

A PRIMER OF SIGNAL DETECTION THEORY

D.McNicol Professor Emeritus, University of Tasmania, Australia

With a New Foreword by Brian C.J.Moore



LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS Mahwah, New Jersey London Originally published 1972.

Copyright © 2005 by Lawrence Erlbaum Associates, Inc.

All rights reserved. No part of this book may be reproduced in any form, by photostat, microform, retrieval system, or by any other means, without the prior written permission of the publisher.

Lawrence Erlbaum Associates, Inc., Publishers 10 Industrial Avenue Mahwah, New Jersey 07430

This edition published in the Taylor & Francis e-Library, 2009.

To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk.

Library of Congress Cataloging-in-Publication Data McNicol, D. A primer of signal detection theory/D.McNicol. p. cm. Originally published: London : Allen & Unwin, 1972. With new foreword. Includes bibliographical references and index. ISBN 0-8058-5323-5 (pbk.: alk. paper) 1. Signal detection (Psychology) 2. Psychometrics. 1. Title. BF441.M22 2004 152.8—dc22 2004056290

ISBN 1-4106-1194-9 Master e-book ISBN

Foreword

Signal Detection Theory (SDT) has had, and continues to have, an enormous impact on many branches of psychology. Although its initial applications were in the interpretation of sensory processes, its domain has since widened considerably. For example, concepts derived from SDT are widely used in memory research and in studies of the processing of verbal information. SDT has been called by many a revolution and I do not think that is an exaggeration. A basic understanding of SDT has become essential for anyone with a serious interest in experimental psychology.

The classic work on SDT is *Signal Detection Theory and Psychophysics* by D.M.Green and J.A.Swets, originally published by Wiley in 1966 and reprinted with corrections by Kreiger in 1974. This remains a very useful source for advanced researchers and those with mathematical sophistication. However, for many readers, the descriptions and derivations will be beyond their grasp. A more recent and more user-friendly text is *Detection Theory: A User's Guide* by N.A. Macmillan and C.D. Creelman, originally published in 1991. The second edition of this book has just been published by Lawrence Erlbaum Associates. The Macmillan and Creelman book still assumes a good deal of mathematical and statistical sophistication, but it makes more use of visual analogies, and is intended especially as a practical guide to those actively involved in areas of research that depend on a good understanding and appreciation of SDT. In their preface, Macmillan and Creelman state "It could be the basic text in a one-semester graduate or upper level undergraduate course." However some undergraduates may find it too detailed if they simply want to get a basic understanding of SDT.

When I first had to teach SDT to undergraduates in psychology, I was delighted to come across *A Primer of Signal Detection Theory* by D.McNicol, published by George Allen and Unwin in 1972. This was the only text book I could find that covered SDT at an introductory level, and that assumed only limited skills in algebra and statistics. I used this book with success as my recommended text for several my two personal copies of the book both "went missing," and all copies in our library also mysteriously disappeared. I was left "in limbo" for many years. It is with relief and pleasure that I now write this foreword to a new printing of the book. I can strongly recommend the book as an introduction to SDT and its applications. It is suitable for use as a student text book, but will also be useful for teachers and researchers in psychology who need to acquire a working understanding of SDT.

—Brian C.J.Moore, FmedSci, FRS Department of Experimental Psychology University of Cambridge

Preface

There is hardly a field in psychology in which the effects of signal detection theory have not been felt. The authoritative work on the subject, Green's & Swets' *Signal Detection Theory and Psychophysics* (New York: Wiley) appeared in 1966, and is having a profound influence on method and theory in psychology. All this makes things exciting but rather difficult for undergraduate students and their teachers, because a complete course in psychology now requires an understanding of the concepts of signal detection theory, and many undergraduates have done no mathematics at university level. Their total mathematical skills consist of dim recollections of secondary school algebra coupled with an introductory course in statistics taken in conjunction with their studies in psychology. This book is intended to present the methods of signal detection theory to a person with such a mathematical background. It assumes a knowledge only of elementary algebra and elementary statistics. Symbols and terminology are kept as close as possible to those of Green & Swets (1966) so that the eventual and hoped for transfer to a more advanced text will be accomplished as easily as possible.

The book is best considered as being divided into two main sections, the first comprising Chapters 1 to 5, and the second, Chapters 6 to 8. The first section introduces the basic ideas of detection theory, and its fundamental measures. The aim is to enable the reader to be able to understand and compute these measures. The section ends with a detailed working through of a typical experiment and a discussion of some of the problems which can arise for the potential user of detection theory.

