2

07814-3-F

Report Title

INTER-OBSERVER AGREEMENT AND MODELS OF MONAURAL

AUDITORY PROCESSING IN DETECTION TASKS

Author

Gordon Wells Wilcox

The University of Michigan

October, 1968

Project Title

Optimal Learning in Detection Situations

Principle Investigator

Wilson P. Tanner, Jr.

Michigan Mathematical Psychology Program : : MMPP 68 - 4

The Sensory Intelligence Laboratory

ORA Project 07814

February 1969

Through

Office of Research Administration

Ann Arbor

This work was supported by Grant NGR-23-005-159 of the National Aeronautics and Space Administration. This report was also submitted as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy in The University of Michigan.

ABSTRACT

INTER-OBSERVER AGREEMENT AND MODELS

OF MONAURAL AUDITORY PROCESSING

IN DETECTION TASKS

by

Gordon Wells Wilcox

Previous studies have shown that the output of an electronic energy detector is correlated with the responses of human observers in monaural auditory detection tasks. However, an experiment performed by Ahumada showed that when the signal (pure tone) is present in the observation interval the correlation is greatest if the filter in the energy detector has a considerably narrower bandwidth than is necessary for maximum correlation when only noise is presented. Since this result is inconsistent with a simple energy detection model of auditory processing, Ahumada proposed a filter-bank model to account for his findings.

The present study presents a quantitative development of the linear-uncertain model which is a generalization of the energy model. The form of the linear-uncertain model is derived from the assumption that the observer is uncertain regarding the exact specification of the signal he is trying to detect. The decision variable of this model is the weighted sum of a linear function and a quadratic function of the input waveform. The relative weight of each component is determined by the observer's level of uncertainty regarding the signal. It is shown that the linear-uncertain model includes as special cases the linear, energy, and envelope models for auditory processing and like the filter-bank model can give an explanation of Ahumada's findings.

Predictions of the level of correlation between observers derived from the linear-uncertain model and the filter-bank model are compared for several experimental conditions. In the experiment the decisions of human observers are compared with the outputs of two electronic devices. The first device is an analog multiplier which computes the cross-correlation between the signal and the noise waveform sample on each trial. The second device is an energy detector which computes the energy of the noise waveform sample during the presentation interval in a narrow frequency band centered at the signal frequency.

The linear-uncertain model predicts that the correlation between the human observers and the cross-correlation is not zero and should increase when a continuous sinusoid is added to the background noise. Neither of these predictions is verified. Since the energy detector receives only noise at its input in this experiment, the model also predicts that the correlation between the observers and the energy detector should be less on trials when signal is present than when it is not present. The results show the correlations to be weakly significant in the opposite direction. It is concluded that the linear-uncertain model and its special cases, the linear, energy, and envelope models, represent an inadequate approximation to the actual form of human monaural auditory processing in detection tasks.

Predictions from the filter-bank model agree with the above results, but cannot account for an observed decrease in inter-observer correlations when the signal presentation interval is shorted. A modification of the filter-bank model is suggested to account for this discrepancy.

A final result remains unexplained by any of the models considered. A decrease in inter-observer correlations is found when a continuous sinusoid is added to the background noise for either of two signal durations. It is emphasized that this unexpected finding implies that there is a serious deficiency in current models of monaural auditory processing in detection tasks.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I

DIRECT COMPARISONS BETWEEN OBSERVERS

1.1 Introduction

Many auditory detection experiments have investigated changes in observer performance as the result of variation of the physical parameters of the input waveform or characteristics of the observer's task. Measures of performance which have been used include indices of the quality of the observer's decisions and studies of the shape of psychometric functions and receiver operating characteristic curves. The physical parameters that have been varied include frequency, phase, amplitude, and duration of a signal, spectral characteristics of background noise, signal-to-noise ratio, and addition of pulsed carriers and continuous waves to the noise. Task characteristics which have been studied include single, double, etc., and random presentation intervals, pulsed versus continuous background noise, binary and multiple response categories, payoff structures, and probability of signal occurence (cf. Swets, 1964; Green and Swets, 1966).

The performance measures in these studies I shall call *indirect comparisons* between observers. Probability of a correct response and the area under the ROC curve compare the observer's performance with the specification of experimenter defined hypotheses. d' compares observer performance with an optimum observer for the same task. Studies of the shape of the psychometric function, ROC curve, or "iso-bias" curve all compare observer performance with a model-specified in a similar task (a Glossary of Symbols and Terms appears in Appendix III). These comparisons are indirect because they do not compare observers on a trial-by-trial basis.

Direct comparisons between observers do compare inter- or intra-observer performance on individual trials. Green (1964b) has termed attempts to *predict* single trial performance "molecular psychophysics" as contrasted with molar psychophysics. In this sense, direct comparisons can be either molar or molecular depending upon the use to be made of the performance measures.

The purpose of using different kinds of comparisons is the same:

> ... the ultimate aim of detection theory is to discover the exact form of the distributions [of sensory events] using (1) the physical parameters of the signal and noise and (2) the character of the sensory detector...The major obstacle of this endeavor is...our ignorance about the nature of the sensory detector (Green and Swets, 1966, pp. 53-54).

1.2 Comparison of Observer Performance with an Electronic Detector

The first direct comparisons between human observers in an auditory detection task and that of an electronic device were made by Sherwin, Kodman, Kovaly, Prothe, & Melrose (1956). They recorded on magnetic tape 50 samples of a 1 kHz burst of several durations (0.03, 0.10, 0.30, and 1.0 seconds) in four different listening conditions. The tone bursts occurred at randomly spaced intervals in a background of noise (0-4 kHz). The tapes were played back through earphones to observers and simultaneously through a 60 Hz wide passive filter, centered at 1 kHz. The output of the filter served as the input to a square-law detector and exponentially decaying integrator. The time constant of the integrator was set at one-half the tone duration for approximately optimum detection (sic). The final output of the electronic detector was recorded on a pen recorder which marked the times that signal was present.

The amplitude of the signal was adjusted so that the observers had a hit rate ("correct detections") of approximately 60% at each condition of signal duration. A criterion for the output of the electronic device could be set by the experimenters so that it also had a 60% hit rate, thus matching the performance of the device with the human observers during the presence of a signal. They found that when this match was made, the false-alarm rates ("incorrect detections") for all observers were lower at each signal duration than for the electronic device, with the smallest difference at a duration of 0.30 seconds. It was also for this duration that the sample distributions of the output of the device, conditional upon observer hits and misses, respectively, were most highly separated. That is, at 0.30 second duration the hits and misses of the electronic device had the highest correlation with the hits and misses of the observers. Unfortunately, a similar statement could not be made for intervals during which the signal was not present. In fact there appeared to be no association between the observers' responses and the output of the device during noise intervals, which is partly in evidence from the fact that the observers maintained lower false alarm rates than the device. The authors suggest that using a filter about 30% narrower that the one they employed would have led to approximately the same performance for the device and the observers. This in itself, of course, would not demonstrate a closer association between an observer's decision variable and the average power statistic computed by the device. It is possible that improving the performance of the device would not improve the correlation during noise intervals.

1.3 Association of Observer Performance with an Energy Detector

The next experiment involving direct comparisons was reported in Watson's Ph.D. thesis (1962). The experimental method made several improvements over the procedure used by Sherwin, *et al.* Discrete observation intervals 0.25 seconds long containing pulsed noise waveform samples were used. A signal, again a 1 kHz tone, was added to the noise on approximately one-half of the observation intervals. The observers could make one of four responses on each trial indicating their confidence that signal had or had not been presented. There was no immediate feedback as to whether the response was correct or not. The input to the observers was simultaneously recorded on magnetic tape for further analysis. This analysis was made by passing the pulsed waveforms through a filter approximately 100 Hz wide centered at 1 kHz and then measuring the number of times voltage peaks from each burst exceeded each of eleven different levels. A "voltage contour" was defined as the percentage of times voltage peaks exceeded a given voltage level plotted versus the voltage level. The area under the voltage contour was interpreted as an approximate index of the energy in a 100-cycle band centered on the signal frequency. Voltage contours could then be plotted conditional upon the observers' responses and upon whether or not signal was presented, averaging across all bursts in one of these eight conditions. The results showed that the area under a voltage contour was monotone increasing with the rating response on both trials containing signal and those which did not contain signal. No attempt was made to predict trial by trial responses of the observers based on the energy statistic. However, two observations about the data were made. First, consistent with the assumptions of the theory of signal detectability an observer's decision variable could be ordered on the basis of a physical parameter of the stimulus. Second, the results showed that there was association between the responses of an observer and the physical parameter on both signal and noise-alone trials, contrary to the findings of Sherwin, *et al.*

Watson had stated in his introduction "... the energy within the critical band has been shown in theory and by experiment to be a primary parameter of the auditory stimulus [in masking experiments]". Nevertheless, it is apparent from Watson's data that the observers' responses were not determined by the area under the voltage contour alone, since every response occured at every voltage level. It could be postulated, as did Sherwin, *et al.*, that the relevant physical parameter had been determined, but that the observers also had fluctuating criteria for making their responses. Equally well, in terms of these two experiments, it might be hypothesized that association between the measured physical statistic and the responses was due to the fact that both the statistic and the responses were associated with a physical parameter *not* measured in the experiment. Such a parameter might account for the trial by trial responses of the individual without assuming a fluctuating criterion.

1.4 Intra-Observer Consistency

Green (1964b) fully realized this latter possibility and set out to determine the level of inconsistency of an observer's response *regardless of the relevant physical parameters* of the stimulus. Pairs of noise samples without signal in either interval were recorded and interspersed with pairs containing signal in one of the intervals. In this 2-interval forced-choice task, percentage self-agreement scores were determined on the no-signal pairs by having each of these pairs presented twice in the experimental session.

In Green's first experiment the "signal" was an increment of power in the wide-band background noise. Three signal-to-noise ratios were used. The results indicated that percentage agreement of responses to identical pairs of no-signal samples was approximately 65% for each observer regardless of the percentage of correct detection responses which averaged 60%, 74%, and 91% for the three S/N ratios, respectively. In a second experiment the signal was a 0.10 second gated sinusoid at a frequency of 250, 500, 1000, or 2000 Hz in four separate conditions. The average percentage self-agreement scores across observers and tapes was approximately 70% at all frequencies except for the 1000 Hz condition, for which the average was 78%. Several other experimental conditions were studied with the same general result: self-agreement scores differed little from 70% although there was a small but consistent trend for tapes using sinusoidal signals to generate somewhat higher scores than tapes for which the signal was an increment in the background noise. Green conjectured, but could not measure in his experiments, the possibility that the inconsistency could be attributed to several causes. The first would be response bias--either pure, i.e., preference for the first or second interval, or sequential, i.e., a bias depending upon the previous response or previous feedback. Another possible cause is "internal noise" which has been variously defined as a fluctuation in the observer's criterion or noise in his sensory apparatus. Green made several calculations which indicated that response bias effects should be small, and therefore proceeded to estimate the level of internal noise which would lead to the observed percentage agreement scores. His calculations suggested that, as a first approximation and minimum estimate, the ratio of external to internal noise is about 1.0. No thesis was advanced to explain the difference in agreement scores as a function of the kind of signal presented. Green concluded that this 70% consistency, since it appears independent of the sensory task, represents an upper bound on any attempt to predict trial-by-trial human detection responses.

Apparently following a suggestion made by Green in his article, Pfafflin and Mathews (1966) studied the consistency of observer responses to computer generated no-signal noise pairs, some of which had an identical noise sample as each member of the pair. In addition, they added a 312.5 Hz tone to one or the other member of some of the pairs. The use of computer-generated and-controlled stimuli provided relatively accurate reproduction of waveform samples and permitted the presentation of stimulus pairs in a large number of different orders, with the hope of greatly reducing any possible effect of response biases. A spectral analysis was made on each of the 12 samples of noise used in the experiments. The spectral analysis determined the relative energy in frequency bands approximately 100 Hz wide as a function of the center frequency. It was thus possible to determine the relative energy difference in a frequency band 100 Hz wide centered at the signal frequency of 312.5 Hz between the members of each pair of waveforms.

The probability of correct detection tended to increase, with considerable scatter, for three observers as a function of the relative energy difference for pairs containing a signal, whether or not the individual noise samples in the pair were identical. However, some inversions (a response preference for the member of the pair with the lower relative energy level) occurred, most often for pairs with low relative energy difference. The authors could not find inter-observer agreement on inversions, nor could they find relevant physical parameters that might account for inversions of a single observer. Another experiment was performed using the same noise samples, where this time the observers were asked to judge if either interval contained a signal. Since the preference probability tended to increase with the probability that a pair was judged to contain a signal for signalless different-noise pairs, it was concluded that at least for these different-noise pairs, the observer must have been trying to detect a signal in the previous experiment. Corresponding results were not conclusive for signalless identical-noise pairs. This could have been due to the procedure under which the observers knew that some pairs contained no signal. The authors concluded that the energy increment produced by the signal, or some quantity closely related to it, is the chief physical parameter relevant to detection behavior. They thus agreed with Watson on the nature of the sensory processing.

Pfafflin and Mathews' major contribution to Green's original effort on intra-observer consistency would seem to be the demonstration that the percent agreement for signalless pairs is dependent on the particular noise pair and perhaps, although this is not demonstrated conclusively, dependent upon some physical parameter of the noise samples. Estimating by eye, it would appear from Pfafflin and Mathews' Figure 4 that the average percent-agreement for signalless different-noise pairs is about 60-70% for each observer, in agreement with Green's results.

1.5 Correlation of Observer Decisions with a Variable-Bandwidth Energy Detector

The most recent study involving direct comparisons is reported in Ahumada's Ph.D. thesis (1967). In a context of attempting to measure critical bandwidths directly, Ahumada had observers respond ("Yes" or "No") to single 100 msec noise bursts which sometimes contained a superimposed tone, as in Watson's experiment. An electronic detector with specifiable bandwidth was simulated on a digital computer as in the Phafflin and Mathews study. The observers' average responses from 5 replications to the same stimulus were correlated (Spearman's rank correlation coefficient corrected for ties) with the outputs of the simulated energy detectors. The correlation was computed separately for signal-plus-noise and noise-alone trials to determine the bandwidth for which there was maximum correlation. The bandwidth associated with the maximum correlation should provide a fairly direct estimate of the critical bandwidth. A surprising result was found. A filter with a 10 or 20 Hz bandwidth had maximum correlation with the observers' responses for signal-plus-noise trials. But on noise trials a wider filter with 100 to 200 Hz bandwidth had maximum correlation. This result, of course, can not be predicted from the simple energy-detection model. Such a model predicts that the same width filter should correlate best with responses to signal-plus-noise and noise-alone stimuli. Ahumada's finding deals a serious blow to the conclusion that the energy in a narrow band is *the* primary physical parameter. The most that can be concluded is that the energy in a given band is associated with the physical determinants of detection behavior.

Ahumada suggested a "filter-bank" model to account for his data. According to this model the observer can monitor the output of a number of narrow-band filters with a total bandwidth of about 150 to 200 Hz. The observer makes the detection response when any of the individual outputs exceeds some critical value. On signal trials the output of the narrow filter centered on the signal frequency almost always has maximum output, whereas on noise-alone trials any member of the bank has equal likelihood of exceeding the criterion.

Ahumada's filter bank model is not contradicted by the data of Sherwin, *et al.*, Watson, Green, or Pfafflin and Mathews. In the case of Sherwin, *et al.*, who used a 60 Hz wide filter, the lack of association of its output with false-alarms intervals could be explained by the fact that the filter was too narrow. Incidentally, this would predict that narrowing the filter as the authors suggest should decrease the association with observer responses in false-alarm intervals. In Watson's case the filter was 100 Hz wide, or in the middle range, so that one would expect about equally poor (or good) association with responses on signal and noise trials, as is consistent with an examination of his data. Also, the same conclusion appears verified from Ahumada's data. With respect to Green's data, a variance in the observer's criterion would still decrease the percentage self-agreement. In fact, if the observer is limited to monitoring the output of a 200 Hz filter-bank, then a fluctuation in the center frequency of the bank would account for the increased decrement in percent agreement for signals which are increments in the power of wide-band noise, as Green found. Such a fluctuation would not cause a further decrement when the signal is a sinusoid. Finally, since the maximum output from a single narrow-band filter should be only poorly correlated with the total energy in a 100 Hz band, Ahumada's model can give a post-hoc explanation for the fact that Pfafflin and Mathews found inversions in preference for signalless different-noise pairs with low relative energy difference.

It would appear that the filter-bank model with its associated decision rule is adequate in a qualitative fashion to account for the data from direct comparisons between human observers and electronic devices currently available.

1.6 Objectives of the Present Work

The filter-bank model has a certain appeal by its analogy to the physiology of auditory system. (This analogy is briefly discussed by Ahumada. However, the details do not concern us here.) On the other hand, the simple energy detection model has had an appeal in auditory psychophysics at least partially because it represents an optimum mode of processing under certain conditions (c.f. Green, 1960; Pfafflin and Mathews, 1962; Swets, 1966, Ch. 8).

In this thesis I shall develop (in Chapter III) the linear-uncertain model which dictates an optimum mode of processing for an observer with uncertainty regarding the specification of the signal. The linear-uncertain model turns out to be a slight generalization of the optimum observer with a noisy stored reference signal derived by Birdsall (1960). The energy model and several other processing models are found to be special cases of the linear-uncertain model. Moreover, the linear-uncertain model, unlike the energy model, is not completely frustrated by Ahumada's findings (Chapter IV). An experimental attempt is made to discriminate between the filter-bank model and the linear-uncertain model based on the predictions that these models make for the degree of concordance between observers in several tasks (Chapter V).

CHAPTER II

REPRESENTATION THEORY

2.1 *Introduction*

Auditory detection and recognition tasks in the laboratory often involve the presentation of complex waveforms to human observers. It is assumed that the observer makes a judgmental response based upon certain operations which he performs on the input to his ears. Models which attempt to describe the observer's judgmental performance must also describe the operations performed on the input which could give rise to the observer's performance. There is a need, then, to obtain a description of the impinging stimulation or input process itself.

A description of or representation theory for auditory waveforms can be developed at many different mathematical levels, depending on the rigor desired and the degree of error that can be tolerated. Fortunately, in psychoacoustical tasks a relatively simple representation theory is often adequate. In particular, acoustically presented waveforms are usually of sufficient bandwidth and duration that a finite Fourier series approximation to a sample waveform is sufficiently accurate for many purposes. Green and Swets (1966) hold this view and have presented such a model for the representation of sample waveforms generated by a real-time stochastic process.

This chapter presents a representation model for waveforms in the frequency domain which is identical to the model used by Green and Swets. However, the temporal representation model presented by Green and Swets is inconsistent. This fact is discussed in some detail because we wish to make extensive use of the geometrical properties of representation models.

