Signal Detection Theory

J. D. Balakrishnan and J. A. MacDonald
Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907, USA

1. INTRODUCTION

Many of the jobs carried out by human operators in the workplace have a "detection" component to them. Various kinds of watchkeeping tasks (e.g. radar monitoring, airport security, product quality control) are some of the more obvious examples. Other seemingly more complicated tasks, such as medical diagnosis and personnel evaluation, can also be reduced in many instances to a set of all-or-none decisions about the need for an intervention of some kind. In virtually all of these cases, the potential for human error exists, and a critical question for the human factors specialist is whether some or all of these errors can be attributed to operator training, system design, or other aspects of the job that could be retooled in some way.

To answer this question, many researchers prefer to use well-controlled experiments and objective performance indices, if they are feasible. In addition to a rigorous measure of the difficulty of an operator's task and a numerical scale for comparing different approaches to a problem, these indices make it possible for the researcher to examine how factors such as attention and strategy may interact with system design or other factors to determine the overall performance pattern of the operator. Unfortunately, none of the performance indices currently available to researchers are "theory-independent." Essentially, they are byproducts of a relatively simple, skeletal model of the operator's behavior in a goal-directed setting. To be used appropriately, therefore, their assumptions and limitations need to be properly understood. Here we briefly review the several measures and experimental techniques that originated with one of the most successful performance theories in psychology, the so-called Theory of Signal Detection (or, Signal Detection Theory; e.g. Green and Swets, 1966). Elsewhere we consider some alternatives to signal detection theory and some recent empirical results that may fundamentally change the way operator performance is assessed in the future.

2. FORMAL MODELS OF DETECTION

In most real world tasks, the operator can and will make different kinds of errors, with different consequences, and the rate at which some kinds of errors are committed will often be inversely related to the rate of other kinds of errors, presumably because of attention shifts or other changes in the operator's approach to the task. In other words, there will be trade-offs. Changes in the environment or system design presumably can affect the decision-making strategy of the operator, exclusively or in addition to their effects on the quality of the information exchange between humans and machines. Formal models such as signal detection theory represent attempts to provide measures of information processing capacity that are not confounded by the operator's decision-making biases.

The first step in the development of such a model is to analyze the logical structure of the operator's task. With few exceptions, human errors in the workplace are incorrect decisions about the state of a system or real world event, or about the best possible course of action in response to a situation. Signal detection theory encompasses both kinds of errors by dividing the decision-making process into two discrete, non-overlapping stages. In the first stage, the operator collects information from the outside world, and in the second stage the operator applies a decision-making strategy or rule to the information state to arrive at a decision about the nature of the circumstances and/or the best course of action. Of course, in most workplace environments, the situations faced by the operator are constantly changing and so this encoding/decision-making sequence must continuously repeat itself. The absence of any overt action on the part of the operator at any given point in time presumably represents a decision (conscious or otherwise) not to react on the basis of current information about the environment.

The terms "collect information" and "apply a rule" are very general, which is one reason why signal detection theory can be so widely used. Any kind of information processing task will involve information collection of some kind followed by the selection of an appropriate action. Errors can be due to poor information (encoding errors) or faulty decision-making (response selection) strategies. The main objective of the signal detection theory analysis is to separate and quantify these two kinds of errors so that their relative frequencies and their dependence on the properties of the system can be studied.

To do this, the theory borrows some fundamental concepts from the statistical decision-making literature. First, the information that reaches the decision-maker is assumed to be ambiguous or "noisy": the same information state can be produced by more than one of the physical events that need to be discriminated. The decision-maker's problem is, therefore, to determine which event is more likely to have produced this information state and from this the action that is most likely to be correct. In this way, the decision-making process is analogous to a statistical hypothesis test. The operator makes an error when the information output of his/her encoding process is sufficiently misleading, or when the statistical decision rule is misconstrued.

