ENGINEERING RESEARCH INSTITUTE
UNIVERSITY OF MICHIGAN
ANN ARBOR

THE EVIDENCE FOR A DECISION-MAKING THEORY OF VISUAL DETECTION

Technical Report No. 40
Electronic Defense Group
Department of Electrical Engineering

By: J. A. Swets
W. P. Tanner, Jr.
T. G. Birdsall

Approved by: A. B. Macnee

Project 2262

TASK ORDER NO. EDG-3
CONTRACT NO. DA-36-039 sc-63203
SIGNAL CORPS, DEPARTMENT OF THE ARMY
DEPARTMENT OF ARMY PROJECT NO. 3-99-04-042
SIGNAL CORPS PROJECT NO. 194B

February, 1955

The experiments reported in this paper were sponsored jointly by the Electronic Defense Group under Contract No. DA-36-039 sc63203 and the Vision Research Laboratory under Contract No. DA-36-039 sc52654 and Contract No. Nobs 53242. Therefore, this report appears also as a portion of Report 2144-20-T from the Engineering Research Institute, a report from the Optics and Vision Group of Project Michigan. Report 2144-20-T differs from the present report principally in that the relevance of the research to a classified military problem is indicated.

ACKNOWLEDGEMENT

The authors wish to acknowledge the valuable contributions of some of our friends and associates without whose assistance this work would have been impossible. Dr. H. R. Blackwell, director of the Vision Research Laboratory, contributed freely of his time in discussions during every stage of the work. In permitting the use of his laboratory, he made possible the successful completion of the experiments. Dr. W. M. Kincaid, of the staff of the Vision Research Laboratory, exhibited a lively interest throughout the progress of this work. He was always available for consultation, and contributed many valuable suggestions. The authors take this opportunity to express their gratitude to Dr. Blackwell and Dr. Kincaid for their assistance.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

ABSTRACT

This is one of a series of papers concerned with the psychophysical application of the mathematical theory of signal detectability. This paper brings together all of the data on visual detection collected to date that bear directly on the case of the signal-known-exactly as treated by the theory of signal detectability. The general conclusion drawn from the experimental results reported is that the model provided by the theory of signal detectability or, more generally, by the theory of statistical decision, is applicable to the visual detection behavior of the human observer. That is to say, the human observer is capable of ordering values of the variable upon which detection depends well below the threshold level as this level is conventionally conceived. Further, the experiments show the observer to be capable of behaving in accordance with several kinds of optimum decision as defined within the theory of signal detectability. The general implications of the applicability of the model for sensory theory and psychophysical methods are discussed.

THE EVIDENCE FOR A DECISION-MAKING THEORY OF VISUAL DETECTION

1. INTRODUCTION

This is one of a series of papers concerned with the psychophysical ɔplication of the mathematical theory of signal detectability. This introductory ɜction contains a short discussion of the references in which the theory of sig-ɜl detectability is presented, and also contains references to other papers in he series on the psychophysical application of the theory of signal detectability. here follows, in this introductory section, a brief description of the theory of ignal detectability and of the theory of visual detection based on it, and an utline of the scope of discussion of the present paper.

.1 Related Articles

"The Theory of Signal Detectability," by Peterson and Birdsall, appears s Technical Report No. 13 of the Electronic Defense Group of the University of ichigan (Ref. 15). Technical Report No. 19, by Fox (Ref. 8), contains a non-athematical discussion of the material of Technical Report No. 13 on the general heory of detectability, as well as a discussion of the methods employed in se-uential observation processes. Birdsall's Technical Report No. 20 (Ref. 2) pre-ents an application of game theory to signal detectability. The material of eports 13 and 19 is also available in the "Transactions of the I.R.E. Profes-ional Group on Information Theory", (Ref. 16).

The application of the theory of signal detectability to visual detec-ion by the human observer was first reported by Tanner and Swets in Technical eport No. 18 of the Electronic Defense Group (Ref. 22). This topic is also dis-ussed elsewhere (Refs. 19, 23, 24). The most readily available source containing a report of some of this work is the Psychological Review of November, 1954 (Re: 25). A paper in the "Transactions of the I.R.E. Professional Group on Informat: Theory" (Ref. 21) contains a report of the work of Tanner and Norman relating tl theory of signal detectability to auditory detection.

Several other papers in this series are currently being prepared. Te(nical Report No. 30 by Tanner, Swets, and Green (Ref. 26) is another paper relat ing the theory to data from auditory experiments. Technical Report No. 42 by Tanner, Birdsall, and Swets (Ref. 20) deals with the implications of this research for methods and analysis procedures in psychophysical testing.

The present paper brings together the data on visual detection collect to date that bear directly on the case of the signal-known-exactly as treated by the theory of signal detectability. Some of these data have been reported previously only in relatively inaccessible sources (Refs. 19, 22, 23, and 24); some of the data that were reported in a more available source (Ref. 25) were only briefly described there. In addition, some of the data presented below have not been previously reported in any form. It is the primary purpose of this paper t present all relevant available data; although the theory of signal detectability and the theory of visual detection based on it are reviewed, they are not discussed completely, and whereas the general implications for sensory theory, methods, and data analysis are drawn, they are not treated here in full detail.

1.2 A Decision-Making Theory of Visual Detection

In this section, the theory of signal detectability is reviewed as it relates to the problem of visual detection by the human observer. It should be pointed out, by way of introduction, that the theory of signal detectability is derived directly from the theory of statistical decision, the theory of testing statistical hypotheses. This latter theory is presented most completely by Wald ef. 28); important papers published earlier include those by Wald (Ref. 27), yman (Ref. 13), and Neyman and Pearson (Ref. 14).

1.2.1 The Fundamental Problem. The fundamental problem in signal detecbility involves a fixed observation interval; the observer is presented a eceiver input", a function of time for T seconds. He is then asked to decide, the basis of the observation, whether that receiver input arose from noise one, or from signal plus noise, where the signal is known to be from a certain semble of signals.

This problem has an exact parallel in visual psychophysics. The obser-r in the most common visual psychophysical experiment, an experiment employing e so-called "yes-no" method, is asked to observe, at a particular time and a rticular location in the visual field, a signal of a particular size, duration, d intensity. Ordinarily the size and duration of the signal, as well as its cation and time of occurrence, are known by the observer; in certain instances, formation about the intensity of the signal and the *a priori* probability of gnal occurrence are also provided the observer. He is then asked to state, on e basis of the observation, whether or not a signal was presented.

1.2.2 Assumptions Made in Applying the Theory of Signal Detectability the Behavior of the Human Observer. Given certain assumptions, asking the ob-rver in a psychophysical experiment to state whether or not a signal was pre-nted is equivalent to asking him to decide whether his observation arose from gnal plus noise or from noise alone, or stated another way, to decide whether or t to accept the hypothesis that a signal existed. The primary assumption is that e value of the variable upon which detection depends, presumably neural activity, ries from instant to instant in a random fashion when no signal is present, and at the value of this variable produced by a signal of given strength is also randomly distributed. Although there is no need to conceptualize this variable in neurophysiological terms, it may be helpful, and the assumption stated is su gested by present knowledge of neurophysiology. So it may be maintained on *a priori* grounds, that if the problem of visual detection is the detection of signals having randomly distributed neural effects, in the presence of a background of random interference, then the theory of statistical decision, or in p ticular, the theory of signal detectability, constitutes a model of possible re vance to visual detection.

An attempt to apply the theory of signal detectability to human behavior, therefore, implies the assumptions that the sensory systems function pr marily to transmit information, and that the sensory systems are noisy channels An additional assumption, made apparent in the following pages, is that the cen tral mechanisms involved in decision making are capable of making optimal use o the information transmitted by sensory paths.

It is assumed, then, that the observer operates with a continuous variable, the values of which constitute "observations," and that any value of this variable may arise either from noise alone or from signal plus noise. It assumed that, when the signal ensemble is known to the observer, the probabilit that a given observation represents noise alone, and the probability that this value arose from the signal-plus-noise distribution, can be estimated by him.

Thus, it is assumed that the observer in a "yes-no" experiment must establish a level of confidence, or criterion, and base his decision on the rel tion of the observation to this criterion.

1.2.3 The Definition of Criterion and Likelihood Ratio. According t the theory, the observer chooses a set of observations (the criterion A) such t an observation in this set will lead him to Accept the existence of a signal, t

, the observer accepts the hypothesis that signal plus noise existed during the servation interval. All other observations are in the Complement of the cri- rion, CA; these are regarded by the observer as representing noise alone. SN ll be used to denote signal plus noise and N will denote noise alone. If there e only a countable number of possible observations, each observation, x, having obability $P_{SN}(x)$ of occurrence if there is signal plus noise present, and prob- ility $P_N(x)$ of occurrence if noise alone is present, then the likelihood ratio defined as $\ell(x) = P_{SN}(x)/P_N(x)$. Here, x will be considered continuous, and obability density functions (frequency functions) $f_{SN}(x)$ and $f_N(x)$ are used; cordingly $\ell(x) = f_{SN}(x)/f_N(x)$. It is assumed that for every x the observer can timate $\ell(x)$ which is the relative likelihood that x arose from signal plus noise compared to the possibility that x arose from noise alone.

A criterion may be evaluated in terms of the integrals of the density nctions over the criterion A, since the integral of $f_{SN}(x)$ over A is the con- tional probability of detection, $P_{SN}(A)$, and the integral of $f_N(x)$ over A is the nditional probability of a false alarm (a Type I error in statistical parlance), $_N(A)$.

1.2.4 The Essence of the Theory of Signal Detectability. The essence the theory of signal detectability is the definition of a class of criteria in rms of likelihood ratio. Under each of several definitions of the optimum cri- rion, the optimum is found to be in this class of likelihood-ratio criteria. A iterion in this class is denoted $A(\beta)$; that is, the criterion A contains all servations with likelihood ratio greater than or equal to β, and none of those th likelihood ratio less than β. The solution, then, with respect to a given finition of optimum, is the exact value of β to be used. β is defined as the erating level of the likelihood-ratio receiver.

It should be noted that the theory of signal detectability specifies as the optimum receiver that receiver whose output is either likelihood ratio or some monotonic function of likelihood ratio. If the output of the receiver is likelihood ratio, then the solution, for each definition of optimum, is the criterion with the proper operating level β. If the receiver's output is some decision function other than likelihood ratio, but a monotonic function of likelihood ratio, then the optimum operating level is the value of the monotonic function at β. That is, if the receiver output is some function $d(x) = F\left[\ell(x)\right]$, where F is strictly monotonic, then the optimum criterion is specified by $\beta' = F(\beta)$, such that $d(x) \geq \beta' \leftrightarrow \ell(x) \geq \beta$. Thus, the theory of signal detectability may describe the behavior of the human observer if the human observer operates with a continuous variable, or decision function, that is either likelihood ratio or some monotonic function of likelihood ratio.

1.3 Scope of Discussion

Peterson, Birdsall, and Fox (Ref. 16) advance six definitions of optimum and their solutions. Experiments have been performed, and are reported in the next three sections of this paper, to test the ability of the human observer to act in accordance with three of these definitions of optimum. These three definitions of optimum and their respective solutions are listed here.

1. Expected-Value Criterion — that criterion that maximizes the total expected value, where the individual values are:

$V_{SN\cdot A}$ = the value of a detection

$V_{N\cdot CA}$ = the value of a correct rejection

$K_{SN\cdot CA}$ = the cost of a miss

$K_{N\cdot A}$ = the cost of a false alarm lution: $A(\beta) = \frac{P(N)}{P(SN)} \cdot \frac{V_{N \cdot CA} + K_{N \cdot A}}{V_{SN \cdot A} + K_{SN \cdot CA}}$

ere $P(N)$ and $P(SN)$ are the a priori probabilities.

2. The Neyman-Pearson Criterion — that criterion such that $P_{SN}(A)$ is maximum, while $P_N(A) \leq k$.

lution: $A(\beta)$ where $P_N\left[A(\beta)\right] = k$

3. A Posteriori Probability — not a criterion but the best estimate of e probability that the observation arose from signal plus noise.

$$P_x(SN) = \frac{\ell(x)\, P(SN)}{\ell(x)\, P(SN) + P(N)}$$

her definitions of optimum, not considered explicitly here, for which the soluon has been provided (Ref. 16) include the Weighted-Combination criterion, the iterion that maximizes $P_{SN}(A) - w\, P_N(A)$; Siegert's Ideal Criterion, the criterion at minimizes total error; and the Information Criterion, the criterion that ximizes the reduction in uncertainty, in the Shannon sense (Ref. 18), as to ether or not a signal was sent. It should be pointed out, however, that the ighted-Combination Criterion is the abstract criterion, of which the Expected-lue Criterion and Ideal Criterion are special cases. The Expected-Value Cririon is identical to the Weighted-Combination Criterion for the case where = β and is identical to the Ideal Criterion for the case where

$$\frac{V_{N \cdot CA} + K_{N \cdot A}}{V_{SN \cdot A} + K_{SN \cdot CA}} = 1 \quad .$$

The fifth section of this paper contains the results of an experiment designed to determine the congruence of the behavior of the human observer and the definition of optimum behavior for the forced-choice situation, a definitior of optimum that is not treated explicitly in the theory of signal detectability.

1.4 The Incompatibility of a Decision-Making Theory of Visual Detection and Conventional Sensory Theory

Anticipating to an extent that will facilitate subsequent description: the primary result of all the experiments performed is that the theory of signal detectability, or more generally, the theory of statistical decision, is applicable to the behavior of the human observer. That is to say, the experiments demonstrate that the human observer operates with a decision function that is either likelihood ratio or some monotonic function of likelihood ratio, and that the human observer tends to behave optimally.

This result implies, of course, the ability to discriminate among obse vations, or values of the decision function, that may result from noise alone. Thus, if a threshold (fixed operating level) exists, this threshold is low enoug to be exceeded by noise alone an appreciable portion of the time. The concept o a threshold that is exceeded more than very rarely by noise alone is quite diffe ent from the concept of threshold that is an integral part of conventional senso theory.

The primary purpose of this paper is the presentation of data; hence, this is not the place to attempt a detailed discussion of the notion of a thresh old as it appears in sensory theory. It may be well, however, to adduce one or two instances of current methods in sensory experimentation in order to support the statement that the threshold as conceived in sensory theory is a relatively fixed level that is rarely, if ever, exceeded by noise alone.

Many studies of sensory processes do not employ trials in which no sig- al is presented. In these cases, it is clearly not regarded as important to ssess the probability that the fixed level be exceeded by noise alone. In other tudies, the experimenter may occasionally insert one or two trials in which no ignal is presented (or occasionally turn off a continuous signal) in order to etect what are regarded as spurious responses, so that he may caution the obser- ers against such responses. Such trials have been referred to as vexirfehlen, term which may be reasonably translated as "catch signals". A refinement of the rocedure in which catch signals are sporadically presented, one that is used ather frequently, is to fail to present signals on a larger proportion of trials 1 order to assess quite accurately the extent of "yes" responses on such trials. nen, however, not only is this information subtracted from the data, but the xtent of "yes" response to catch signals is used in estimating the amount of spurious" responses made to actual signals so that these, too, may be eliminated rom the data. This use of the correction for chance in psychophysical experi- ents has been described in Ref. 25. Regarding the totality of "yes" responses ɔ catch signals as spurious, and regarding such spurious responses to be equally ikely for all values of signal intensity, is valid only if the threshold level s such that it is exceeded by noise alone on a negligible proportion of the rials.

The validity of the application of the chance correction to psychophysi- al data depends upon the validity of the assumption that "yes" responses to catch ignals are spurious responses, or random guesses, and that these responses are ndependent of "sensory-determinate" responses.[1] The theory advanced here, on the ther hand, assumes a dependence between the conditional probability that an

---

I. R. Blackwell, unpublished manuscript.

observation arising from signal plus noise will be in the criterion and the conditional probability that an observation arising from noise alone will be in the criterion.

The data presented in this paper, to the extent that they are analyzed in this paper, do not indicate how far down into the noise the observers actuall ordered their observations, that is, how low relative to the distribution of observations arising from noise alone a fixed threshold must be to be compatible with the data. It may be the case, then, that further analysis will show the present data or new data to be consistent with a fixed threshold, say, at the me of the noise distribution, a threshold exceeded by noise alone approximately 50 percent of the time. It can be stated, on the basis of completed analyses of th present data which are to be reported in a paper which considers and evaluates several alternative models (Ref. 20), that a fixed threshold, if one exists, mus be lower than plus one sigma from the mean of the noise distribution, that is, must be exceeded by noise alone more than 16 percent of the time. It is importa that the reader note that, unless specifically stated otherwise, further references in this paper to the threshold refer to a threshold in the conventional sense, that is, to a threshold exceeded, to be conservative, less than 5 percent of the time by noise alone. The experiments were designed to detect a threshold at approximately this level if one existed.

