Lake Victoria. Many of these recently evolved species differ mainly in male coloration and mate choice by females. Increased turbidity

of some parts of the lake in recent decades, as a result of sedimentation from human activities, is associated with a loss of many species (Seehausen *et al.,* 1997). Apparently, the species-specific colorations of the males no longer contrast enough to allow females to differentiate them.

Contrast applies to complex signals as well as to signals with a single characteristic feature. As with simpler signals, there has been little investigation of complex signals in the presence of background stimulation. One exception is human speech. Early experiments showed that human subjects have trouble understanding one person speaking in the presence of others, the so called ''cocktail-party problem'' (Cherry, 1953; Cherry and Taylor, 1954). Similar tasks requiring discrimination of one conspecific's vocalizations from those of other conspecifics in the background recur in many natural situations, for instance in choruses of frogs or insects, colonies of seabirds, and dawn choruses of birds or primates. Detection and discrimination in these situations have received little attention. One such study in a colony of king penguins (*Aptenodytes patagonicus*) confirmed that the presence of large numbers of conspecifics increased attenuation and degradation of the adults' calls that allow chicks to recognize their parents (Aubin and Jouventin, 1998). The situation is particularly difficult because the noise has nearly the same spectral distribution as the signals of interest to a chick. Nevertheless, these chicks can recognize their parents' calls even when the overall signal-to-noise ratio is less than 1. In such ''cocktail-party'' situations, birds as well as humans use cues for spatial localization to increase the effective signal-to-noise ratio of signals in more evenly distributed noise (Cherry, 1953; Cherry and Taylor, 1954; Dooling, 1982). In this case, contrast between signals consists mostly of differences in location.

B. REDUNDANCY

Redundancy results from predictable relationships between the parts of a stimulus, either in time or space. It takes two forms, both of which improve detectability of a signal. Sequential redundancy consists of fixed temporal relationships between components of a signal. Repetition of a signal, the simplest form of sequential redundancy, increases its detectability (Swets and Birdsall, 1978; Swets *et al.*, 1959). In fact, the detectability of tones increases with the square root of the number of presentations. This result is consistent with an assumption that each instance of a stimulus is assessed independently (Swets *et al.*, 1959).

All psychophysical experiments on detectability use an alerting signal to tell the subject when to respond. An alerting signal, one with high contrast and low uncertainty, accompanying a more informative signal is a special

case of redundancy. Although many natural signals might include alerting components (Richards, 1981a), this possibility has received little attention.

Simultaneous redundancy consists of fixed relationships between concurrent dimensions of a signal. Simultaneous redundancy can take the form of multiple components with fixed spatial relationships in a visual signal, multiple molecular components in an olfactory signal, or multiple components with fixed spectral relationships in an acoustic signal. Such a multidimensional stimulus is more detectable than one with a single feature. The increase in detectability with the number of features characterizing a stimulus again suggests that human observers assess each feature independently (Macmillan and Creelman, 1991; Mulligan and Shaw, 1980; Shaw, 1982).

An interesting twist on redundancy involves predictable relationships within the background noise rather than within the signal of interest. If different frequencies in noise are subject to synchronized amplitude modulation (called comodulation), then it is possible to use the properties of noise in one band of frequencies to improve detection of a signal in another band. This "comodulation masking release" has been demonstrated in both humans and birds (Klump and Langemann, 1995; Langemann and Klump, 2001; Nieder and Klump, 2001).

C. UNCERTAINTY AND UNFAMILIARITY

Uncertainty about signals takes two forms, each of which decreases detectability. Intrinsic uncertainty occurs when a subject lacks prior information about a signal's features, including the interval of time and location in which it might occur. Extrinsic uncertainty occurs when a subject must respond to many different signals. A subject can have prior information about the features of each signal but still face uncertainty about which signal will occur. Multiplicity of signals reduces the detectability of each of them.

Uncertainty about the features of signals reduces their detectability (Green, 1961; Pelli, 1985). Detectability also decreases when observers are uncertain about the locations or intervals of time in which signals might occur (Egan *et al.*, 1961b; Watson and Nichols, 1976; Starr *et al.*, 1975; Swensson and Judy, 1981). These latter situations are in fact special cases of the detection of signals with uncertain features.

Uncertainty about which of several signals might occur also reduces their detectability. For instance, if human observers are asked to report any of several possible signals, the overall detectability of the signals decreases as the number of alternatives increases (Cary and Reder, 2003; Nolte and Jaarsma, 1967).

Human performance in detecting multiple signals again implicates independent perceptual channels. It is as if a separate channel assesses each

signal's characteristic feature, and the subject decides that a signal has oc-
curred when the criterion in any channel is met (Cohn, 1978; Green and
Birdsall, 1978). These conclusions rest on a comparison of the ROCs of subjects
detecting different numbers of signals. This analysis also confirms that the
reduction in detectability of signals in this situation results from the uncertainty
of the task, not from any change in the observers' criterion for response.
Thus detection of signals from a repertoire of possibilities is inherently more
difficult than detection of a single signal specified in advance.