At least in principle, this theoretical framework can be adapted to arbitrarily complex decision-making tasks, involving many different situations and many possible decisions that might need to be taken (and to some extent, Thurstonian Scaling is an example of this; Torgerson 1958). However, most of the efforts to develop the theory have concentrated on relatively simple decision-making tasks in which there are only two possible states and, therefore, two possible classification responses. In this two-choice classification (or discrimination) task, the two possible states of the world can be arbitrarily labeled "noise" and "signal," and there are two possible correct responses (correct signal responses or "hits" and correct "noise" responses, or "correct rejections") and two kinds of errors (incorrect "signal" responses to noise or "false alarms" and incorrect "noise" responses to signals, or "misses"). Signal detection theory assumes that the information state of the operator can be represented by a point on a one-dimensional, bipolar continuum or "decision axis." Values on the one end of this continuum represent instances during the experiment in which the operator is highly certain that the correct response decision should be "noise," and values the other end represent high confidence "signal" responses. Somewhere between

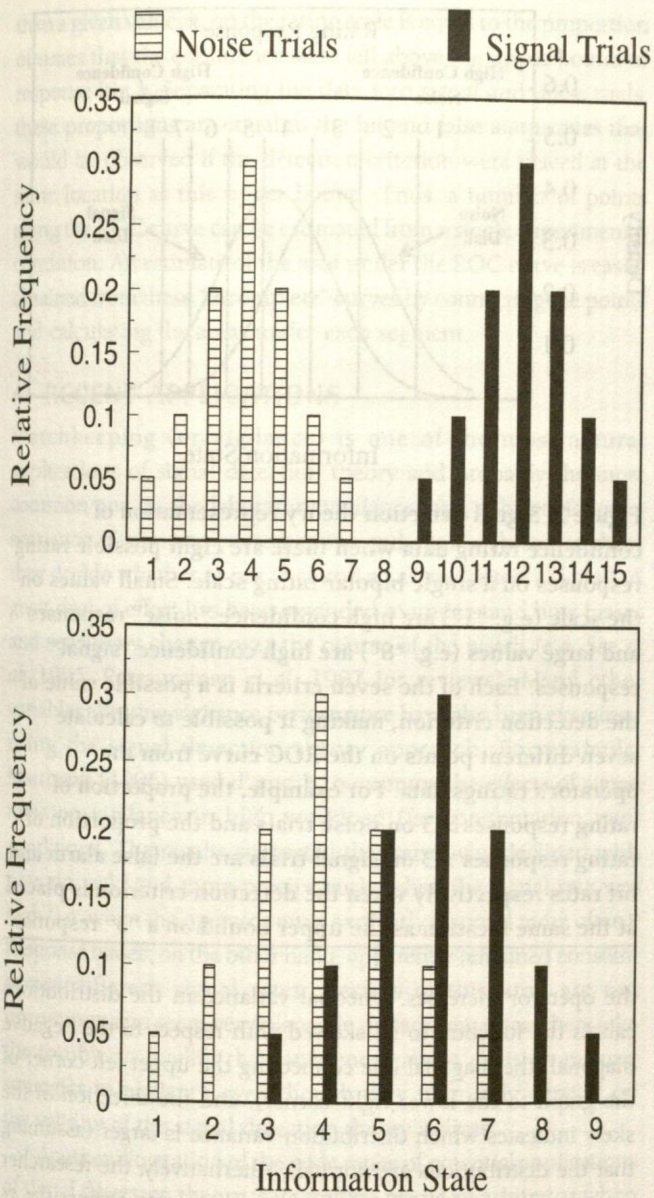

Figure 1. Concept of sensitivity in signal detection theory. (upper) The presentation of a signal puts the operator into one of seven states (9–15 on the abscissa). Noise alone never causes any of these information states, but instead may cause one of a different set (1–7). In this case, the operator can (with an appropriate decision rule) perfectly discriminate between the two events. (lower) Some of the information states (3–7) are caused by both signal and by noise alone (i.e. the distributions overlap), making discrimination errors inevitable for all possible decision rules.

these two extremes is the "complete uncertainty" or "indifference" point.

2.1. Encoding Distributions and Detection Criteria

If the presentation of a signal (or noise alone) does not always have the same effect on the operator, then it follows that the information state of the operator has some univariate probability (relative frequency) distribution. The difference between this distribution of states on noise trials from its distribution on signal trials presumably depends on the physical differences between the two events and the "sensitivity" of the operator to these differences. In fact, the degree overlap of these two distributions would be the degree to which the noise and signal stimuli are confusable when they are presented equally often. This property is illustrated in Figure 1. In the upper panel, the distributions are completely non-overlapping. In this case, the decision-maker can always correctly identify the stimulus because the information states caused by the noise stimulus are never caused by the signal stimulus, and vice versa. In the lower panel, the distributions overlap. In this case, information states lying close to the middle of the continuum have similar relative frequencies under the two stimulus conditions, and perfectly equivalent relative frequencies for one value (information state "5"). No matter which response is assigned to information states "3" through "7," the operator's decision will sometimes be incorrect.