Finally, this chapter establishes a notation which will be exploited throughout the remainder of this work.

2.2 Representation of Waveforms in the Frequency Domain

Following Green and Swets (1966) we assume that a finite Fourier series approximation to a waveform x(t) with no d.c. component in the interval $0 \leq t \leq T$ is sufficiently accurate for our model construction. This approximation to x(t) is given by

$$\hat{x}(t) = \sum_{i=1}^{WT} [a_i c_i(t) + b_i s_i(t)], \tag{2.1}$$

where W is highest frequency component in the series,

$$c_i(t) = \sqrt{\frac{2}{T}} \cos \frac{2\pi i t}{T} \tag{2.2a}$$

$$s_i(t) = \sqrt{\frac{2}{T}} \sin \frac{2\pi i t}{T}\ , \tag{2.2b}$$

and

$$a_i = \sqrt{\frac{2}{T}} \int_0^T x(t) c_i(t)\ dt \tag{2.3a}$$

$$b_i = \sqrt{\frac{2}{T}} \int_0^T x(t) s_i(t)\ dt, \tag{2.3b}$$

for $i = 1, 2, \ldots, WT$.

The functions $c_i(t)$ and $s_i(t)$ form an orthonormal set over the interval $[0,T]$, i.e.,

$$\int_0^T c_i^2(t)\ dt = \int_0^T s_i^2(t)\ dt = 1 \tag{2.4a}$$

and

$$\int_0^T c_i(t)s_i(t)\ dt = \int_0^T c_i(t)c_j(t)\ dt = \int_0^T s_i(t)s_j(t)\ dt = 0 \tag{2.4b}$$

for $i \neq j$ and $i,j = 1, 2, \ldots, WT$, as is well known and easily verified. A less often noticed fact is that the functions $2c_i(t)$ and $2s_i(t)$ form an orthonormal set over the interval $[0,T/2]$. We shall use both of these orthonormal sets of functions below.

The form of Equation 2.1 indicates that x(t) may be identified with the (column) vector

$$x_f = [a_1, a_2, \ldots, a_{WT}, b_1, b_2, \ldots, b_{WT}]' \tag{2.5}$$

in a 2WT dimensional vector space (the prime " ' " denotes the transpose of a matrix). The set $\{c_1(t), \ldots, c_{WT}(t), s_1(t), \ldots, s_{WT}(t)\}$ is an orthonormal set of basis vectors for the space. This 2WT dimensional vector space, denoted F, will be called the *frequency representation space*.

We follow Green and Swets (1966) for our model of noise waveforms. Bandlimited Gaussian *noise* is a real time-parameter stochastic process $\{n(t,\tilde{n})\}$ where the vector $\tilde{n}$ of random variables has a multivariate normal distribution $N(\mu_n, \Sigma_n)$ with mean vector μ_n and dispersion (variance-covariance matrix Σ_n. For a particular sample $n = [a_1, \ldots, a_{WT}, b_1, \ldots, b_{WT}]'$ of $\tilde{n}$ the corresponding noise waveform sample $n(t) = n(t,n)$ is approximated in the interval $0 < t < T$ by the series

$$\hat{n}(t) = \sum_{i=1}^{WT} [a_i c_i(t) + b_i s_i(t)] \tag{2.6}$$

where W is the bandwidth of the noise process.

The noise is said to be *white* when

$$\not{\Sigma}_n = (N_0/2)I \tag{2.7}$$

where I is an identity matrix, $N_0 = N/W$ is the *noise power density* and N is the average noise power. Also, note that $\det \not{\Sigma}_n = (N_0/2)^{2WT}$ and that $\not{\Sigma}_n^{-1} = (2/N_0)I$.

The multivariant normal density function of $\tilde{n}$ for a bandlimited white Gaussian noise process may be written explicitly as

$$f(n) = (2\pi)^{-WT}(\det \not{\Sigma}_n)^{-1/2} e^{-1/2[n'\not{\Sigma}_n^{-1}n]}$$

$$= (\pi N_0)^{-WT} e^{-[n'n/N_0]} . \tag{2.8}$$

2.3 Representation of Waveforms in the Temporal Domain

Green and Swets (1966) have stated that there exist a set of *interpolation functions* $\{\psi_j(t)\}$, $j = 1, \ldots, 2WT$ such that

i) the set $\{\sqrt{2W}\ \psi_j(t)\}$ is orthonormal over the interval $[0,T]$,

ii) $\hat{x}(t) = \sum_{j=1}^{2WT} \hat{x}(\frac{j}{2W})\psi_j(t)$, and

iii) $\psi_j(t) = \sum_{i=1}^{WT} [a(j)_i c_i(t) + b(j)_i s_i(t)]$,

where $a(j)_i$ and $b(j)_i$ are constants.

The assertion that (i), (ii), and (iii) hold simultaneously is not strictly true. I shall show conditions under which the frequency and temporal representations give nearly equivalent results.

According to (ii) we require functions $\psi_j(t)$ such that

$$\sum_{i=1}^{WT} [a_i c_i(t) + b_i s_i(t)] = \sum_{j=1}^{2WT} \left\{ \sum_{i=1}^{WT} [a_i c_i(\tfrac{j}{2W}) + b_i s_i(\tfrac{j}{2W})] \right\} \psi_j(t)$$

$$= \sum_{i=1}^{WT} [\, a_i \sum_{j=1}^{2WT} c_i(\tfrac{j}{2W})\psi_j(t) + b_i \sum_{j=1}^{2WT} s_i(\tfrac{j}{2W})\psi_j(t)].$$

According to (i) we may multiply each side of this expression by $\psi_j(t)$ and integrate with respect to t over [0,T] to obtain

$$\sum_{i=1}^{WT} [a_i \int_0^T c_i(t)\psi_j(t)\,dt + b_i \int_0^T s_i(t)\psi_j(t)\,dt]$$
$$= \sum_{i=1}^{WT} [a_i c_i(\tfrac{j}{2W})(\tfrac{1}{2W}) + b_i s_i(\tfrac{j}{2W})(\tfrac{1}{2W})].$$

Equating coefficients of a_i and b_i we have

$$\int_0^T c_i(t)\psi_j(t)\,dt = \frac{1}{2W} c_i(\tfrac{j}{2W})$$

and

$$\int_0^T s_i(t)\psi_j(t)\,dt = \frac{1}{2W} s_i(\tfrac{j}{2W}) .$$

But according to (iii) these integrals are the coefficients $a(j)_i$ and $b(j)_i$, respectively, in the finite Fourier series representation of $\psi_j(t)$. Therefore, we may write

$$\psi_j(t) = \frac{1}{2W} \sum_{i=1}^{WT} [c_i(\tfrac{j}{2W})c_i(t) + s_i(\tfrac{j}{2W})s_i(t)]. \qquad (2.9)$$

(Green and Swets omitted reporting this explicit form for $\psi_j(t)$.)

However, consider the integral

$$\int_0^T \psi_m(t)\psi_n(t) = (\tfrac{1}{2W})^2 \sum_{i=1}^{WT} \sum_{j=1}^{WT} \int_0^T [c_i(\tfrac{m}{2W})c_i(t) + s_i(\tfrac{n}{2W})s_i(t)]$$
$$\cdot [c_j(\tfrac{m}{2W})c_j(t) + s_j(\tfrac{n}{2W})s_j(t)]\,dt$$

$$= (\tfrac{1}{2W})^2 \sum_{i=1}^{WT} [c_i(\tfrac{m}{2W})c_i(\tfrac{n}{2W}) + s_i(\tfrac{m}{2W})s_i(\tfrac{n}{2W})] \qquad (2.10)$$

$$= \left(\frac{1}{2W}\right)^2 \sum_{i=1}^{WT} \left[c_m\left(\frac{i}{2W}\right)c_n\left(\frac{i}{2W}\right) + s_m\left(\frac{i}{2W}\right)s_n\left(\frac{i}{2W}\right)\right]$$

According to (i) this must be equal to $1/2W$ when $m = n$ and 0 when $m \neq n$. This will be approximately true under certain conditions stated below.

The following lemma is a consequence of "sampling theorems" for bandlimited waveforms (Middleton, 1960, Ch. 4).

Lemma 2.1. If $x(t)$ and $s(t)$ are two waveforms, Fourier transform band-limited to the same frequency interval $[-W,W]$ and the integer $2WT >> 1$, then their *cross correlation*

$$\int_0^T x(t)s(t)\,dt \simeq \frac{1}{2W} \sum_{j=1}^{2WT} x\left(\frac{j}{2W}\right)s\left(\frac{j}{2W}\right), \tag{2.11}$$

where the error in the approximation to the integral for fixed W is of order $1/T$.

Over the half interval $[0,T/2]$ Equation 2.11 may be expressed as

$$\int_0^{T/2} x(t)s(t)\,dt \simeq \frac{1}{2W} \sum_{j=1}^{WT} x\left(\frac{j}{2W}\right)s\left(\frac{j}{2W}\right) \tag{2.12}$$

where the error is of order $2/T$. We may apply this latter expression to the following integral:

$$\int_0^{T/2} [c_m(t)c_n(t) + s_m(t)s_n(t)]dt \simeq \frac{1}{2W} \sum_{i=1}^{WT} \left[c_m\left(\frac{i}{2W}\right)c_n\left(\frac{i}{2W}\right) + s_m\left(\frac{i}{2W}\right)s_n\left(\frac{i}{2W}\right)\right]$$

However, when $m = n$ the left hand side is simply $1/2 + 1/2 = 1$ and zero otherwise. We have proved

Theorem 2.1. When $2WT >> 1$

$$\sum_{i=1}^{WT} [c_m(\frac{i}{2W})c_n(\frac{i}{2W}) + s_m(\frac{i}{2W})s_n(\frac{i}{2W})] \approx \begin{cases} 2W & \text{if } m = n \\ 0 & \text{if } m \neq n \end{cases} \tag{2.13}$$

where the error in the approximation to the sum is of order $2/T$.

The theorem is not new (see Goldman, 1953, Appendix VI).

Corollary 2.1.1. The functions $\sqrt{2W}\psi_j(t)$ given by (2.9) form an approximately othonormal set over the interval $0 < t < T$ when $2WT >> 1$.

We have shown that the three conditions which Green and Swets impose on the functions $\psi_j(t)$ are not strictly satisfied simultaneously. We can find $\psi_j(t)$ which do in fact form an orthonormal set over the interval. In this case it will be impossible to express the functions exactly by a finite Fourier series as required by condition (iii). Alternatively, we can require condition (iii) in which case orthogonality of the interpolation functions will not be strictly satisfied. In neither case will condition (ii) be strictly satisfied. Of course, the differences between the frequency and temporal representations of $x(t)$ become small when $2WT$ is sufficiently large. It is instructive to pursue this approximate equivalence somewhat further, although doing so represents a digression from our main purpose.

2.4 On the Relation Between the Temporal and Frequency Representations when $2WT >> 1$.

We may apply Lemma 2.1 to Equations (2.3) to obtain

$$a_i \simeq \hat{a}_i = \sqrt{\frac{2}{T}}\,\frac{1}{2W}\sum_{j=1}^{2WT} x\left(\frac{j}{2W}\right)c_i\left(\frac{j}{2W}\right) \tag{2.14a}$$

and

$$b_i \simeq \hat{b}_i = \sqrt{\frac{2}{T}}\,\frac{1}{2W}\sum_{j=1}^{2WT} x\left(\frac{j}{2W}\right)s_i\left(\frac{j}{2W}\right) \tag{2.14b}$$

Thus as an approximation to $\hat{x}(t)$ we may write

$$\begin{aligned}\hat{x}(t) \simeq \hat{\hat{x}}(t) &= \sum_{i=1}^{WT}\left[\hat{a}_i c_i(t) + \hat{b}_i s_i(t)\right] \\ &= \frac{1}{2W}\sqrt{\frac{2}{T}}\sum_{i=1}^{WT}\sum_{j=1}^{2WT} x\left(\frac{j}{2W}\right)\left[c_i\left(\frac{j}{2W}\right)c_i(t)\right. \\ &\qquad\qquad \left. + s_i\left(\frac{j}{2W}\right)s_i(t)\right] \\ &= \sqrt{\frac{2}{T}}\sum_{j=1}^{2WT} x\left(\frac{j}{2W}\right)\psi_j(t).\end{aligned} \tag{2.15}$$

where $\psi_j(t)$ is defined by (2.9).

The vector of weighted samples of $x(t)$ we denote by

$$\mathcal{x} = \frac{1}{\sqrt{2W}}\left[x\left(\frac{1}{2W}\right), \ldots, x\left(\frac{2WT}{2W}\right)\right]'. \tag{2.16}$$

We also let

$$\hat{x}_f = \sqrt{\frac{T}{2}}\left[\hat{a}_1, \ldots, \hat{a}_{WT}, \hat{b}_1, \ldots, \hat{b}_{WT}\right]' \tag{2.17}$$

and define the $2WT \times 2WT$ matrix

$$C = \frac{1}{\sqrt{2W}} \cdot \begin{bmatrix} c_1\left(\frac{1}{2W}\right) & c_2\left(\frac{1}{2W}\right) & \cdots & c_{WT}\left(\frac{1}{2W}\right) & s_1\left(\frac{1}{2W}\right) & \cdots & s_{WT}\left(\frac{1}{2W}\right) \\ c_1\left(\frac{2}{2W}\right) & c_2\left(\frac{2}{2W}\right) & \cdots & c_{WT}\left(\frac{2}{2W}\right) & s_1\left(\frac{2}{2W}\right) & \cdots & s_{WT}\left(\frac{2}{2W}\right) \\ \vdots & & \cdots & & \vdots & & \vdots \\ c_1\left(\frac{2WT}{2W}\right) & & \cdots & c_{WT}\left(\frac{2WT}{2W}\right) & s_1\left(\frac{2WT}{2W}\right) & \cdots & s_{WT}\left(\frac{2WT}{2W}\right) \end{bmatrix} \tag{2.18}$$

Then Equations 2.14 may be expressed in matrix notation as

$$\hat{x}_f = C'x. \tag{2.14'}$$

Theorem 2.2. The matrix C is approximately orthogonal for $2WT \gg 1$.

Proof. A necessary and sufficient condition that C be orthogonal is that $CC' = I$ where I is a $2WT \times 2WT$ identity matrix. By picking the m^{th} row of C' and the n^{th} column of C and adding the product of corresponding terms, one of the following expressions is obtained:

$$\frac{1}{2W}\sum_{j=1}^{2WT} c_m\left(\frac{j}{2W}\right)c_n\left(\frac{j}{2W}\right),$$

$$\frac{1}{2W}\sum_{j=1}^{2WT} c_m\left(\frac{j}{2W}\right)s_n\left(\frac{j}{2W}\right), \quad \text{or}$$

$$\frac{1}{2W}\sum_{j=1}^{2WT} s_m\left(\frac{j}{2W}\right)s_n\left(\frac{j}{2W}\right).$$

But according to Lemma 2.1 each of these expressions is approximately equivalent to a corresponding integral. Examination of Equation 2.4 identifies these integrals and completes the proof of the theorem.

Thus, when $2WT$ becomes large $\hat{\hat{x}}(t)$ approaches $\hat{x}(t)$ and the frequency and temporal representation spaces F and T, respectively, become equivalent differing only by a choice of axes in the space. We shall, however, take as our primary mode of approximation the frequency representation space of Section 2.2.

2.5 *Representation of Certain Linear Filters*

The action of a linear filter Λ on a waveform $x(t)$ may be conveniently represented by a linear transformation on $x_f \in F$.

Let A_f be a (possibly singular) linear transformation on F. Then A_f is a *projection operator* if (i) $A_f^2 = A_f$ and (ii) $A_f = A_f'$, that is, if A_f is an idempotent and symmetric matrix.

Projection operators on F are of particular interest because they may be used to represent idealized square-bandpass filters. A $2W_\alpha T$ dimension subspace F_α of F is generated by the projection operator A_f if for any $x_f \in F$, $A_f x_f \in F$. The rank of A_f is $2W_\alpha T$.

The *output* y_f of a filter A is defined by

$$y_f = A_f x_f. \tag{2.19}$$

More complex (and more realistic) representations for filters may be constructed. The objective here, however, is not to find a representation for such physically realizable filters as might be used electronically in real time. Rather, the purpose is to indicate idealized operations which might be performed on a waveform by a device with memory which can record waveforms for short periods of time. For such a device the operation indicated in Equation 2.19 could be performed.

When the input to a fixed filter A represented by A_f in the frequency domain, is a noise vector n_f, the output An_f of the filter also has a multivariate normal distribution, although degenerate if the rank of A_f is $2W_\alpha T < 2WT$. Artificial difficulties in describing the distribution of the output of the filter may be overcome when the action of the filter is a projection operation.

Suppose that $\{s_1, s_2, \ldots, s_k\}$ is a set of k orthonormal vectors in F. Let B be a $2WT \times k$ matrix with columns $s_1, s_2, \ldots, s_k$ i.e.,

$$B = [s_1 s_2 \ldots s_k]. \tag{2.20}$$

It is easily verified that $A_f = BB'$ is a projection operator on F.

Furthermore, since

$$A_f s_i = BB's_i = s_i, \quad i = 1, \ldots, k,$$

$\{s_1, s_2, \ldots, s_k\}$ may be considered an orthonormal set of basis vectors for the k-dimensional subspace F^* generated by the action of A_f on vectors of F. In fact, for any $x_f \in F$,

$$y = A_f x_f = BB'x_f = B \cdot \begin{bmatrix} s_1'x_f \\ s_2'x_f \\ \vdots \\ s_n'x_f \end{bmatrix} \tag{2.21}$$

$$= (s_1'x_f)s_1 + (s_2'x_f)s_2 + \ldots + (s_k'x_f)s_k.$$

With $y^* = B'x_f$, Equation 2.21 establishes an isomorphism between y and y^* given B.

Definition. A *rectangular* filter is a k-dimensional projection operator on F which has the form $A_f = BB'$ where B is a matrix whose k columns form an orthonormal set of vectors.

Theorem 2.3. The output of a rectangular filter with bandwidth W_α, whose input is white Gaussian nosie with bandwidth $W \geq W_\alpha$, is white Gaussian noise with bandwidth W_α.

Proof. Let $\tilde{n}_f$ have the density of Equation 2.8, and let $A_f = BB'$ be a rectangular filtering operation where B is the matrix of Equation 2.20. Then the output of the filter is represented, up to isomorphism by the random vector

$$\tilde{n}^* = B'\tilde{n}_f. \tag{2.22}$$

Since the rank of B is k, n^* has a multivariate normal distribution with mean vector $\mu(n^*) = E[B'\tilde{n}_f] = B'E[\tilde{n}_f] = B'\mu_n$, and dispersion matrix

$$\Sigma_{n^*} = E[B'\tilde{n}\tilde{n}'B] = B'\Sigma_n B$$
$$= B'(N_0/2)IB = (N_0/2)B'B$$
$$= (N_0/2)I_k$$

where I_k is a $k \times k$ identity matrix. The proof is completed by letting $2W_\alpha T = k$.