2. THE EXPECTED-VALUE OBSERVER

An experiment concerning the ability of the human observer to maximiz the total expected value is the one experiment discussed elsewhere in a readily accessible source (Ref. 25). This section supplements the discussion in Referen 25 in that it presents in more detail the data that were necessarily described only briefly there. In addition, this section contains the previously unpublis esults of further experimentation on the Expected-Value Observer that is superior n certain respects.

The analysis of data, in terms of the theory of signal detectability, akes the form of plots of what are called ROC curves — Receiver Operating naracteristics. An ROC curve is a plot of the conditional probability of detecion, $P_{SN}(A)$, against the conditional probability of false alarm, $P_N(A)$. A comlete curve is obtained if all possible values of operating level, or β, are conidered. Peterson and Birdsall (Ref. 15) have demonstrated that the optimum perating level is represented by a point on the ROC curve where its slope is β.

The typical ROC curve, one that has occurred frequently in this work, s shown in Fig. 1. Frequently, as in this figure, a family of ROC curves is lotted with signal strength as the parameter. In this particular case, the arameter is d', the index used for the analysis of data in terms of the theory f signal detectability.[1] It is defined as the difference between the means of he noise and signal-plus-noise distributions normalized to the standard deviation f the noise distribution that is, $d' = \frac{M_{SN} - M_N}{\sigma_N}$. Thus, d' can be conceived f as a standard score, an $\frac{x}{\sigma}$ measure, or again, it may be thought of as a output) signal-to-noise ratio. A more complete description of d' and of the ROC urves may be found elsewhere (Refs. 15, 16, 20).

If a threshold exists, this fact is immediately apparent from the ROC urves. The conventional notion of a fixed criterion or threshold, in the termiology of this paper, implies the existence of a set of observations leading

d' is essentially a dependent variable; the paper on psychophysical methods (Ref. 20) treats explicity the reasons for preferring it to the usual dependent variable in psychophysical experiments; namely, the calculated threshold or minimum detectable signal. The primary purpose of the present paper is to present the data upon which the recommendations made in the paper on psychophysical methods are based.

1970 A-G3-124 JEM 4-21-54

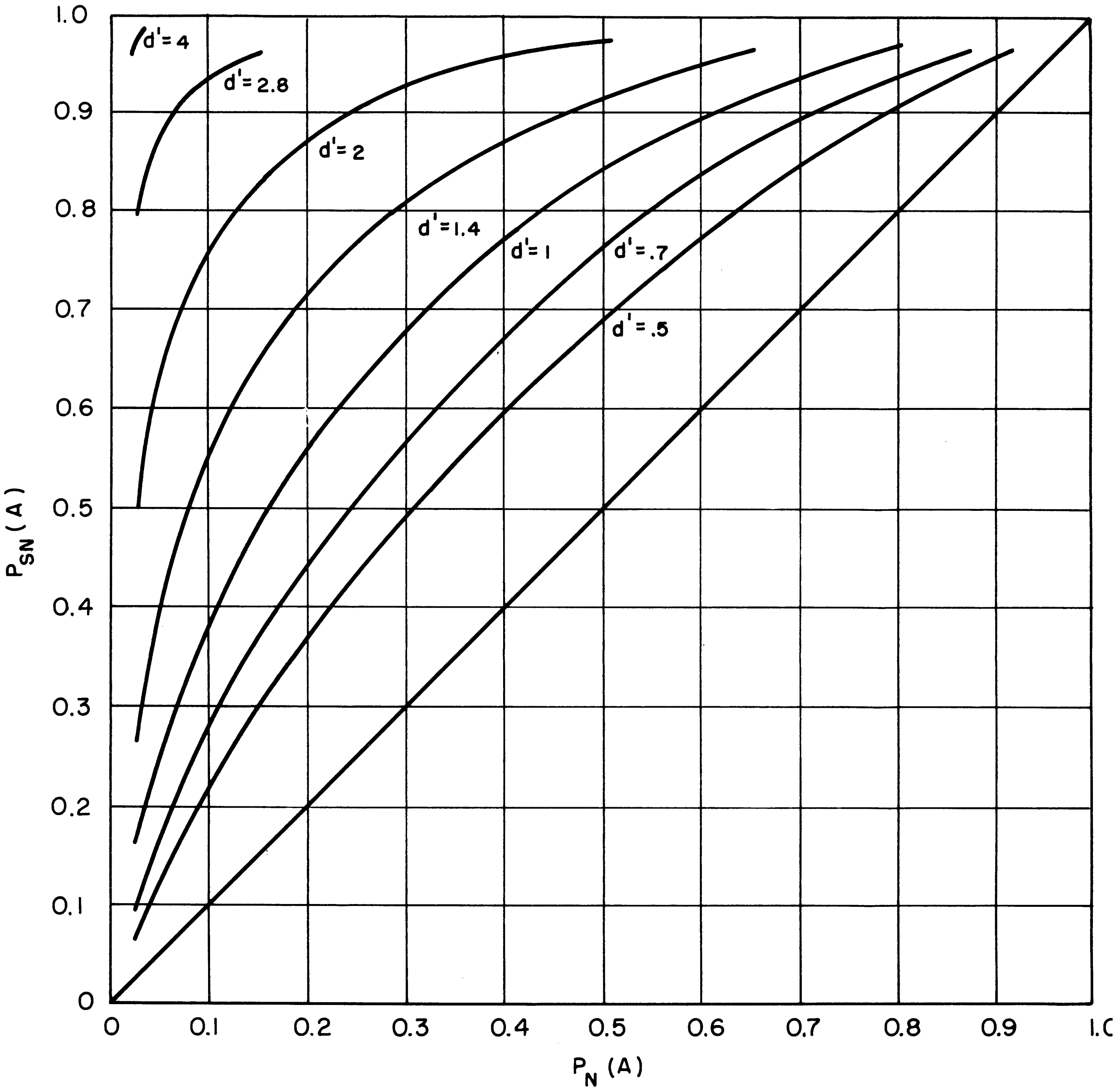

FIG. 1. P_{SN}(A) VS. P_N(A) WITH d' AS THE PARAMETER .

ɔ an extremely small probability of false alarm. The observer's criterion inludes this set and may or may not include a random selection of other observations ıat is, "guesses". The ROC curve in this case, as pointed out elsewhere (22, 25), s a straight line from the left-hand vertical axis to the upper right-hand corner.

.1 Data from the First Experiment

The ROC curves obtained from the first experiment on the Expected-Value bserver, the experiment discussed in reference 25, are displayed here in Figures , 3, and 4. Five values of signal intensity, including the zero-intensity sigal or "blank", were used in this experiment. It may be seen from an examination f these plots that the data of the first experiment do not provide an adequate efinition of the entire curve. As a matter of fact, even in the region where oints have been obtained (roughly from $P_N(A) = 0.0$ to $P_N(A) = .50$), the curve is ot well defined. This latter inadequacy may be attributed, in part, to the small umber of observations per point. A second factor which operated to cause disersion of these points was the day-to-day variation in signal and background ntensities which, unlike in the calculation of contrast thresholds, is not taken nto account in the present analysis. The effects of both of these sources of ariance were reduced, to a large extent, in the second Expected-Value experiment hich is reported below.

Although entire ROC curves were not precisely defined by the data from ıe first experiment, these data are entirely adequate for the purpose of disinguishing between the predictions of a theory based on the model provided by ıe theory of signal detectability and the predictions which follow from the conept of a threshold or fixed criterion, the purpose for which they were used in eference 25. It is clear, for example, that the straight lines fitted to the

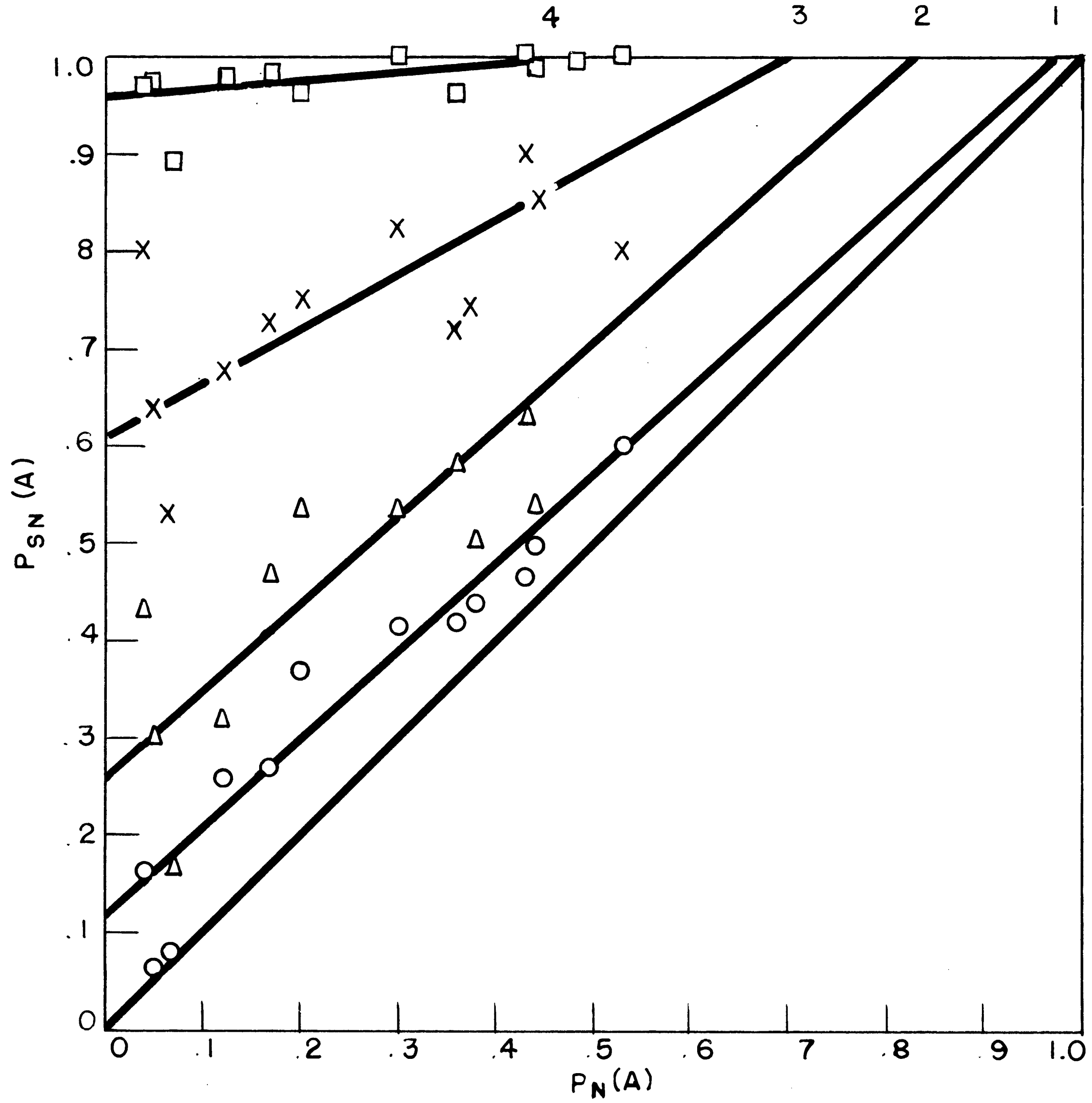

FIG. 2. $P_{SN}(A)$ VS. $P_N(A)$ FOR OBSERVER I IN THE FIRST EXPECTED-VALUE EXPERIMENT.

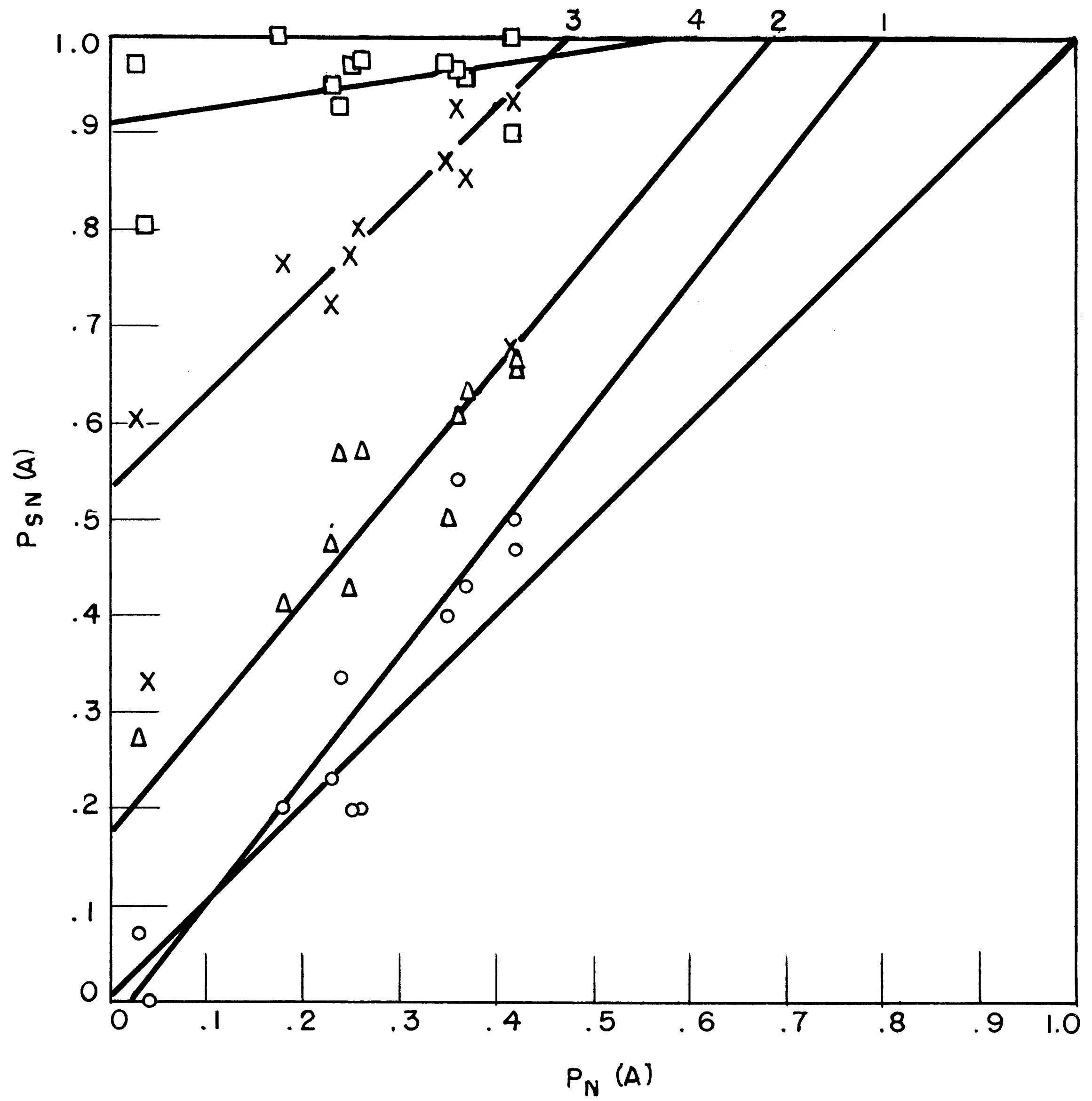

FIG. 3. $P_{SN}(A)$ VS. $P_N(A)$ FOR OBSERVER 2 IN THE FIRST EXPECTED-VALUE EXPERIMENT.