Using a probability distribution to describe the effect a stimulus on the operator was standard practice long before signal detection theory was developed. The new contribution of this theory was in its emphasis on the sophisticated decision-making processes that should be (and presumably are) applied by human operators to minimize the problems caused by noisy encoding. In addition to the relative frequency of a state on signal trials (or on noise trials), the statistical decision-maker would also consider the relative frequencies, or "base rates," of the signal and noise events during the experiment, the relative costs of the two possible errors (false alarms and misses), and the rewards for the two possible correct decisions (hits and correct rejections). For example, if the signal occurs very infrequently (i.e. its base rate is low), then a state that occurs with moderate frequency on signal trials and moderator frequency on noise trials would actually constitute fairly strong evidence in favor of a "noise" response. Presumably, the operator combines knowledge about the probability distributions with the base rates and payoffs to select an appropriate "criterion" or "decision boundary" to divide the information state continuum into two response regions. As the base rate of the signal event increases, for example, the operator presumably shifts the criterion to the left, increasing the size of the signal response region and hence the relative frequency of the signal response.

3. PERFORMANCE INDICES

Different assumptions about the shapes of the information state distributions partition signal detection theory into several different detection models. The most popular of these assumes that the distributions are normal with equal variance. In this case, the distance between the means of the two distributions in standard deviation units, or *d'*, is a suitable index of the degree to which the distributions overlap ("sensitivity"), and the ratio of the two distributions at the criterion (b) is an index of preference for one of the two responses ("response bias"). Thus, b < 1 would indicate a bias towards the signal response and > 1 would indicate a bias towards the noise response. The decision rule is unbiased (b = 1) when the criterion is set at the point of intersection between the two distributions, which occurs at the midpoint between their means. Both *d'* and b can be calculated from the hit and false alarm rates of the operator.

If the distributions are normal but have different variances, then an additional parameter is needed to fix the sensitivity scale. Unfortunately, these indices cannot be obtained from a single pair of hit and false alarm rates, which presumably explains why they are not widely used, despite the fact that the equal variance assumption is generally untenable. To fit the unequal variance normal model to data, estimates of the hit and false alarm rates must be obtained for several different values of the decision criterion.

3.1. Criterion Shifts and the ROC Curve

Running the same experiment under several different base rate or payoff conditions is one way to obtain the extra data needed to allow for unequal variance. Another method is simply to instruct the operator to favor one type of response over the other to some degree. Each manipulation should cause the operator to shift the detection criterion. The different pairs of hit and false alarm rates can then be graphed together in a single parametric plot (or scatterplot) called the "receiver operating characteristic" (ROC) curve. The points on the graph are the hit rates (on the ordinate) corresponding to each observed false alarm rate (on the abscissa). Two examples are shown in Figure 2. For several reasons the shape of the ROC curve is the most powerful diagnostic from the point of view of signal detection theory. The area under the curve, for example, increases as the sensitivity of

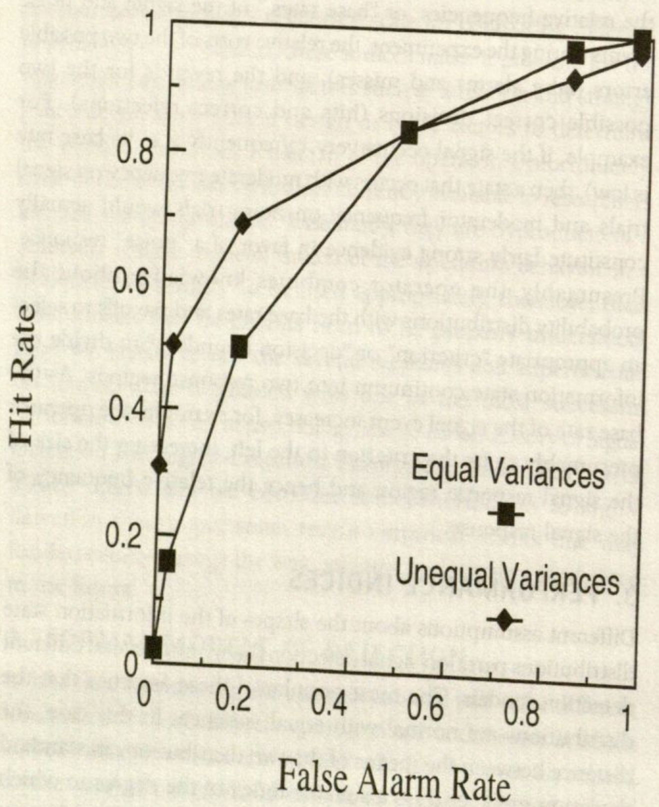

Figure 2. Two possible ROC curves when the distributions are normal with equal variance (squares) or with larger variance in the signal distribution (diamonds). Each of the six different points on the curve corresponds to a different placement of the detection criterion. Area under the complete ROC curve (when the hit and false alarm rate are calculated for every possible location of the detection criterion) can be estimated by computing the area under the these incomplete functions. In this case, however, the operators' placement of the criteria will have some effect on the measure.