It should be noted that passing from the space F to its subspace F^* is not, in general, a reversible process. Given only the output vector n^*, there exist infinitely many vectors n_f such that $n^* = B'n_f$ when B has rank less than 2WT.

The next chapter develops a general class of models for human monaural auditory processing in detection and recognition tasks based on the representation theory developed here.

CHAPTER III

MODELS OF MONAURAL AUDITORY PROCESSING

3.1 Introduction

Several models have been proposed in the signal detection literature to account for the performance of human observers in monaural detection tasks. The models postulate a processing mechanism by which the observer derives information from the input acoustic waveform regarding which of several alternative experimenter-defined hypotheses is presented. Recently the "linear", "energy", and "envelope" models have been extensively reviewed in Green and Swets (1966). These particular models have in common the fact that they describe sensory operations which would give rise to an optimum decision variable for *some* task (not necessarily the one in which the observer finds himself).

The primary objective of the present chapter is to develop a new model, the linear-uncertain model, which includes the linear, envelop and energy models as special cases. The theory is presented in a form which requires the observer to use all the information available to him in a way which is optimum given a residual uncertainty regarding parameters of the signal. A discussion of empirical predictions from these models is presented in the next chapter.

3.2 Task Considerations

It is assumed that the acoustic input to the observer is Gaussian noise, to which occasionally a constant waveform is added. The constant waveform is the *signal* and its presence has a one-to-one correspondence with the experimenter-defined input hypotheses. Thus, the observer's task is one in which the signal is specified exactly. It is not, of course, a foregone conclusion that the observer knows precisely the experimenter's exact signal specification.

Denoting the input waveform sample by x(t) and the signal waveform by s(t), $0 \leq t \leq T$, the alternatives presented to the observer are

$$\begin{aligned} H_1&: \; x(t) = n(t) + s(t) \, ; \\ H_0&: \; x(t) = n(t) . \end{aligned} \tag{3.1}$$

The task also requires that the observer make a judgmental response r indicating which hypothesis alternative was actually presented. For the task to be effective, the observer's ability to discriminate between H_1 and H_0 through r must depend only upon the information contained in the input waveform. This task requirement may be stated in terms of the conditional independence of the response from the presented hypothesis, given the presence of the input waveform.

Assumption 3.1. For every response r in the observer's repertoire and input waveform x

$$P(r|x,H) = P(r|x) \tag{3.2}$$

where H is either H_1 or H_0.

In our presentation of models for sensory processing we will use the representation theory for waveforms developed in Chapter II. Thus, the sample space X may be interpreted as either F or T and x(t) may be represented by x defined at Equation 2.5 or 2.16 when 2WT is large. We also restrict the space of all possible signal vectors to X. Thus, for example, a particular signal vector s identifies a one-dimensional subspace of X, namely, the set of vectors in X proportional to s.

In terms of the representation theory the experimenter-defined hypotheses of Equation 3.1 are

$$\begin{aligned} H_1&: \quad x = n + s; \\ H_0&: \quad x = n. \end{aligned} \tag{3.3}$$

As in Chapter II we shall not assume that the mean of the noise μ_n is zero and shall continue to adopt the more general position that the multivariate normal density of the noise vector is

$$f(n) = N(\mu_n, \Sigma_n), \tag{3.4}$$

where as before the dispersion matrix $\Sigma_n = (N_0/2)I$.

The conditional densities of the input vector x are, from (3.3) and (3.4),

$$\begin{aligned} H_1&: \quad f_1(x) = f(x - s) = N(\mu_n + s, \Sigma_n) \\ H_0&: \quad f_0(x) = f(x) = N(\mu_n, \Sigma_n). \end{aligned} \tag{3.5}$$

It is assumed that every observer knows the distribution of the noise, including the mean of the noise. Some observers may have, however, uncertainty regarding the signal, since it is not always present in the background of noise.

If the signal s is only known to an observer through an a priori distribution $G(s)$ of possible signals, then the unconditional distributions of the input *known* to the observer are

$$\begin{aligned} H_1&: \quad h_1(x) = \int_X f(x - s)\,dG(s), \quad \text{and} \\ H_0&: \quad h_0(x) = f(x) \end{aligned} \tag{3.6}$$

respectively, where f is the normal density of (3.4).

It is well known that the likelihood ratio $\ell(x)$, or some monotone function of it, z, is the optimal decision variable for discrimination between hypotheses H_1 and H_0 (Peterson, Birdsall, and Fox, 1954). From (3.6) the likelihood ratio $\ell(x)$ is,

$$\begin{aligned} \ell(x) &= h_1(x)/h_0(x) \\ &= \int_X f(x - s)dG(s)/f(x) \qquad (3.7) \\ &= \int_X \frac{f(x - s)}{f(x)} dG(s) \\ &= \int_X \ell(x|s)dG(s) \end{aligned}$$

where

$$\ell(x|s) = f_1(x)/f_0(x) = f(x - s)/f(x) \qquad (3.8)$$

is the conditional likelihood ratio given that s is known exactly.

3.3 *Signal Uncertainty*

Models of the ideal observer have been constructed for various tasks by evaluating (3.7) for the unconditional likelihood ratio function of the input process. This is done by assuming that in the task the signal is not specified exactly, but rather has a distribution $G(s)$. A resulting model of the ideal observer is then taken as a model of human observer performance in a task in which the signal is specified exactly. The rationale for this approach to model construction appears to be based on the argument that prior uncertainty regarding parameters, which are actually constant in the task and characterize the signal, should result in a response performance which is the same as would arise from an observer with precise knowledge of the actual uncertainty of signal parameters. However, a close look at Equation 3.6 shows that one should not expect this assumed equivalence. The H_1 conditional density $h_1(x)$ is the density of the input process assumed *known* to the observer. The density $h_1(x)$ does not necessarily represent the *knowable* distribution density of the input to a perfectly informed observer. For the latter observer, $G(s)$ represents the actual or experimenter-defined signal uncertainty in the task.

The assumption which will be made here, perhaps gratuitiously, is that the processing operation performed on the input by a human observer is determined by the likelihood ratio Equation 3.7, where the distribution G(*s*) represents the "internal" observer-specified uncertainty regarding the signal. In tasks where the signal is fixed the distribution of the processing function or decision variable determined by (3.7) is completely determined by the distributional characteristics of external and internal noise.

Assumption 3.2. An observer's prior uncertainty regarding the signal *s*, in tasks in which *s* is specified exactly, may be represented by a multivariate normal density function

$$g(s) = dG(s)/d \;\; = N(\mu_\alpha, \Sigma_s) \tag{3.9}$$

with mean vector μ_α and dispersion matrix Σ_s. The random vector *s* is independent of the noise vector *n*.

It will be shown that (3.9) in connection with (3.7) is quite general enough to specify a processing function which includes as special cases most previously proposed models of human monaural auditory processing.

3.4 A General Structure for Processing Models

Assumption 3.2 allows evaluation of the observer-specified distributions of the input defined at (3.6). Under input hypothesis H_1 the convolution integral may be found using (3.5) and (3.9), so that

$$\begin{aligned} h_1(x) &= N(\mu_n + \mu_\alpha, \Sigma_n + \Sigma_s) \\ &= N(\mu_m, \Sigma_m), \end{aligned} \tag{3.10}$$

where

$$\mu_m = \mu_n + \mu_\alpha, \qquad (3.11)$$
$$\Sigma_m = \Sigma_n + \Sigma_s.$$

Under H_0, as before,

$$h_0(x) = N(\mu_n, \Sigma_n). \qquad (3.12)$$

It is convenient to define the *precision* matrices

$$Q_m = \Sigma_m^{-1} = (\Sigma_n + \Sigma_s)^{-1} \qquad (3.13)$$

and

$$Q_n = \Sigma_n^{-1}. \qquad (3.14)$$

The likelihood ratio of (3.7) is then found directly by using (2.8):

$$\ell(x) = h_1(x)/h_0(x)$$
$$= \frac{(2\pi)^{-WT}(\det Q_m)^{\frac{1}{2}} e^{-\frac{1}{2}(x-\mu_m)'Q_m(x-\mu_m)}}{(2\pi)^{-WT}(\det Q_n)^{\frac{1}{2}} e^{-\frac{1}{2}(x-\mu_n)'Q_n(x-\mu_n)}}. \qquad (3.15)$$

Or, simplifying in terms of the logarithm of the likelihood ratio,

$$z^* = \ln[\ell(x)]$$
$$= \frac{1}{2}\ln\left(\frac{\det Q_m}{\det Q_n}\right) - \frac{1}{2}(x - \mu_m)'Q_m(x - \mu_m)$$
$$+ \frac{1}{2}(x - \mu_n)'Q_n(x - \mu_n)$$
$$= \frac{1}{2}\ln\left(\frac{\det Q_m}{\det Q_n}\right) + \frac{1}{2}(x - \mu_n)'(Q_n - Q_m)(x - \mu_n) \qquad (3.16)$$
$$+ \mu_\alpha' Q_m(x - \mu_n) - \frac{1}{2}\mu_\alpha' Q_m \mu_\alpha.$$

In the following the constant term of (3.16) will be of no interest and may be ignored without loss of generality. It may also be seen that $(x - \mu_n)$ is an invariant translation of the input, regardless of the specification of the parameters Q_m and μ_α. Thus, we shall take the following form as the general structure for our processing models:

$$z = \frac{1}{2}y'(Q_n - Q_m)y + \mu_\alpha' Q_m(y - \frac{1}{2}\mu_\alpha) \tag{3.17}$$

where

$$y = x' - \mu_n, \tag{3.18}$$

and (3.11) and (3.16) have been used.

The decision variable* z of any particular observer within the general class of observers covered by assumption 3.2 may be evaluated by specifying Q_m and μ_α in (3.17)

3.5 The Model of the Ideal Observer.

As an application of the foregoing discussion we may obtain the decision variable of the ideal observer for the case of signal known exactly (Peterson, Birdsall, and Fox, 1954).

Since the signal is specified exactly in the listening task the ideal observer knows the specification exactly and it follows that the ideal observer's prior specification of the signal is the signal itself; i.e., $\mu_\alpha = s$. Since the specification is exact, $\Sigma_s = 0$. It then follows that $\Sigma_m = \Sigma_n + \Sigma_s = \Sigma_n$ and $Q_m = Q_n = (2/N_0)I$. The desired values of Q_n and Q_m having been obtained, they may be inserted into (3.17) to give

$$\begin{aligned} z_\iota &= 0 + s'(\frac{2}{N_0} I)(y - \frac{1}{2}s) \\ &= \frac{2}{N_0}(s'y - \frac{1}{2}s's) \end{aligned} \tag{3.19}$$

as the decision variable of the ideal observer. Of course, there are other functions, monotone with z_ι, which would serve as well.

It may be seen from Chapter II that $s's = E_s$ is the energy of the input signal (since s is represented exactly by a finite Fourier series). Defining the dimensionless quantity

$$d = 2E_s/N_0, \tag{3.20}$$

the decision variable may be written in the form

$$z_\iota = \frac{2s'y}{N_0} - \frac{1}{2}d \tag{3.21}$$

which has a straightforward geometric interpretation in F.

---

* Formally, we may identify the judgmental responses r with ordered subsets of the range of z.

The projection of y onto the line determined by s in F is the vector $\frac{s'y}{||s||}s$. Thus z_ι is a linear function of the magnitude of y in the direction of s in F.

In the usual analysis of the two alternative single-interval task (the "Yes-No" experiment) an observer sets a cutoff value for z, say z_β, such that when the modified input y produces a value of z greater than z_β, the observer makes an R_1 response, i.e., "Yes -- signal was present". The set $R_1 = \{y|z(y) \geq \beta, \ y \in F\}$ is called the *criterion region*. In the case of the ideal observer it is easy to determine a criterion region based on the decision variable z_ι of Equation 3.21.

A fixed value β of z_ι determines a hyperplane in F perpendicular to the line of s. This is illustrated in Figure 3.1. The plane of the figure is taken to be the plane in F determined by y and s. The dashed line represents the intersection of the hyperplane perpendicular to s and the plane of the paper..

As was seen in Chapter II certain linear operations in F may be viewed as filtering the input. Such is the case here for the action of the ideal observer in producing the decision variable z_ι. The subspace of F onto which the modified input y is projected is simply the line of s. The filtering is perfectly "matched" in frequency and phase to the signal. By definition, the *decision space* of an observer is the range of his decision variable. In this case the range is isomorphic to the line of s. (The isomorphism results from the fact that the metric of F carries over to the decision space of the ideal observer.)

3.6 A Model for Human Monaural Processing: The Linear Uncertain Model

The general structure developed in Section 3.4 for a linear observer with uncertainty could be taken directly as a model of human monaural

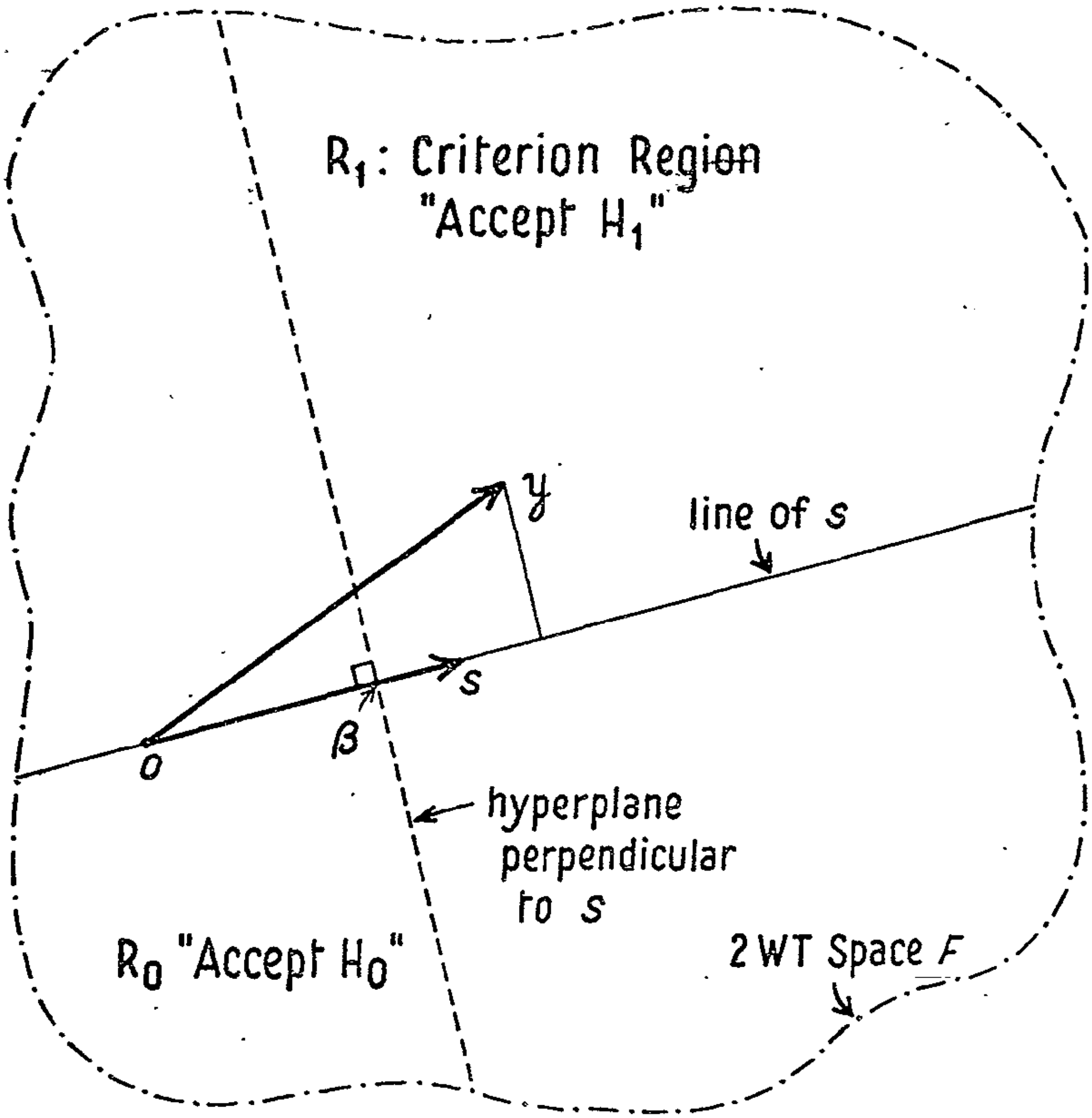

Fig. 3.1. Representation space of the ideal observer for signal specified exactly. The paper represents the plane in F determined by the modified input y and the signal vector s. A criterion region R_1 in F is determined by the cutoff value β on the decision axis which is isomorphic with the line of s. In the case illustrated, y lies in the region R_1 which is equivalent to the fact that its projection onto the line of s is above the cutoff β.

in the task with signal specified exactly. However, without additional simplifying assumptions the model is too general in the sense that as parameters we would have all the entries of μ_α and Σ_s. Most of this freedom is quite unnecessary to obtain strong predictions from the model. I will assume that the human observer knows a region of F occupied by the true signal s, but is uncertain about the exact magnitude of the components of s within that region. This assumption is made more explicit in the following.

Assumption 3.3. The dispersion matrix Σ_s of the prior distribution for an observer α is proportional to a projection operator D_α on F. Furthermore, the subspace F_α generated by the action of D_α on vectors of F contains μ_n, μ_α and s.

The assumption implies that

$$\Sigma_s = \frac{M_0}{2} D_\alpha, \tag{3.22}$$

where $M_0/2$ is the constant of proportionality, and that D_α has rank $2W_\alpha T$ which is less than or equal to 2WT, the dimensionality of F. Since D_α is indempotent, it preserves the magnitude of every vector in F_α. In particular,

$$D_\alpha \mu_n = \mu_n, \tag{3.23a}$$

$$D_\alpha \mu_\alpha = \mu_\alpha, \tag{3.23b}$$

and

$$D_\alpha s = s. \tag{3.23c}$$

D_α represents a rectangular linear filter in F(see Chapter II), and corresponds to idealized square band-pass filtering of the modified input.