Unfamiliarity also makes signals more difficult to detect. For instance,
the frequencies of words in common usage influence their thresholds for
visual recognition (Pierce, 1963). Other studies have confirmed that high-
frequency words are more detectable than low-frequency ones (although
memory of high-frequency words presented previously is less accurate)
(Broadbent, 1967; Glanzer and Adams, 1985; Glanzer *et al.*, 1993; Pollack
et al., 1959). Thus greater familiarity with a stimulus increases its detect-
ability, just as greater uncertainty reduces it.

Human performance during vigilance fits the same pattern. The greater
the uncertainty about the features, timing, or location of possible signals,
the lower the efficiency of the observer (Davies and Parasuraman, 1982;
Davies and Tune, 1970; Loeb and Alluisi, 1977; Warm, 1977). Studies of
vigilance have not provided definitive evidence that detectability changes,
as opposed to the subject's criterion, because such studies, as explained
earlier, do not allow analysis of the ROC. Nevertheless, these results
resemble those of studies with a full analysis of detectability and thus
reinforce the conclusion that uncertainty about a stimulus, in any form,
reduces its detectability.

VII. Classification of Signals in Addition to Detection

Although in many situations it is reasonable to assume that an animal's
task involves no more than detection of an appropriate signal, in others
some classification of a stimulus is essential. Detection, for example, is
involved when an individual responds to a suitable mate or to its offspring
or chooses a diet based on profitability of prey. Classification, on the other
hand, is required when it recognizes several social partners or chooses a
diet with an optimal mixture of nutrients. The discussion so far has focused
on detection of a signal in noise. This section considers the use of
signal detection theory to understand a receiver's performance when
classification is as important as detection.

An experiment to show detection of a signal is designed so that the
subject must make a binary decision about the occurrence of the signal,

"yes" or "no," go or no-go. An experiment to show discrimination likewise requires only a single binary decision, either "signal 1" or "signal 2." Other situations, however, require both detection and subsequent classification of signals. Detection plus classification requires one of at least three responses ("no," "1," or "2") as a result of at least two binary decisions ("yes" or "no;" if "yes" then "1" or "2").

Detection plus classification is the basis for recognition or identification, as these terms are often used. In some discussions, however, recognition means detection of multidimensional signals or detection of signals with uncertain features, situations discussed in the previous section. These cases require single binary responses to a multiplicity of possible signals. The distinguishing feature of a classification of signals, in contrast, is the multiplicity of possible responses.

A few experiments confirm that classification in addition to detection is a more difficult task for receivers than detection alone. For instance, the task of identifying a stimulus as familiar or not requires less attention during previous exposures to the stimulus than does recollecting specific associations of a stimulus (Dobbins *et al.*, 2004). Female frogs (*Hyla ebraccata*) detect a conspecific male's calls in background noise from a natural chorus at signal-to-noise ratios above 3 dB. Yet they express a preference for those calls with lower fundamental frequencies only at signal/ratios greater than 9 dB (Wollerman and Wiley, 2002). At intermediate signal-to-noise ratios, females did not discriminate between otherwise preferred and nonpreferred males' calls, even though she could detect these calls.

Classification in addition to detection has surprisingly complex influences on a receiver's performance. To analyze these complexities and to assess their influence on receivers, we first consider a basic experiment. This approach leads to more complex ones and ultimately to a theoretical justification for a general principle: a receiver's performance in a task requiring classification is inevitably lower than in a comparable task requiring only detection.

To investigate detection plus classification, an experiment might present background alone and background in combination with each of two signals. With human subjects, we can simply ask for two responses, first "yes" or "no" for the presence of a stimulus, then "1" or "2" for the class of stimulus, provided one has been detected. Because classification presupposes correct detection of signals, the probability of correct classification can never exceed the P_{CD}. Some evidence for "subliminal" classification does not alter the situation significantly (Macmillan and Creelman, 1991, p. 255).

One approach in a study of this sort is to calculate both an ROC and an analogous *identification operating characteristic* (IOC). To construct the ROC for this case, one measures P_{CD} as the probability of a correct "yes"

response when either signal occurs and P_{FA} as the probability of a "yes" response when no signal occurs. For the analogous IOC, one measures P_{CD} as the probability of correct identification of a signal when it occurs; P_{FA} is still the probability of a "yes" response when no signal occurs (Benzschawel and Cohn, 1985; Green and Birdsall, 1978; Green et al., 1977; Macmillan and Creelman, 1991). The IOC, thus defined, can be derived from the ROC for detection of uncertain signals discussed earlier. Despite this theoretical advantage, the IOC fails to consider errors of classification once a signal is detected and thus provides an unrealistic measure of a receiver's performance.

A better approach in a study of detection plus classification is to consider a bivariate plot of PDFs, with one axis for a measure of the characteristic feature of each stimulus (Fig. 5). If the characteristic features of the two signals are orthogonal (in other words, if they vary independently), the PDFs for background only and for each signal in combination with background lie along two perpendicular axes. A receiver's performance then depends on three thresholds: two that separate background from each signal in combination with background (T1 and T2) and a third that separates the two signals (T3, Fig. 5). This third threshold differentiates the two signals based on the ratio of measures of their respective characteristic features. The slope of threshold T3 changes, as the receiver alters its criterion for classifying the signals.