The second section considers three more advanced topics. The first of these, which is treated thoroughly elsewhere in the literature, is threshold theory. However, because this contender against signal detection theory has been so ubiquitous in the literature of experimental psychology, and so powerful in its influence both in the construction of theories and the design of experiments, it is discussed again. The second topic concerns the extension of detection theory, which customarily requires experiments involving recognition tests, to experiments using more open-ended procedures, such as recall; and the third topic is an examination of Thurstonian scaling procedures which extend signal detection theory in a number of useful ways.

An author needs the assistance of many people to produce his book, and I have been no exception. I am particularly beholden to David Ingleby, who, when he was working at the Medical Research Council Applied Psychology Unit, Cambridge, gave me much useful advice, and who was subsequently most generous in allowing me to read a number of his reports. The reader will notice frequent reference to his unpublished Ph.D. thesis from which I gained considerable help when writing Chapters 7 and 8 of this book. Many of my colleagues at Adelaide have helped me too, and I am grateful to Ted Nettelbeck, Ron Penny and Maxine Shephard, who read and commented on drafts of the manuscript, to Su Williams and Bob Willson, who assisted with computer programming, and to my Head of Department, Professor A.T.Welford for his encouragement. I am equally indebted to those responsible for the production of the final manuscript which was organised by Margaret Blaber ably assisted by Judy Hallett. My thanks also to Sue Thom who prepared the diagrams, and to my wife Kathie, who did the proof reading.

The impetus for this work came from a project on the applications of signal detection theory to the processing of verbal information, supported by Grant No A67/16714 from the Australian Research Grants Committee. I am also grateful to St John's College, Cambbridge, for making it possible to return to England during 1969 to work on the book, and to Adelaide University, which allowed me to take up the St John's offer.

A final word of thanks is due to some people who know more about the development of this book than anyone else. These are the Psychology III students at Adelaide University who have served as a tolerant but critical proving ground for the material which follows.

Adelaide University September 1970 D.MCNICOL

Contents

	Foreword by Brian C.J.Moore	
	Preface	
1	WHAT ARE STATISTICAL DECISIONS?	1
	An example	1
	Some definitions	2
	Decision rules and the criterion	5
	Signal detection theory and psychology	8
2	NON-PARAMETRIC MEASURES OF SENSITIVITY	14
	The yes-no task	14
	The rating scale task	20
	Area estimation with only a single pair of hit and false alarm rates	25
	The forced-choice task	33
	An overall view of non-parametric sensitivity measures	36
3	GAUSSIAN DISTRIBUTIONS OF SIGNAL AND NOISE WITH EQUAL	
	VARIANCES	41
	The ROC curve for the yes-no task	41
	Double-probability scales	44
	The formula for d'	47
	The criterion	48
	Forced-choice tasks	53
4	GAUSSIAN DISTRIBUTIONS OF SIGNAL AND NOISE WITH UNEQUAL VABLANCES	66
	ROC curves for unequal variance cases	66
	Sensitivity measures in the unequal variance case	72
	Measuring the signal distribution variance	76
	The measurement of response bias	76
5	CONDUCTING A RATING SCALE EXPERIMENT	82
	Experimental design	82
	Analysis of data	86
	Measures of sensitivity	94
	Measures of bias	99

6 CHOICE THEORY APPROXIMATIONS TO SIGNAL DETECTION	
THEORY	109
The logistic distribution	111
Determining detection measures from logistic distributions	113
The matrix of relative response strengths	116
Open-ended tasks	117
A summary of the procedure for an open-ended task	123
7 THRESHOLD THEORY	131
Classical psychophysics	131
High threshold theory and the yes-no task	135
Forced-choice tasks	144
Other evidence favouring detection theory	151
8 THE LAWS OF CATEGORICAL AND COMPARATIVE JUDGEMENT	155
Antecedents of signal detection theory	155
Category rating tasks	156
Forced-choice tasks and the Law of Comparative Judgement	174
Bibliography	182
Appendix 1 Answers to problems	187
Appendix 2 Logarithms	191
Appendix 3 Integration of the expression for the logistic curve	193
Appendix 4 Tables	195
Index	205

Chapter 1 WHAT ARE STATISTICAL DECISIONS?

AN EXAMPLE

Often we must make decisions on the basis of evidence which is less than perfect. For instance, a group of people has heights ranging from 5 ft 3 in. to 5 ft 9 in. These heights are measured with the group members standing in bare feet. When each person wears shoes his height is increased by 1 inch, so that the range of heights for the group becomes 5 ft 4 in. to 5 ft 10 in. The distributions of heights for members of the group with shoes on and with shoes off are illustrated in the histograms of Figure 1.1.