As we shall see, it is convenient to think of the constant in (3.22) as representing the observer's uncertainty (or imprecision, or variance) regarding the signal amplitude, per unit bandwidth of the signal. It is only appropriate to scale such uncertainty relative to physical measurements; in this case the uncertainty in the input is the variance N_0 of the noise process per unit cycle. The ratio of internal uncertainty to the total uncertainty, therefore, is defined to be

$$\lambda = \frac{M_0}{M_0 + N_0} . \tag{3.24}$$

It follows that λ is a nonsingular transformation of the scaled value of M_0, i.e.,

$$M_0/N_0 = \frac{\lambda}{1 - \lambda} .$$

Now Q_m can be found directly:

$$\begin{aligned} Q_m &= (\Sigma_n + \Sigma_s)^{-1} \\ &= (\frac{N_0}{2} I + \frac{M_0}{2} D)^{-1} \\ &= \frac{2}{N_0} (I + \frac{M_0}{N_0} D)^{-1} \\ &= \frac{2}{N_0} (I - \lambda D) . \end{aligned} \tag{3.25}$$

The last step is easily verified since D is indempotent. As a consequence,

$$\begin{aligned} Q_n - Q_m &= \frac{2}{N_0} I - \frac{2}{N_0} (I - \lambda D) \\ &= \frac{2\lambda}{N_0} D. \end{aligned} \tag{3.26}$$

With the explicit representation for the precision matrices in (3.25) and (3.26), the decision variable z_α for observer α is found by substitution into Equation 3.17:

$$z_\alpha = \frac{1}{2}y'(\frac{2}{N_0} D_\alpha)y + \mu'_\alpha[\frac{2}{N_0}(I - \lambda D_\alpha)](y - \frac{1}{2}\mu_\alpha),$$

or, since $\mu'_\alpha I = \mu'_\alpha = \mu'_\alpha D_\alpha$ this reduces to

$$z_\alpha = \lambda \frac{y' D_\alpha y}{N_0} + (1 - \lambda) \frac{2\mu_\alpha' D_\alpha (y - \frac{1}{2}\mu_\alpha)}{N_0} . \quad (3.27)$$

The decision variable z_α of the linear-uncertain observer is thus a convex combination of the standardized energy of the projection of the modified input y onto F and a linear term depending upon the cross correlation of y with the supposed signal μ_α. From Equation 3.24 it is seen that as the prior internal specification of signal becomes poor, that is the uncertainty M_0 becomes large relative to N_0, the uncertainty parameter λ approaches 1 and α will behave like a pure energy observer. On the other hand as the observer α becomes increasingly certain of the correctness of the internal specification signal μ_α, λ approaches zero and α behaves like a pure linear observer. When λ lies between the extremes of 0 and 1, the decision variable z_α may be seen to be a linear transformation of the squared radius β of the hyperspheroidal cylinder

$$(y + \frac{(1-\lambda)}{2\lambda} \mu_\alpha)' D_\alpha (y + \frac{(1-\lambda)}{2\lambda} \mu_\alpha) = \beta$$

with center at $-\frac{(1-\lambda)}{2\lambda} D_\alpha^{1/2} \mu_\alpha$ in F.

The decision regions for α are depicted in Figure 3.2 (compare with Figure 3.1). Here the plane of the paper is determined by the true signal s and the prior specified signal μ_α. The center of the region R_0 lies on the line of μ_α, but on the opposite side of the origin. The modified input y is projected onto the subspace F_α. If the image of y lies outside the hypersphere then $z_\alpha > a\beta + b$ (where β is the squared radius and a and b are appropriate constants) and observer α accepts the hypothesis that the input contains the signal. Otherwise, the image is in R_0 and the observer rejects that hypothesis. As λ increases the center moves toward the origin and the direction (phase)

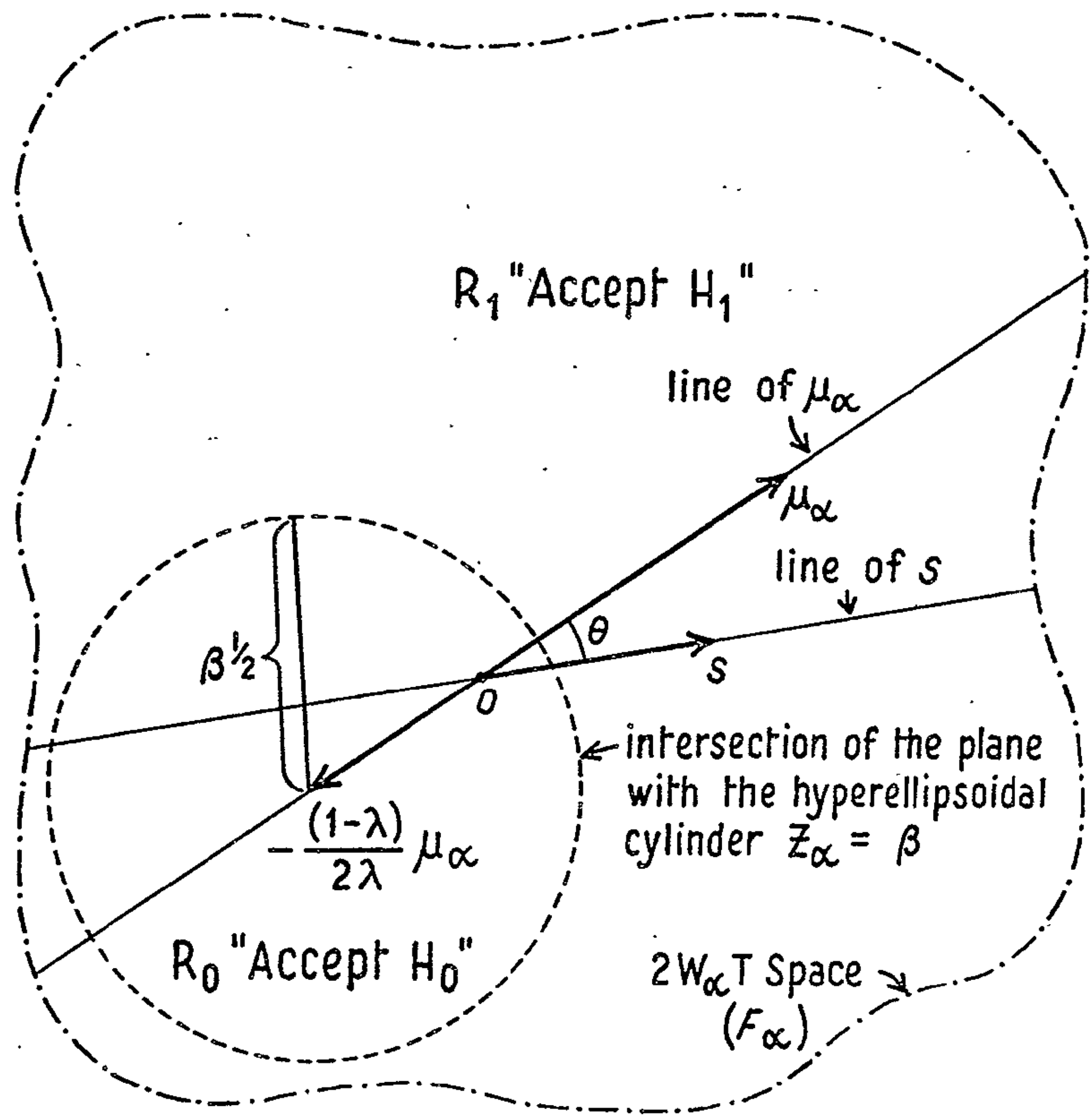

Fig. 3.2. Representation space of the linear-uncertain observer. The paper represents the plane in F_α determined by the signal vector s and the observer's mean representation of the signal μ_α. A criterion region R_1 is determined by a cutoff value β on the decision axis which is isomorphic to the squared radius of a hyperspheroid in F_α. As $\lambda \to 1$ the center moves on the line of μ_α to the origin O of the space and the observer becomes a simple energy detector.

of μ_α becomes irrelevant. As $\lambda \to 0$ the center moves away from the tip of μ_α and the decision boundary becomes a perpendicular hyperplane at the point of intersection with the line of μ_α. The observer's decision variable then behaves like the ideal observer's decision with variable z_ι, but with μ_α substituted for s.

It is perhaps surprising that the linear-uncertain model of the observer's decision variable is a generalization of most previously proposed detection theory based models for human monaural auditory processing. This will be demonstrated in the next section.

3.7 Special Cases of the Linear-Uncertain Model

The models of this section are obtained from the fundamental Equation 3.27 for the decision variable of the linear-uncertain observer.

The linear observer. When $\lambda = 0$ in (3.27)

$$z_\alpha = \frac{2}{N_0}\,\mu'_\alpha D_\alpha (y - \tfrac{1}{2}\mu_\alpha) \tag{3.28}$$

and the observer performs purely linear operations on the input waveform. If $\mu_\alpha = s$ and $D_\alpha = I$, the linear observer is optimum since $z_\alpha = z_\iota$ of Equation 3.19. The observer is even optimum if $D_\alpha \neq I$ since according to Assumption 3.3 D_α cannot degrade s. Thus, non-optimal performance of the linear observer must be due to the fact that $\mu_\alpha \neq s$ or be due to additional internal noise or criterion variability.

The energy observer. When $\lambda = 1$ in (3.27)

$$z_\alpha = \frac{y' D_\alpha y}{N_0} \tag{3.29}$$

and therefore the decision variable of observer α is based only on the energy of the modified input passing through an idealized square band-pass filter with bandwidth W_α. Other forms of an energy model have been proposed. A more usual assumption is that it is the input x itself whose energy is found at the output of the filter, i.e.,

$$z_\alpha = \frac{x'D_\alpha x}{N_o} . \tag{3.30}$$

The two models are identical only when mean of the noise is the zero vector. When the mean is not zero, these energy models make different predictions for observer performance for some tasks, as will be seen in the next chapter.

The envelope observer is a very special case of the Equation 3.29. Let $s_\perp(t)$ be the *Hilbert transformation* of the signal waveform $s(t)$, $0 \le t \le T$ (cf. Hancock and Wintz, 1966). If $|x(t)|$ is the instantaneous modulus of $x(t)$, then in quite general terms the function

$$E_n(t) = (|s/t)| + |s_\perp(t)|)^{1/2}, \quad 0 \le t \le T, \tag{3.31}$$

is the (instantaneous) *envelope* of $s(t)$. For example, if $s(t) = A\sin 2\pi\omega t$, then $s_\perp(t) = A\cos 2\pi\omega t$ and $E_n(t) = A$, as one might expect. The Hilbert transform is orthogonal to the original waveform, i.e.,

$$\int_0^T s(t)s_\perp(t) = 0. \tag{3.32}$$

If $s_\perp$ is the vector of F corresponding to $s_\perp(t)$, then (3.32) implies that

$$s's_\perp = 0. \tag{3.33}$$

Now if $s(t)$ is a sinusoid, then $s(t)$ and $s_\perp(t)$ differ only in phase. If observer α knows s but does not know the phase of s then it may be shown that α should let

$$D_\alpha = ss'/||s|| + s_\perp s_\perp'/||s_\perp|| \tag{3.34}$$

in Equation 3.29 (cf. Wainstein and Zubako, 1962, Sec. 33). It may also be shown that D_α is then a projection operator by the methods of Section 2.4.

The idempotency of D_α follows with help of (3.33). When the substitution of D_α of (3.34) into (3.29) is made it results in the decision variable of an envelope observer:

$$z_\alpha = \frac{(s'y)^2/||s|| + (s_\perp' y)^2/||s_\perp||}{N_0}. \tag{3.35}$$

It is clear from (3.35) that z_α for the envelope observer is the squared radius of a circle lying in the plane of s and $s_\perp$ in F with center at the origin. This is illustrated in Figure 3.3. The modified input y is projected onto the plane of s and $s_\perp$. If the length of the projection is greater than $\beta^{1/2}$, then it lies outside a circle $z_\alpha = \beta$ in the criterion region R_1. The circle defines a hypercylinder in F whose inside is R_0.

It has been pointed out elsewhere that the energy model and the envelope model make similar predictions in a variety of tasks (Green and Swets, 1966). In a general way this is apparent from the geometry for the two observers. The primary distinguishing characteristic of the envelope observer would appear to be his narrow and precise bandwidth as an energy observer. This conclusion is also born out by a comparison of the statistics of $\sqrt{z_\alpha}$ with the statistics of the output of a narrow-band filter excited by Gaussian noise (Green and Swets, 1966, Section 6.5.2, and references there cited).

3.8 The Noisy Linear-Uncertain Model

An assumption of noise-free processing has been implicit in the presentation of the linear-uncertain model in the previous two sections. For any realizable system, e.g., a human processor, this is certainly a generous assumption. Two kinds of internal noise have often been suggested. The first is that noise is added to the input x before it is processed. The

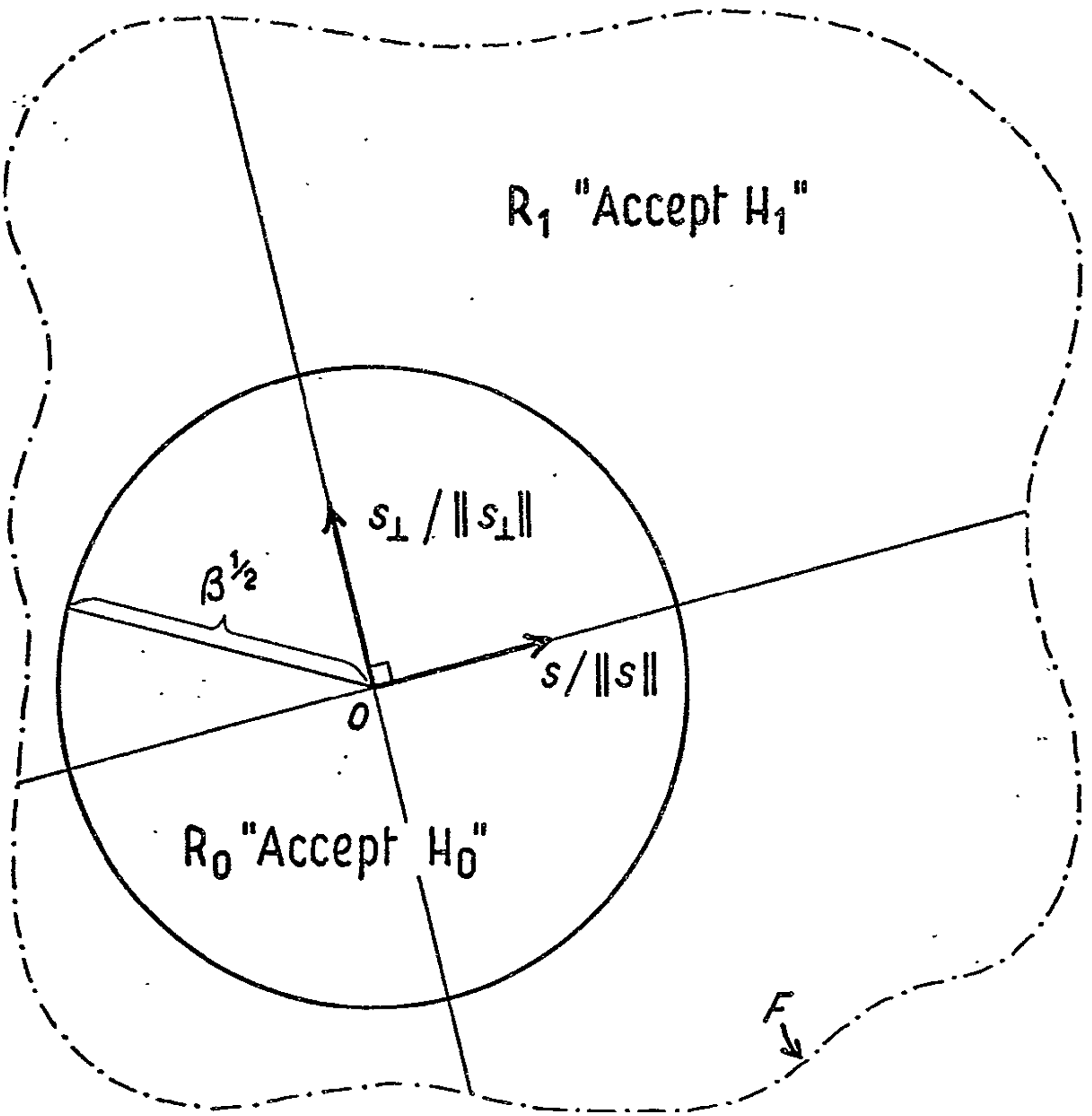

Fig. 3.3. The representation space of an envelope observer. The paper represents the plane in F determined by the orthogonal vectors s and $s_\perp$. A criterion region R_1 is determined by a cutoff value β on the decision axis which is isomorphic to the squared radius of a circle with center at the origin in the plane. If y projects outside the circle in the region R_1, the observer accepts the hypothesis H_1.

second is that the response is a noisy transformation of the decision variable z_α. Less often it has also been suggested that the observer may have a noisy memory for the specification of signal parameters relevant to good detection performance. A noisy memory would lead to a fluctuation in the processing of the input. The details of this latter source for degradation of observer performance will be considered first.

Noisy memory models. It is possible that the observer-specified signal μ_α varies in a random manner which is not under observer control. If this were then the case, we might suppose that observer specified signal vector could be adequately characterized as

$$\mu^*_\alpha = \mu_\alpha + e \tag{3.36}$$

where as before μ_α is the mean of the prior distribution for s and e is a random error term, perhaps with a multivariate normal distribution density with mean vector 0 and dispersion matrix Σ_e.

If the linear-uncertain observer is unaware of the variability in the mean of his prior distribution, then his decision variable would remain in the form given by z_α in (3.27), but with μ^*_α of (3.36) replacing μ_α. It is important to realize that the distributional character of z_α is thereby changed as well.

Another possibility may be considered by assuming that the observer specifies the signal from memory on each trial, but that there is random error in this specification. The error could be due to changing uncertainty regarding signal parameters, for example. If the observer is aware of the fluctuation in his memory, then it may be shown that the observer should be a linear-quadratic processor of Equation 3.17, where Σ_s is the dispersion matrix of the memory signal vector μ_α whose mean is (presumably) the true signal s (Birdsall, 1960). This version of the noisy memory model differs formally from the linear-uncertain model only in that $\mu_\alpha = s$, and thus is another special case of the latter.*

The sensory-noise model. If we suppose that the sensory encoding produces random fluctuations in the input waveform, this could be represented by replacing the input x in the linear-uncertain model by

$$x^* = x + e. \tag{3.37}$$

Birdsall (1960) has shown that a linear processor with Gaussian internal noise added to the input has quite different psychometric functions from his noisy memory model mentioned above (for the special case where $2WT = 1$). This is to be expected since x^* is the effective input to the observer and would not change his mode of processing when e is normally distributed.

The noisy decision variable model. Perhaps the most common assumption made in one form or another is to consider that the decision variable itself is noisy; i.e., that

$$z^*_\alpha = z_\alpha + e \tag{3.38}$$

replaces the decision variable z_α. Strangely enough, however, only until very recently have some of the implications of this plausible model been investigated (Wickelgren, 1968).