This experiment thus allows measurement of three d' values (Macmillan, 2002; Macmillan and Creelman, 1991; Tanner, 1956): between background (B) and background plus one of the signals (B + S1), between B and B + S2, and between B + S1 and B + S2. Suppose the receiver processes the characteristic features of the two signals independently, as predicted for orthogonal features, and the variances of the three PDFs are equal, as predicted for constant signals added to background, with equal variance in each signal's characteristic feature. Then these three d' values have a Pythagorean relationship, $d_3' = \sqrt{(d_1'^2 + d_2'^2)}$ as seen by geometry in Fig. 5, in which each d' is proportional to the distance between the means of the respective PDFs.

An even more robust experiment would include a fourth stimulus, background in combination with both signals at once, B + S1 + S2. The six d' values in this case specify the nature of any interaction in processing the features of the two signals (masking of one signal by the other, inhibitory interaction between channels, correlation of the background in the two channels) (Klein, 1985; Thomas, 1985).

To understand the consequences of detection plus classification for a receiver's overall performance, we can compare P_{CD} and P_{FA} for detection plus classification with those for simple detection. The probability of
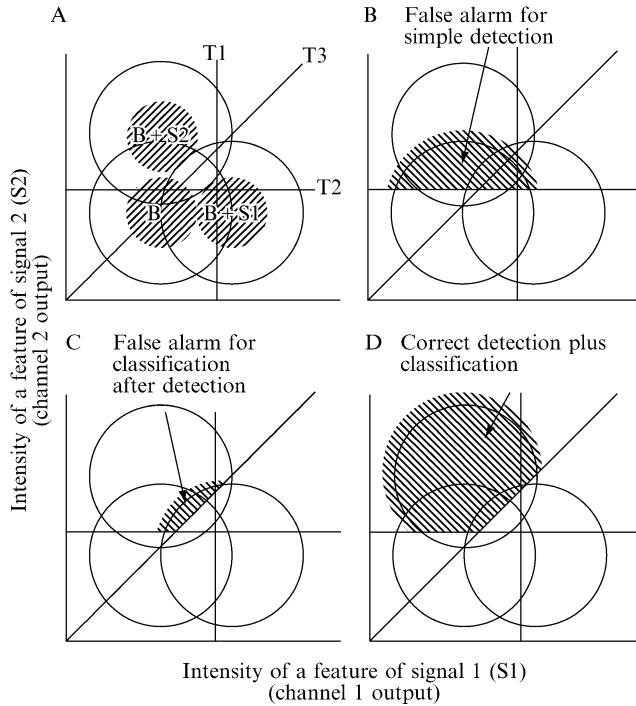
FIG. 5. (A) A bivariate plot of probability densities for combined detection plus classification shows the PDFs (now represented topographically by circles of equal probability density) for background stimulation alone, B, and in combination with each of two signals, S1 and S2. Decisions in this case require three thresholds: T1 for detection of B + S1 from B; T2 for detection of B + S2 from B; and T3 for classification of a signal once detected. (B) Threshold T2 results in a P_{FA} (shaded) for responses appropriate for S2 when background alone occurs. (C) Threshold T3 results in a P_{FA} (shaded) for responses appropriate for S2 when S1 occurs. (D) A combination of thresholds T1 and T3 results in a P_{CD} (shaded) for correct detection and classification of S2.

correct response to a particular signal (P_{CD} for detection plus classification) is always less than or equal to that for simple detection. As the threshold for classification, T3, decreases in slope, P_{CD} for detection plus classification increases from near 0 to a value approaching P_{CD} for simple detection (Fig. 5).

The situation for P_{FA} is more complex, because it involves two kinds of false alarm responding when only background occurs or when the alternative signal occurs. Because classification must follow detection, the two kinds of false alarm are not independent. Consequently, to combine the P_{FA} for simple detection and the P_{FA} for detection plus classification

requires information about the relative frequencies of these two situations. A full analysis of this situation is not yet available.

Analysis of this situation is simplified by considering only false alarms for detection. False alarms in this narrow sense include only responses to background stimulation and thus include only false alarms for detection and exclude those for classification. For any level of false alarm in this narrow sense, classification in addition to detection reduces correct responses to signals in comparison to detection alone. Conversely, for any level of correct detections of signals, classification in addition to detection increases false alarms in this narrow sense (Macmillan, 2002; Starr *et al.*, 1975). Classification plus detection, in comparison to simple detection, thus inevitably results in more false alarms by a receiver, even in the narrow sense. Classification thus inevitably reduces a receiver's performance in comparison to detection alone.

VIII.   COMPLEX PATTERNS: EXTENSION OF THE CONCEPT OF CHANNELS

Signal detection theory, as we have seen, describes decisions based on the outputs of perceptual channels. Detection and discrimination, the focus of discussion so far, suggest that the perceptual channels under consideration are sensory receptors and their immediate neural connections. Peripheral mechanisms of perception have been the main concern of many applications of signal detection theory, especially in studies of hearing.