FIGURE 1.1

You can see that the two histograms are identical, with the exception that s, the 'Shoes on' histogram, is 1 in. further up the *X*-axis than n, the 'Shoes off histogram.

Given these two distributions you are told that a particular person is 5 ft 7 in. tall and from this evidence you must deduce whether the measurement was taken with shoes on or with shoes off. A look at these histograms in Figure 1.1 shows that you will not be able to make a decision which is certain to be correct. The histograms reveal that 3/16ths of the group is 5 ft 7 in. tall with shoes off and that 4/16ths of the group is 5 ft 7 in. tall with shoes on. The best bet would be to say that the subject had his shoes on when the measurement was taken. Furthermore, we can calculate the odds that this decision is correct. They will be (4/16)/(3/16), that is, 4/3 in favour of the subject having his shoes on.

You can see that with the evidence you have been given it is not possible to make a completely confident decision one way or the other. The best decision possible is a statistical one based on the odds favouring the two possibilities, and that decision will only guarantee you being correct four out of every seven choices, on the average.

2 A primer of signal detection theory

It is possible to calculate the odds that each of the eight heights of the group was obtained with shoes on. This is done in Table 1.1. The probabilities in columns 2 and 3 have been obtained from Figure 1.1.

	Probability of obtain		
Height in inches x	Shoes off $(n) P(x n)$	Shoes on (s) $P(x s)$	Odds favouring $s l(x)$
63	1/16	0	0
64	2/16	1/16	1/2
65	3/16	2/16	2/3
66	4/16	3/16	3/4
67	3/16	4/16	4/3
68	2/16	3/16	3/2
69	1/16	2/16	2/1
70	0	1/16	

TABLE 1.1 The odds favouring the hypothesis 'Shoes on' for the eight possible heights of group members.

P(x|n) and P(x|s) are called 'conditional probabilities' and are the probabilities of x given n, and of x given s, respectively.

l(x) is the symbol for the 'odds' or likelihood ratio.

For the sake of brevity we will refer to the two states of affairs 'Shoes on' and 'Shoes off as states *s* and *n* respectively.

It can be seen that the odds favouring hypothesis *s* are calculated in the following way:

For a particular height, which we will call x, we take the probability that it will occur with shoes on and divide it by the probability that it will occur with shoes off. We could, had we wished, have calculated the odds favouring hypothesis n rather than those favouring s, as has been done in Table 1.1. To do this we would have divided column 2 entries by column 3 entries and the values in column 4 would then have been the reciprocals of those which appear in the table.

Looking at the entries in column 4 you will see that as the value of x increases the odds that hypothesis *s* is correct become more favourable. For heights of 67 in. and above it is more likely that hypothesis s is correct. Below x=67 in. hypothesis n is more likely to be correct. If you look at Figure 1.1 you will see that from 67 in. up, the histogram for 'Shoes on' lies above the histogram for 'Shoes off. Below 67 in. the 'Shoes off histogram is higher.

SOME DEFINITIONS

With the above example in mind we will now introduce some of the terms and symbols used in signal detection theory.

The evidence variable

In the example there were two relevant things that could happen. These were state s (the subject had his shoes on) and state n (the subject had his shoes off). To decide which of these had occurred, the observer was given some evidence in the form of the height, x, of the subject. The task of the observer was to decide whether the evidence favoured hypothesis s or hypothesis n.

As you can see we denote evidence by the symbol x.¹ Thus x is called the *evidence* variable. In the example the values of x ranged from x=63 in. to x=70 in. In a psychological experiment x can be identified with the sensory effect produced by a stimulus which may be, for example, a range of illumination levels, sound intensities, or verbal material of different kinds.

Conditional probabilities

In the example, given a particular value of the evidence variable, say x=66 in., Table 1.1 can be used to calculate two probabilities:

(a) P(x|s): that is, the probability that the evidence variable will take the value x given that state s has occurred. In terms of the example, P(x|s) is the probability that a subject is 66 in. tall given that he is wearing shoes. From Table 1.1 it can be seen that for x=66 in., $P(x|s) = \frac{3}{16}$

(b) P(x|n): the probability that the evidence variable will take the value x given that state *n* has occurred. Table 1.1 shows that for x=66 in., $P(x|n) = \frac{4}{10}$

P(x|s) and P(x|n) are called *conditional probabilities* because they represent the probability of one event occurring conditional on another event having occurred. In this case we have been looking at the probability of a person being 66 in. tall given that he is (or conditional on him) wearing shoes.

The likelihood ratio

It was suggested that one way of deciding whether state *s* or state *n* had occurred was to first calculate the odds favouring *s*. In signal detection theory, instead of speaking of 'odds' we use the term *likelihood ratio*. 'Odds' and likelihood ratio' are synonymous The likelihood ratio is represented symbolically as l(x).