There are, of course, many other ways of degrading an observer's decision variable. The uncertainty parameter λ could be made a random variable. The center frequency of the pass-band could fluctuate (a possibility recently investigated by Henning (1967)). The quality of the decision variable could vary, and so on. For the present, however, let us be content with noise

---

* Birdsall's work (1960) provided much of the incentive for developing the linear-uncertain model.

added to the input (e_1), unknown fluctuation in the observer's specification in his prior mean (e_2), and perturbations of the decision variable itself (e_3).

The decision variable of the noisy linear-uncertain observer is, then,

$$z_\alpha^* = \frac{1}{N_0}\left[\lambda(y + e_1)'D_\alpha(y + e_1) + \right. \tag{3.39}$$

$$\left. + (1 - \lambda)(\mu_2 + e_2)'D_\alpha[2(y + e_1) - (\mu_\alpha + e_2)]\right] + e_3.$$

In order to be concrete in the following, we suppose that e_3 is normally distributed with mean zero and variance V_3, and that e_2 and e_3 are each multivariate normal with mean vector zero and dispersion matrices

$$\Sigma_{e_1} = \frac{N_0V_2}{2} I \quad \text{and} \quad \Sigma_{e_2} = \frac{N_0V_3}{2} I,$$

respectively. Further, e_1, e_2, e_3 and the noise vector n are assumed mutually independent.

CHAPTER IV

MEASURES OF CONCORDANCE BETWEEN OBSERVERS

4.1 *Introduction*

The previous chapter outlines the basic theory of a general model for monaural auditory processing, the linear-uncertain model. This chapter investigates a variety of predictions of the model. Since the linear-uncertain model includes as special cases a number of previously proposed models for auditory processing, the methods of this chapter apply equally well to them. The derivation of the level of association between observers appears here for the first time. Possible methods for estimating parameters for the linear-uncertain observer using non-parametric measures of association are considered. The psychometric function for the linear uncertain observer is approximated, and the relation between measures of performance and concordance are investigated.

4.2 *Observer Performance*

In a two-alternative task an observer's decision variable z may be considered as having the distributions $F_0(z)$ and $F_1(z)$ conditional upon H_0 and H_1, respectively. If the observer uses the rule "say H_1 if $z \geq z_c$, otherwise H_0" then the observer's *hit rate* would be

$$P(H) = \int_{z_c}^{\infty} dF_1(z) \tag{4.1}$$

and his *false-alarm* rate would be

$$P(F) = \int_{z_c}^{\infty} dF_0(z). \tag{4.2}$$

An observer's *receiver operating characteristic* (ROC) curve is the set of ordered pairs [P(F),P(H)] parameterized by z_c (Peterson, Birdsall, and Fox, 1954).

If the observer uses the rule (2) "say r_c if $z = z_c$", where r_c is a strictly increasing function of z_c, the probabilities P(H) and P(F) remain defined for each value of v_c but have no special significance except as coordinates of the ROC curve. Clearly the ROC curves determined by z using either decision rule are identical.

The *ideal observer* may be characterized as that observer whose hit-rate is at least as great as the hit-rate of any other observer with the same false-alarm rate (Birdsall, 1966). This characterization is more general than that given in the last chapter where it is assumed that the signal is specified exactly.

It is clear from the definitions that the area under the ROC curve for ideal observer must be at least as great as the area under the ROC curve for any other observer in the same task. An important result regarding the area was obtained by Green (1964a).

Lemma 4.1. If $z_{\alpha 1}$ and $z_{\alpha 0}$ are two independent samples of an observer's decision variable, conditional upon H_1 and H_0, respectively, then the area P_A under the ROC curve generated by the decision variable z (using either rule (1) or rule (2) above) is given by

$$P_A = \int_{-\infty}^{\infty} F_1(z_c) dF_1(z_c) = P(z_{\alpha 1} > z_{\alpha 0}). \tag{4.3}$$

A proof is given in Green and Moses (1966). I have generalized somewhat the language over the original statement of the result since Green interprets the probability on the right as the probability of making a correct response in a two-interval forced-choice task in which the observer knows that exactly one of the intervals contains signal-plus-noise and it will be shown that the ROC area is related to non-parametric measures of association between observers.

4.3 *Nonparametric Measures of Association*

We review two popular nonparametric measures of association which will be useful in the discussion of observer performance and concordance.

Following Kendall (1962), for two random variables z_α and z_β and two independent, joint samples $(z_{\alpha i}, z_{\beta i})$ and $(z_{\alpha j}, z_{\beta j})$ of the variables, define

$$a_{ij} = \mathrm{sgn}(z_{\alpha i} - z_{\alpha j}) = \begin{cases} 1 & \text{if } z_{\alpha i} - z_{\alpha j} > 0 \\ 0 & \text{if } z_{\alpha i} - z_{\alpha j} = 0 \\ -1 & \text{if } z_{\alpha i} - z_{\alpha j} < 0. \end{cases} \tag{4.4}$$

Similarly,

$$b_{ij} = \mathrm{sgn}(z_{\beta i} - z_{\beta j}). \tag{4.5}$$

Then, *tau* is defined by

$$\tau \equiv \mathrm{Corr}(a_{ij}, b_{ij}). \tag{4.6}$$

If (z_α, z_β) has a continuous bivariate distribution,

$$\tau = E(a_{ij} b_{ij}), \tag{4.7}$$

since in this case $E(a_{ij}) = E(b_{ij}) = 0$ and $\mathrm{Var}(a_{ij}) = \mathrm{Var}(b_{ij}) = 1$.

Another way of looking at tau is in terms of the probability of agreement and disagreement of the sign of the difference between samples of the decision variables. With

$$P(S) = P(a_{ij} b_{ij} > 0)$$

and

$$P(D) = P(a_{ij} b_{ij} < 0), \tag{4.8}$$

it is readily shown that

$$E(a_{ij} b_{ij}) = P(S) - P(D). \tag{4.9}$$

Goodman and Kruskal (1954) were interested in indices of association for bivariate, ordered contingency tables. In this case the random variables are not continuous. The probability of a tie in one or the other of the variables is

$$P(T) = P(a_{ij}b_{ij} = 0), \tag{4.10}$$

so that

$$P(S) + P(D) + P(T) = 1.$$

The Goodman-Kruskal coefficient *gamma* is based on only the untied pairs of the variables:

$$\gamma \equiv [P(S) - P(D)]/[1 - P(T)] \tag{4.11}$$

For continuous variables $P(T) = 0$, so that $\gamma = \tau$ in this case.

It is also possible to view gamma as a conditional correlation between signs.

Proposition 4.1.

$$\gamma = \mathrm{Corr}(a_{ij}, b_{ij} | a_{ij} \neq 0, \ b_{ij} \neq 0). \tag{4.12}$$

4.4 The ROC Area as a Measure of Association Between Decision Variables

We may define the *perfect* decision variable by

$$z_p = \begin{cases} 1 & \text{if } H_1 \\ 0 & \text{if } H_0. \end{cases} \tag{4.13}$$

This is the experimenter's "decision variable" and, of course, does not depend upon the input x.

Theorem 4.1. Let P_A be the area under the ROC curve for observer α with a continuous decision variable z_α, and γ the Goodman-Kruskal coefficient between the perfect decision variable z_p and z_α. Then,

$$P_A = (\gamma + 1)/2. \tag{4.14}$$

Proof. Let (z_α, z_p) and (z'_α, z'_p) be two independent joint samples of the decision variables. Then, since the value of z_p specifies the hypothesis,

$$P[(z_\alpha - z'_\alpha)(z_p - z'_p) > 0] = P(z_{\alpha 1} > z_{\alpha 0}).$$

The right hand side is P_A by Lemma 4.1. The left hand side is apparently $P(S)/[1 - P(T)]$. But, $P(D)/[1 - P(T)] = 1 - P(S)/[1 - P(T)]$, so that $P(S)/[1 - P(T)] = (\gamma + 1)/2$, completing the proof.

The theorem's importance lies in the fact that it suggests a way to make reasonable estimates of the ROC area, even when the *observed* decision variable is not continuous. In Appendix I it is shown that if $P(T_\alpha)$ is the probability of a tie in the observed decision variable $\hat{z}_\alpha$, an appropriate estimate of the area under the ROC curve is given by

$$\text{est } P_A = [\text{est Corr}(a_{ij}, b_{ij} \mid b_{ij} \neq 0) + 1]/2$$

$$= \frac{1}{2}\left[\frac{P(S) - P(D)}{\sqrt{1 - P(T_\alpha \mid \overline{T}_p)\,[1 - P(T_p)]}} + 1\right] \tag{4.15}$$

where $\overline{T}_p$ is the event $b_{ij} \neq 0$. A computing formula is also given.

When both decision variables, z_α and z_β, are considered continuous the estimate of tau is given by

$$\tau_{\alpha\beta} = \text{est Corr}(a_{ij}, b_{ij})$$

$$= \frac{P(S) - P(D)}{\sqrt{[1 - P(T_\alpha)][1 - P(T_\beta)]}} \tag{4.16}$$

4.5 The Relation Between Parametric and Nonparametric Measures of Association

The exact distribution theory for the noisy linear-uncertain model is extremely difficult and is apparently unsolved. (In the internal noise-free case, however, the distributions can be shown to be non-central chi-square.) However, the moments of z_α^* are relatively less difficult to determine. This opens the possibility of trying to approximate the theoretical value of τ between z_α and z_β using only the moments of the joint distribution of z_α and z_β. It is at least plausible that sample estimates of τ could then be used to make estimtes of the unknown parameters entering into an observer's decision variable such as his uncertainty parameter λ_α or bandwidth-time product $\delta_\alpha = W_\alpha T$. We shall call such a procedure "nonparametric estimation".

Perhaps a somewhat less ambitious approach than nonparametric estimation would be to try to find experimental situations which would discriminate between alternative models using nearly any measure of concordance. If such experiments could be found, then statistics such as the linear correlations could serve at the theoretical level, and tau could serve at the empirical level. This latter procedure is the one followed here. It is worthwhile to give a more complete justification, however.

Greiner's relation. The ordinary product-moment (linear) correlation between two decision variables z_α and z_β will be denoted by

$$\rho_{\alpha\beta} = \mathrm{Corr}(z_\alpha, z_\beta).$$

If z_α and z_β have a bivariate normal distribution with correlation $\rho_{\alpha\beta}$, then it is known (Kendall, 1962; Greiner, 1909) that

$$\rho_{\alpha\beta} = \sin(\pi\tau_{\alpha\beta}/2). \tag{4.17}$$

When the stated assumptions are tenable this relation provides a consistent estimate of $\rho_{\alpha\beta}$ from an estimate of $\tau_{\alpha\beta}$.

Griener's relation is a special case of a more general approximation to τ based on the joint moments for non-normal variation of random variables. Kendall (1949) assumed that the joint distribution may be closely approximated by a truncated Gram-Charlier series. With standardized moments defined by

$$\mu'_{ij} = E\left\{\left[\frac{z_\alpha - E(z_\alpha)}{[\mathrm{Var}(z_\alpha)]^{1/2}}\right]^i \left[\frac{z_\beta - E(z_\beta)}{[\mathrm{Var}(z_\beta)]^{1/2}}\right]^j\right\}$$

and $\rho = \rho_{\alpha\beta} = \mu'_{11}$, Kendall's approximation* to τ is given by

$$\begin{aligned}\tau_{\alpha\beta} \approx \frac{2}{\pi}\sin^{-1}\rho + \frac{1}{24\pi(1-\rho^2)^{3/2}}\,&[(\mu'_{40} + \mu'_{04} - 6)(3\rho - 2\rho^3) \\ &- 4(\mu'_{31} + \mu'_{13} - 6\rho) + 6\rho(\mu'_{22} - 2\rho^2 - 1)]\end{aligned} \tag{4.18}$$

* Kendall's original expression, his Equation (25), was given in terms of the cumulants of the joint distribution.

If the decision variables are jointly normal the correction term (in braces) to Greiner's relation is zero.

A small amount of empirical data presented by Kendall suggests that the unfortunate complexity of the correction term to Greiner's relation cannot be safely ignored for precise results in many situations. Unfortunately little more seems to be known regarding estimating ρ from τ.

For the linear-uncertain model the correction term is zero only if both observers are linear, i.e., $\bar{\lambda}_\alpha = \bar{\lambda}_\beta = 1$ $(\bar{\lambda} = 1 - \lambda)$. Otherwise the correction term is not only non-zero, but contains moments of the physical waveforms not ordinarily measured, e.g.,

$$\int_0^T s^4(t)\, dt,$$

for a constant signal waveform $s(t)$. (Interesting results can be obtained when the signal is a sample of Gaussian noise, which however does not concern us here.) In the general case the correction term is extremely complicated for linear uncertain observers and it would require advanced computer techniques in non-linear estimation to obtain estimates of model parameters. I am forced to conclude that, although feasible, non-parametric estimation is not presently practical.

The alternative mentioned above to non-parametric estimation depends upon an approximately monotone relation between τ and ρ. An examination of Kendall's approximation to τ is not very revealing in this regard. The difficulty is that factors which affect ρ may also affect the value of the other joint moments in some unknown fashion. In our situation, however, we know that the association between linear-uncertain decision variables depends upon common elements of the noise process at the input. As was seen in

Chapter III the decision variable of the linear-uncertain observer is a linear transformation of the squared radius of a hyperspheroidal cylinder in the representation space. By assumption two observers share a common subspace in the frequency representation space. Thus, at least in the case where the cosine of the angle between μ_α and μ_β is not negative, the radii will tend to increase and decrease together, i.e., will be to some extent concordant. Both τ and ρ are measures of concordance in this sense (Kruskal, 1958) and therefore may be expected to be highly correlated with one another.

4.6 *The Correlation between Observers*

The results of this section make the implicit assumption that to a first order of approximation the joint distribution of the decision variables may be considered bivariate normal. Somewhat more accurate results for some purposes, with a corresponding increase in technical difficulty, might be expected if marginal monotone transformations of the variables are made prior to the computation of the moments. Some possibilities for normalizing transformaitons are considered in Lamphier and Birdsall (1960).

We consider the general case where the modified input to observer α is $y_\alpha = y + s_\alpha$ and the input to observer β is $y_\beta = y + s_\beta$. It is convenient to have the following definitions (we assume the observers' filters overlap in F):

$$\mu_{s_\alpha} = s_\alpha \tag{4.19}$$

$$R_{\alpha\beta} = \mu'_\alpha D_{\alpha\beta}\mu_\beta, \quad r_{\alpha\beta} = 2R_{\alpha\beta}/N_0 \tag{4.20}$$

$$E_\alpha = R_{\alpha\alpha}, \quad d_\alpha = 2E_\alpha/N_0 \tag{4.21}$$

Thus, for example, $r_{\alpha s_\beta}$ means $2\mu'_\alpha D_{\alpha\beta}s_\beta/N_0$, the standardized cross-correlation in the joint subspace determined by $D_{\alpha\beta} = D_\alpha D_\beta$ between the prior mean vector μ_α and signal vector s_β presented to observer β. Of course, if s_β is the null signal 0, then $r_{\alpha s_\beta} = 0$ and $r_{s_\alpha s_\beta} = 0$ regardless of the value of μ_α or s_α, respectively.

Theorem 4.2. The covariance between noisy linear-uncertain observers with decision variables z^*_α and z^*_β as given by Equation 3.39 is given by

$$\begin{aligned}\mathrm{Cov}(z^*_\alpha, z^*_\beta) = {} & \lambda_\alpha \lambda_\beta [\delta_{\alpha\beta} + r_{s_\alpha s_\beta}] + \bar{\lambda}_\alpha \lambda_\beta r_{\alpha s_\beta} \\ & + \lambda_\alpha \bar{\lambda}_\beta r_{s_\alpha \beta} + \bar{\lambda}_\alpha \bar{\lambda}_\beta r_{\alpha\beta}.\end{aligned} \tag{4.22}$$

Further, the variance of z^*_α is given by

$$\begin{aligned}\mathrm{Var}(z^*_\alpha) = {} & \lambda^2_\alpha [\delta_\alpha (1 + V_{\alpha 1})^2 + d_{s_\alpha}(1 + V_{\alpha 1})] + 2\bar{\lambda}_\alpha \lambda_\alpha r_{\alpha s_\alpha}(1 + V_{\alpha 1}) \\ & + \bar{\lambda}^2_\alpha \{\delta_\alpha V_{\alpha 2}[2(1 + V_{\alpha 1}) + V_{\alpha 2}] \\ & + d_\alpha (1 + V_{\alpha 1} + V_{\alpha 2}) + (d_{s_\alpha} - r_{\alpha s_\alpha})V_{\alpha 2}\} + V_{\alpha 3}.\end{aligned} \tag{4.23}$$

where $\delta_{\alpha\beta} = \min\{\delta_\alpha, \delta_\beta\}$, $\delta_\alpha = W_\alpha T$ and $\bar{\lambda} = 1 - \lambda$.
The proof of the theorem is given in Appendix II.

The linear correlation between noisy linear-uncertain observers is thus given by

$$\rho_{\alpha^*\beta^*} = \mathrm{Corr}(z^*_\alpha, z^*_\beta) = \frac{\mathrm{Cov}(z^*_\alpha, z^*_\beta)}{[\mathrm{Var}(z^*_\alpha)\mathrm{Var}(z^*_\beta)]^{1/2}} \tag{4.24}$$

There are, of course, many special cases which could be considered by assigning parameter values. There are over 2000 such cases for extreme values of parameters of which perhaps 100 might be considered "interesting" for some purpose. We shall consider only several of these interesting cases.

Theorem 4.3. The necessary and sufficient conditions that two noisy linear uncertain observers have identical correlation on signal-plus-noise trials (H_1) and noise-alone trials (H_0) are that

i) both observers are linear processors, i.e., $\lambda_\alpha = \lambda_\beta = 0$, and that

ii) neither observer has a noisy memory-specified reference signal, i.e., $V_{\alpha 2} = V_{\beta 2} = 0$.

Proof. Under H_1, $s_\alpha = s_\beta = s$ so that $r_{s_\alpha s_\beta} = d_{s_\alpha} = d_{s_\beta} = d_s$. $\rho^{(1)}$ is obtained by inserting these values into (4.24). Likewise under H_0, $s_\alpha = s_\beta = 0$, so that $r_{s_\alpha s_\beta} = r_{\alpha s_\beta} = r_{\alpha s_\alpha} = r_{s_\alpha \beta} = r_{s_\beta \beta} = d_{s_\alpha} = d_{s_\beta} = 0$, which when inserted into (4.24) gives $\rho^{(0)}$. Requiring that $\rho^{(1)} = \rho^{(0)}$ for any (non-trivial) values of μ_α and μ_β can then be seen to be equivalent to conditions (i) and (ii).