Nevertheless, the theory applies equally well to more cognitive aspects of nervous systems. A channel can in fact represent any step in the hierarchy of perceptual analysis of a signal. It could represent ''detection'' of a species-specific song, for instance, when the issue is not whether or not a listening bird can hear each of the component frequencies but whether or not the entire pattern fits some criterion for a decision to respond.

Such pattern detection has all the same general properties as feature detection. A channel for pattern detection produces an output that reflects the presence of components with particular sequential or simultaneous relationships. Irrelevant background stimulation can include similar relationships, differing in unpredictable ways from those in the signal, and the mechanism of the channel can itself include some unpredictability. A criterion for a decision to respond based on the output from such a pattern-detecting channel inevitably results in false alarms and missed detections, just as from a feature-detecting channel.

Thus all of the preceding discussion of signal detection theory applies equally well to complex, as well as simple, perception. It applies to recognition of conspecific songs, to recognition of the vocalizations of mates,

offspring, or neighboring individuals, to mate choice based on complex repertoires, and to interpretation of subtle innuendos in the close-range vocalizations of group-living animals—signal detection theory applies to all communication.

IX. EVOLUTION OF SIGNALING AND RECEPTION

Signal detection theory suggests ways that receivers and signalers could coevolve (Wiley, 1994). We can understand many features of this coevolution by applying principles of signal detection first to optimizing receivers' performance and then to optimizing signalers' behavior. Because receivers provide the essential power for responses, their adaptation is primary. Nevertheless, signal detection theory shows that receivers do not necessarily get what they want. Because of the inevitable limitations on their performance, receivers can attain optimal, but not ideal, performance. Signalers can then evolve in response to the conditions set by their intended receivers. If changes in signalers' behavior alter the features or frequency of signals, receivers might evolve new optima for their own performance. Then signalers might evolve new features of signals. It seems probable that this form of coevolution could either reach an equilibrium or propagate perpetual lags between the adaptations of signalers and receivers.

Receivers can optimize the net utility of their decisions to respond or not by adjusting their criteria for response. The net utility for a receiver's decision depends on the probabilities and payoffs (net gains, positive or negative) of correct detections, missed detections, false alarms, and correct rejections (for details, see Wiley, 1994). The payoffs from these four possible outcomes must be measured in units relevant to natural selection. The probabilities of these outcomes, we have seen, depend on the discriminability of signals and the receiver's criterion. Depending on these payoffs and probabilities, the optimal criterion for a receiver can lie anywhere between *adaptive gullibility* (a low criterion for response when missed detections are especially costly) and *adaptive fastidiousness* (a high criterion for response when false alarms are especially costly). Gullability of receivers should result in the evolution of dishonest signals, fastidiousness in the evolution of exaggerated signals (Wiley, 1994).

Signaling should evolve to increase the predictability of responses from intended receivers. As a result, signals should often evolve to improve detectability (Wiley, 1983, 1994), so receivers can in turn evolve criteria that permit high P_{CD} and low P_{MD}. Greater contrast and redundancy and less uncertainty about a signal's features, including its timing and location,

all increase detectability and thus ultimately increase the probability of responses by receivers. Signal detection theory can explain why these properties of a stimulus affect detection and consequently learning and memory (the ''receiver psychology'' of Guilford and Dawkins, 1991, 1993). The widely reported phenomenon of peak shift in discrimination learning (Enquist and Arak, 1993; Guilford and Dawkins, 1991, 1993; ten Cate and Bateson, 1988; Weary *et al.*, 1993) follows from maximizing the net utility of a receiver's criterion for response (Lynn *et al.,* 2005). When false alarms are more costly than missed detections, it pays for a receiver to adapt a strict criterion for response. Because the adaptive solution is to respond to extremes of signal properties in one direction rather than the other, in order to minimize false alarms, peak shift is the result.

On the other hand, unintended receivers (eaves-dropping predators and parasites or conspecific rivals, for instance) can reduce the advantages of highly detectable signals. Properties that improve detectability, such as redundancy and predictability, also limit possibilities for encoding of complex information, which requires variation rather than constancy in signals (Wiley, 1994). Signals might thus evolve a compromise between advantages of detectability and advantages of privacy or complex coding.

X. Interpretation of Playback Experiments in Terms of Signal Detection Theory

Experimental studies of communication depend on presentations of signals to subjects in order to record their responses. Signal detection theory suggests new approaches for designing and interpreting such experiments. First of all, it calls into question the use of clear signals. Because the ability of animals to detect or to discriminate any signals depends on background stimulation, experiments with intense signals and weak background stimulation often have little relevance to communication in natural situations. Signal detection theory, however, does not simply suggest cautious interpretation of playback experiments. It also identifies two distinct reasons why results should depend on background stimulation: both the features of effective signals and a receiver's criterion for response should change with the level of background stimulation.