From the foregoing discussion it can be seen that in this example the likelihood ratio is obtained from the formula¹

¹ Another symbol used by Green & Swets (1966) for evidence is *e*.

¹ More correctly we should write $l_{sn}(x_i) = P(x_i|s)/P(x_i|n)$, with the subscripts *i*, *s* and *n*, added to (1.1). The subscript *i* denotes the likelihood ratio for the ith value of *x* but normally we will just write x with the subscripts implied. The order of the subscripts s and *n* tell us which of $P(x_i|s)$ and $P(x_i|n)$ is to act as the denominator and numerator in the expression for the likelihood ratio. The likelihood ratio $l_{sn}(x_i)$ is the ratio of $P(x_i|s)$ to $P(x_i|n)$ where $P(x_i|s)$ serves as the numerator. On the other hand the likelihood ratio $l_{ns}(x_i)$ is the ratio of $P(x_i|n)$ to $P(x_i|n)$ where $P(x_i|s)$ where $P(x_i|n)$ serves as the numerator. As all likelihood ratios in this book will use probabilities involving *s* as the numerator and probabilities involving *n* as the denominator the *s* and *n* subscripts will be omitted.

$$lx = \frac{P(x \mid s)}{P(x \mid n)}.$$
(1.1)

Thus from Table 1.1 we can see that

$$l(x = 64) = \frac{1/16}{2/16}, \quad l(x = 66) = \frac{3/16}{4/16}, \text{ etc.}$$

Hits, misses, false alarms and correct rejections

We now come to four conditional probabilities which will be often referred to in the following chapters. They will be defined by referring to Table 1.1.

First, however, let us adopt a convenient convention for denoting the observer's decision.

The two possible stimulus events have been called s and n. Corresponding to them are two possible responses that an observer might make; observer says 's occurred' and observer says 'n occurred'. As we use the lower case letters s and n to refer to stimulus events, we will use the upper case letters S and N to designate the corresponding response events. There are thus four combinations of stimulus and response events. These along with their accompanying conditional probabilities are shown in Table 1.2.

TABLE 1.2 The conditional probabilities, and their names. which correspond to the four possible combinations of stimulus and response events. The data in the table are the probabilities for the decision rule: 'Respond S if x > 66 in.; respond N if $x \le 66$ in.'

		Respon	Row sum	
		S	Ν	
	s	'Hit' P(S s) = (4+3+2+1)/16	'Miss' P(N s) = (3 + 2 + 1)/16	1.0
event		'Faise alarm'	'Correct rejection'	
	n	$P(S \mid n) = (3+2+1)/16$	$P(N \mid n) = (4+3+2+1)/16$	1.0

The meanings of the conditional probabilities are best explained by referring to an example from Table 1.1. An observer decides to respond *S* when x>66 in. and *N* when x<66 in. The probability that he will say *S* given that *s* occurred can be calculated from column 3 of the table by summing all the P(x|s) values which fall in or above the category x=66 in., namely, (4+3+2+1)/16 = 10/16. This is the value of P(S|s), the hit rate or hit probability. Also from column 3 we see that P(N|s), the probability of responding *N* when *s* occurred is (3+2+1)/16 = 6/16. From column 2 P(N|n), the probability of responding *N* when *n* occurred, is 10/16, and P(S|n), the false alarm rate, is 6/16. These hits, misses, false alarms and correct rejections are shown in Table 1.2.

DECISION RULES AND THE CRITERION

The meaning of β

In discussing the example it has been implied that the observer should respond N if the value of the evidence variable is less than or equal to 66 in. If the height is greater than or equal to 67 in. he should respond S. This is the observer's *decision rule* and we can state it in terms of likelihood ratios in the following manner:

'If l(x) < 1, respond N; if $l(x) \ge 1$, respond S.'

Check Table 1.1 to convince yourself that stating the decision rule in terms of likelihood ratios is equivalent to stating it in terms of the values of the evidence variable above and below which the observer will respond S or N.

Another way of stating the decision rule is to say that the observer has set his *criterion* at β =1. In essence this means that the observer chooses a particular value of l(x) as his criterion. Any value falling below this criterion value of l(x) is called *N*, while any value of l(x) equal to or greater than the criterion value is called *S*. This criterion value of the likelihood ratio is designated by the symbol β .

Two questions can now be asked. First, what does setting the criterion at $\beta=1$ achieve for the observer? Second, are there other decision rules that the observer might have used?