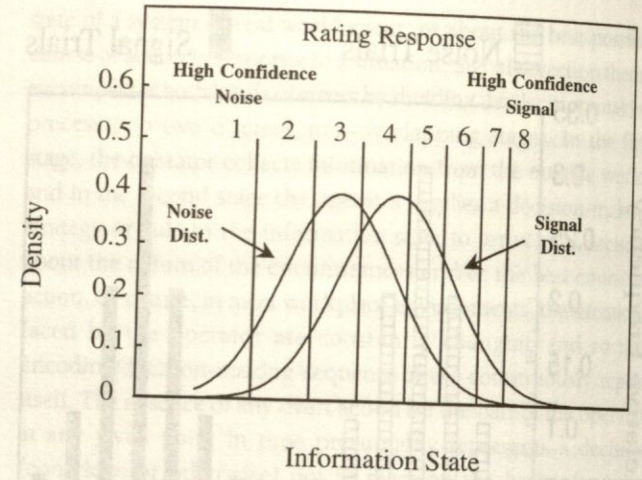

Figure 3. Signal detection theory representation of confidence rating data when there are eight possible rating responses on a single bipolar rating scale. Small values on the scale (e.g. "1") are high confidence "noise" responses and large values (e.g. "8") are high confidence "signal" responses. Each of the seven criteria is a possible value of the detection criterion, making it possible to calculate seven different points on the ROC curve from the operator's ratings data. For example, the proportion of rating responses > 3 on noise trials and the proportion of rating responses > 3 on signal trials are the false alarm and hit rates respectively when the detection criterion is placed at the same location as the upper bound on a "3" response.

the operator increases. Unequal variance in the distributions causes the function to be skewed with respect to the negative diagonal (the diagonal line connecting the upper left corner of the graph to the lower right corner), and the direction of the skew indicates which distribution variance is larger (assuming that the distributions are normal). Alternatively, the researcher can plot the z-transform of the false alarm rate against the z-transform of the corresponding hit rate and observe the shape of this "z-ROC" scatterplot. If the normality assumption is satisfied, this function will be linear with a slope equal to the ratio of the standard deviation of the noise distribution to the standard deviation of the signal distribution.

3.2. Ratings Method

Empirical studies of the shape of a ROC curve can be time consuming and expensive if each point of the curve must be estimated from a different experimental condition. An alternative approach that is perfectly consistent with the signal detection theory assumptions, and at the same time considerably more efficient in terms of costs, is to elicit confidence rating responses from the operator. According to signal detection theory, the operator's information is "graded" and stochastic, varying from strong and reliable on some occasions to weak and uncertain on others. Asking the operators to choose one of two responses is equivalent to asking them to separate their confidence states into two "response bins": "more confident signal" and "more confident noise." Asking them to also rate their confidence in their detection judgment is equivalent to asking them to separate their confidence states into more than two confidence bins. The results are illustrated in Figure 3. The proportion of rating responses larger than a given value, k, on the rating scale is equal to the proportion of times that the confidence state fell above the upper bound of response bin k. Separating the data into signal and noise trials, these proportions are equal to the hit and false alarm rates that would be observed if the detection criterion were placed at the same location as this upper bound. Thus, a number of points along the ROC curve can be estimated from a single experimental condition. An estimate of the area under the ROC curve is easily obtained from these "incomplete" curves by connecting the points and calculating the areas under each segment.

4. RECENT APPLICATIONS

Watchkeeping (or vigilance) is one of the most natural applications of signal detection theory and probably the most common one. In the laboratory vigilance task, subjects follow a sequence of stimulus presentations and after each one of these they decide whether or not to sound an alarm (detect a signal). A great deal of effort has been extended to understand how biases and sensitivity change over the course of the watch (e.g. See *et al.* 1995, Parasuraman *et al.* 1987 for reviews). Many other variables affecting vigilance performance have also been examined using the signal detection theory approach. For example, Matthews (1996) used d' and b to examine the effects of signal rates on vigilance in high workload (fast presentation rate) conditions. The results suggested that sensitivity declined with time on task, and more precipitously when the signal rate was high (i.e. when the operator must explicitly respond more often). Response biases, on the other hand, apparently remained constant across different signal rates. Results of this kind are not uncommon and seem very plausible. In fact, few researchers take the trouble to qualify their inferences about decision-making strategies in vigilance even though they ultimately depend on the validity of the signal detection theory indices.