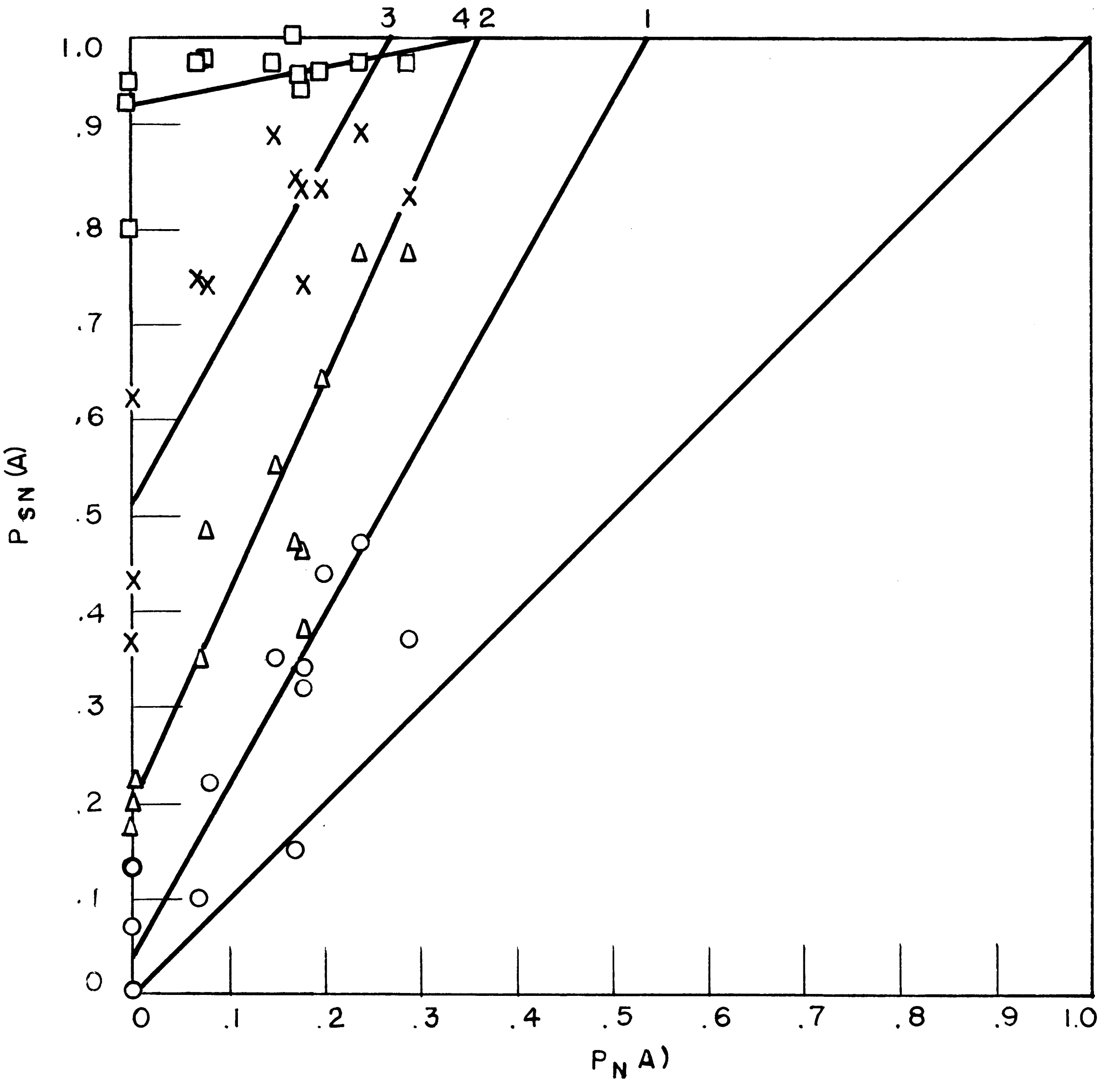

FIG. 4. $P_{SN}(A)$ VS. $P_N(A)$ FOR OBSERVER 3 IN THE FIRST EXPECTED-VALUE EXPERIMENT.

data in Figures 2, 3, and 4 do not intersect the upper right-hand corner of the graph, as required by the concept of a fixed threshold. These straight lines appear, rather, to be arcs of the type of ROC curve predicted by the theory of signal detectability. The data displayed graphically in Figures 2, 3, and 4 are reported more precisely in Tables 1, 2, and 3.

On days 1 through 8, there were 50 presentations per day of each value of signal intensity including the zero-intensity signal. On days 9 through 12, there were 60 presentations per day at each intensity. On days 13 through 16, there were 30 presentations per day of each value of signal intensity greater than zero, and 180 blanks per day. This experiment, like the other experiments reported in this paper, employed a circular target, 30 minutes of visual angle in diameter, with a duration of 1/100 second, on a ten foot-lambert background. The more general aspects of the procedure and apparatus involved in this experiment and the other experiments reported in this paper are discussed in Reference 25 and in an article by Blackwell, Pritchard, and Ohmart (Ref. 4).

Tables 1, 2, and 3 also contain the data which serve as a basis for the coefficients of correlation that are reported in References 22 and 25 and again below, between $P_N(A)$ and calculated threshold. The implication of these correlations is the same as the implication of the straight lines fitted to the data of Figures 2, 3, and 4, namely, that a dependence exists between the conditional probability that an observation arising from SN will be in the criterion and the conditional probability that an observation arising from N will be in the criterion — a condition that, as described elsewhere (Refs. 22, 24, 25), is inconsistent with the concept of threshold, or fixed criterion.

The product-moment correlations for the three observers between $P_N(A)$ and calculated threshold, based on the twelve sessions involving a payoff matrix

DAY	P(SN)	$V_{SN \cdot A}$	$K_{SN \cdot CA}$	$K_{N \cdot A}$	$V_{N \cdot CA}$	$P_N(A)$	$P_{S_1N}(A)$	$P_{S_2N}(A)$	$P_{S_3N}(A)$	$P_{S_4N}(A)$	Contrast Threshold
1	0.8	---	---	---	---	0.50	0.62	0.76	0.90	0.94	.68
2	.8	---	---	---	---	.44	.46	.74	.84	.94	.73
3	.4	---	---	---	---	.18	.40	.64	.76	1.00	.68
4	.4	---	---	---	---	.29	.32	.48	.80	.96	.81
5	.8	+1	-1	-2	+2	.44	.50	.54	.86	.98	.74
6	.8	+1	-1	-2	+2	.12	.26	.32	.68	.98	.73
7	.8	+1	-1	-2	+2	.36	.42	.58	.72	.96	.71
8	.8	+1	-1	-2	+2	.38	.44	.50	.74	.84	.73
9	.8	+1	-1	-2	+2	.30	.42	.53	.82	1.00	.78
10	.8	+1	-1	-2	+2	.20	.37	.53	.75	.97	.75
11	.8	+1	-1	-3	+3	.17	.27	.47	.73	.98	.78
12	.8	+1	-1	-4	+4	.07	.08	.17	.53	.89	.91
13	.4	+1	-1	-2	+2	.05	.07	.30	.63	.97	.86
14	.4	+1	-1	-1	+1	.04	.17	.43	.80	.97	.74
15	.4	+2	-2	-1	+1	.43	.47	.63	.90	1.00	.76
16	.4	+3	-3	-1	+1	.54	.60	.80	.80	1.00	.80

TABLE 1

DAY	P(SN)	$V_{SN \cdot A}$	$K_{SN \cdot CA}$	$K_{N \cdot A}$	$V_{N \cdot CA}$	$P_N(A)$	$P_{S_1N}(A)$	$P_{S_2N}(A)$	$P_{S_3N}(A)$	$P_{S_4N}(A)$	Contrast Threshold
1	.8	---	---	---	---	.38	.56	.76	.86	.98	.62
2	.8	---	---	---	---	.46	.72	.64	.88	.98	.66
3	.4	---	---	---	---	.17	.24	.56	.64	1.00	.76
4	.4	---	---	---	---	.15	.12	.60	.72	.96	.77
5	.8	+1	-1	-2	+2	.36	.54	.60	.92	.96	.63
6	.8	+1	-1	-2	+2	.18	.20	.40	.76	1.00	.77
7	.8	+1	-1	-2	+2	.24	.34	.56	.76	.92	.66
8	.8	+1	-1	-2	+2	.42.	.50	.66	.68	.90	.68
9	.8	+1	-1	-2	+2	.23	.23	.47	.72	.95	.84
10	.8	+1	-1	-2	+2	.37	.43	.63	.85	.95	.76
11	.8	+1	-1	-3	+3	.42	.47	.65	.93	1.00	.74
12	.8	+1	-1	-4	+4	.25	.20	.42	.77	.97	.81
13	.4	+1	-1	-2	+2	.04	---	---	.33	.80	1.07
14	.4	+1	-1	-1	+1	.03	.07	.27	.60	.97	.85
15	.4	+2	-2	-1	+1	.26	.20	.57	.80	.97	.77
16	.4	+3	-3	-1	+1	.35	.40	.50	.87	.97	.84

TABLE 2
YES-NO DATA FOR OBSERVER 2 IN THE
FIRST EXPECTED-VALUE EXPERIMENT

DAY	P(SN)	$V_{SN \cdot A}$	$K_{SN \cdot CA}$	$K_{N \cdot A}$	$K_{N \cdot CA}$	$P_N(A)$	$P_{S_1N}(A)$	$P_{S_2N}(A)$	$P_{S_3N}(A)$	$P_{S_4N}(A)$	Contrast Threshold
1	.8	---	---	---	---	.42	.54	.66	.86	1.00	.72
2	.8	---	---	---	---	.40	.30	.62	.90	1.00	.74
3	.4	---	---	---	---	.08	.08	.28	.60	.96	.84
4	.4	---	---	---	---	.03	.24	.32	.48	.96	.78
5	.8	+1	-1	-2	+2	.20	.44	.64	.84	.96	.61
6	.8	+1	-1	-2	+2	.18	.34	.38	.84	.96	.70
7	.8	+1	-1	-2	+2	.08	.22	.48	.74	.98	.63
8	.8	+1	-1	-2	+2	.18	.32	.46	.74	.94	.63
9	.8	+1	-1	-2	+2	.17	.15	.47	.85	1.00	.76
10	.8	+1	-1	-2	+2	.15	.35	.55	.90	.97	.66
11	.8	+1	-1	-3	+3	.07	.10	.35	.75	.97	.81
12	.8	+1	-1	-4	+4	---	.07	.22	.62	.95	.83
13	.4	+1	-1	-2	+2	---	.13	.17	.42	.93	.89
14	.4	+1	-1	-1	+1	---	---	.20	.37	.80	1.00
15	.4	+2	-2	-1	+1	.24	.47	.77	.90	.97	.54
16	.4	+3	-3	-1	+1	.29	.37	.77	.83	.97	.67

TABLE 3
YES-NO DATA FOR OBSERVER 3 IN THE

)ays 5-16), are -.37(p = .245), -.60(p = .039) and -.81(p = .001). For the three ıbjects combined, p = .0008.

The dependence of false-alarm rate and calculated threshold is also il- ıstrated, though somewhat crudely, in Figures 5 and 6. These plots represent the ıta of days 5-16 for all three observers. The portion of data comprising each of ıe curves was selected to be relatively homogeneous with respect to $P_N(A)$. Figure shows the raw data; the proportion of positive responses is plotted as a function : ΔI, the signal intensity. Figure 6 shows the same data after application of the ıance correction; here the calculated threshold (the value of ΔI corresponding to ıe corrected proportion of .50) is seen to be dependent upon $P_N(A)$ in the pre- ıcted direction. The observer, apparently, can adjust his criterion. If he)erates with a lower value of β, the proportion of correct detections will be ıcreased by an amount that is not completely eliminated by the coincident increase ı the size of the chance correction factor, $P_N(A)$.

.2 Data from the Second Experiment

Since the data from the first experiment did not suffice to trace out a)mplete ROC curve, a second experiment was conducted to obtain a broader range of ılues of $P_N(A)$. A different set of observers was used in this experiment, and ıly one value of signal intensity was employed. The data for the four observers :e shown in Figures 7 through 10. Each point represents a two-hour observing :ssion including 200 presentations of signal and 200 observations in which no ıgnal was presented; that is, P(SN) was held at .50 throughout this second experi- :nt. Changes in β, and thus in $P_N(A)$, were effected entirely by changes in the :lative values and costs with which an Expected-Value Observer operates. The data :om which Figures 7 through 10 were plotted are presented in Table 4. The column

2262 A-G3-175 JEM 1-3-55

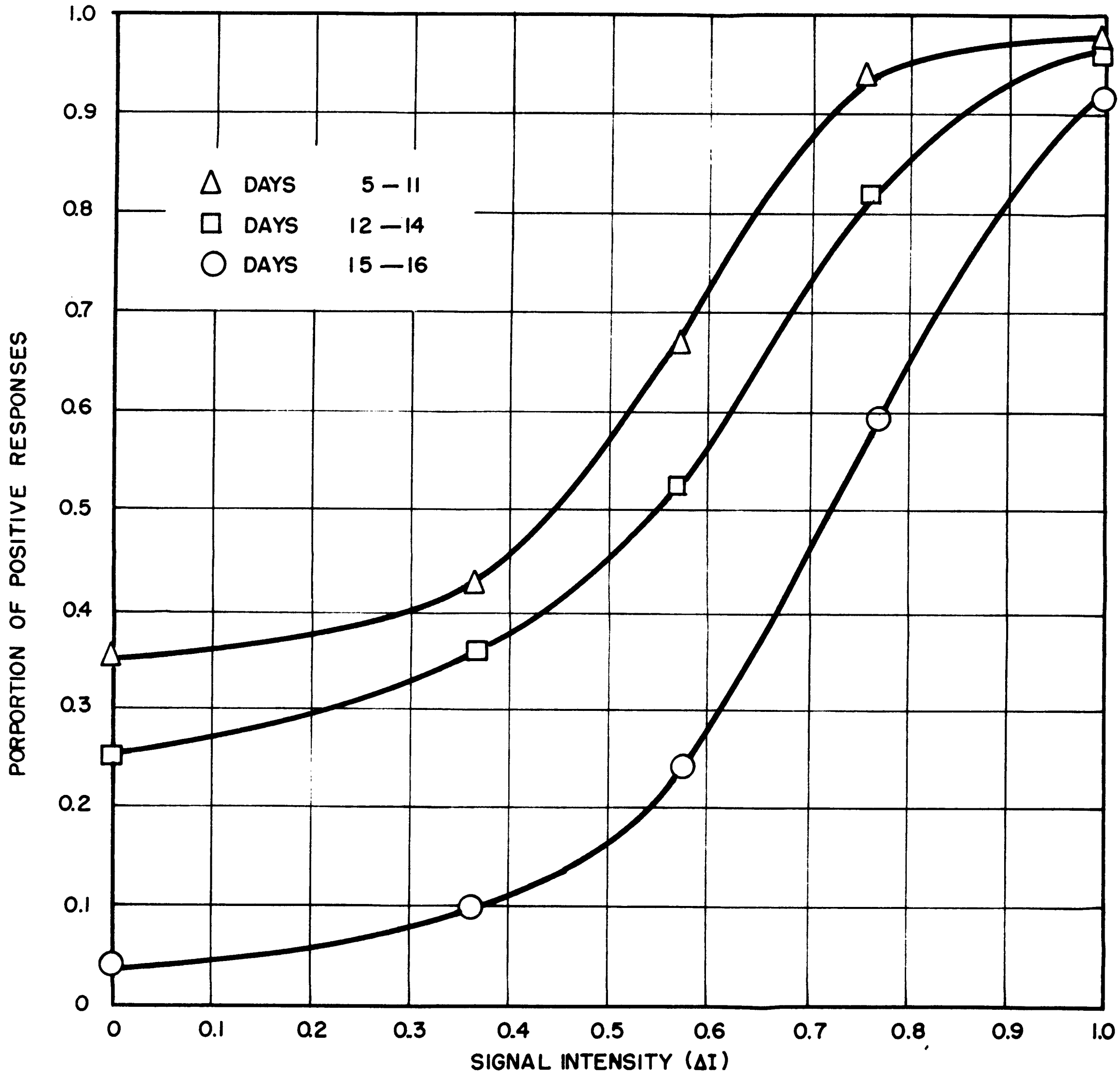

FIG. 5. RAW DATA, AVERAGE FOR ALL OBSERVERS IN THE FIRST EXPECTED-VALUE EXPERIMENT.