Maximizing the number of correct responses

If, in the example, the observer chooses the decision rule: 'Set the criterion at $\beta=1$ in., he will make the maximum number of correct responses for those distributions of *s* and *n*. This can be checked from Table 1.1 as follows:

If he says N when l(x) < 1 he will be correct 10 times out of 16, and incorrect 6 times out of 16. If he says S when $l(x) \ge 1$ he will be correct 10 times out of 16, and incorrect 6 times out of 16. Overall, his chances of making a correct response will be 20/32 and his chances of making an incorrect response will be 12/32.

Can the observer do better than this? Convince yourself that he cannot by selecting other decision rules and using Table 1.1 to calculate the proportion of correct responses. For example, if the observer adopts the rule: 'Say N if $l(x) < \frac{1}{4}$ and say S if $l(x) \ge \frac{1}{4}$,' his chances of making a correct decision will be 19/32, less than those he would have had with $\beta=1$.

It is a mistake, however, to think that setting the criterion at $\beta=1$ will always maximize the number of correct decisions. This will only occur in the special case where an event of type *s* has the same probability of occurrence as an event of type n, or, to put it in symbolic form, when P(s)=P(n). In our example, and in many psychological experiments, this is the case.

When s and n have different probabilities of occurrence the value of β which will maximize correct decisions can be found from the formula

$$\beta = P(n)/P(s) \tag{1.2}$$

We can see how this rule works in practice by referring to the example in Table 1.1.

Assume that in the example $P(s) = \frac{1}{2}P(n)$. Therefore by formula (1.2) $\beta=2$ will be the criterion value of l(x) which will maximize correct responses. This criterion is twice as strict as the one which





Number of incorrect responses (out of 48) = $6 + (6 \times 2) = 18$

maximized correct responses for equal probabilities of *s* and *n*. First we can calculate the proportion of correct responses which would be obtained if the criterion were maintained

at β =1. This is done in Table 1.3. As *n* events are twice as likely as s events, we multiply entries in row *n* of the table by 2.

The same thing can be done for $\beta=2$. Table 1.1 shows that $\beta=2$ falls in the interval x=69 in. so the observer's decision rule will be: 'Respond S if $x \ge 69$ in., respond N if x < 69 in. Again, with the aid of Table 1.1, the proportion of correct and incorrect responses can be calculated. This is done in Table 1.4.

TABLE 1.4 The number of correct and incorrect responses for $\beta = 2$ when $P(s) = \frac{1}{2} P(n)$.



It can be seen that $\beta=2$ gives a higher proportion of correct responses than $\beta=1$ when $P(s) = \frac{1}{2}P(n)$. There is no other value of β which will give a better result than 33/48 correct responses for these distributions of *s* and *n*.

Other decision rules

One or two other decision rules which might be used by observers will now be pointed out. A reader who wishes to see these discussed in more detail should consult Green & Swets (1966) pp. 20–7. The main purpose here is to illustrate that there is no one correct value of l(x) that an observer should adopt as his criterion. The value of β he should select will depend on the goal he has in mind and this goal may vary from situation to situation. For instance the observer may have either of the following aims.

(a) *Maximizing gains and minimizing losses*. Rewards and penalties may be attached to certain types of response so that

V_S=value of making a hit,

 $C_{s}N=$ cost of making a miss,

 $C_n S$ =cost of making a false alarm,

 $V_n N$ =value of making a correct rejection.

In the case where P(s)=P(n) the value of β which will maximize the observer's gains and minimize his losses is

$$\beta = \frac{V_{nN} + C_{nS}}{V_{sS} + C_{sN}}.$$
(1.3)

It is possible for a situation to occur where P(s) and P(n) are not equal and where different costs and rewards are attached to the four combinations of stimuli and responses. In such a case the value of the criterion which will give the greatest net gain can be calculated combining (1.2) with (1.3) so that

$$\beta = \frac{(V_{nN} + C_{nS}) \cdot P(n)}{(V_{sS} + C_{sN}) \cdot P(s)}.$$
(1.4)

It can be seen from (1.4) that if the costs of errors equal the values of correct responses, the formula reduces to (1.2). On the other hand, if the probability of *s* equals the probability of *n*, the formula reduces to (1.3).

(b) Keeping false alarms at a minimum: Under some circumstances an observer may wish to avoid making mistakes of a particular kind. One such circumstance with which you will already be familiar occurs in the conducting of statistical tests. The statistician has two hypotheses to consider; H_0 the null hypothesis, and H_1 , the experimental hypothesis. His job is to decide which of these two to accept. The situation is quite like that of deciding between hypotheses *n* and *s* in the example we have been discussing.

In making his decision the statistician risks making one of two errors:

Type I error: accepting H_1 when H_0 was true, and

Type II error: accepting H_0 when H_1 was true.