Many investigations of the features of signals that make them effective in eliciting responses have employed clear signals and minimal background stimulation. This approach is unlikely to provide a full understanding of communication because, as the preceding review has indicated, the features of effective signals, those that optimize receivers' performance, differ in the presence of high and low background stimulation. Signals effective

when background stimulation is low could prove much less so when background stimulation is high. In the latter case, we should expect greater emphasis on features that contribute to detectability of signals (contrast, redundancy, low uncertainty, familiarity). Experiments with playbacks have so far never considered the possible effects of background stimulation on detectability of signals.

The interpretation of responses is also complicated by the possibility of confounding detectability of signals with criteria for responses. In studies of animal communication, experiments are usually interpreted in terms of the subjects' attitude toward the experimental signals. For instance, do subjects have a lower threshold for a particular response to one type of signal in comparison to another? Yet the probability of response depends both on the listener's attitude (its threshold or criterion) and on the level of the signal in relation to background stimulation as perceived by the listener (the detectability of the signal). A few studies of responses to bird songs in the field have considered both of these possibilities (Brenowitz, 1982; Richards, 1981b), but all have so far relied on indirect evidence to separate them. Even differences in responses to loud, repeated, clean signals might reflect differences in detectability of signals rather than differences in receivers' criteria for response. When it is important to be sure that the receivers' attitude (criterion) differs, only an ROC analysis can separate these possibilities.

Signal detection theory also shows how to characterize the general properties of perceptual channels by comparing responses to at least three types of signals. Each pair of signals elicits responses that depend on outputs from a perceptual channel or combination of channels. Although only neurophysiology can determine the neural components and mechanisms of these channels, we can nevertheless learn something about their overall properties even without knowing the details of their mechanisms. For instance, are the pattern-detecting channels for each of the three possible pairs of signals independent (A-B, B-C, A-C)? Measuring the discriminabilities for the three possible pairs of signals can provide an answer. As explained earlier, discriminabilities that summed would indicate completely shared channels; discriminabilities with Pythagorean relationships would indicate completely independent channels; intermediate results would suggest partially correlated channels.

XI. Practicalities of Experiments in Natural Situations

To take advantage of these possibilities, we must measure detectabilities and discriminabilities in the field. To accomplish this task, we have to broaden the way we think about experiments with playbacks. Presentation of

loud, repeated, clear signals close to subjects provides little information for comparisons of detectability or discriminability of signals. Instead, for this purpose, it would be better for each trial to present a brief (perhaps a single) stimulus in combination with background stimulation. Furthermore, the nature of the background must become part of the experimental design.

To determine the detectability of a single stimulus, we can use background stimulation as a null stimulus (background only) for comparison with the signal (background plus signal). To determine the discriminability of two signals, the problem of a null stimulus does not arise. Nevertheless, including a null stimulus in the experimental design adds the possibility of a full analysis of detection plus classification, as described earlier. An ROC can then allow evaluation of normality and variance in the outputs of the channels involved and thus choice of an appropriate measure of detectability or discriminability.

To construct an ROC from field studies of animals, a rating scale is likely to be the method of choice. To do so, we must first determine the distribution of some measure of response (perhaps the first principal component of all behavioral measures) across all trials. Depending on sample sizes, we can assign scores, for instance, to quartiles or deciles of this distribution. These scores provide nonverbal ratings of the subjects' levels of confidence in discriminating between the two signals. The distributions of scores for each signal then generate pairs of P_{CD} and P_{FA} for the construction of an ROC.

A practical problem in measuring ROCs in the field is the limited number of trials. Experiments with animals in the field can rarely expect, as psychophysical experiments do, to present signals hundreds of times to each subject and then to examine each subject's ROC separately. Field studies will probably have to combine data from different subjects and thus determine only characteristics of populations, ideally ones as homogeneous as possible. Nevertheless, practical numbers of trials could yield useful measures of detectability in experiments with rating scales (McNicol, 1972, Chapter 5).

Once an ROC is constructed, we can apply standard procedures for calculating detectability (or discriminability) of the signals. Furthermore, each pair of scores used to construct the ROC reveals the subjects' average criterion under particular conditions. Procedures for calculating detectability or discriminability from a rating scale, summarized earlier, are thoroughly reviewed by McNicol (1972). The location of the criterion for response under particular conditions is best specified by its absolute location with respect to the underlying PDFs. Macmillan and Creelman (1990) recommend simple measures, such as $(P_{CD} + P_{FA})/2$ or $-[z(P_{CD}) + z(P_{FA})]/2$. With these procedures, the application of signal detection theory

to field studies of animal communication seems unlikely to encounter insurmountable problems.

XII. Summary

Signal detection theory involves a level of abstraction unfamiliar in field studies of animal communication. Mastering its implications, however, leads to some strong predictions about the evolution of signals and responses and to some new procedures for investigating animal communication.

A consequence of this approach to communication is the fundamental conclusion that a receiver cannot independently adjust its P_{CD} and P_{FA}. The only exception is the limiting case in which the output of a channel in the presence of a signal is perfectly distinct from the output in its absence, so $P_{FA} = 0$. Otherwise, no matter how the criterion for response changes, any change in P_{CD} is accompanied by a corresponding change in P_{FA}.