2262 A-63-176 JEM 1-3-55

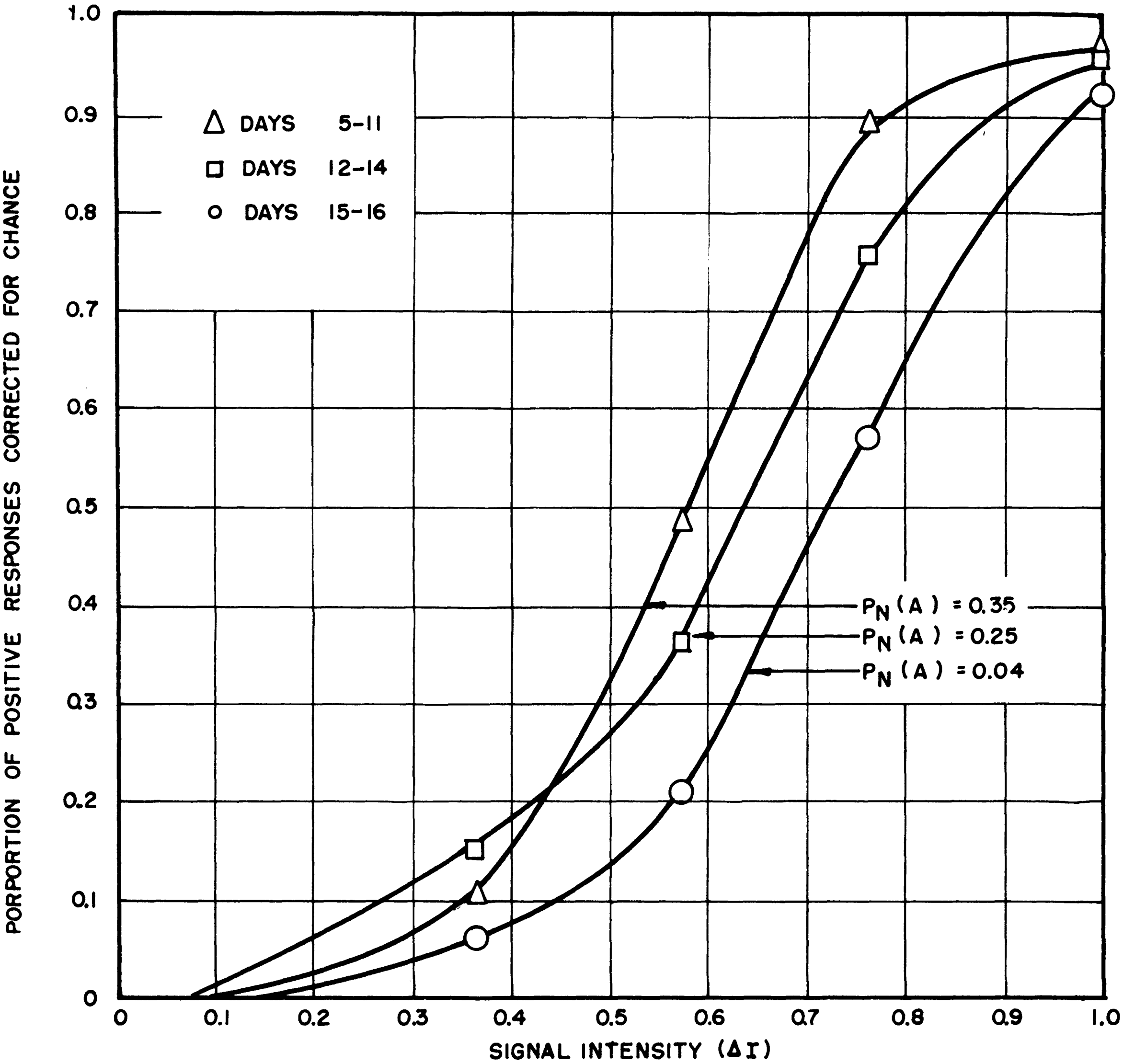

FIG. 6. RAW DATA CORRECTED FOR CHANCE, AVERAGE FOR ALL OBSERVERS IN THE FIRST EXPECTED-VALUE EXPERIMENT.

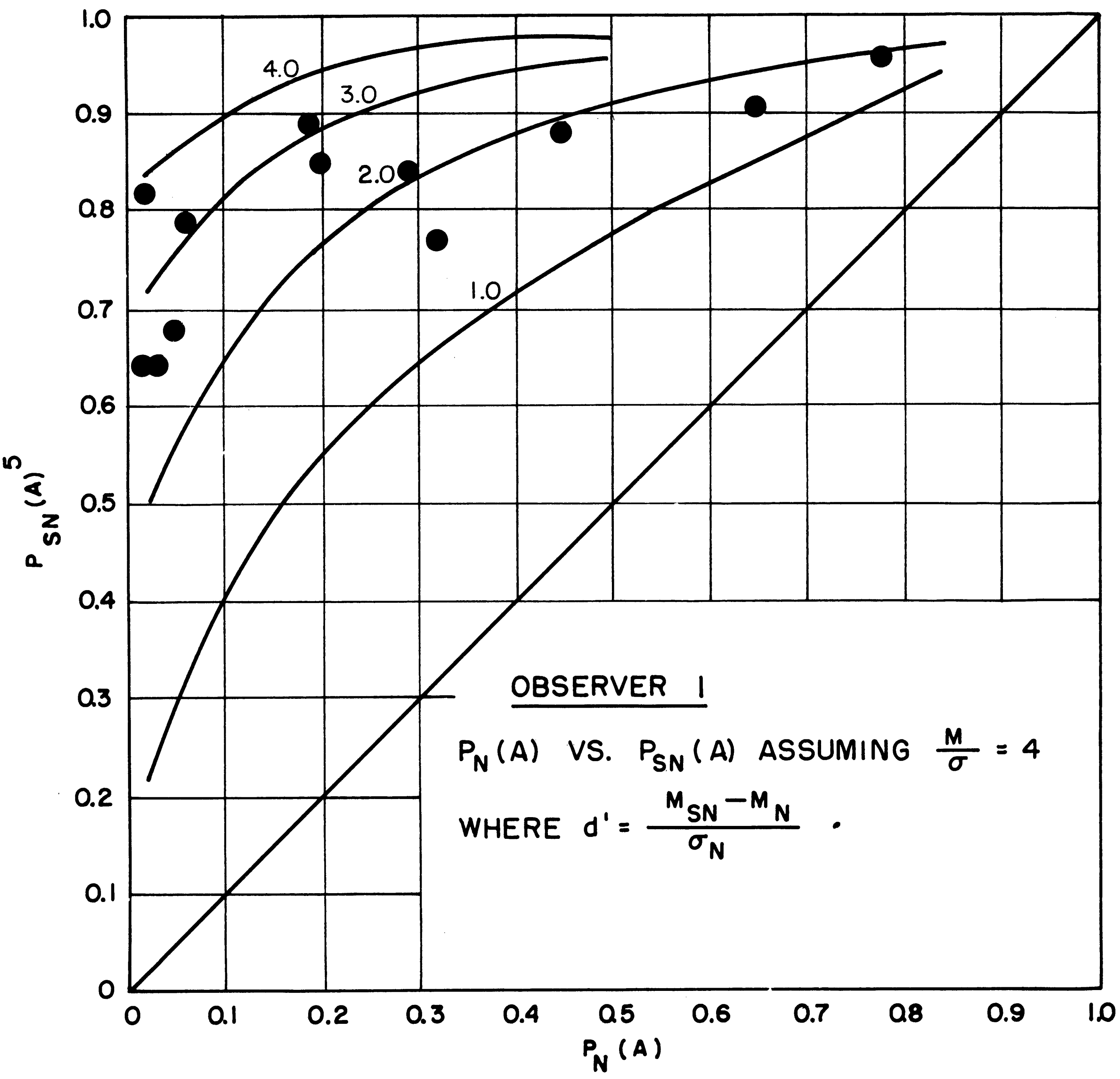

FIG. 7. $P_{SN}(A)$ VS. $P_N(A)$ FOR OBSERVER I IN THE SECOND EXPECTED-VALUE EXPERIMENT.

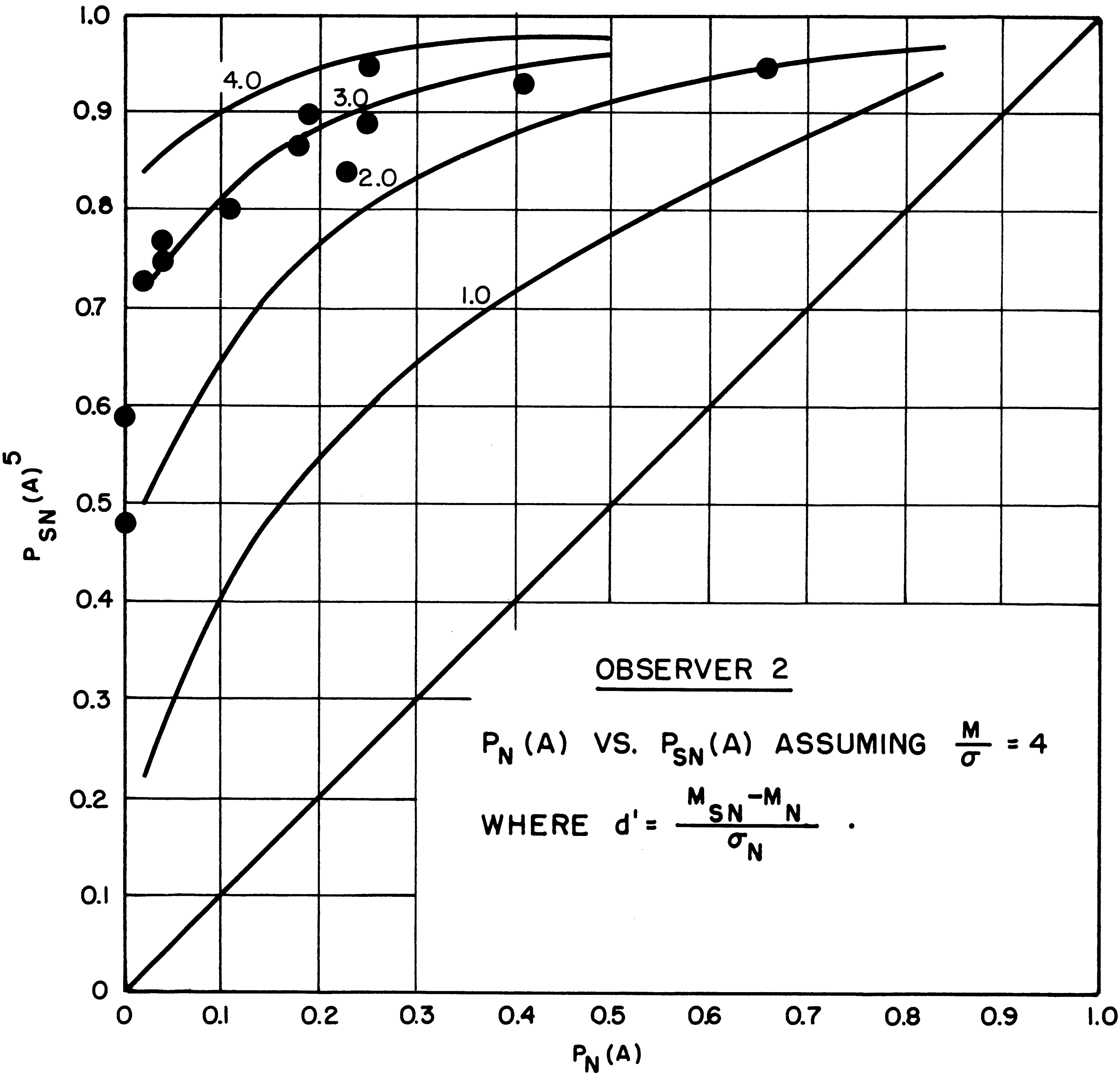

FIG. 8. $P_{SN}(A)$ VS. $P_N(A)$ FOR OBSERVER 2 IN THE SECOND EXPECTED-VALUE EXPERIMENT.

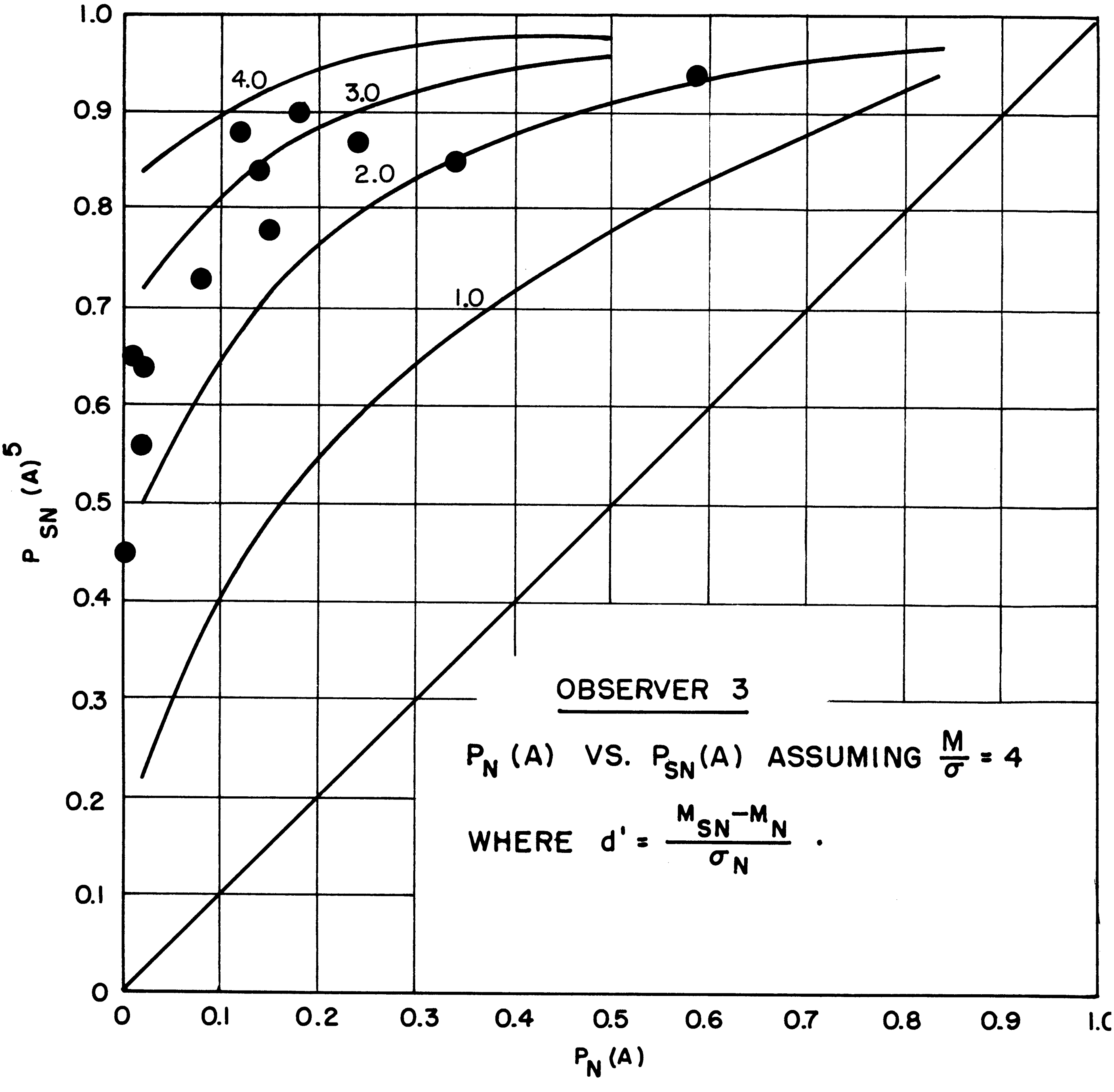

FIG. 9. $P_{SN}(A)$ VS. $P_N(A)$ FOR OBSERVER 3 IN THE SECOND EXPECTED-VALUE EXPERIMENT.

2262 A-63-148 JEM 9-8-54

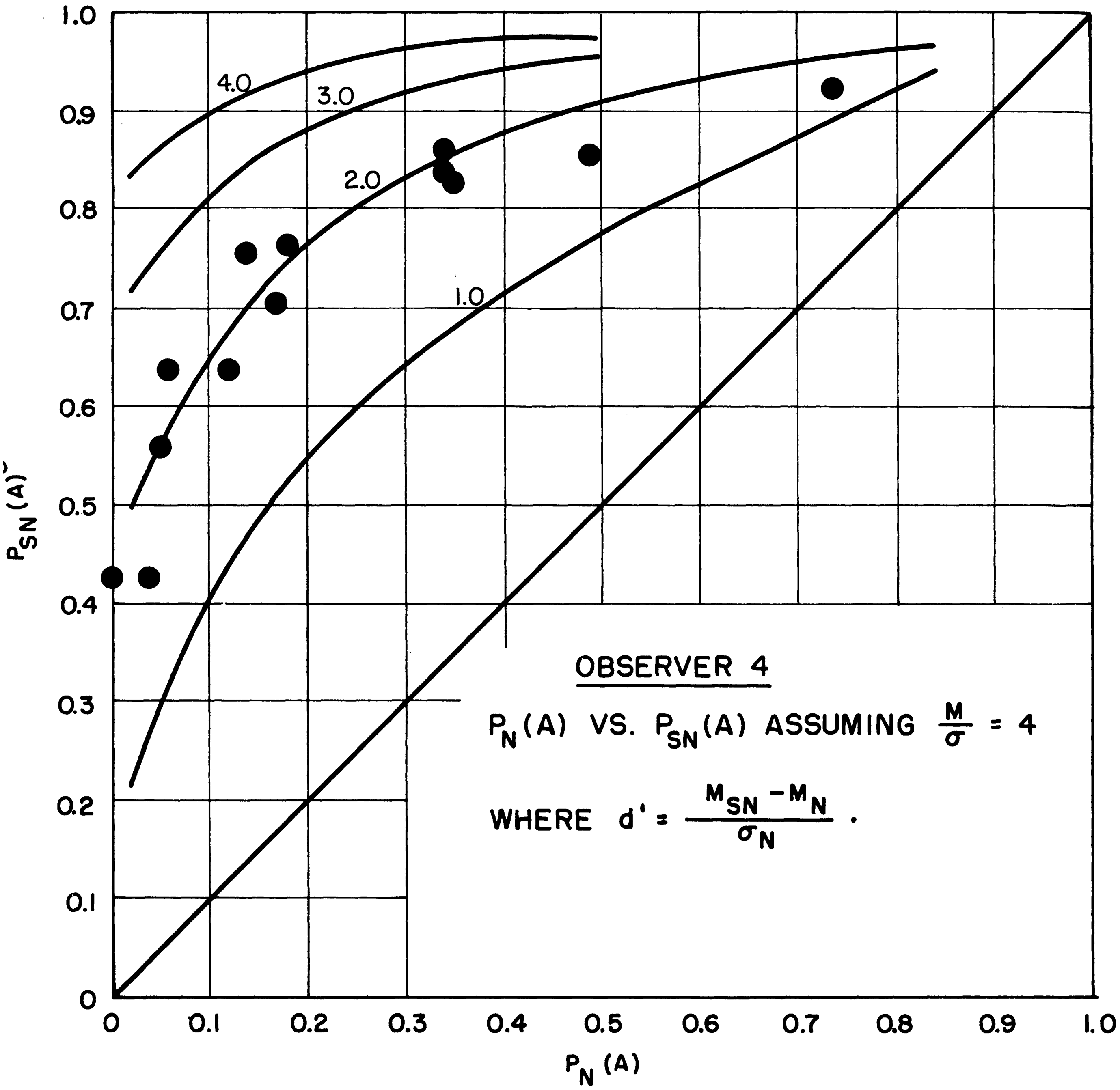

FIG. 10. $P_{SN}(A)$ VS. $P_N(A)$ FOR OBSERVER 4 IN THE SECOND EXPECTED-VALUE EXPERIMENT.