Corollary 4.3.1. Under conditions (i) and (ii) of the theorem the correlation is

$$\rho = \frac{r_{\alpha\beta}}{\{[d_\alpha(1 + V_{\alpha 1}) + V_{\alpha 3}][d_\beta(1 + V_{\beta 1}) + V_{\beta 3}]\}^{1/2}} \tag{4.25}$$

and Greiner's relation holds so that

$$\tau^{(1)} = \tau^{(0)} = \tau = \frac{2}{\pi}\sin^{-1}\rho. \tag{4.26}$$

It would be desirable to have the theorem stated in terms of τ rather than ρ. The sufficiency of the conditions for equal τ values on H_1 and H_0 is given in the preceeding corollary. That equal τ values imply the conditions remains a reasonable conjecture.

An interesting situation arises if one of the observers is an electronic energy detector which receives only noise at the input on both H_1 and H_0 trials. In this case $\lambda_\beta = 1$, $V_{\beta 1} \simeq 0$, $V_{\beta 3} \simeq 0$, and $\mu_\beta = s_\beta = 0$ under both H_1 and H_0. Further, we let

$$\begin{aligned} &\mathrm{Var}_1(z_\alpha^*) = \mathrm{Var}(z_\alpha^* \mid s_\alpha = s) \\ \text{and} \quad &\mathrm{Var}_0(z_\alpha^*) = \mathrm{Var}(z_\alpha^* \mid s_\alpha = 0) \end{aligned} \tag{4.27}$$

from (4.23). It is clear on inspection that $\mathrm{Var}_1(z_\alpha^*)$ is always greater than or equal to $\mathrm{Var}_0(z_\alpha^*)$ regardless of model parameter values. But, from (4.22) we have for both H_1 and H_0 that $\mathrm{Cov}(z_\alpha^*, z_\beta^*) = \lambda_\alpha \delta_{\alpha\beta}$. This proves

Theorem 4.4. The correlation between a noisy linear-uncertain observer and an electronic energy detector which receives only noise is never less on noise-alone trials than on signal-plus- noise trials.

The importance of the theorem stems from the fact that the opposite result is expected from the filter bank model discussed in Chapter I. The filter bank model assumes that the human observer responds to the maximum of outputs from narrow-band filters in the filter bank. A relatively narrow band electronic filter will respond to components of the noise in a narrow frequency band around the center frequency of the signal and so should be only poorly correlated with the observer's response on noise-alone trials. On the other hand, when signal is presented to the observer the maximum output of his filters will nearly always occur in the passband of the energy detector and thus increase the communality of the two observers.

It may also be argued that if the signal is presented to a narrow band energy detector on signal-plus-noise trials then both the filter bank model and the linear-uncertain model predict increased correlation on signal trials. Thus, that experiment does not provide the comparison between the models afforded by giving the energy detector only noise. An experiment using the special conditions of Theorem 4.4 would be necessary only if the linear-uncertain model can predict the results of Ahumada's experiment reported in Chapter I. We now show that it can.

The squared correlation between the linear-uncertain observer and an energy detector with identical inputs on signal-plus-noise trials is

$$\rho_1^2 = \frac{[\lambda_\alpha(\delta_{\alpha\beta} + d_s) + \overline{\lambda}_\alpha r_{\alpha s}]^2}{\mathrm{Var}_1(z_\alpha^*)(\delta_\beta + d_s)} \tag{4.28}$$

On noise-alone trials this reduces to

$$\rho_0^2 = \frac{(\lambda_\alpha \delta_{\alpha\beta})^2}{\mathrm{Var}_0(z_\alpha^*)\delta_\beta} \tag{4.29}$$

Since $\delta_{\alpha\beta} = \min(\delta_\alpha, \delta_\beta)$ and $\mathrm{Var}_0(z_\alpha^*)$ does not depend upon δ_β, ρ_0^2 is maximized by setting $\delta_\beta = \delta_\alpha$. For ρ_1^2 it is clear on inspection that the expression will be maximized by taking $\delta_\beta \leq \delta_\alpha$. So we may set $\delta_{\alpha\beta} = \delta_\beta$. To investigate further we differentiate the logarithm of ρ_1^2 with respect to δ_β (ignoring the constant $\mathrm{Var}_1(z_\alpha^*)$):

$$\frac{\partial}{\partial \delta_\beta} \{2\, \ln[\lambda_\alpha(\delta_\beta + d) + \bar{\lambda}_\alpha r_{\alpha s}] - \ln(\delta_\beta + d_s)\} =$$

$$\frac{2\lambda_\alpha}{\lambda_\alpha(\delta_\beta + d_s) + \bar{\lambda}_\alpha r_{\alpha s}} - \frac{1}{\delta_\beta + d_s} .$$

By setting the expression equal to zero and manipulating, we obtain

$$\delta_\beta = \frac{\bar{\lambda}_\alpha}{\lambda_\alpha} r_{\alpha s} - d_s$$

as the value of δ_β which maximizes ρ_1^2 (as long as the right hand term is less than δ_α). If $r_{\alpha s} = d_s$, as in the noisy memory model, then the solution is reasonable if the linear-uncertain observer has somewhat less memory noise M_0 than N_0, since $\bar{\lambda}_\alpha/\lambda_\alpha = N_0/M_0$.

4.7 The Performance of Linear-uncertain Observers

It has been found empirically that the ROC curves of human observers often appear rather like straight lines when plotted on double-probability paper, that is when the coordinates P(H) and P(F) are transformed to deviation scores for a standard normal distribution. Such ROC curves would be exactly straight when there exists a monotone transformation ψ of z_α such that the conditional distributions $F_0[\psi(z_\alpha)]$ and $F_1[\psi(z_\alpha)]$ are both normal. It has been shown that in a task with signal specified exactly, strictly speaking, there exists no such transformation ψ unless the observer is equivalent to a linear processor of the input (Wilcox, 1967).

To a first order of approximation, however, we may take the observer's decision variable z_α as being normally distributed with mean $E(z_\alpha)$ and variance $V(z_\alpha)$. Generally the means and variances will be different on H_1 and H_0 trials.

Theorem 4.5. Let an observer have a decision variable z which is distributed normally with mean μ_1 and variance σ^2 on H_1 trials and with mean μ_0 and variance σ_0^2 on H_0 trials. Then, the area under the ROC curve generated by z is given by

$$P_A = \int_{-\infty}^{d'_z/\sqrt{2}} \phi(t)\ dt, \tag{4.30}$$

where $\phi(t)$ is the standard normal density function and

$$d'_z = \frac{\sqrt{2}(\mu_1 - \mu_0)}{(\sigma_0^2 + \sigma_1^2)^{1/2}} \tag{4.31}$$

The proof is made straight-forward by inserting the appropriate normal densities into Equation 4.3 and making a simple change of variables in order to reverse the order of integration.

It may be noted that if the observer is the ideal observer we find that

$$d'_z = \sqrt{\frac{2E_s}{N_0}} = d'_s = d_s^{1/2} \tag{4.32}$$

as is well known.

We shall call d'_z obtained from (4.30) the *sensitivity index** That is if P_A is known

---

* Jeffress (1967) has defined index d_z which is d'_z of Equation 4.31. I prefer the more general definition of d'_z of Equation 4.33, of which d_z is a special case.

$$d'_z = \sqrt{2}\Phi^{-1}(P_A) \tag{4.33}$$

where Φ is the standard normal distribution function.

Theorem 4.6. The sensitivity index of the noisy linear-uncertain obeservers with decision variable z^*_α is given approximately by

$$d'_{z^*_\alpha} = \frac{\sqrt{2}[\lambda_\alpha d_s/2 + \bar{\lambda}_\alpha r_{\alpha s}]}{[\mathrm{Var}_0(z^*_\alpha) + \mathrm{Var}_1(z^*_\alpha)]^{1/2}} \tag{4.34}$$

where Var_0 and Var_1 are given by (4.27).

The theorem follows from the fact, proved in Appendix II, that

$$\begin{aligned} E(z^*_\alpha) = \lambda_\alpha[\delta_\alpha(1 + V_{\alpha 1}) + d_{s_\alpha}/2] \\ + \bar{\lambda}_\alpha[r_{\alpha s_\alpha} - d_\alpha/2 - \delta_\alpha V_{\alpha 2}] \end{aligned} \tag{4.35}$$

The *efficiency* η_α of an observer α is defined as a ratio of signal energies (Tanner and Birdsall, 1958):

$$\eta_\alpha = E_{z_\alpha}/E_s \tag{4.36}$$

where E_{z_α} is the signal energy necessary for the ideal observer to perform at the same overall level as observer α. We shall take this to mean that E_{z_α} is the signal energy necessary for the area under the ROC curve of the ideal observer to be equal to the area under the ROC curve of an observer α who has signal energy E_s in his detection task. Then, for the case in which signal is specified exactly, E_{z_α} may be found from

$$d'_{z_\alpha} = \sqrt{2E_{z_\alpha}/N_0} \ .$$

Thus,

$$\eta_\alpha = (d'_{z_\alpha}/d'_s)^2. \tag{4.37}$$

Plots of P(C) (the probability of a correct response inaa two-interval forced-choice task), P_A, log η_α, or log $(d'_{z_\alpha})^2$ versus log d_s are examples of psychometric functions. The first is most commonly used, although the latter two are definitely superior for comparing processing models.

Birdsall (1960) has investigated the shape of psychometric functions of noisy-memory observer in the case $2W_\alpha T = 1$. McGill (1967) has obtained the psychometric functions for the internal-noise-free energy model for several values of $2W_\alpha T$.

4.8 The Relation Between Performance and Concordance

Theorem 4.7. Let η_α be the efficiency of a linear observer α with no reference signal noise ($V_{\alpha 2} = 0$) and let ρ be the correlation between α and the ideal observer. Then,

$$\eta_\alpha = \rho^2. \tag{4.38}$$

Proof. ρ is given by (4.25) with $d_{s_\beta} = d_s$, $V_{\beta 1} = V_{\beta 3} = 0$. η is given by the square of (4.34) divided by d_s and with $\lambda_\alpha = 0$, $V_{\alpha 2} = 0$. The equality (4.38) follows.

Apparently the efficiency of the noisy linear-uncertain observer in the general case cannot be expressed solely in terms of the linear correlation with the ideal observer.

Theorem 4.8. Let $\rho_i = \mathrm{Corr}(z^*_\alpha, z^*_\beta | H_i)$, $V_{\alpha i} = \mathrm{Var}(z^*_\alpha | H_i)$, $i = 0,1$. Then

$$\rho_1\sqrt{V_{\alpha 1}V_{\beta 1}} - \rho_0\sqrt{V_{\alpha 0}V_{\beta 0}} = \lambda_\beta\left[\frac{V_{\alpha 0} + V_{\beta 0}}{2}\right]^{1/2} d'_{z_\alpha} + \lambda_\alpha\left[\frac{V_{\beta 0} + V_{\beta 1}}{2}\right]^{1/2} d'_{z_\beta} \tag{4.39}$$

The theorem is proved by writing out (4.39) in detail for noisy linear-uncertain observers. The theorem provides a way to predict inter-observer correlation differences from the observer sensitivity indices and observer correlations with the ideal observer. First note that $V_{\alpha 0}/V_{\alpha 1} = v_\alpha$ is the slope of the ROC curve plotted on double probability paper, as long as the straight line approximation is good (Green and Swets, 1966, Ch. 3). By dividing both sides of (4.39) by $\sqrt{V_{\alpha 1}V_{\alpha 0}}$ we obtain

$$\rho_1 - \rho_0\sqrt{v_\alpha v_\beta} = \frac{\lambda_\beta}{\sqrt{V\beta_1}}\left[\frac{v_\alpha + 1}{2}\right]^{1/2} d'_{z} + \frac{\lambda_\alpha}{\sqrt{V_{\alpha 1}}}\left[\frac{v_\beta + 1}{2}\right]^{1/2} d'_{z_\beta} \qquad (4.40)$$

Since d'_{z_α}, d'_{z_β}, v_α and v_β may be estimated from performance data, the remaining unknowns are a single parameter in the form $\lambda_\alpha/\sqrt{V_{\alpha 1}}$ for each observer. Suppose that observer β is the ideal observer, which may be simulated by a cross-correlator when signal is specified exactly. Then $\lambda_\beta = 0$, $v_\beta = 1$ and $d'_{z_\beta} = \sqrt{d_s}$. In this case (4.40) can be written in the form

$$\frac{\lambda_\alpha}{\sqrt{V_{\alpha 1}}} = \frac{\rho_1 - \rho_0\sqrt{s_\alpha}}{\sqrt{d_s}} \qquad (4.41)$$

which provides an estimate of $\lambda_\alpha/\sqrt{V_{\alpha 1}}$ from the correlations of observer α with the ideal observer. These estimates may be inserted into (4.40) to provide a prediction of the quantity $\rho_1 - \rho_0\sqrt{v_\alpha v_\beta}$ for each pair of observers. Obviously such predictions should be considered only as good as the estimates of v_α and v_β.

CHAPTER V

AN EXPERIMENTAL INVESTIGATION OF INTER-OBSERVER CONCORDANCE

5.1 Predictions of the Models

An experiment was conducted to examine the effects of several input conditions on the level of concordance between observers in a single-interval detection task. The observers were three human observers and two electronic devices. The first device was the "ideal observer" (CC), that is, it computed the cross-correlation of the signal waveform with the noise waveform present in the presentation interval. The second device, the energy detector (ED), computed the energy of the filtered noise waveforms in the presentation interval. The human observers gave responses indicating their confidence that signal was present in the presentation interval. A nonparametric measure of concordance was computed on both H_1 and H_0 trials between all pairs of the five observers in each of the conditions of the experiment.

As discussed in Chapter IV, several models of monaural auditory processing make different predictions for the levels of concordance between observers. The major predictions for these experiments are:

i) Linear observers with non memory-specified reference-signal noise have equal correlations on H_1 and H_0 trials.

ii) If the noisy linear-uncertain observer has a linear component in his decision variable ($\lambda \neq 1$) then the absolute value of the correlation between the observer and the cross-correlator (with a noise-only input) is not zero and is greater on H_1 trials than on H_0 trials.

iii) Under the same conditions as (ii) the insertion of CW (carrier wave) into the noise cannot decrease the correlation between the observer and the cross-correlator (with noise-only input).

iv) The linear correlation between a filter-bank observer and the cross-correlator is zero on both H_1 and H_0 trials.

v) The correlation between a noisy linear-uncertain observer and the energy detector (with a noise-only input) must be greater on H_0 trials than on H_1 trials.

vi) The correlation between filter-bank observers and the energy detector should be greater on H_1 trials than on H_0 trials.

vii) Insertion of a continuous sinusoid (CW) into the noise background with the same frequency and phase as the signal should increase the correlation between filter-bank observers.

5.2 Method

The experiment was conducted at the Sensory Intelligence Laboratory, The University of Michigan. It involved the presentation of a 1000 Hz tone pulse, the signal, in a background of Gaussian noise. Three observers listened monaurally (one ear) to identical inputs through earphones. On each trial a random selector determined whether signal was to be present (H_1) or not (H_0) in the noise background. The observers were asked to report on each trial their confidence that the hypothesis H_1 was correct. During the presentation interval the noise waveform was also gated to two electronic devices. The first device computed the cross-correlation of the noise with the signal. The second device computed the energy of the noise in a 50 cycle band centered at the signal frequency. The outputs of these devices were compared with observer reports on H_1 and H_0 trials separately.

Apparatus. The experiment was programmed on a system known as N.P. Psytar which has been extensively described elsewhere (Green, Birdsall, and Tanner, 1957). N.P. Psytar contains a white noise source, ocillators for producing tones, an automatic random number generator, amplifiers, attenuators, audio gates, and the necessary logic and timing circuits to completely automate the presentation of acoustic waveforms and record observer responses. N.P. Psytar was augmented in the present experiment by two analog multipliers with outputs fed to gatable analog integraters, a digital voltmeter (Hewlitt Packard) with printed output (Hewlett Packard) and response sliders one foot long attached by a dial cord to a ten-turn linear potentiometer. The trial type, H_1 or H_0, was punched automatically on computer cards.

A block diagram of the analog multiplier and integrator circuits is shown in Figure 5.1. The signal and noise sources shown were also used to generate the inputs to the earphones. The analog gates to the integrators were closed (ungrounded) simultaneously with the presentation interval. The gates across the integrator capacitors were opened at the onset of the presentation interval and remained open until the offset of the observers' response interval during which the digital voltmeter recorded the stored charge on the capacitors.

Following the response interval the digital voltmeter also recorded the position of the sliders by measuring the voltage drop across the slider potentiometers. The input to the digital voltmeter was determined by a stepping relay which was reset after each trial. The voltmeter was allowed 250 ms to stabilize on each reading before the result was printed.

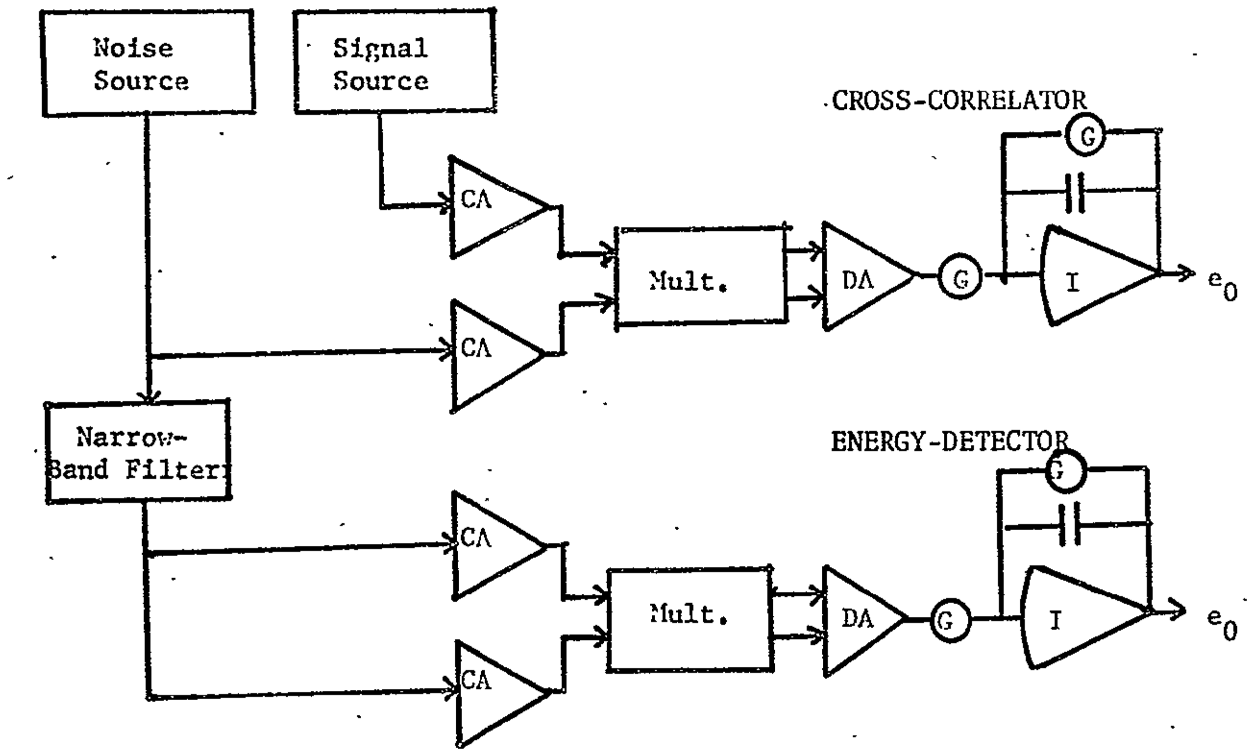

Fig. 5.1. Block diagram of Electronic Observers. CA is a current amplifier, DA a differential amplifier, I an operational amplifier in an integrator configuration, and G is an analog gate circuit. The differential output voltage of the multiplier is proportional to the product of the input currents.