The Type I errors are analogous to false alarms and the Type II errors are analogous to misses. The normal procedure in hypothesis testing is to keep the proportion of Type I errors below some acceptable maximum. Thus we set up confidence limits of, say, p=0.05, or, in other words, we set a criterion so that P(S|n) does not exceed 5%. As you should now realize, by making the criterion stricter, not only will false alarms become less likely but hits will also be decreased. In the language of hypothesis testing, Type I errors can be avoided only at the expense of increasing the likelihood of Type II errors.

SIGNAL DETECTION THEORY AND PSYCHOLOGY

The relevance of signal detection theory to psychology lies in the fact that it is a theory about the ways in which choices are made. A good deal of psychology, perhaps most of it, is concerned with the problems of choice. A learning experiment may require a rat to choose one of two arms of a maze or a human subject may have to select, from several nonsense-syllables, one which he has previously learned. Subjects are asked to choose, from a range of stimuli, the one which appears to be the largest, brightest or most pleasant. In attitude measurement people are asked to choose, from a number of statements, those with which they agree or disagree. References such as Egan & Clarke (1966), Green & Swets (1966) and Swets (1964) give many applications of signal detection theory to choice behaviour in a number of these areas.

Another interesting feature of signal detection theory, from a psychological point of view, is that it is concerned with decisions based on evidence which does not unequivocally support one out of a number of hypotheses. More often than not, real-life

decisions have to be made on the weight of the evidence and with some uncertainty, rather than on information which clearly supports one line of action to the exclusion of all others. And, as will be seen, the sensory evidence on which perceptual decisions are made can be equivocal too. Consequently some psychologists have found signal detection theory to be a useful conceptual model when trying to understand psychological processes. For example, John (1967) has proposed a theory of simple reaction times based on signal detection theory; Welford (1968) suggests the extension of detection theory to absolute judgement tasks where a subject is required to judge the magnitude of stimuli lying on a single dimension; Boneau. & Cole (1967) have developed a model for decision-making in lower organisms and applied it to colour discrimination in pigeons; Suboski (1967) has applied detection theory in a model of classical discrimination conditioning.

The most immediate practical benefit of the theory, however, is that it provides a number of useful measures of performance in decision-making situations. It is with these that this book is concerned. Essentially the measures allow us to separate two aspects of an observer's decision. The first of these is called *sensitivity*, that is, how well the observer is able to make correct judgements and avoid incorrect ones. The second of these is called *bias*, that is, the extent to which the observer favours one hypothesis over another independent of the evidence he has been given. In the past these two aspects of performance have often been confounded and this has lead to mistakes in interpreting behaviour.

Signal and noise

In an auditory detection task such as that described by Egan, Schulman & Greenberg (1959) an observer may be asked to identify the presence or absence of a weak pure tone embedded in a burst of *white noise*. (Noise, a hissing sound, consists of a wide band of frequencies of vibration whose intensities fluctuate randomly from moment to moment. An everyday example of noise is the static heard on a bad telephone line, which makes speech so difficult to understand.) On some trials in the experiment the observer is presented with noise alone. On other trials he hears a mixture of tone+noise. We can use the already familiar symbols *s* and *n* to refer to these two stimulus events. The symbol *n* thus designates the event 'noise alone' and the symbol *s* designates the event 'signal (in this case the tone)+noise'.

The selection of the appropriate response, S or N, by the observer raises the same problem of deciding whether a subject's height had been measured with shoes on or off. As the noise background is continually fluctuating, some noise events are likely to be mistaken for signal+ noise events, and some signal+noise events will appear to be like noise alone. On any given trial the observer's best decision will again have to be a statistical one based on what he considers are the odds that the sensory evidence favours s or n.

Visual detection tasks of a similar kind can also be conceived. The task of detecting the presence or absence of a weak flash of light against a background whose level of illumination fluctuates randomly is one which would require observers to make decisions on the basis of imperfect evidence.

Nor is it necessary to think of noise only in the restricted sense of being a genuinely

random component to which a signal may or may not be added. From a psychological point of view, noise might be any stimulus not designated as a signal, but which may be confused with it. For example, we may be interested in studying an observer's ability to recognize letters of the alphabet which have been presented briefly in a visual display. The observer may have been told that the signals he is to detect are occurrences of the letter 'X' but that sometimes the letters 'K', 'Y' and 'N' will appear instead. These three non-signal letters are not noise in the strictly statistical sense in which white noise is defined, but they are capable of being confused with the signal letter, and, psychologically speaking, can be considered as noise.