This compromise leads ultimately to a prediction that receivers evolve to optimize the net utility of their responses. The optimum might lie anywhere between extremes of gullibility or fastidiousness. In turn, signalers should evolve to balance the often incompatible advantages of increased detectability of signals, increased complexity of encoding, and restriction of signals to intended receivers.

A second consequence of signal detection theory is the fundamental distinction between the detectability of a signal and the receiver's criterion for a response. Detectability depends on the contrast of the signal impinging on the subject and on the selectivity of the subject's perceptual channels. A receiver's criterion for response depends on its attitude toward the output of its perceptual channels, as a result of a decision to accept particular P_{FA} and P_{CD}.

Because any receiver's responses to stimulation depend on both detectability of the stimulus and criterion for response, a definitive interpretation of responses requires attention to both. For a full interpretation of a receiver's performance, it is necessary to include null presentations in experiments in order to measure false alarms as well as correct detections.

Signal detection theory thus suggests new ways to design and to interpret experiments that compare responses to stimulation. Although some practical difficulties face any application of signal detection theory to field studies, none seems insurmountable.

With this approach, we stand to learn more about (1) the adaptations for communication in situations with high background stimulation, such as in

choruses or complex social groups or at long range, (2) the effects of contrast, redundancy, reduced uncertainty, and familiarity on receivers' abilities to detect and discriminate signals, and (3) the evolution of exaggeration or dishonesty in signals as a consequence of the evolution of receivers' performance. In all of these ways, signal detection theory can advance our understanding of both the physiology and the evolution of communication.

Acknowledgments

References

Aubin, T., and Jouventin, P. (1998). Cocktail-party effect in king penguin colonies. *Proc. R. Soc. Lond. B* **265,** 1665–1673.

Benzschawel, T., and Cohn, T. E. (1985). Detection and recognition of visual targets. *J. Opt. Soc. Am. A* **2,** 1543–1550.

Brenowitz, E. A. (1982). The active space of red-winged blackbird song. *J. Comp. Physiol. A* **147,** 511–522.

Broadbent, D. E. (1958). ''Perception and Communication.'' Pergamon Press, London.

Broadbent, D. E. (1967). Word-frequency effect and response bias. *Psychol. Rev.* **74,** 1–15.

Brumm, H. (2004). The impact of environmental noise on song amplitude in a territoral bird. *J. Anim. Ecol.* **73,** 434–440.

Brumm, H., and Todt, D. (2002). Noise-dependent song amplitude regulation in a territorial songbird. *Anim. Behav.* **63,** 891–897.

Brumm, H., Voss, K., Köllmer, I., and Todt, D. (2004). Acoustic communication in noise: Regulation of call characteristics in a New World monkey. *J. Exp. Biol.* **207,** 443–448.

Cary, M., and Reder, L. M. (2003). A dual-process account of the list-length and strength-based mirror effects in recognition. *J. Mem. Lang.* **49,** 231–248.

Cherry, E. C. (1953). On the recognition of speech with one, and with two ears. *J. Acoust. Soc. Am.* **25,** 975.

Cherry, E. C., and Taylor, W. K. (1954). Some further experiments upon the recognition of speech with one, and with two ears. *J. Acoust. Soc. Am.* **26,** 554–559.

Cohn, T. (1978). Detection of 1-of-M orthogonal signals: Asymptotic equivalence of likelihood ratio and multiband models. *Opt. Lett.* **3,** 22–23.

Cynx, J., Lewis, R., Tavel, B., and Tse, H. (1998). Amplitude regulation of vocalizations in noise by a songbird, *Taeniopygia guttata. Anim. Behav.* **56,** 107–113.

Davies, D. R., and Parasuraman, R. (1982). ''The Psychology of Vigilance.'' Academic Press, London.

Davies, D. R., and Tune, G. S. (1970). ''Human Vigilance Performance.'' Staples, London.

Dobbins, I. G., Kroll, N. E. A., and Yonelinas, A. P. (2004). Dissociating familiarity from recollection using rote rehearsal. *Mem. Cognition* **32,** 932–944.

Dooling, R. J. (1982). Auditory perception in birds. *In* ''Acoustic Communication in Birds'' (D. E. Kroodsma and E. H. Miller, Eds.), Vol. 1, pp. 95–130. Academic Press, New York.

Egan, J. P., Schulman, A. I., and Greenberg, G. Z. (1959). Operating characteristics determined by binary decisions and by ratings. *J. Acoust. Soc. Am.* **31,** 768–773.

Egan, J. P., Greenberg, G. Z., and Schulman, A. I. (1961a). Operating characteristics, signal detectability and the method of free response. *J. Acoust. Soc. Am.* **33,** 993–1007.

Egan, J. P., Greenberg, G. Z., and Schulman, A. I. (1961b). Interval of time uncertainty in auditory detection. *J. Acoust. Soc. Amer.* **33,** 771–778.

Endler, J. A. (1990). On the measurement and classification of color in studies of animal colour patterns. *Biol. J. Linn. Soc. London* **41,** 315–352.

Endler, J. A. (1993). The color of light in forests and its implications. *Ecol. Monogr.* **63,** 1–27.