The offset voltages of the integrators varied slightly from day to day but were quite stable during each two hour experimental session.

The narrow-band filter used in the energy detector was a single tuned passive filter with center frequency at 1000 Hz and 3 db power points at 974 Hz and 1025 Hz. Therefore, the 3 db bandwidth was $W_{3db} = 51$ Hz and the equivalent square band-pass was approximately $W = (\pi/2)W_{3db} = 80$ Hz.

Signal and noise levels. The detectability of a signal specified exactly in a background of white, Gaussian noise is appropriately measured by the index $d = 2E/N_0$ which is the square of the sensitivity index of an optimum observer in the task. E is the signal energy and N_0 is the noise power per unit bandwidth. The method used to measure the ratio is described in Green and Swets (1966, Appendix III). In the conditions with a signal duration of 100 ms, d was equal to 28.8, and for signal durations of 40 ms, $d = 28.5$. In conditions with CW added to the background noise the CW had a level of 14 db above the average noise power density N_0. The CW does not affect the computation of d since it is ignored by the optimum observer.

Subjects. Three female undergraduate students served as observers in the experiment. Observer 1 (OB 1) had served in a previous experiment which required confidence judgments. OB 2 was also an experienced observer, but had not previously used the confidence mode of response. OB 3 had had no previous experience as an observer in psychoacoustic experiments. The observers were paid at a base-rate per hour commensurate with their previous experience. In addition, bonus points were computed on each trial and were converted to money in such a way that for average performance an observer could, in effect, increase her hourly wage by 50%. However, all bonus payments were made after the completion of the experiment contingent upon continued attendance.

Payoffs. The bonus points were computed on each trial according to the following formula.

H_1 trials: $B = 1 - (1 - r^2)$;

H_0 trials: $B = (1 - r^2)$,

where B is the number of bonus points and r is the confidence rating as a proportion of distance along the slider scale. With this payoff scheme observers should maximize their subjective expected bonus by reporting their "true" subjective probability that signal was present on that trial. The bonus points for a day's session were usually reported to the observers on the following day before their next session.

Preliminary training. The observers were given four sessions each with 500 to 700 trials in which two response buttons were used. The responses were labeled "Yes" and "No" regarding signal occurrence. The two-button sessions were followed by eleven sessions in which four response buttons were used. The observers were instructed to use the buttons to indicate (1) "I am quite sure signal was presented", (2) "I am not certain, but I think the signal was presented", (3) "I am not certain, but I think the signal was not presented", and (4) "I am quite certain that signal was not presented". In order to acquaint the observers with the way bonus points would be computed when they used the slider a modified bonus scheme was used in the four-button sessions. Also, the observers were told to keep in mind that later on in the experiment they would be using a continuous slider to indicate their confidence and that they should think of the four buttons as being approximately evenly spaced along the scale of the slider. These sessions continued until it was ascertained that the *a posteriori* probability of H_1 given a response was a monotone function of the button number for each observer for four days in a row. The observers had some experience with each of the four experimental conditions described below. Following the four button sessions, the observers were given two sessions with the slider response before the data reported here were obtained.

Procedure. A trial, from an observer's point of view, consisted of four intervals in time, each marked by a separate neon indicator. The duration of a trial was about 5.6 seconds. The first 500 ms was a "get ready" period. The presentation interval of 100 ms in conditions I and III or 40 ms in conditions II and IV immediately followed. A two second response interval followed during which the observer was to position the slider to indicate her confidence that signal was presented.* After eleven slider sessions the observers complained that the response interval was too long. For the remaining five sessions the response interval was decreased to 1 sec so that the total trial duration became about 4.6 sec.

Each day's session consisted of five to seven blocks of trials. Each block consisted of 100 trials after which the observers were given a one to two minute rest. Halfway through each session observers were given a ten to fifteen minute break. Each session lasted approximately 1 1/2 hours.

Sixteen experimental sessions were conducted using the slider response. The first two sessions were for practice (from the experimenter's point of

---

* Use of a mechanical analog to a continuous rating response has been used previously by Watson, Rilling, and Bourbon (1964).

view). The data of two other sessions were omitted from analysis because of equipment problems. The remaining twelve sessions were divided into three sessions each for four experimental conditions.

Condition I had a presentation interval of 100 ms and no continuous sinusoid ($\overline{CW}$) added to the background noise. Condition II had a 40 ms presentation interval with $\overline{CW}$. Condition III had $T = 100$ ms and CW. Condition IV had $T = 40$ ms and CW. The three sessions of a condition were run on sequential days except for session 3 of condition I which was the last of the experimental sessions.

Preliminary data analysis. The trial-by-trial responses of each observer and device, which were read by the digital voltmeter, were punched on computer cards and analysed by a preliminary data analysis program on an IBM 7090 computer. This program converted the response values to standard scores for each observer and device. These scores were separated into two groups corresponding to H_1 trials and H_0 trials, respectively. The scores for the three sessions for each condition were then merged. The program generated many cut-points for the data and determined the frequency distribution of scores for each observer and device. From these distributions it could be determined which cut-points would give approximately ten equally probable categories of response. The program then obtained the ten joint frequency distributions for each pair of observers and devices for each of the four conditions on H_1 and H_0 trials separately. Thus, a total of $80 = 10 \times 8$ joint frequency tables were obtained in this way. The joint frequency tables were used to determine the level of concordance between observers and devices.

In order to obtain empirical ROC curves the cut-points were determined which would give approximately ten equally probable categories of response regardless of H_1 or H_0. The program then determined the $2 \times m$ ($m = 9$ or 10) frequency table of hypothesis (H_1 or H_0) versus response category for each observer and each device for each of the four conditions. From these data tables the empirical ROC curves, the area estimate and several other performance measures could be determined.

5.3 *Results*

Observer performance. The ROC curves for each of the four conditions are presented separately for each observer in Figures 5.2, 5.3 and 5.4. Each curve is based on approximately 1,800 trials. The ROC curves for the two electronic devices are not shown since they fell nearly perfectly along the chance line as expected. (Both devices, it will be recalled, were presented only the noise waveform sample on each trial.)

The shape of the observer ROC curves indicate that they would be only poorly approximated by a straight line on double probability paper. If any generalization can be made it would be that the curves differ from straight lines with unity slope by having a slightly smaller slope and are somewhat concave toward the chance line. Also the curves appear quite similar in shape across conditions. This conclusion is most pronounced for OB3 who had the highest efficiency in all conditions.

The performance of all observers is better when CW is present. Observers 1 and 3 show better performance with $T = 40$ ms when no CW is present, but show better performance at $T = 100$ ms when CW is present. There appears to be little effect of duration on the performance of observer 2.

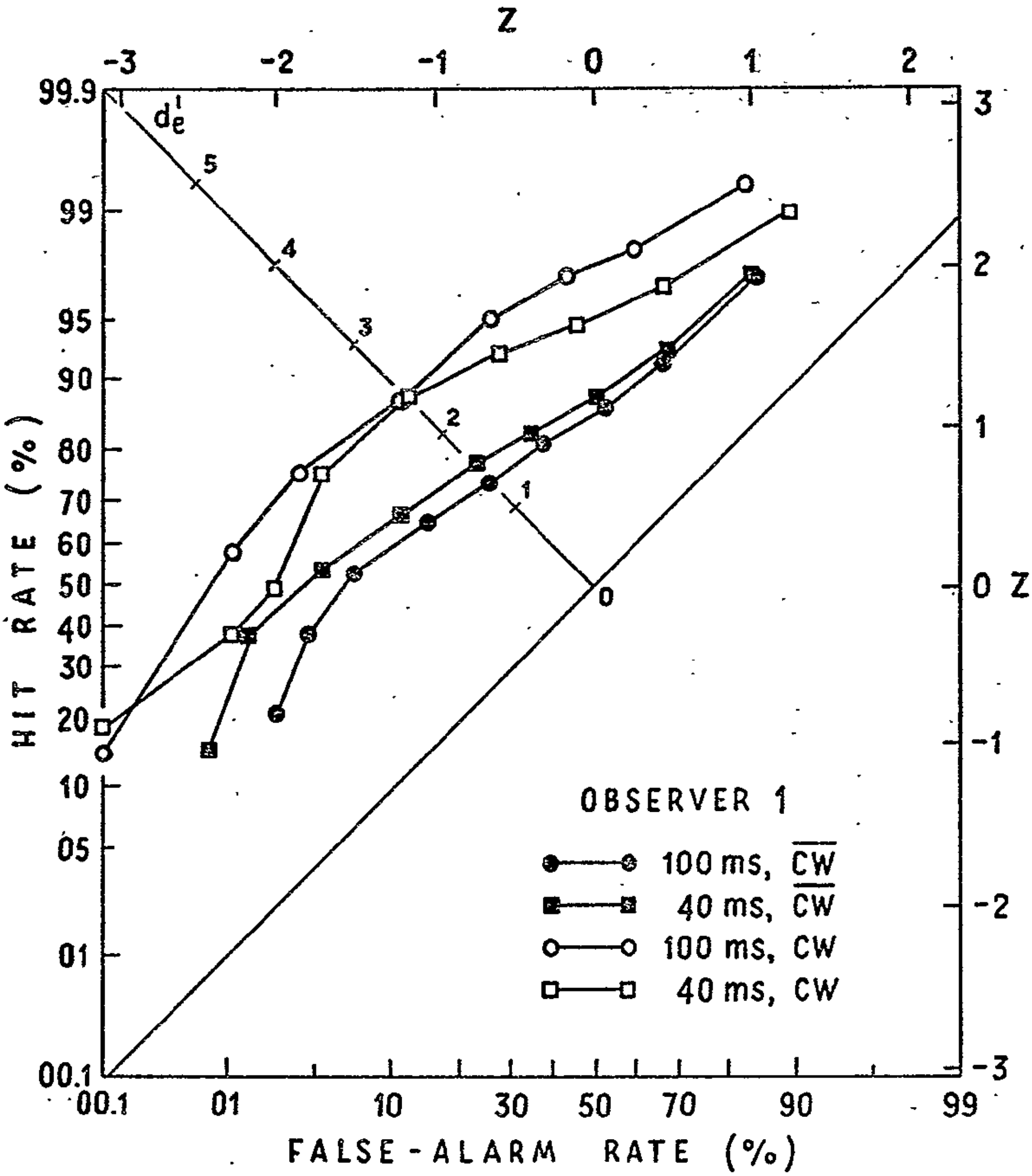

Fig. 5.2. Empirical ROC curves for observer 1. The curves are based on 1,800 trials each and are plotted on double normal-probability paper. The line of chance performance (P(H) = P(F)) in the task runs from the lower left-hand corner to the upper right of the figure.

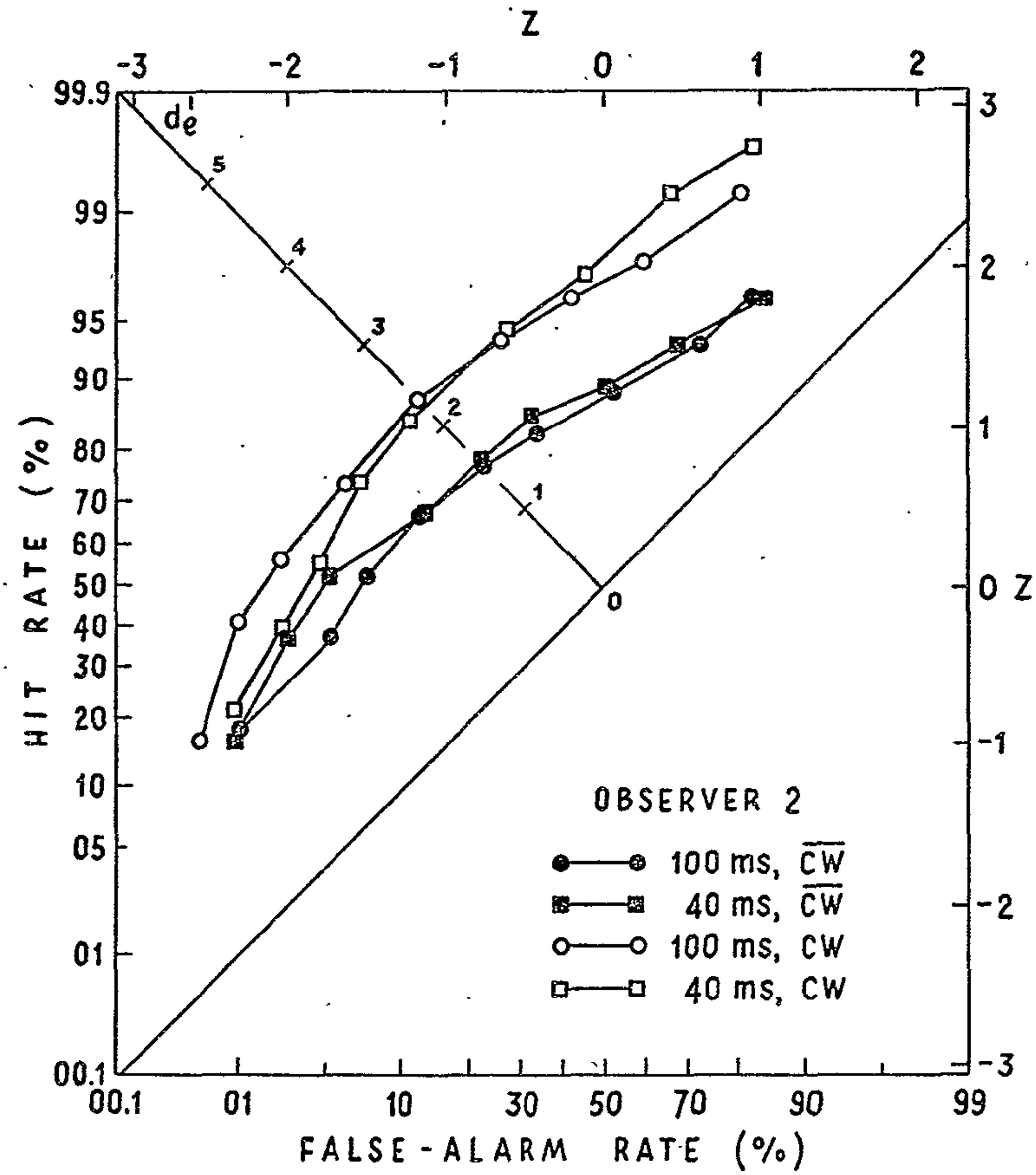

Fig. 5.3. Empirical ROC curves for observer 2.

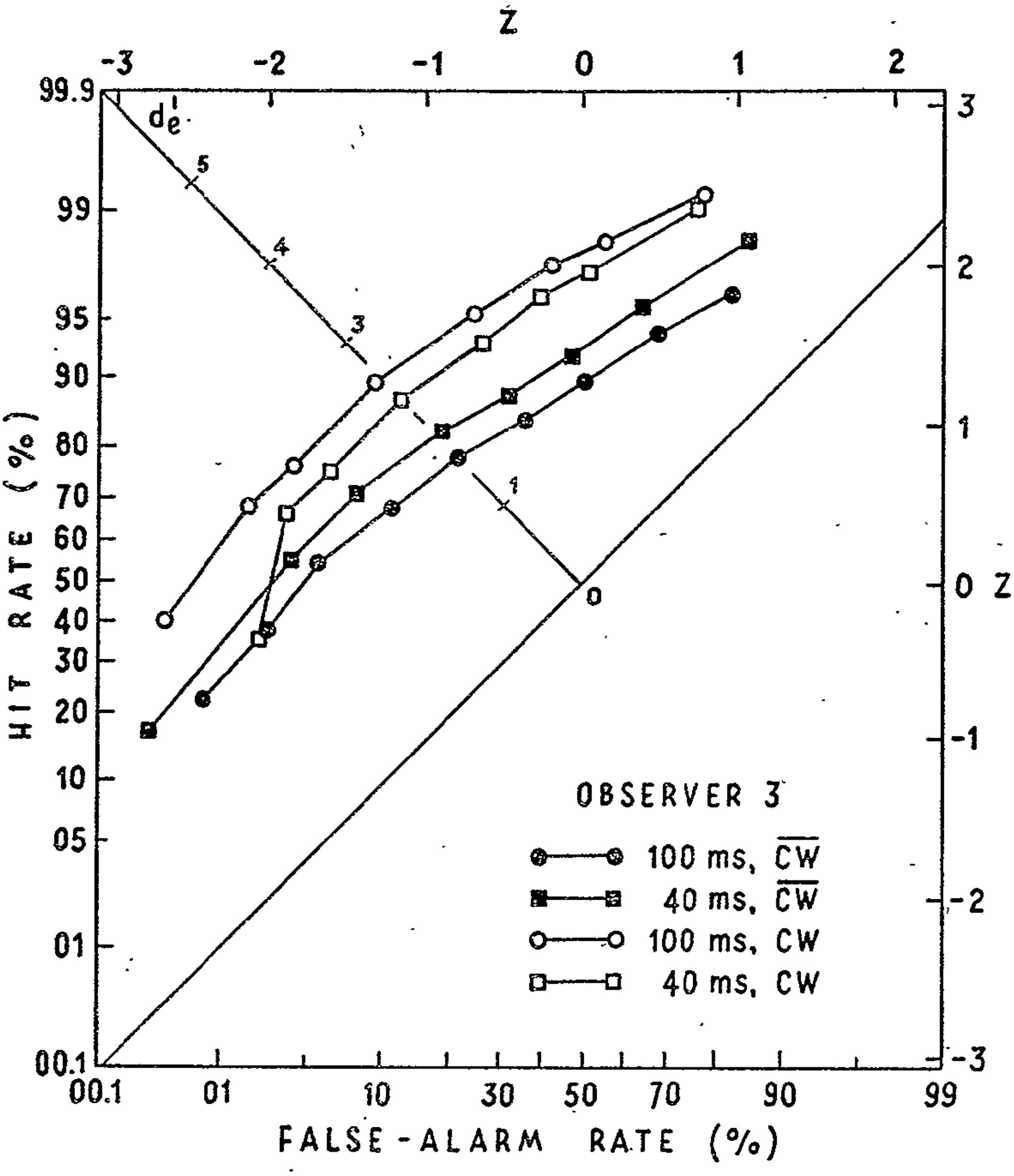

Fig. 5.4. Empirical ROC curves for observer 3.