Another example of this extended definition of noise may occur in the context of a memory experiment. A subject may be presented with the digit sequence '58932' and at some later time he is asked: 'Did a "9" occur in the sequence?', or, alternatively: 'Did a "4" occur in the sequence?' In this experiment five digits out of a possible ten were presented to be remembered and there were five digits not presented. Thus we can think of the numbers 2, 3, 5, 8, and 9, as being signals and the numbers 1, 4, 6, 7, and 0, as being noise. (See Murdock (1968) for an example of this type of experiment.)

These two illustrations are examples of a phenomenon which, unfortunately, is very familiar to us—the fallibility of human perception and memory. Sometimes we 'see' the wrong thing or, in the extreme case of hallucinations, 'see' things that are not present at all. False alarms are not an unusual perceptual occurrence. We 'hear' our name spoken when in fact it was not; a telephone can appear to ring if we are expecting an important call; mothers are prone to 'hear' their babies crying when they are peacefully asleep.

Perceptual errors may occur because of the poor quality or ambiguity of the stimulus presented to an observer. The letter 'M' may be badly written so that it closely resembles an 'N'. The word 'bat', spoken over a bad telephone line, may be masked to such an extent by static that it is indistinguishable from the word 'pat'. But this is not the entire explanation of the perceptual mistakes we commit, Not only can the stimulus be noisy but noise can occur within the perceptual system itself. It is known that neurons in the central nervous system can fire spontaneously without external stimulation. The twinkling spots of light seen when sitting in a dark room are the result of spontaneously firing retinal cells and, in general, the continuous activity of the brain provides a noisy background from which the genuine effects of external signals must be discriminated (Pinneo, 1966). FitzHugh (1957) has measured noise in the ganglion cells of cats, and also the effects of a signal which was a brief flash of light of near-threshold intensity. The effects of this internal noise can be seen even more clearly in older people where degeneration of nerve cells has resulted in a relatively higher level of random neural activity which results in a corresponding impairment of some perceptual functions (Welford, 1958). Another example of internal noise of a rather different kind may be found in schizophrenic patients whose cognitive processes mask and distort information from the outside world causing failures of perception or even hallucinations.

The concept of internal noise carries with it the implication that all our choices are based on evidence which is to some extent unreliable (or noisy). Decisions in the face of uncertainty are therefore the rule rather the exception in human choice behaviour. An experimenter must expect his subjects to 'perceive' and 'remember' stimuli which did not occur (for the most extreme example of this see Goldiamond & Hawkins, 1958). So, false alarms are endemic to a noisy perceptual system, a point not appreciated by earlier psychophysicists who, in their attempts to measure thresholds, discouraged their subjects from such 'false perceptions'. Similarly, in the study of verbal behaviour, the employment of so-called 'corrections for chance guessing' was an attempt to remove the effects of false alarms from a subject's performance score as if responses of this type were somehow improper.

The fact is, if noise does play a role in human decision-making, false alarms are to be expected and should reveal as much about the decision process as do correct detections. The following chapters of this book will show that it is impossible to obtain good measures of sensitivity and bias without obtaining estimates of both the hit and false alarm rates of an observer.

A second consequence of accepting the importance of internal noise is that signal detection theory becomes something more than just another technique for the special problems of psychophysicists. All areas of psychology are concerned with the ways in which the internal states of an individual affect his interpretation of information from the world around him. Motivational states, past learning experiences, attitudes and pathological conditions may determine the efficiency with which a person processes information and may also predispose him towards one type of response rather than another. Thus the need for measures of sensitivity and response bias applies over a wide range of psychological problems.

Egan (1958) was first to extend the use of detection theory beyond questions mainly of interest to psychophysicists by applying it to the study of recognition memory. Subsequently it has been employed in the study of human vigilance (Broadbent & Gregory, 1963a, 1965; Mackworth & Taylor, 1963), attention (Broadbent & Gregory, 1963b; Moray & O'Brien, 1967) and short-term memory (Banks, 1970; Murdock, 1965; Lockhart & Murdock, 1970; Norman & Wickelgren, 1965; Wickelgren & Norman, 1966). The effects of familiarity on perception and memory have been investigated by detection theory methods by Allen & Garton (1968,1969) Broadbent (1967) and Ingleby (1968). Price (1966) discusses the application of detection theory to personality, and Broadbent & Gregory (1967), Dandeliker & Dorfman (1969), Dorfman (1967) and Hardy & Legge (1968) have studied sensitivity and bias changes in perceptual defence experiments.

Nor has detection theory been restricted to the analysis of data from human observers. Suboski's (1967) analysis of discrimination conditioning in pigeons has already been mentioned, and Nevin (1965) and Rilling & McDiarmid (1965) have also studied discrimination in pigeons by detection theory methods. Rats have received similar attention from Hack (1963) and Nevin (1964).