Day	β	Announced Optimum $P_N(A)$	Observer 1		Observer 2		Observer 3		Observer 4	
			$P_N(A)$	$P_{SN}(A)$	$P_N(A)$	$P_{SN}(A)$	$P_N(A)$	$P_{SN}(A)$	$P_N(A)$	$P_{SN}(A)$
1	1.00	.23	.32	.77	.23	.84	.18	.90	.18	.77
2	.50	.42	.50	.88	.25	.95	.12	.88	.35	.83
3	.25	.58	.65	.91	.41	.93	.33	.85	.49	.87
4	.16	.68	.78	.96	.66	.95	.59	.94	.74	.93
5	2.00	.12	.05	.63	.11	.80	.01	.65	.17	.71
6	4.00	.05	.06	.79	.03	.75	.02	.64	.12	.64
7	1.50	.15	.20	.85	.19	.90	.15	.78	.14	.76
8	6.00	.02	.02	.82	.02	.73	---	---	.06	.64
9	8.00	.00	.02	.64	.00	.48	.00	.45	.00	.43
10	2.50	.09	---	---	.04	.71	.08	.73	.05	.56
11	.75	.28	.19	.89	.18	.87	.14	.84	.34	.84
12	.75	.28	.29	.84	.25	.89	.24	.87	.34	.86
13	8.00	.00	.03	.64	.00	.59	.02	.56	.04	.43

TABLE 4
DATA FROM 4 OBSERVERS
IN SECOND EXPECTED-VALUE EXPERIMENT

n Table 4 labeled "Announced Optimum $P_N(A)$" will be discussed later in this sec-ion of the paper.

It may be seen in Figures 7 through 10 that the experimentally determined oints are reasonably well fitted by the type of ROC curve predicted by the theory ubscribed to here. It is equally apparent that points are not well fitted by a traight line intersecting the point $P_N(A) = P_{SN}(A) = 1$. It is important to note, n this connection, that the theoretical ROC curves plotted in Figures 7 through 10 re not exactly the same as those plotted in Figure 1. The exact form of the theo-etical ROC curves portrayed in Figure 1, as pointed out in Reference 25, is depend-nt upon the assumption that the distributions of N and S + N are Gaussian and of qual variance. It was apparent at the time of writing of Reference 25 that the ssumption of equal variance was not entirely adequate; the nature of the variance ssumption, however, was not critical for the purposes of that paper, and the equal-ariance assumption was accepted to facilitate analysis. The theoretical curves hown here in Figures 7 through 10, which are fitted quite well by the data plot-ed there, were drawn under the assumption that the ratio of the increment in the ean to the increment in the standard deviation is equal to 4. The implication for nalysis procedures of a dependence between signal strength and variance is dis-ussed in the paper on methods (Ref. 20).

.3 The Approach to the Optimum Behavior

The fact that the data obtained in the two Expected-Value experiments is itted more adequately by the type of ROC curve consistent with the theory of sig-al detectability than by the straight-line ROC curve consistent with the fixed hreshold notion implies that the human observer can vary his criterion. The uestion remains as to how closely he approaches the optimum criterion.

To establish the applicability of the theory of signal detectability, it is necessary to demonstrate only that the observer's operating level is some monotonic function of β. When sampling error is taken into account, the restriction is more lenient; in this case, it is sufficient to demonstrate a significant correlation between the observer's operating level and β. It is interesting, however, to determine how closely the observer's operating level approaches that specified by β.

In the first experiment, the observers were informed of the various values taken on by the elements of the β-equation (see page 6); i.e., the a priori probabilities, the values and costs. They were not made aware of the way in which these elements determine a value of β, nor did they know that the single number, β, taken in conjunction with an ROC curve, yields the optimum false-alarm rate. They were informed, after each experimental session, of their false-alarm rate, the proportion of correct detections of each of the four signal intensities, and the payoff. The correlations obtained in this study between β and false-alarm rate indicate that these observers tended toward the optimum setting of the criterion. For the three observers, the rank-order correlations were .70, .46, and .71; a correlation of .68 is significant at the .01 level of confidence.

The first study demonstrated that the human observer quite naturally adjusts his criterion in a way approaching optimum. With this as background, the second experiment was performed with the observers having a fairly complete knowledge of its purposes. The second study attempted to determine how closely the observers could approximate the optimum false-alarm rate, as specified by β, given a knowledge of it. Hence the column in Table 4 labelled "Announced Optimum $P_N(A)$". The experimenter's ability to announce a value of $P_N(A)$ approaching the optimum value, before the ROC curves of the observers had been determined directly epends upon the fact that, as reported in references 22 and 25, forced-choice and es-no response procedures yield consistent values of d'. The observers in the econd study, like those in the first, had been trained under the forced-choice rocedure, so estimates of d' were available.

A comparison of the column in Table 4 labelled "Announced Optimum $P_N(A)$" nd the columns containing the $P_N(A)$ obtained from each observer indicate how well he observers reproduced the value announced as optimum. The rank-order correla- ions between the announced optimum and obtained values of $P_N(A)$ are .94, .97, .86, nd .98. Two of the observers served in twelve sessions; the other two served in hirteen sessions. The rank-order coefficient associated with a probability of 01, given twelve pairs of measures, is .68. In this study, the observers were nformed of their proportion of false alarms after each group of fifty trials. he study of the Neyman-Pearson Observer, reported in the next section of this aper, provides additional evidence for the rather remarkable ability of the obser- er to reproduce a given false-alarm rate.

3. THE NEYMAN-PEARSON OBSERVER

.1 The Approach to the Announced Optimum Value of $P_N(A)$

A different set of four observers served in this experiment. The obser- ers were informed of the *a priori* probability of signal occurrence ($P(SN) = .72$ as held throughout the experiment), but instead of operating in terms of values nd costs, they attempted to satisfy a restriction placed on the proportion of alse alarms.

The restriction on the proportion of false alarms took the form of a tipulation for the observers of the acceptable number of "yes" responses to the ourteen "no-signal" presentations in each block of fifty presentations. The observers were instructed to respond positively to approximately n or n + 1 of th fourteen "no-signal" presentations (n, for the four successive conditions of the experiment, equalled 3, 0, 6, and 9 respectively), so that any proportion of fals alarms, across several blocks of fifty presentations, within a given range of .08 satisfied the restriction. A given restriction on $P_N(A)$ was effective across eighteen blocks of fifty trials. There were then four conditions, each with a different acceptable range for $P_N(A)$; thus, the primary data consist of four valu of $P_N(A)$ for each observer, with each of these four values based on 252 "no-signa presentations. The acceptable range for $P_N(A)$ for the four conditions are shown as column headings in Table 5; the false-alarm rates obtained from the four obser vers appear as cell entries. Note that the largest deviation from the range announced as acceptable is .04. These data, then, also suggest that the observer i able to vary, and quite precisely, the cutoff point on the continuum of observations.

It is true that the data presented in the previous paragraph, to the extent that they were reported there, could have been obtained if the threshold concept were valid. If the observer were given immediate knowledge of correctnes of response, any false-alarm rate could be approximated, for example, by saying "yes" until the given proportion of false alarms was achieved and then saying "no on the rest of the presentations. This would entail, however, a severe depressio of d'. In this study, the observers were given knowledge of correctness of response only after each block of 50 presentations, and the values of d' were not depressed.

3.2 Other Analyses of the Data of this Study

In this study, twelve values of signal intensity were used, in addition to the "blank" or "zero-intensity signal". This rather large number of signal

TABLE 5 -- NEYMAN - PEARSON OBSERVER

Restriction on $P_N(A)$ / Observer	.21-.29	0.0-.07	.43-.50	.64-.71
1	.28	.03	.42	.74
2	.19	.02	.42	.61
3	.21	.05	.39	.63
4	.22	.02	.40	.60

values was employed in an attempt to define more adequately the shape of the psychophysical curve. The results of the analysis of curve shape will be discuss in the paper dealing with psychophysical methods (Ref. 20).

The nature of the ROC curves resulting from this study cannot be adequately determined since only four points were obtained for each curve; that is, for each value of d' or signal intensity. Neither can the degree of correlation between $P_N(A)$ and calculated threshold be estimated since only four pairs of measures were obtained. Thus, the results of this study could not stand alone as evidence for the existence of a variable criterion, or stated more generally, as evidence that the observer's decision function is a monotonic function of likelihood ratio. It is believed, however, that the other four experiments reported herein are sufficient to establish this point.

4. THE *A POSTERIORI* PROBABILITY OBSERVER

The same four observers who served in the second Expected-Value experiment served in this study of the ability of the human observer to report *a posteriori* probability. In this experiment, P(SN) = .50. The task posed for the observers was to place each observation in one of six categories of *a posteriori* probability. Here again, the ability to order the observations down into the noise is required.

One analysis of the data is reported in Table 6. The categories of *a posteriori* probability with which the observers worked head the columns. The boundaries of the categories were chosen in conference with the observers; they believed that they would be able to operate reasonably with this particular schem The cell entries show the proportion of the observations placed in a given catego that were, in fact, observations of signal-plus-noise; note that each of the

TABLE 6 -- A POSTERIORI PROBABILITY OBSERVER

Announced Categories

	0.0 - .04	.05 - .19	.20 - .39	.40 - .59	.60 - .79	.80 - 1.00
Observer 1	.20	.28	.45	.49	.76	.97
	.23	.24	.42	.46	.84	.95
	.20	.23	.28	.61	.77	.95
Observer 2	.11	.28	.47	.70	.92	1.00
	.14	.16	.50	.86	.96	1.00
	.18	.19	.38	.64	.91	.96
Observer 3	.16	.31	.76	.56	.65	.94
	.22	.38	.34	.24	.70	.92
	.21	.21	.42	.43	.36	.82
Observer 4	.26	.71	.88	1.00	1.00	1.00
	.27	.77	.91	1.00	1.00	1.00
	.33	.53	.75	.78	.95	.95

observers served in three experimental sessions. Thus, the entry in the upper left-hand corner indicates that 20 percent of the observations placed in the lowest probability category by Observer 1, in the first session, were observations o signal-plus-noise; 28 percent of the observation placed in the next category were observations of signal-plus-noise, and so forth. In nine of twelve cases, i.e., in nine of the twelve rows of Table 6, a rank-order correlation of unity exists between the estimated a posteriori probability and the relative frequency of obse vations arising from signal-plus-noise, indicating an ordering of observations. Each of these correlations has an associated probability of less than .05. All of the three cases showing a correlation of less than unity are attributable to a single observer, Observer 3.

4.1 The Relationship between d(x) and ℓ(x)

It is possible, using the data from this experiment, to determine the relation existing between the observer's decision function, d(x) and likelihood-ratio, $\ell(x)$. For this experiment, with P(SN) = .50, the equation for a posterior probability, given in the introductory section of this paper, reduces to $P_x(SN) = \frac{\ell(x)}{\ell(x) + 1}$. Since $P_{SN}(A)$ and $P_N(A)$ can be estimated for the boundaries o the six categories, d' can be determined. A knowledge of $P_N(A)$ (relative to each boundary) and d' allows a graphic determination of the value of $\ell(x)$ correspondin to each boundary since the critical value of $\ell(x)$ for a given value of $P_N(A)$ equals the slope of the ROC curve at the point which corresponds to that value of $P_N(A)$. Thus $P_x(SN) = \frac{\ell(x)}{\ell(x) + 1}$ can be determined for each operating level employed by the observer. The values of $\frac{\ell(x)}{\ell(x) + 1}$ obtained in this way for the observer will correspond directly to the probability values marking off the categories — if the observer is operating with a decision function, d(x), that i equal to $\ell(x)$, and according to the optimum relation between $P_x(SN)$ and $\ell(x)$.

Figures 11, 12, 13, and 14 show plots of the probability values corresponding to boundary categories versus $P_x(SN) = \frac{\ell(x)}{\ell(x) + 1}$ determined from the data for the four observers. Observers 1 and 2 appear to be operating with d(x) similar to ℓ(x) and approximately according to knowledge of the optimum relation between P_x(SN) and ℓ(x). This result should probably not be generalized beyond the condition of P(SN) = P(N) = .50. It should be remembered, however, that the observers were exposed to this task for only four[1] experimental sessions, with feedback after each session limited to the information reported in Table 6. With respect to Observers 3 and 4, it is not clear whether they are operating with a d(x) quite unlike ℓ(x), or with an imperfect approximation to the optimum relation between P_x(SN) and d(x), or both.

It is possible to determine more exactly the relation between ℓ(x) and d(x) for the four observers. The observer's decision function, d(x), can be set equal to $\frac{\ell(x)}{w}$ and the value of w determined. Then the cutoff values of P_x(SN) used by the observers are described by

$$P_x(SN) = \frac{\frac{\ell(x)}{w}}{\frac{\ell(x)}{w} + 1}, \text{ or } P_x(SN) = \frac{\ell(x)}{\ell(x) + w}, \text{ and } w = \ell(x)\,\frac{1-p}{p},$$

where p is set equal to the boundaries of the categories. One interesting question is whether the weights, w's, are constant for each observer. The values of w corresponding to each category boundary, for each session and for each observer, are given in Table 7.

1. Observers 1, 3, and 4 reported *a posteriori* probability in a single session preceding the three sessions reported in Table 6. Since different categories were used, this session was considered as practice, and the data excluded from the analysis.

2262 A-G3-171 JEM 12-31-54

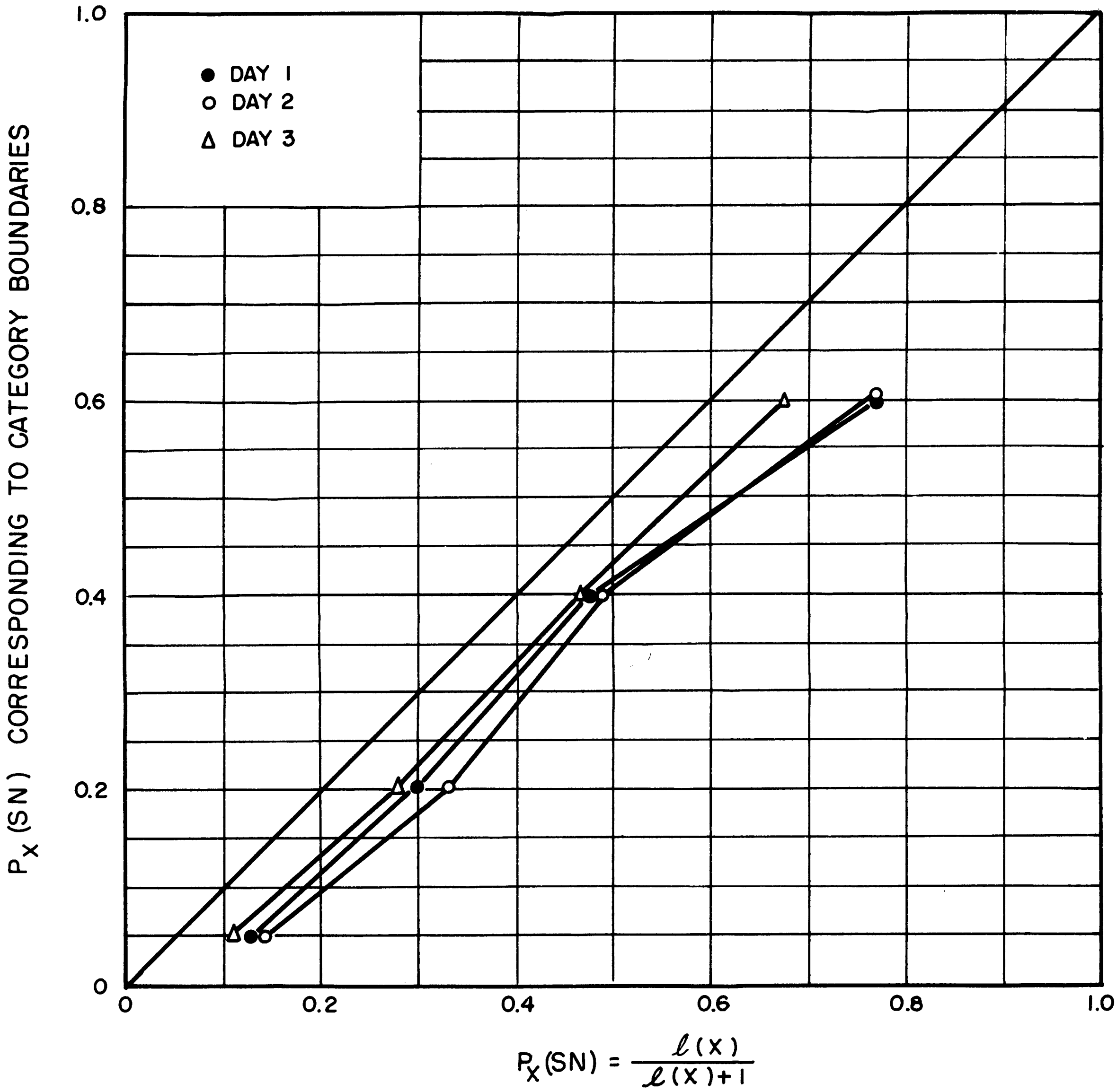

FIG. 11. OBJECTIVE PROBABILITY VS. SUBJECTIVE PROBABILITY FOR OBSERVER 1.