Another illustration of the wide range of potential applications of signal detection theory is its routine use as an arbitrator when two relatively informal theories or hypotheses make different predictions about the effects of a factor on the operator's performance. If one theory predicts that there should be no effect of this factor or the opposite kind of effect, then the d' measure is typically recruited to verify that observed differences in overall percent correct or in the hit and false alarm rates are not merely due to effects of the factor on response bias. The direction and size of the operator's response bias are not usually specifically predicted by theories of operator performance (response biases are, by definition, "subject-controlled" factors) and so the d' statistic is usually the measure of most interest. Patterns of bias, however, can sometimes inform the researcher about variables such as cognitive style or motivation. For example, in a recent study on aircraft recognition and recognition training, Goettl (1996) looked at two memory models and their predictions about aircraft recognition under different learning conditions. Recognizing whether an aircraft is or is not a member of a predefined class (e.g. commercial versus noncommercial) is a two-response discrimination task, making it possible to apply a signal detection theory analysis. The signal detection theory measures were used to estimate memory strength (sensitivity) and the results indicated that for male subjects, there was a difference between two different types of learning schedules on memory strength and not merely on bias, but for the females,the differences between the two schedules could be attributed to bias effects alone.

Another natural application for signal detection theory has been in the various kinds of expert decision-making problems involved in medical and clinical diagnosis. ROC curves are used fairly routinely in these areas to characterize the extent to which some quantifiable property of a medical image informs the physician about the presence or absence of a pathology. Researchers also make use of these measures to compare experienced and inexperienced diagnosticians and to show how new imaging technologies can combine with or replace traditional methods of diagnosis (e.g. Jiang *et al.* 1999, Tsuda *et al.* 1999). Quite often, new approaches to diagnosis do not increase both the "hit" and "correct rejection" rates, making bias effects a very important factor to consider.

In each of these examples, and in many others, some interesting and important discoveries about human performance, including losses of sensitivity with time on task, idiosyncratic response biases, and other well-documented phenomena, would not be possible without the benefit of a formal measurement system. However, these discoveries are only "conditional" statements of fact because of their dependence on the specific assumptions of signal detection theory. Other interpretations of the data are possible, and occasionally other indices are adopted in addition to or in place of the signal detection theory measures. In the next chapter, we look at the evidence in support of signal detection theory and then discuss some of the alternatives and their motivations.

ACKNOWLEDGEMENTS

NASA Dryden Flight Research Center Grant NCC2-374 supported the authors.

REFERENCES

GOETTL, B.P., 1996, The spacing effect in aircraft recognition. *Human Factors*, 38, 34–49.

GREEN, D.M. and SWETS, J.A., 1966, *Signal Detection Theory and Psychophysics* (New York: Wiley; repr. Huntington: Krieger, 1974).

JIANG, Y., NISHIKAWA, R.M., SCHMIDT, R.A., METZ, C.E., GIGER, M.L. and DOI, K., 1999, Improving breast cancer diagnosis with computer-aided diagnosis. *Academic Radiology*, 6, 22–33.

MATTHEWS, G., 1996, Signal probability effects on high-workload vigilance tasks. *Psychonomic Bulletin and Review*, 3, 339–43.

PARASURAMAN, R., WARM, J.S. and DEMBER, W.N., 1987, Vigilance: taxonomy and utility. In L.S. Mark, J.S. Warm and R.L. Huston (eds), *Ergonomics and Human Factors: Recent Research* (New York: Springer).

SEE, J.E., HOWE, S.R., WARM, J.S. and DEMBER, W.N., 1995, Meta-analysis of the sensitivity decrement in vigilance. *Psychological Bulletin*, 117, 230–49.

TORGERSON, W.S., 1958, *Theory and Methods of Scaling* (New York: Wiley).

TSUDA, K., YU, K.K., COAKLEY, F.V., SRIVASTAV, S.K., SCHEIDLER, J.E. and HRICAK, H., 1999, Detection of extracapsular extension of prostate cancer: role of fat suppression endorectal MRI. *Journal of Computer Assisted Tomography*, 23, 74–8.