Problems

The following experiment and its data are to be used for problems 1 to 6.

In a card-sorting task a subject is given a pack of 450 cards, each of which has had from 1 to 5 spots painted on it. The distribution of cards with different numbers of spots is as follows:

Number of spots on card	Number of cards in pack
1	50
2	100

12 A primer of signal detection theory

3	150
4	100
5	50

Before giving the pack to the subject the experimenter paints an extra spot on 225 cards as follows:

Original number of spots on card Number of cards in this group receiving an extra spot

1	25
2	50
3	75
4	50
5	25

The subject is then asked to sort the cards in the pack into two piles; one pile containing cards to which an extra spot has been added and the other pile, of cards without the extra spot.

- 1. What is the maximum proportion of cards which can be sorted correctly into their appropriate piles?
- 2. State, in terms of x, the evidence variable, the decision rule which will achieve this aim.
- 3. If the subject stands to gain 1ϕ for correctly identifying each card with an extra spot and to lose 2ϕ for incorrectly classifying a card as containing an extra spot, find firstly in terms of β , and secondly in terms of x, the decision rule which will maximize his gains and minimize his losses.
- 4. What proportions of hits and false alarms will the observer achieve if he adopts the decision rule $\beta = \frac{3}{2}$?
- 5. What will P(N|s) and β be if the subject decides not to allow the false alarm probability to exceed $\frac{1}{3}$?
- 6. If the experimenter changes the pack so that there are two cards in each group with an extra spot to every one without, state the decision rule both in terms of x and in terms of β which will maximize the proportion of correct responses.
- 7. Find the likelihood ratio for each value of *x* for the following data:

x	1	2	3	4
P(x n)	0.2	0.4	0.5	0.6
P(x s)	0.5	0.7	0.8	0.9

- 8. At a particular value of x, l(x)=0.5 and the probability of x given that n has occurred is 0.3. What is the probability of x given that s has occurred?
- 9. If P(S|s)=0.7 and P(N|n)=0.4, what is P(N|s) and P(S|n)?
- 10. The following table shows P(x|n) and P(x|s) for a range of values of x.

x	63	64	65	66	67	68	69	70	71
P(x n)	1/16	2/16	3/16	4/16	3/16	2/16	1/16	0	0
P(x s)	1/16	1/16	2/16	2/16	4/16	2/16	2/16	1/16	1/16

Draw histograms for the distributions of signal and noise and compare your diagram with Figure 1.1. What differences can you see?

Find l(x) for each x value in the table. Plot l(x) against x for your data and compare it with a plot of l(x) against x for the data in Table 1.1. How do the two plots differ?

If P(s) were equal to 0.6 and P(n) to 0.4 state, in terms of x, the decision rule which would maximize correct responses:

(a) for the problem data,

(b) for the data in Table 1.1.

(The issues raised in this problem will be discussed in Chapter 4.)

Chapter 2

NON-PARAMETRIC MEASURES OF SENSITIVITY

In Chapter 1 it was said that signal detection theory could be used both as the basis for a general theory of behaviour and for devising methods for measuring performance in experimental tasks. For the time being we will concern ourselves with this latter use and proceed to describing the basic types of detection task. There are three main types of experimental situation in which detection data can be collected; the yes-no task, the rating scale task and the forced-choice task.

THE YES-NO TASK

(Green & Swets, 1966, 32, ff.; Swets, Tanner & Birdsall, 1961)

The following is an example of a yes-no task. An observer watches a television screen on which, at regular intervals, some information appears. This information takes one of two forms. On half the occasions noise alone is shown. On other occasions noise plus a weak signal (a circular patch of light in the centre of the screen) is shown. Noise, and signal+noise trials occur at random. After each presentation the observer must say whether it was a signal+noise trial or just noise alone.

After a number of trials it is possible to construct a stimulus-response matrix summarising the results of the experiment. The form of the matrix is shown in Table 2.1. In it appear the four conditional probabilities, calculated from the subject's raw data, which were introduced in Table 1.2.

From the stimulus-response matrix we ought to be able to get some measure of how well the observer discriminates signal events from those consisting of noise alone. At first sight it might be thought that P(S|s), the hit rate, is a good index of the observer's sensitivity to signals. However it seems reasonable that some account should also be taken of P(S|n), the false alarm rate. An observer who never looked at the display would be able to give a perfect hit rate by responding *S* to all trials. At the same time he would produce a false alarm rate equal to the hit rate and we would be unwilling to believe that such a perverse observer was showing any real sensitivity to signals.