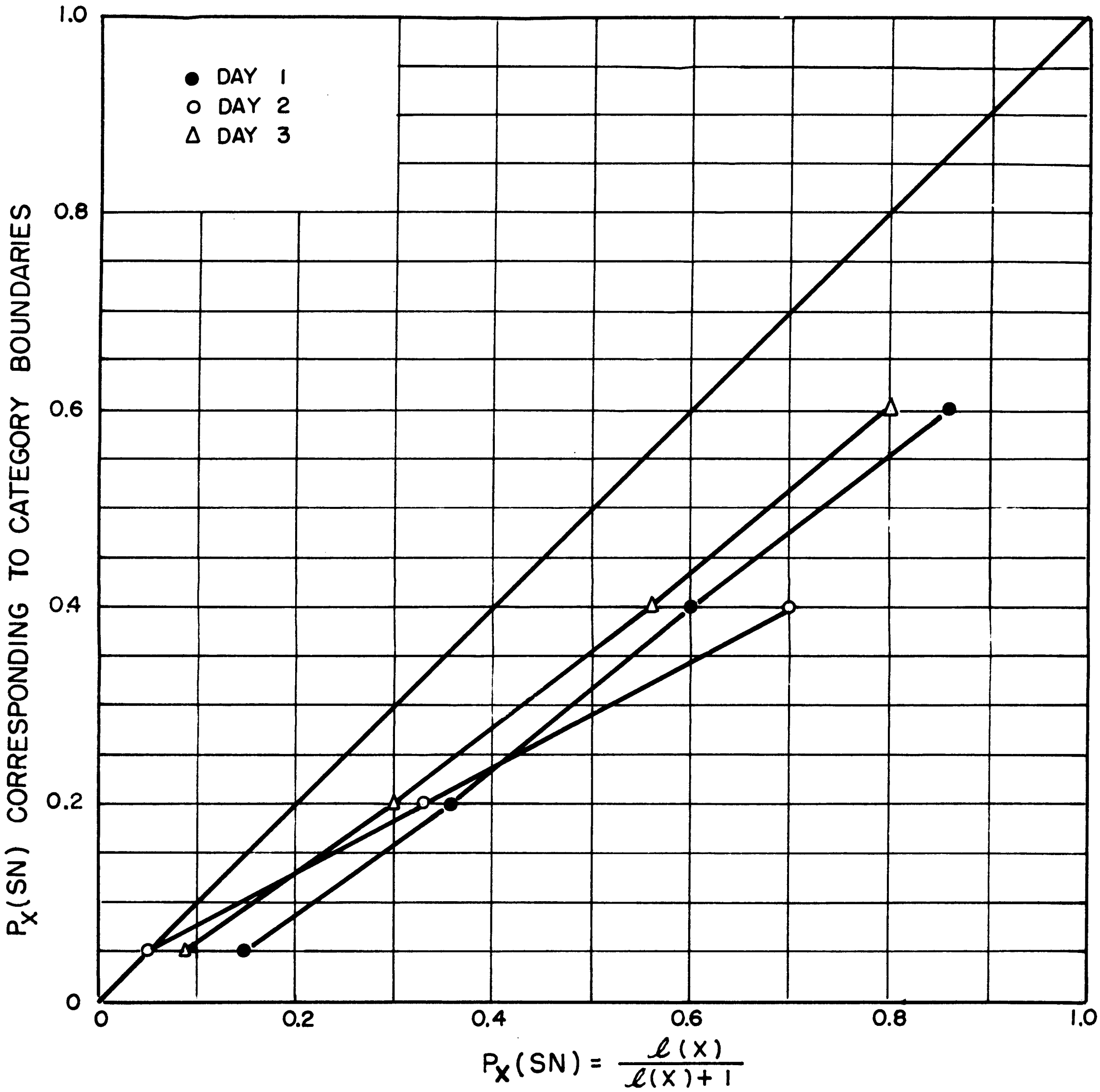

FIG. 12. OBJECTIVE PROBABILITY VS. SUBJECTIVE PROBABILITY FOR OBSERVER 2.

2262 A-G3-173 JEM 12-31-54

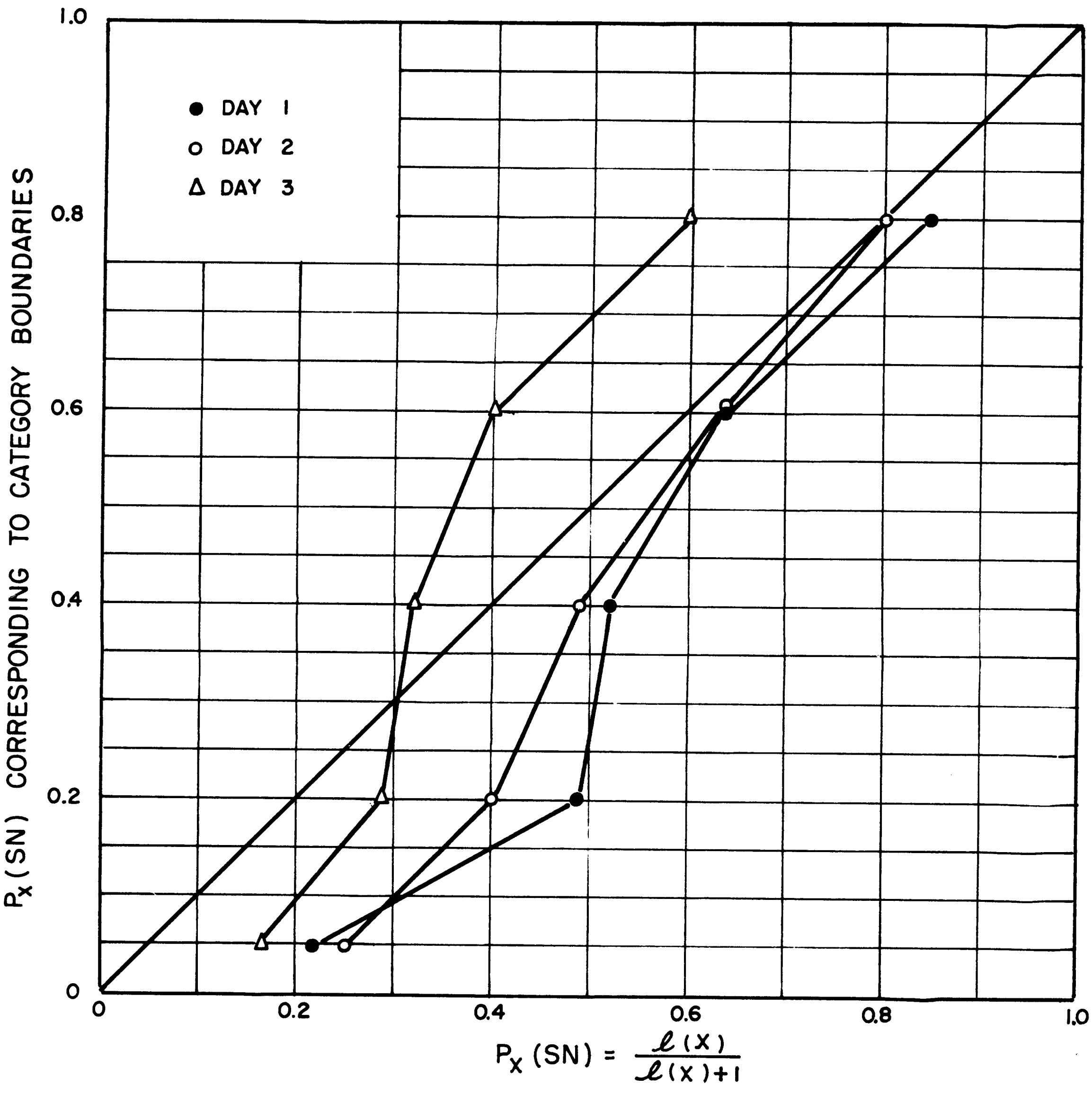

FIG. 13. OBJECTIVE PROBABILITY VS. SUBJECTIVE PROBABILITY FOR OBSERVER 3.

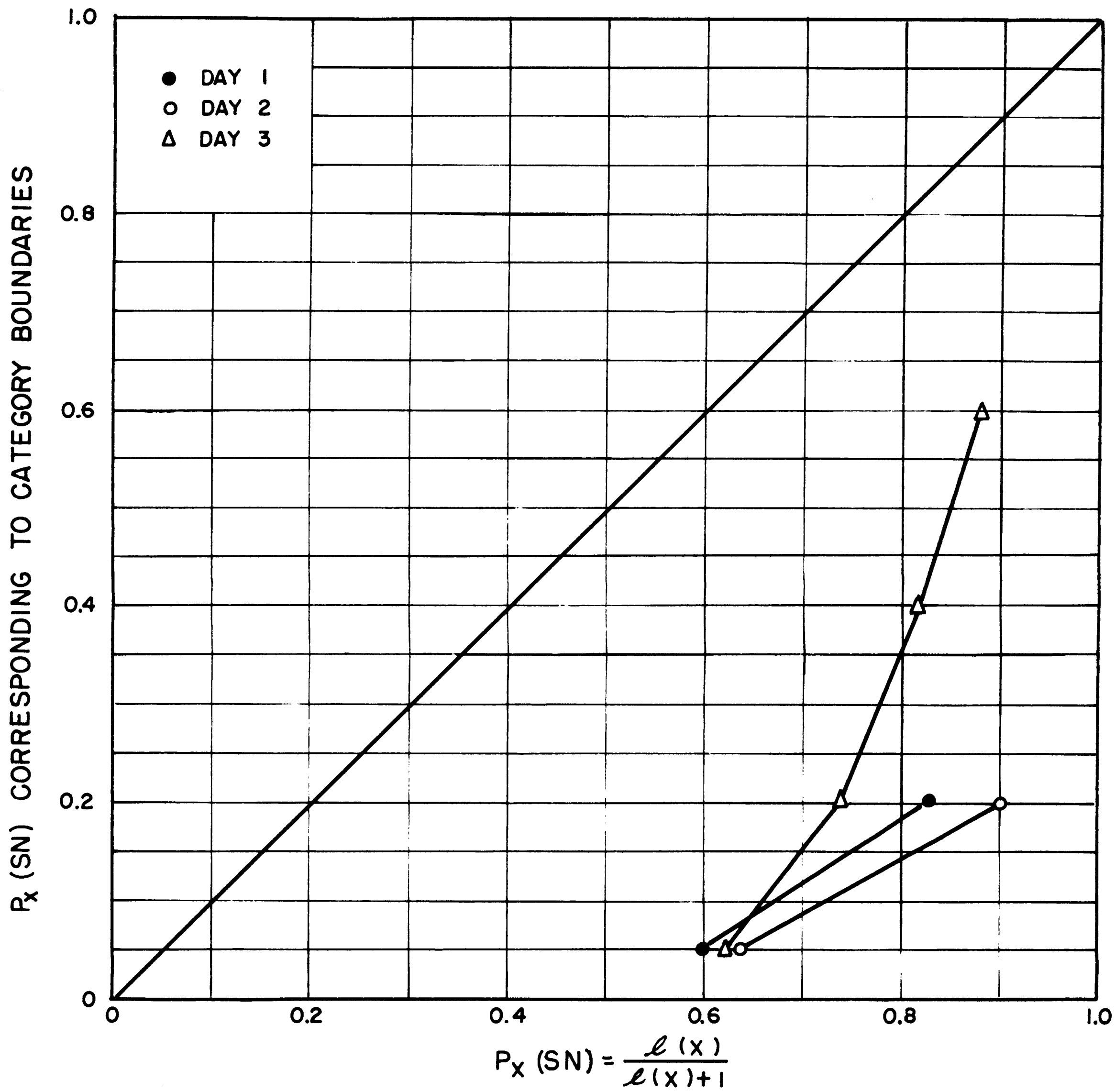

FIG. 14. OBJECTIVE PROBABILITY VS. SUBJECTIVE PROBABILITY FOR OBSERVER 4.

TABLE 7 -- THE RELATIONSHIP BETWEEN d(x) and $\ell(x)$[+]

Observer and Day \ Category Boundaries		.05	.20	.40	.60	.80
	Day 1	2.9	1.7	1.3	2.2	---
Obs. 1	Day 2	3.2	2.0	1.4	2.2	---
	Day 3	2.5	1.6	1.3	1.4	---
	Day 1	3.2	2.2	2.3	4.0	---
Obs. 2	Day 2	.95	2.0	3.5	---	---
	Day 3	1.9	1.7	1.9	2.7	---
	Day 1	5.3	3.8	1.7	1.2	1.4
Obs. 3	Day 2	6.4	2.6	1.4	1.2	1.0
	Day 3	4.0	1.6	.7	..4	.3
	Day 1	28.5	20.0	---	---	---
Obs. 4	Day 2	34.2	36.0	---	---	---
	Day 3	30.4	11.6	6.9	4.7	---

[+] The weights for some of the higher categories are indeterminate since the critical value of $\ell(x)$ approaches infinity as $P_N(A)$ approaches 0.0.

To determine the degree to which w approximates a constant for each ession and each observer, a given fixed value of w can be chosen for each obser- er, and used to determine $P_x(SN) = \frac{\ell(x)}{\ell(x) + w}$. If the "true" values of w are pproximately constant for a given observer, then a plot of the points, represent- ng category boundaries versus $P_x(SN) = \frac{\ell(x)}{\ell(x) + w}$ (with w fixed), should be ade- uately fitted by a straight line intersecting the points (0,0) and (1.00, 1.00). hat is, each of the lines in Figures 11 through 14 should correct to the diagonal. his, however, is not the result. Using fixed values of w resulted in a plot of oints for Observers 1 and 2 that are only slightly better fitted by the diagonal ine than are the points shown in Figures 11 and 12. A fixed value of w greatly mproves the correspondence of the diagonal line and the plotted points for bserver 4 — for Days 1 and 2. The correspondence for Observer 3 is not improved oticeably by the transformation.

It may be noted, however, that there is the suggestion (in Table 7) of consistent pattern between "subjective probability" and "objective probability" hat is, of a pattern in the relation between w and the category boundaries. The 's for the last two sessions for Observer 1 show the same rank-order. For Obser- er 3, the values of w rank from left to right for each of the three sessions with single exception.

5. THE SECOND-CHOICE EXPERIMENT

In this experiment, a variation of the forced-choice method of response as employed to demonstrate the ability of the observer to order observations.[1]

. This experiment was suggested by R. Z. Norman, formerly a member of the Electronic Defense Group, now at Princeton University.