Endler, J. A., and Théry, M. (1996). Interacting effects of lek placement, display behavior, ambient light, and color patterns in three Neotropical forest-dwelling birds. *Am. Nat.* **148,** 421–452.

Enquist, M., and Arak, A. (1993). Selection of exaggerated male traits by female aesthetic senses. *Nature* **361,** 446–448.

Glanzer, M., and Adams, J. K. (1985). The mirror effect in recognition memory. *Mem. Cognit.* **13,** 8–20.

Glanzer, M., Adams, J. K., Iverson, G. J., and Kim, K. (1993). The regularities of recognition memory. *Psych. Rev.* **99,** 546–567.

Gomez, D., and Théry, M. (2004). Influence of ambient light on the evolution of colour signals: Comparative analysis of a Neotropical rainforest bird community. *Ecol. Lett.* **7,** 279.

Green, D. M. (1961). Detection of auditory sinusoids of uncertain frequency. *J. Acoust. Soc. Am.* **33,** 897–903.

Green, D. M., and Birdsall, T. G. (1978). Detection and recognition. *Psychol. Rev.* **85,** 192–206.

Green, D. M., and Swets, J. A. (1966). ''Signal Detection Theory and Psychophysics.'' Wiley & Sons, New York.

Green, D. M., Weber, D. L., and Duncan, J. E. (1977). Detection and recognition of pure tones in noise. *J. Acoust. Soc. Am.* **62,** 948–954.

Guilford, T., and Dawkins, M. S. (1991). Receiver psychology and the evolution of animal signals. *Anim. Behav.* **42,** 1–14.

Guilford, T., and Dawkins, M. S. (1993). Receiver psychology and the design of animal signals. *Trends Neurosci.* **16,** 430–436.

Heindl, M., and Winkler, H. (2003). Vertical lek placement of forest-dwelling manakin species (Aves, Pipridae) is associated with vertical gradients of ambient light. *Biol. J. Linn. Soc.* **80,** 647–658.

Klein, S. A. (1985). Double-judgment psychophysics, problems and solutions. *J. Opt. Soc. Am. A* **2,** 1560–1585.

Klump, G. M. (1996). Bird communication in the noisy world. *In* ''Ecology and Evolution of Acoustic Communication in Birds'' (D. E. Kroodsma and E. H. Miller, Eds.), pp. 321–338. Cornell University Press, Ithaca.

Klump, G. M., and Langemann, U. (1995). Comodulation masking release in a songbird. *Hearing Res.* **87,** 157–164.

Langemann, U., and Klump, G. M. (2001). Signal detection in amplitude-modulated maskers: I. Behavioral auditory threshold in a songbird. *Eur. J. Neurosci.* **13,** 1025–1032.

Langemann, U., Gauger, B., and Klump, G. M. (1998). Auditory sensitivity in the great tit perception of signals in the presence and absence of noise. *Anim. Behav.* **56,** 763–769.

Leonard, M. L., and Horn, A. G. (2005). Ambient noise and the design of begging signals. *Proc. Biol. Sci.* **272,** 651–656.

Loeb, M., and Alluisi, E. A. (1977). An update of findings regarding vigilance and a reconsideration of underlying mechanisms. *In* "Vigilance: Theory, Operational Performance, and Physiological Correlates" (R. R. Mackie, Ed.), pp. 719–749. Plenum Press, New York.

Luce, R. D. (1963). A threshold theory for simple detection experiments. *Psych. Rev.* **70,** 61–79.

Luce, R. D., and Green, D. M. (1974). Detection, discrimination and recognition. *In* "Handbook of Perception" (E. C. Carterette and M. P. Friedman, Eds.), Vol. 2. Academic Press, New York.

Lynn, S. K., Cnaani, J., and Papaj, D. R. (2005). Peak shift discrimination learning as a mechanism of signal evolution. *Evolution* **59,** 1300–1305.

Mackie, R. R. (Ed.) (1977). "Vigilance: Theory, Operational Performance, and Physiological Correlates." Plenum Press, New York.

Macmillan, N. A. (2002). Signal detection theory. *In* "Stevens' Handbook of Experimental Psychology, Third Edition. Methodology in Experimental Psychology" (H. Pashler and J. Wixted, Eds.), Vol. 4, pp. 43–90. John Wiley & Sons, New York.

Macmillan, N. A., and Creelman, C. D. (1990). Response bias, characteristics of detection theory, threshold theory, and "nonparameteric" indexes. *Psychol. Bull.* **107,** 401–413.

Macmillan, N. A., and Creelman, C. D. (1991). "Detection Theory: A User's Guide." Cambridge University Press, Cambridge.

Macmillan, N. A., and Kaplan, H. L. (1985). Detection theory analysis of group data, estimating sensitivity from average hit and false-alarm rates. *Psychol. Bull.* **98,** 185–199.

Mulligan, R., and Shaw, M. L. (1980). Multimodal signal detection, independent decisions vs. integration. *Percept. Psychophys.* **28,** 471–478.

McNicol, D. (1972). "A Primer of Signal Detection Theory." Allen and Unwin, London.