In the version of the forced-choice method most commonly used, the observer knows that on each trial the signal will occur in one of four short, successive time intervals, and he is forced to choose in which of these intervals he believes the signal occurred. If the observer can order observations, then to behave optimally he must select the interval with the greatest associated observation. If the observer behaves optimally, then the probability that a correct answer will result for a given value of d', for the four-choice or four-interval situation, is given by the equation:

$$P(c) = \int_{-\infty}^{+\infty} \left[F(x)\right]^3 g(x)dx \qquad (1)$$

where F(x) is the area of the noise distribution and g(x) is the ordinate of the signal-plus-noise distribution. The value of P(c) corresponding to each value of d' is shown by the middle curve of Fig. 15. This curve is plotted under the assumption that the distributions of N and S + N are Gaussian and of equal variance

5.1 The Rationale for the Second-Choice Experiment

Consider the situation where the observer is required to indicate a second choice as well as a first choice. What is the probability of a correct second choice on those trials on which the first choice is incorrect?

The conventional notion of a threshold — as pointed out above and, in more detail, in Reference 25 — is essentially that the mechanism of detection is one that triggers when the observation exceeds a critical amount, and loses all discrimination among observations falling short of this amount. For the (four-choice) forced-choice situation where the observer is required to indicate a second choice as well as a first choice, this conception of the mechanism leads to the prediction that, when the first choice is incorrect, the probability that the second choice will be correct is .33. According to this iew, in other words, the second choice is made from among three intervals on a hance basis.

On the other hand, according to the conception of the human observer as perating in terms of likelihood ratio, the observer is capable of ordering the our observations associated with the four intervals. If this is the case, the roportion of correct second choices, on those trials in which an incorrect first hoice is made, should be greater than .33. The relationship between this pre-icted probability and d' is given by the expression

$$\frac{3\int_{-\infty}^{+\infty}\left[F(x)\right]^2 \left[1 - F(x)\right] g(x)\, dx}{1 - \int_{-\infty}^{+\infty}\left[F(x)\right]^3 g(x)\, dx} \qquad (2)$$

here the symbols have the same meaning as in Equation 1 above. This relationship s plotted in Figure 15 under the assumption that the distribution of N and S + N re Gaussian and of equal variance.

.2 Results

Data were collected from four observers; two of them had served pre-iously in the second Expected-Value experiment whereas the other two had only eceived routine forced-choice training. Each of the observers served in three xperimental sessions. Each session included 150 trials in which both a first nd second choice were required. The resulting twelve proportions of correct econd choices are plotted against d' in Figure 15. The values of d' were deter-ined by using the proportions of correct first choices as estimates of P(c) and eading the corresponding values of d' from the middle curve of Fig. 15. Although single value of signal intensity was used, the values of d' differed sufficiently rom one observer to another to provide an indication of the congruence of the

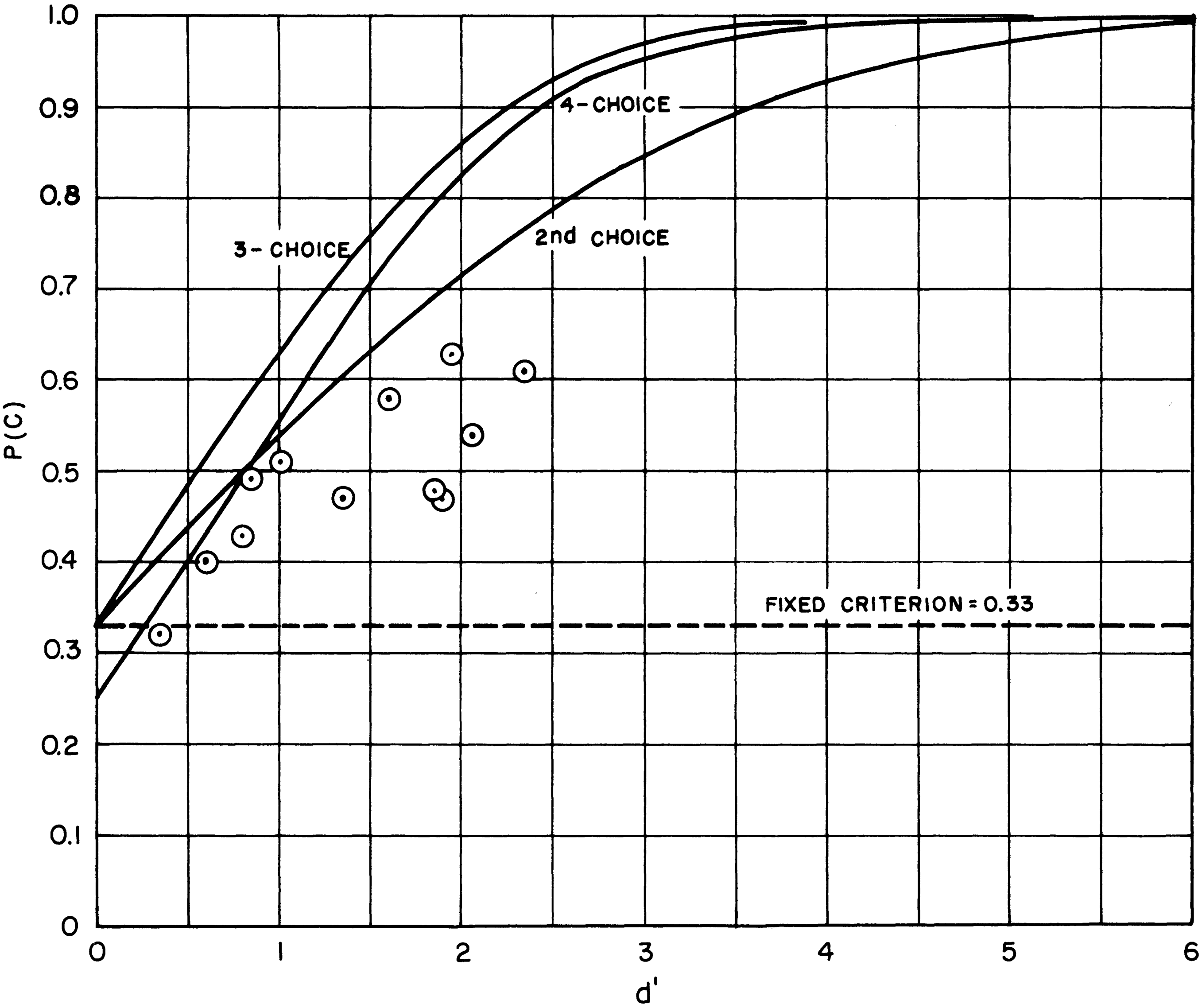

FIG. 15. THE SECOND-CHOICE DATA AND PREDICTIONS FOR THE SECOND-CHOICE SITUATION UNDER THE ASSUMPTION THAT THE N AND S+N DISTRIBUTIONS ARE GAUSSIAN AND OF EQUAL VARIANCE.

lata with the predicted functions. Additional variance in the estimates of d' resulted from the fact that, for two observers, a constant distance from the signal was not maintained. (The function predicted by the theory of signal detectability for the proportion of correct first choices in a three-choice, or three interval, situation is included in Fig. 15 to indicate that this function is not the same as the predicted function of the probability of a correct second choice, given an incorrect first choice, for the four-choice situation).

A systematic deviation of the data from a proportion of .33 clearly exists. Considering the data of the four observers combined, the proportion of correct second choices is .46. The deviation of this obtained proportion from .33 is highly significant, the χ^2 obtained (43.66) is more than twice the χ^2(19.0) associated with a probability of .00001.

Two control conditions aid in interpreting these data. The first of these allowed for the possibility that requiring the observer to make a second choice might depress his first-choice performance. During the experiment, blocks of 50 trials in which only a first choice was required were alternated with blocks of 50 trials in which both a first and second choice were required. Pooling the data from the four observers, the proportions of correct first choices for the two conditions are .650 and .651, a difference that is obviously not significant.

A preliminary experiment in which data were obtained from a single observer for five values of signal intensity also serves as a control. This experiment substantiates the predicted correlation between the probability of a correct second choice and signal intensity that derives from the theory of signal detectability. 150 observations were made at each value of signal intensity. The relative frequencies of correct second choices for the lowest four values of signal intensity were, in increasing order of signal intensity, 26/117 (.22), 33/95 (.35), 30/75 (.40), and 20/30 (.67). For the highest value of signal intensity, none of five second choices were correct. Thus, the proportion of correct second choices is seen to be correlated with a physical measure of signal intensity as well as with the theoretical measure (d') — this eliminates the possibility that the correlation found with a constant value of signal intensity, involving d' as one of the variables (Figure 15), is an artifact of theoretical manipulation. The second-choice data, then, demonstrate clearly the untenability of the assumption of a fixed criterion or threshold.

It may be seen from Fig. 15 that second-choice data also deviate systematically from the predicted function derived from the theory of signal detectability. This discrepancy results from the inadequacy of the assumption — of equal variance of the noise and signal-plus-noise distributions — upon which the predicted functions in Fig. 15 are based. It was pointed out above and in Reference 25 that the equal-variance assumption was accepted in the early stages of data collection in order to facilitate analysis, in spite of existing indications of its inadequacy. It was also pointed out above that more recent data on the Expected-Value Observer (see Figures 7-10 and accompanying text) indicate that a valid assumption would be the assumption that the ratio of the increment in the mean to the increment in the standard deviation, $\frac{\Delta M}{\Delta \sigma}$, is equal to 4. Figure 16 shows the second-choice data and the predicted four-choice and second-choice curves derived from the theory of signal detectability under the assumption that $\frac{\Delta M}{\Delta \sigma} = 4$. In view of the variance associated with each of the points (each first-choice d' was estimated on the basis of 300 observations and each second-choice proportion on less than 100 observations) the congruence of the data and the predicted function shown in Fig. 16 is quite remarkable.

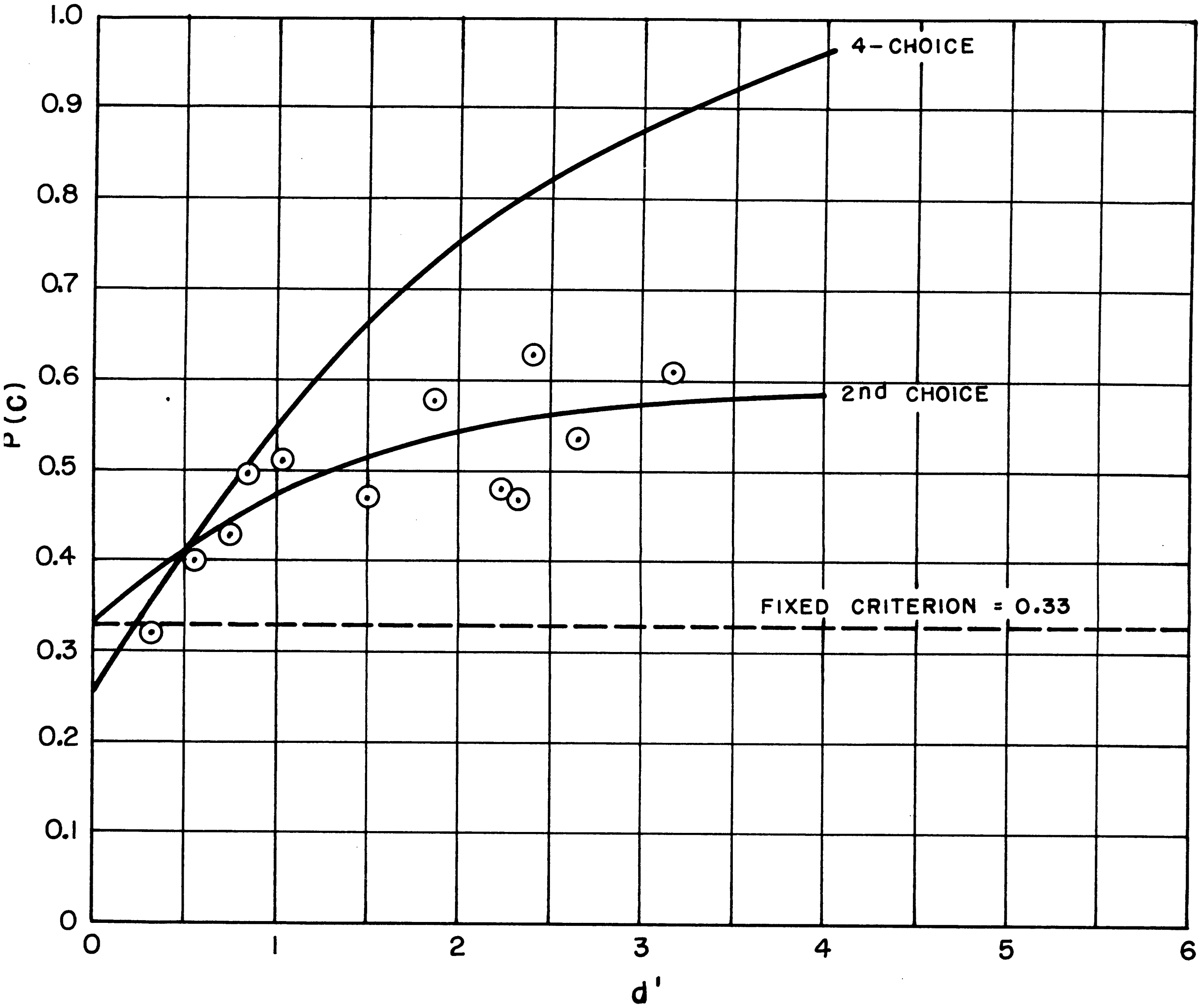

FIG. 16. THE SECOND-CHOICE DATA AND PREDICTIONS FOR THE SECOND-CHOICE SITUATION UNDER THE ASSUMPTION THAT $\frac{\Delta M}{\Delta \sigma} = 4$.

5.3 A Note on the Variance Assumption

The particular assumption made in this paper about the variance of the noi and signal-plus-noise distributions, namely that $\frac{\Delta M}{\Delta \sigma} = 4$, needs qualification in t respects. First, it is very likely specific to the experimental conditions employed. Second, it is to be regarded as only provisionally applicable to the present data.

If the variance of these sampling distributions is a function of sample size, then it may be presumed that their variances are different for different signal durations, or observation times, and for signals of different sizes. If the variance of the noise and signal-plus-noise distributions decreases with increases in signal size and duration, then an assumption concerning the increase i variance with increasing signal intensity is applicable only to data collected using a single duration and size of signal. Since the various experiments reported here employed identical signal conditions, it was possible to assess the adequacy of a single variance assumption for different forms of data, for "yes-nc data in Section 2 and for forced-choice data in Section 5. Although positive results were obtained from this check of internal consistency, it should not be inferred that the particular assumption will describe the results of experiments in volving different physical parameters.

Also, as indicated, that the assumption that $\frac{\Delta M}{\Delta \sigma} = 4$ derives from the data reported here is advanced with certain provisions. Other assumptions have not be thoroughly explored. It may be that ratios equal to certain other constants will fit the reported data even better, and that several other constant ratios will fi the data as well. The sensitivity of the assumed ratio to the existing data has not as yet been determined. It is likely that more precise data is required for the purpose of determining the relative adequacy of different variance assumption

These problems are presently being explored; the results will be reported in the paper concerned with methods of data analysis (Ref. 20).

6. DISCUSSION

The general conclusion drawn from the experimental results reported ove is that the model provided by the theory of signal detectability (Refs. 5, 16) or more generally, by the theory of statistical decision (Refs. 13, 14, 27, 3) is applicable to the detection behavior of the human observer. This model will roduce data like those observed.

This type of model has come to be called a "computer model". The term computer", in this connection, is meant very generally; it includes anything that rocesses information in a precisely defined way. Quastler's discussion of the ture of the computer model is pertinent here.

> "The computer model may be a system of equations. It may be a black-box diagram with boxes labelled 'receiver', 'memory', 'transducer', 'decision', etc. It may be a piece of hardware. It may have many or few components; it may be determinate or stochastic. Neither the size nor the type nor the physical nature of the model matter. All that does matter is that it should serve as a framework to organize past and future experience" (Ref. 17).

; became clear to the authors, as the series of experiments reported above was ing carried out, that the model described served admirably the function of rganizing past and future experience. The block-diagram representation of the del of the Expected-Value Observer, after Quastler (Ref. 17), is shown in gure 17.

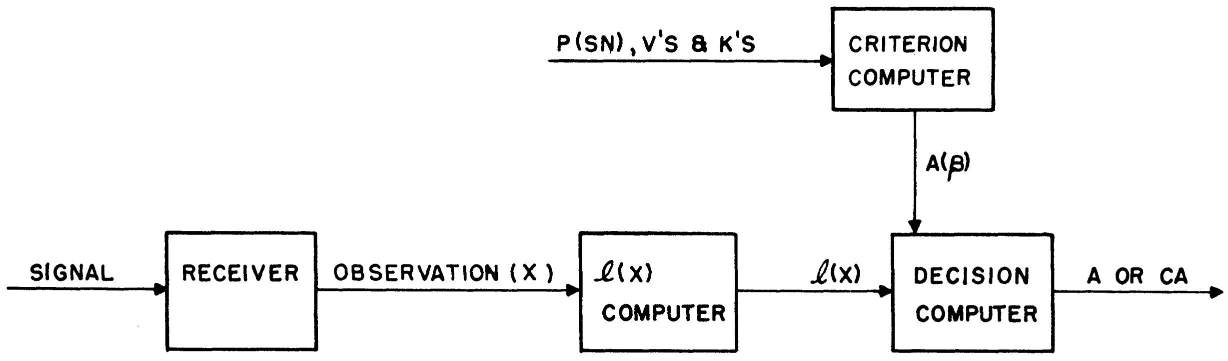

FIG. 17. BLOCK DIAGRAM OF THE EXPECTED-VALUE OBSERVER

The same diagram applies to the Neyman-Pearson Observer when the input to the criterion computer (P(SN), V's, and K's) is replaced by $P_N(A(\beta)) = k$. In the case of the A posteriori Observer, there is, of course , no criterion com puter, and the decision computer is replaced by a computer which makes the trans formation from $\ell(x)$ to $P_x(SN)$, having $\ell(x)$ and P(SN) as input and $P_x(SN)$ as output. To represent the forced-choice situation, the criterion computer is eliminated, and the decision computer merely selects the greatest of the input likeli hood ratios as its output.

Each of the five experiments reported above demonstrates that the huma observer operates with information in the form of likelihood ratio, and tends toward optimum behavior. These experiments provide convincing evidence of the applicability of the proposed model to the problems of visual detection. The re vance of the model is also supported by its congruence with auditory data (Ref. 26). There is, however, still another imposing reason for treating sensory prob lems in terms of the model, namely, that the model provides a unification of the data obtained with forced-choice and yes-no methods of response. If psychophysi data is collected and analyzed in accordance with this model, performance in the forced-choice situation can be predicted from yes-no data, and vice versa. Evidence presented elsewhere (Refs. 22, 25) shows that the estimates of d' from the two response procedures are highly consistent.

It is interesting to note, parenthetically, that the present account i not among the first to model psychophysical theory after developments in the the of statistical decision. Fechner, the founder of psychophysics, was appreciably influenced by Benoulli who first suggested computing expectations in terms of satisfaction units. As Boring (Ref. 5) relates the story, Bernoulli's interest in games of chance led him to formulate the concept of "mental fortune"; a chang n mental fortune he believed to vary with the ratio that the change in physical ortune has to the total fortune. This mathematical relation between mental and hysical terms was the relation that Fechner sought to establish with his psycho-hysics.

.1 The Possibility of a Threshold Theory

It will be well to consider again, after the presentation of the data, he possible validity of a theory incorporating the threshold concept. It is lear that the assumption of a threshold as previously conceived, one that is very arely exceeded by noise alone, is not reasonable. At this time, however, the xistence of a threshold well down into the noise distribution cannot be dis-redited. To take a single example, the data, representing values of $P_N(A)$ above point between .10 and .30, that are fitted to the R.O.C. curves in Figures 7 hrough 10 may be adequately fitted by a straight line through $P_{SN}(A) = P_N(A) =$ .0; these data, then, do not preclude postulation of a threshold in the neighbor-ood of the mean of the noise distribution. Further analyses of the second-choice ata, in relation to thresholds at various levels in the noise distribution, are eing undertaken and will be reported elsewhere (Ref. 20).

It should be noted, however, that determination of the level of the oise distribution at which a threshold may possibly exist is neither critical or useful. A threshold at such a level is not a readily workable concept. The rimary virtue of a threshold that is rarely exceeded by noise alone is that it acilitates mathematical treatment of the data, chiefly by being consistent with he usual correction for chance. It has been demonstrated above, however, that athematical manipulations must not involve assumptions incompatible with a noise istribution much of which exceeds a threshold, if a threshold is to be postu-ated; at this point then, adhering to a threshold concept complicates the mathematics. As a matter of fact, a threshold at a level well within the range of the noise distribution, is, for all practical purposes, not measurable. The forced-choice methodology is a case in point; the observer conveys less informa-tion than he is capable of conveying if only a first choice is required. That t second choice contains a significant amount of information has been demonstrated it is not unlikely that the third choice will convey information. Thus it is very difficult for an experimenter to determine when enough information has beer extracted from forced choices to yield a sufficiently low estimate of the thresh old. In addition, the existence of such a threshold is of no consequence to the application of the theory proposed here; to take an example, "yes-no" data re-sulting from a suprathreshold operating level depends on the operating level but is completely independent of the threshold value.

6.2 Some Implications of the Proposed Model for Psychological Theory

The applicability of the model provided by the theory of signal detect bility stands in opposition to the view that the so-called sensory phenomena are independent of control by general psychological variables, a view that is consis tent with the concept of the threshold. The theory built upon this model takes into account the influence on detection behavior of "non-sensory" central deter-minants.

In conventional theory, the decision concerning the existence of a sig nal is assumed to depend entirely upon the threshold being exceeded, and the threshold level is assumed to be independent of control by other variables that might influence the attitude or set of the observer. A successful application o the theory of signal detectability to psychophysical data leads to the replace-ment of the threshold concept by a concept of criterion range of acceptance; the way in which the control of this range is conceptualized acknowledges the rele-vance of variables that influence set.

In the framework of a statistical theory, it is necessary to make an as- mption which permits the definition of behavior, or the prescription of predic- d behavior. The concept of the threshold has served this purpose. The assump- on that the threshold is relatively invariant permits making predictions which n be tested experimentally. In the theory presented here, where the position of e cutoff between acceptance and rejection of the existence of a signal (the oper- ing level) is assumed to be under the control of the observer, it is necessary define the method of control exerted by the observer on the operating level. e assumption that the observer tends toward the optimum behavior provides the cessary definition. Thus, in the typical yes-no experiment, the observer may be nceived of as regulating the operating level in terms of a priori probabilities d the values and costs associated with the various types of correct and incorrect swers, in such a way as to maximize the total expected gain. (It should, per- ps, be pointed out that these values and costs do exist in the typical yes-no periment, whether or not they are explicitly translated into numerical values). the forced-choice experiment, optimum behavior requires that no operating level assumed, and that the interval with the greatest associated observation be se- cted. Although there has been a general discontent among psychologists with the ncept of a fixed operating level or threshold, there has not been advanced pre- ously a way of defining the mode of control exerted by the individual over a riable operating level. It is the chief virtue of the model provided by the eory of signal detectability that it specifies operating level variability.

3 Some Implications for Practice

The results of these experiments, of the Expected-Value experiments in rticular, give an account for Blackwell's (Ref. 3) finding that forced-choice ta are more reliable than yes-no data. In the yes-no experimental setting, when e usual caution against making false-alarm responses is included, the operating level may vary over a wide range, with the variation having no direct reflectio in the data. False-alarm rates of .01, .001, and .0001, for example, are not d criminable in an experimentally feasible number of observations. This fact may also account for the failure to detect previously the operation of a mechanism with a variable operating level.

The results presented above account also for the often-reported[1] find that the forced-choice procedure yields lower calculated thresholds than does t yes-no procedure. Conventionally, observers are cautioned against making false-alarm responses in yes-no experiments. It is very likely that the stigma attac to "hallucinating" serves to depress $P_N(A)$ in those test settings not including explicit warning to avoid false alarms. The inverse relationship between $P_N(A)$ and calculated threshold predicted by the theory of signal detectability is discussed in Reference 25. Data substantiating this predicted relationship includ the correlations reported above and in Reference 25 between $P_N(A)$ and calculate threshold, and that shown in Figure 5 above. Plots of $P_{SN}(A)$ vs. $P_N(A)$ which d not fit a straight line through $P_{SN}(A) = P_N(A) = 1.00$(such as in Figures 7-10) provide another way of saying the same thing. The relationship between d' and $P_N(A)$ at "threshold" is described in more detail in the forthcoming paper on methods (Ref. 20).

In spite of the view that "sensory" phenomena are peripherally determined, the necessity of assuming a constancy of "set" across people and over ti in psychophysical experiments has been generally accepted. Effecting this constancy is a primary function, at least in the yes-no experiments, of the verbal instructions. The theory of signal detectability and the experiments reported have advanced a check on the assumption of constancy of set with respect to the most important aspect of set in yes-no experiments, namely, the location of the

1. See, for example, Blackwell (Ref. 3) and Goldiamond (Ref. 9). Miller (Ref. reviews earlier literature on "subliminal perception" that is relevant to t point.

toff between acceptance and rejection of the existence of a signal. This re- arch has demonstrated that a measurable estimate of $P_N(A)$ can be, and should be, oduced in yes-no experiments. By the same token, theoretical support has been ovided for the advisability of using the forced-choice procedure whenever pos- ble. With this technique, the observer is not faced with the problem of locat- g and maintaining the stability of an operating level, and thus a source of riance in the data is removed. Since a method for unifying forced-choice and s-no data has been provided (Refs. 22, 25), forced-choice techniques can be used en when an estimate of performance in the yes-no situation is desired. The prac- cal import of this is that data pertinent to yes-no situations, that is more re- able, can be obtained with greater economy. This topic is treated in more de- il in the paper on methods (Ref. 20).

That the condition of the organism affects perception has been demon- rated previous to the studies reported here. The theory of Ames and his co- rkers (Ref. 1), the "new look" theory of Bruner and Postman (Ref. 6), and the eories of Hebb (Ref. 11), Brunswik (Ref. 7), and Woodworth (Ref. 29) have taken to account the effect of "non-sensory" central determinants. The present ac- unt, however, goes beyond the stage of demonstration; it provides operational d theoretical specifications of the "conditions". The present account, then, tisfies the desiderata discussed by Graham (Ref. 10): the conditions of the ganism are specified at other than the conversational level; the conditions are fined in the theory and anchored to operations at both ends. Since the variables ich determine these conditions are expressed quantitatively, the quantification an important group of instruction stimuli has been achieved. Said another way, is research proceeds a step in the direction of specifying conditions of the ganism, due to instructions, as parameters of observable stimulus-response lations.

REFERENCES

1. Ames, A., Jr., "Visual Perception and the Rotating Trapezoidal Window," Psycl Monogr., 65, No. 7, 1951.

2. Birdsall, T. G., "An Application of Game Theory to Signal Detectability," Te‹ Rpt. No. 20, Electronic Defense Group, University of Michigan, Ann Arbor, Mi‹ gan, 1953.

3. Blackwell, H. R., Psychophysical Thresholds: Experimental Studies of Method‹ Measurement. Ann Arbor, University of Michigan Press, 1953 (Eng. Res. Inst. Bulletin No. 36).

4. Blackwell, H. R., Pritchard, B. S., and Ohmart, T. G., "Automatic Apparatus i Stimulus Presentation and Recording in Visual Threshold Experiments," J. Opt. Amer., 1954, 44, 322-326.

5. Boring, E. G., A History of Experimental Psychology. New York: Appleton-Ce Crafts, 1950 (2nd edition).

6. Bruner, J. A., in Blake, R. R. and Ramsey, G. V. (Eds.) Perception: An Appr‹ to Personality. New York: Ronald Press, 1951.

7. Brunswik, E., Systematic and Representative Design of Psychological Experime Berkeley: University of California Press, 1949.

8. Fox, W. C., "Signal Detectability: A Unified Description of Statistical Metl Employing Fixed and Sequential Observation Processes," Tech. Rpt. No. 19, El‹ tronic Defense Group, University of Michigan, Ann Arbor, 1953.

9. Goldiamond, I., "The Relation of Subliminal Perception to Forced Choice and Psychophysical Judgements, Simultaneously Obtained," Amer. Psychol., 9, No. { 1954, (abstract).

10. Graham, C. H. "Visual Perception," in S. S. Stevens (ed.) Handbook of Experi Psychology. New York: Wiley, 1951.

11. Hebb, D. O., The Organization of Behavior. New York: Wiley, 1949.

12. Miller, J. G., Unconsciousness. New York: Wiley, 1942.

13. Neyman, J., "Basic Ideas and Some Recent Results of the Theory of Testing St‹ tistical Hypotheses," J. Roy. Stat. Soc., CV, Part IV, 1942.

14. Neyman, J., and Pearson, E. S., "On the Problem of the Most Efficient Tests ‹ Statistical Hypotheses," Phil. Trans. Roy. Soc., 231, Series A, 1933.

15. Peterson, W. W., and Birdsall, T. G., "The Theory of Signal Detectability,"] Rpt. No. 13, Electronic Defense Group University of Michigan, Ann Arbor, 195:

Peterson, W. W., Birdsall, T. G. and Fox, W. C., "The Theory of Signal Detectability," in the Transactions of the 1954 Symposium on Information Theory held at Massachusetts Institute of Technology, Cambridge, Massachusetts, September 15-17, 1954. Published by the Institute of Radio Engineers, 1 E. 79th St., New York 21, N. Y.

Quastler, H., Notes on the Concluding Discussion of the "Symposium on Methodology in the Estimation of Information Flow," held at the University of Illinois, Urbana, July 5-9, 1954 (mimeo).

Shannon, C. E., "The Mathematical Theory of Communication," Bell System Technical Journal, 1948.

Swets, J. A., "An Experimental Comparison of Two Theories of Visual Detection," unpublished doctoral dissertation, University of Michigan, 1954.

Tanner, W. P., Jr., Birdsall, T. G., and Swets, J. A., "Psychophysical Methods," Tech. Rpt. No. 42, Electronic Defense Group, University of Michigan, Ann Arbor, 1955 (in preparation).

Tanner, W. P., Jr., and Norman, R. Z., "The Human Use of Information. II. Signal Detection for the Case of an Unknown Signal Parameter," in the Transactions of the 1954 Symposium on Information Theory held at Massachusetts Institute of Technology, Cambridge, Massachusetts, September 15-17, 1954. Published by the Institute of Radio Engineers, 1 E. 79th St., New York 21, N. Y.

Tanner, W. P., Jr., and Swets, J. A., "A New Theory of Visual Detection," Tech. Rpt. No. 18, Electronic Defense Group, University of Michigan, Ann Arbor, 1953.

Tanner, W. P., Jr., and Swets, J. A., "A Psychophysical Application of the Theory of Signal Detectability," Electronic Defense Group, Report 1970-5-S reprinted from minutes of Armed Forces NRC Vision Committee Meetings held at Fort Knox, Kentucky, November, 1953.

Tanner, W. P., Jr., and Swets, J. A., "The Human Use of Information, I. Signal Detection for the Case of the Signal-Known Exactly," in the Transactions of the 1954 Symposium on Information Theory, held at Massachusetts Institute of Technology, Cambridge, Massachusetts, September, 15-17, 1954. Published by the Institute of Radio Engineers, 1 E. 79th St., New York 21, N. Y.

Tanner, W. P., Jr., and Swets, J. A., "A Decision-Making Theory of Visual Detection," Psychological Review, Vol. 61, No. 6, 1954.

Tanner, W. P., Jr., Swets, J. A., and Green, D. M., "The General Properties of the Hearing Mechanism," Tech. Rpt. No. 30, Electronic Defense Group, University of Michigan, Ann Arbor, 1955, (in preparation).

Wald, A., "A Contribution to the Theory of Statistical Estimation," Amer. Math. Stat., 10, 1939.

Wald, A., Statistical Decision Functions. New York: Wiley, 1950.

Woodworth, R. S. Psychology. New York: Holt, 1945 (4th Edition).

DISTRIBUTION LIST

1 copy	Director, Electronic Research Laboratory Stanford University Stanford, California Attn: Dean Fred Terman
1 copy	Commanding General Army Electronic Proving Ground Fort Huachuca, Arizona Attn: Director, Electronic Warfare Department
1 copy	Chief, Engineering and Technical Division Office of the Chief Signal Officer Department of the Army Washington 25, D. C. Attn: SIGJM
1 copy	Chief, Plans and Operations Division Office of the Chief Signal Officer Washington 25, D. C. Attn: SIGOP-5
1 copy	Countermeasures Laboratory Gilfillan Brothers, Inc. 1815 Venice Blvd. Los Angeles 6, California
1 copy	Commanding Officer White Sands Signal Corps Agency White Sands Proving Ground Las Cruces, New Mexico Attn: SIGWS-CM
1 copy	Commanding Officer Signal Corps Electronics Research Unit 9560th TSU Mountain View, California
75 copies	Transportation Officer, SCEL Evans Signal Laboratory Building No. 42, Belmar, New Jersey FOR - SCEL Accountable Officer Inspect at Destination File No. 22824-PH-54-91(1701)

1 copy	H. W. Welch, Jr. Engineering Research Institute University of Michigan Ann Arbor, Michigan
1 copy	Document Room Willow Run Research Center University of Michigan Willow Run, Michigan
11 copies	Electronic Defense Group Project File University of Michigan Ann Arbor, Michigan
1 copy	Engineering Research Institute Project File University of Michigan Ann Arbor, Michigan