Naguib, M. (2003). Reverberation of rapid and slow trills: Implications for signal adaptations to long range communication. *J. Acoust. Soc. Am.* **133,** 1749–1756.

Naguib, M., and Wiley, R. H. (2001). Estimating the distance to a source of sound: Mechanisms and adaptations for long-range communication. *Anim. Behav.* **62,** 825–837.

Nieder, A., and Klump, G. M. (2001). Signal detection in amplitude-modulated maskers: II. Processing in the songbird's auditory forebrain. *Eur. J. Neurosci.* **13,** 1033–1044.

Nolte, L. W., and Jaarsma, D. (1967). More on the detection of one of M orthogonal signals. *J. Acoust. Soc. Am.* **41,** 497–505.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *J. Opt. Soc. Am. A* **2,** 1508–1532.

Pierce, J. (1963). Some sources of artifact in studies of tachistoscopic perception of words. *J. Exp. Psychol.* **66,** 363–370.

Pollack, I., Rubenstein, H., and Decker, L. (1959). Intelligibility of known and unknown message sets. *J. Acoust. Soc. Am.* **31,** 273–279.

Richards, D. G. (1981a). Alerting and message components in songs of rufous-sided towhees. *Behaviour* **76,** 223–249.

Richards, D. G. (1981b). Estimation of distance of singing conspecifics by the Carolina wren. *Auk* **98,** 127–133.

Seehausen, O., van Alphen, J. J. M., and Witte, F. (1997). Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* **277,** 1808–1811.

Shaw, M. L. (1982). Attending to multiple sources of information, I. The integration of information in decision making. *Cognit. Psychol.* **14,** 353–409.

Slabbekoorn, H., and Peet, M. (2003). Birds sing at a higher pitch in urban noise. *Nature (Lond.)* **424,** 267.

Starr, S. J., Metz, C. E., Lusted, L. B., and Goodenough, D. J. (1975). Visual detection and localization of radiographic images. *Radiology* **116,** 533–538.

Swensson, R. G., and Judy, P. F. (1981). Detection of noisy visual targets: Models for the effects of spatial uncertainty and signal-to-noise ratio. *Percept. Psychophys.* **29,** 521–534.

Swets, J. A., and Birdsall, T. G. (1978). Repeated observation of an uncertain signal. *Percep. Psychophys.* **23,** 269–274.

Swets, J. A., Shipley, E. F., JMcKee, J. M., and Green, D. M. (1959). Multiple observations of signals in noise. *J. Acoust. Soc. Am.* **31,** 514–521.

Tanner, W., and Swets, J. (1954). The human use of information—I: Signal detection for the case of the signal known exactly. *IRET. Inform. Theor.* **4,** 213–221.

Tanner, W. P., Jr. (1956). Theory of recognition. *J. Acoust. Soc. Amer.* **28,** 882–888.

ten Cate, C., and Bateson, P. P. G. (1988). Sexual selection: The evolution of conspicuous characteristics in birds by means of imprinting. *Evolution* **42,** 1355–1358.

Thomas, J. P. (1985). Detection and identification, how are they related? *J. Opt. Soc. Am. A* **2,** 1457–1467.

Uy, J. A. C., and Endler, J. A. (2004). Modification of the visual background increases the conspicuousness of golden-collared manakin displays. *Behav. Ecol.* **15,** 1003–1010.

Warm, J. S. (1977). Psychological processes in sustained attention. *In* ''Vigilance: Theory, Operational Performance, and Physiological Correlates'' (R. R. Mackie, Ed.), pp. 623–644. Plenum Press, New York.

Watson, C. S., and Nichols, T. L. (1976). Detectability of auditory signals presented without defined observation intervals. *J. Acoust. Soc. Amer.* **59,** 655–668.

Weary, D., Guilford, T. C., and Weisman, R. G. (1993). A product of discriminative learning may lead to female preferences for elaborate males. *Evolution* **47,** 333–336.

Wiley, R. H. (1983). The evolution of communication, information and manipulation. *In* ''Animal Behaviour, Volume 2, Communication'' (T. R. Halliday and P. J. B. Slater, Eds.), pp. 131–181. Blackwell Scientific Publications, Oxford.

Wiley, R. H. (1991). Associations of song properties with habitats for territorial oscine birds of eastern North America. *Am. Nat.* **138,** 973–993.

Wiley, R. H. (1994). Errors, exaggeration, and deception in animal communication. *In* ''Behavioral Mechanisms in Evolutionary Ecology'' (L. R. Real, Ed.), pp. 157–189. University of Chicago Press, Chicago.

Wiley, R. H., and Richards, D. G. (1982). Adaptations for acoustic communication in birds, sound propagation and signal detection. *In* ''Acoustic Communication in Birds'' (D. E. Kroodsma and E. H. Miller, Eds.), Vol. 1, pp. 131–181. Academic Press, New York.

Wollerman, L., and Wiley, R. H. (2002). Background noise from a natural chorus alters female discrimination of male calls in a Neotropical frog. *Anim. Behav.* **63,** 15–22.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *J. Mem. Lang.* **46,** 441–517.