A comparison of measures of observer performance is given in Table 5.1 for each observer under each experimental condition. The estimate of the area under the ROC curve based on the conditional correlation between signs, est P_A, was computed according to Equation A1.17 in Appendix I. Trap P_A is the estimate of the area based on using the trapezoidal rule for integration and was computed using Equation A1.18. d_z' is the sensitivity index computed from est P_A using Equation 4.33. d_e' is an index of performance which may be defined as $2\Phi^{-1}(P)$, where P is the hit-rate at which the ROC curve crosses the negative diagonal. Finally η is the efficiency of an observer computed from d_z' and $d_s = 2E/N_0$ according to Equation 4.37.

For all observers and conditions trap P_A is a little less than est P_A as it should be. For the fairly large number of cut-points used here the trapezoidal estimate is only smaller by about 1%.

The values of d_e' are generally somewhat larger than d_z'. This constitutes partial confirmation of the generalization that the ROC curves are concave downward.

There is an interaction effect of CW with duration. The efficiencies are smaller for $T = 100$ ms than for $T = 40$ ms when no CW is present. This inequality is reversed when CW is present. The introduction of CW also substantially improves the performance of the observers.

It was hypothesized that the use of the continuous rating response might have a depressive effect on observer performance. If so, the observers should have shown somewhat better performance in the preliminary training sessions where only four response buttons were used. No such consistent differences were apparent using the index d_e'.

TABLE 5.1

A COMPARISON OF MEASURES OF OBSERVER PERFORMANCE

	Condition			
	I	II	III	IV
	$\overline{CW}$		CW	
Observer 1	100	40	100	40
est P_A	.818	.850	.953	.930
trap P_A	.807	.839	.945	.923
d'_z	1.28	1.47	2.37	2.08
d'_e	1.31	1.45	2.44	2.32
η	.057	.076	.195	.152
Observer 2				
est P_A	.841	.855	.940	.936
trap P_A	.830	.844	.931	.927
d'_z	1.41	1.50	2.20	2.16
d'_e	1.51	1.61	2.23	2.26
η	.069	.079	.168	.164
Observer 3				
est P_A	.861	.897	.960	.937
trap P_A	.851	.886	.953	.929
d'_z	1.53	1.79	2.48	2.16
d'_e	1.54	1.86	2.54	2.31
η	.081	.113	.213	.164

Observer Concordance. Values of tau, corrected for ties, were computed from Equation A1.13 using the 10×10 joint frequency tables described above. The observed values of tau between the cross-correlation and the other observers are presented in Table 5.2 for each condition. Maximum confidence intervals for each value were computed (Kendall, 1962, Eq. 4.12). In Table 5.2 all but three values have 50% confidence intervals which include zero. For the remaining three values 60% confidence intervals include zero. These intervals are generally considered quite conservative so that it may still be worthwhile to look for systematic trends in the data.

There seems to be a very small but persistent positive concordance between the cross-correlator and the energy detector. An investigation of the noise source showed a slight skewness in the noise distribution around the zero amplitude value. This non-linearity in the noise waveform appeared to be the most likely cause of the slight degree of correlation found.

For observers 2 and 3 the values of tau are a little higher on H_1 trials than H_0 when CW is not present. Even this inequality is reversed for OB2 when CW is present. Thus there appears to be no systematic basis upon which to accept hypothesis (ii) in Section 5.1. Furthermore, there appears to be no clear increase in the correlations when CW is present as predicted in hypothesis (iii). Thus, we may conclude that the data provided by the correlations between the cross-correlator and the human observers can be explained either by the noisy linear-uncertain model with no linear component, i.e., a noisy energy detector, or by the filter bank model.

TABLE 5.2

OBSERVED TAU VALUES BETWEEN THE CROSS CORRELATOR

AND

THE OTHER OBSERVERS

		Condition			
		I	II	III	IV
		$\overline{CW}$		CW	
		100	40	100	40
CC vs ED					
	H_1	-0014	0838*	0242	0085
	H_0	0082	0225	0181	0162
CC vs OB_1					
	H_1	-0131	0052	-0119	-0197
	H_0	-0597*	-0056	0080	0082
CC vs OB_2					
	H_1	0160	0418*	-0086	-0086
	H_0	-0380*	-0160	-0273	-0273
CC vs OB_3					
	H_1	0132	0247	-0034	0288
	H_0	-0045	0015	0094	0293

* These values are significantly different from zero at the 0.5 level of confidence but not at $p < 0.4$.
Decimal points omitted.

The remaining tau values were converted to rough estimates of the linear correlation using Greiner's relation (Equation 4.17). The 50% confidence intervals for the tau values were also converted to correlation confidence intervals. The correlations between the energy detector and the human observers are presented in Table 5.3. Over half the correlations at $T = 40$ ms fail to be significantly different from zero at the 0.5 level of confidence, while none of the correlations for $T = 100$ ms are zero at this level. Thus, there is a pronounced decrease in the correlations for the shorter durations regardless of whether or not CW is present.

With no CW the correlations for H_1 trials are greater than for H_0 trials. The differences are significnat for OB1, but not for OB2 or OB3. There appear to be no other consistent differences between the correlations which are also reliable.

No specific hypothesis was offered in Section 5.1 to attempt to predict the effect of duration. This is because such predictions can only be made for the linear uncertain model by assuming specific parameter values. This topic is further discussed below. The weak evidence that the correlations are greater on H_1 than on H_0 trials, at least when no CW is present, tends to reject the noisy linear-uncertain model in favor of the filter-bank model according to predictions (v) and (vi).

The inter-observer correlations and their 50% confidence intervals are presented in Table 5.4. The relations between the correlations are highly stable, regardless of the pair of observers considered. With a single minor exception (OB1 vs OB3, Condition II vs IV) correlations on H_1 trials indicate *lower* correlations on H_0 trials within comparisons for an observer pair. The

TABLE 5.3

CORRELATIONS BETWEEN THE ENERGY DETECTOR AND THE HUMAN OBSERVERS†

		Condition			
		I	II	III	IV
		$\overline{CW}$		CW	
		100	40	100	40
ED vs OB1					
H_1	$\hat{\rho}$	272	129	325	-017*
	CI	(220,323)	(079,179)	(279,369)	(-068,033)
H_0	$\hat{\rho}$	151	026*	133	045*
	CI	(101,200)	(-024,075)	(085,180)	(-005,095)
ED vs OB2					
H_1	$\hat{\rho}$	204	056	166	-035*
	CI	(150,256)	(006,106)	(118,214)	(-085,016)
H_0	$\hat{\rho}$	166	-036*	094	057
	CI	(116,216)	(-085,013)	(046,141)	(007,107)
ED vs OB3					
H_1	$\hat{\rho}$	228	094	101	001*
	CI	(175,279)	(044,144)	(052,149)	(-049,051)
H_0	$\hat{\rho}$	217	048	164	027*
	CI	(168,265)	(-001,098)	(117,211)	(-023,077)

† Decimal points ommitted. Each estimate based on approximately 1800 trials. 50% confidence intervals (CI) are given in parentheses.

* N.S. for $p \leq 0.5$.

differences between correlations on H_1 and H_0 trials with $\overline{CW}$ are all quite significant ($p \leq .05$). The differences when CW is present are smaller but there are no violations of sign. Thus, the hypothesis (i) of linear observers with no reference signal noise is strongly rejected.

The H_1-H_0 correlation differences are larger with $\overline{CW}$ than with CW. This effect appears primarily due to a decrease in correlations on H_1 trials. However, with a single exception (the same as before) H_0 correlations are higher with CW present for corresponding durations. The decrease on H_1 trials present a serious difficulty for the filter-bank model according to hypothesis (vii). The CW signal lies in the center of the signal spectrum by construction. It could be argued that CW should increase the correlations more on H_0 trials than on H_1 trials since the maximum narrow filter output will be more uniformly distributed in frequency when no CW is present. However, it is inconceivable how introduction of CW could decrease the correlation on H_1 trials.

It may be inquired whether the relation between inter-observer correlations can be predicted from observer sensitivity measure as described in Section 4.8. The procedure described there required non-zero correlations with the cross-correlator which is doubtful considering the present data. It was decided, therefore, not to attempt the prediction. However, the equation (4.39) derived for linear-uncertain observer suggests that the weighted difference for correlations on H_1 and H_0 trials should be related to the weighted sum of the sensitivity indices. Examination of the present data shows that with a single exception (OB2 vs OB3, Condition III), $\rho_{\alpha\beta}^{(1)} - \rho_{\alpha\beta}^{(0)}$ is a monotone decreasing function of the sum $d'_{z_\alpha} + d'_{z_\beta}$ across

TABLE 5.4

INTER-OBSERVER CORRELATIONS*

		Condition I	II	III	IV
		$\overline{CW}$		CW	
		100	40	100	40
OB1 vs OB2					
H_1	$\hat{\rho}$	686	605	472	519
	CI	(650,719)	(567,640)	(431,512)	(477,558)
H_0	$\hat{\rho}$	313	326	381	328
	CI	(265,359)	(279,371)	(337,423)	(281,373)
OB1 vs OB3					
H_1	$\hat{\rho}$	736	691	498	597
	CI	(703,765)	(658,722)	(458,537)	(559,633)
H_0	$\hat{\rho}$	337	400	437	353
	CI	(290,383)	(355,443)	(395,478)	(307,398)
OB2 vs OB3					
H_1	$\hat{\rho}$	660	671	513	495
	CI	(623,695)	(636,703)	(473,551)	(453,536)
H_0	$\hat{\rho}$	250	291	382	373
	CI	(201,298)	(244,337)	(338,424)	(327,417)

* Decimal points omitted. Each estimate based on approximately 1800 trials. 50% confidence intervals (CI) are given parentheses.

conditions for each pair of obervers. I can give no interpretation to this result.

5.4 Discussion

The noisy linear-uncertain model was developed and proposed as a plausible alternative to the filter-bank model. The linear component in the linear uncertain model is a necessary ingredient in order to account for Ahumada's thesis findings as was seen in Chapter IV. Thus, for the linear uncertain model to remain a viable alternative it is of importance to consider the first three predictions of Section 5.1. The first, that correlations between observers are equal on H_1 and H_0 trials, is unequivacably rejected by the data. The conclusion is that human observers do not perform simple linear operations on the input; at least there must be noise in the memory-specified reference signal. The first part of the conclusion has been verified repeatedly using indirect comparisons. However, the possibility that the observer performs noisy linear operations on the input has never previously been investigated using empirical comparisons.

The predictions (ii) and (iii) are an attempt to face directly the possibility of a linear component in the observer's decision variable regardless of whether or not there is reference-signal noise. There appears to be little evidence of correlation with the cross-correlator at all. Thus, the further questions of whether the correlation is greater on H_1 trials than on H_0 trials or whether the correlation is increased by insertion of CW into the noise are irrelevant. A rejection of the hypothesis of a linear component in the decision variable serves to reject the noisy linear-uncertain model as well(in connection with Ahumada's findings). An attempt can be made to explain the small correlations with the cross-correlator by assuming that the linear component is present but has a small weight, i.e., that $\overline{\lambda}_\alpha << \lambda_\alpha$. However, this is not sufficient to save the model. As was seen at the end of Section 4.7 $\overline{\lambda}_\alpha$ should be somewhat greater than λ_α for the explanation of Ahumada's finding to be reasonable. Another possibility for saving the linear component is to assume that the reference-signal noise $V_{\alpha 2}$ is quite large. But $V_{\alpha 2}$ cannot be very large before the correlations among observers are depressed. However, there appears to be no way to positively reject this latter assumption with the present data.

The linear-uncertain model, temporarily preserved by the assumption of considerable reference signal noise, must still meet the fifth prediction, namely, greater correlation on H_0 trials than on H_1 trials with the energy detector. No significant differences were found in this direction although some differences in the opposite direction were weakly significant. Thus, the correlations with the energy detector also provide some evidence for rejecting the linear-uncertain model.

The evidence provided by the correlations with the cross-correlator, the energy detector and Ahumada's variable bandwidth energy detector lead to the conclusion that the noisy linear-uncertain model, including its special cases, is rejected as an adequate model of human monaural auditory processing.

The filter-bank model survives predictions (iv) and (vi) that observer correlation is zero with the cross-correlator and that the correlations with the energy detector are greater on H_1 trials than on H_0 trials, respectively. However, the final prediction, that inter-observer correlations should increase with the insertion of CW into the background noise, is not verified. The insertion of CW does cause a slight increase on H_0 trials but also a pronounced decrease on H_1 trials. Although this is the first clear evidence against the filter-bank model it cannot be taken lightly. The prediction is strong even though formal development of the filter-bank model has not been made.

It might be argued that the decrease in correlations with the insertion of CW is an artifact caused by the improved performance of the observers. Since they were asked to give their confidence that they would be right if they had said H_1, one would expect a greater concentration of responses near the ends of the slider in CW conditions. This concentration could cause a reduction in the association between the observers' responses. An examination of the response distributions across the slider scale did show increased grouping of responses towards the ends of the slider. However, if the drop in association is caused by this then tau computed on 2×2 joint response tables with marginal cut-points at the medians should be relatively unaffected or increase (according to prediction (vii)). These tau values were computed, as usual, correcting for ties. With $\overline{CW}$ there was a slight increase in the range of inter-observer correlations between H_1 and H_0 trials. In the CW conditions, rather than an increase, both correlations on H_1 and H_0 trials decreased slightly with the range staying approximately the same as when computed on the 10×10 joint frequency tables. Thus, it is concluded that the result is not an artifact of response grouping.

There are two experimental results which elude explanation by the models considered so far: (a) the decrease in observer concordance with the energy detector when the signal duration is decreased, and (b) the lower inter-observer concordance on H_1 trials when CW is present.

The first can be given an *ad hoc* explanation in terms of a modification of the simple energy or envelope model. The modification is described by Jeffress (1967). His "leaky integrator model" has a narrow-band filter followed by a square-law detector (or perhaps a linear rectifier). The output is exponentially weighted with a fixed rate of decay and integrated continuously in time. It is presumed that the observer's decision variable is the value of the integral at the termination of the signal presentation interval. Since the rate of decay is constant (about 100 ms according to Jeffress) short signal presentation intervals will cause part of the noise waveform not in the presentation interval to be integrated into the decision variable. However, the energy detector of the present study has an integration time equal to the duration of the presentation interval. Thus, the decrease in the amount of common noise for the observers and the energy detector could cause the decrease in concordance at shorter durations.

The leaky integrator model, of course, suffers the same difficulties as the linear-uncertain model in explaining Ahumada's finding and result that correlations with the energy detector on H_1 trials are greater than on H_0 trials. Further, Jeffress' model apparently cannot account for the decrease in inter-observer correlations with the introduction of CW at constant durations.

The leaky-integrator model and the filter-bank model can be combined. The leaky integrator computes a short-term power-like quantity for the output of a single narrow-band filter. Assuming that the bandwidth is somewhat smaller than the equivalent square band-width of 90 Hz estimated by Jeffress, a bank of such filters could be postulated. The filter-bank in this case is computing a portion of the short-term power spectrum of the input. It is reasonable to extend Ahumada's decision variable to be the maximum output in time as well as frequency. I shall call this model the *leaky filter-bank* model.

The leaky filter-bank model may be viewed as appending a particular decision rule to a processor which continuously computes the short-term, frequency-limited power spectrum of the input. Such a processor could also be implimented by taking the Fourier transform of the short-term auto-correlation function of the input. Thus, the processor of the leaky filter-bank model is quite similar to a suggestion by Licklider (1951) that the ear performs a short-term autocorrelation of the input in monaural listening tasks.

In summary, the noisy linear-uncertain model is unable to account for most of the results from direct comparisons between observers. A leaky form of Ahumada's filter bank model is able to give a qualitative account of most of the data from direct comparisons, but has serious difficulty with the finding that the concordance between observers decreases on signal trials when CW is added to background noise.

5.5 *Summary*

Predictions of the level of correlation between observers derived from the linear-uncertain model and the filter-bank model are compared for several experimental conditions. In the experiment the decisions of human observers are compared with the outputs of two electronic devices. The first device is an analog multiplier which computes the cross-correlation between the signal and the noise waveform sample on each trial. The second device is an energy detector which computes the energy of the noise waveform sample during the presentation interval in a narrow frequency band centered at the signal frequency.

The linear-uncertain model predicts that the correlation between the human observers and the cross-correlation is not zero and should increase when a continuous sinusoid is added to the background noise. Neither of these predictions is verified. Since the energy detector receives only noise at its input in this experiment, the model also predicts that the correlation between the observers and the energy detector should be less on trials when signal is present than when it is not present. The results show the correlations to be weakly significant in the opposite direction. It is concluded that the linear-uncertain model and its special cases, the linear, energy, and envelop models, represent an inadequate approximation to the actual form of human monaural auditory processing in detection tasks.

Predictions from the filter-bank model agree with the above results, but cannot account for an observed decrease in inter-observer correlations when the signal presentation interval is shorted. A modification of the filter-bank model is suggested to account for this discrepancy.

A final result remains unexplained by any of the models considered. A decrease in inter-observer correlations is found when a continuous sinusoid is added to the background noise for either of two signal durations. It is emphasized that this unexpected finding implies that there is a serious deficiency in current models of monaural auditory processing in detection tasks.

APPENDIX I

A NONPARAMETRIC ESTIMATE OF THE AREA UNDER THE ROC CURVE

A1.1 Theory

The notation and definitions of Sections 4.3 and 4.4 will be used here.

With sample decision values or response categories ties can occur. The marginal probabilities of a tie are given by

$$P(T_\alpha) = P(a_{ij} = 0)$$
$$P(T_\beta) = P(b_{ij} = 0).$$

Then it is easily shown that

$$\mathrm{Var}(a_{ij}) = 1 - P(T_\alpha)$$
$$\mathrm{Var}(b_{ij}) = 1 - P(T_\beta),$$

whereas the expression for the covariance is unaffected by ties:

$$\mathrm{Cov}(a_{ij}, b_{ij}) = P(S) - P(D),$$

as in (4.9).

Therefore the estimate of tau (preferred by Kendall, 1962) is given by

$$\begin{aligned} \text{est } \tau &= \text{est Corr}(a_{ij}, b_{ij}) \\ &= \frac{P(S) - P(D)}{\sqrt{[1 - P(T_\alpha)][1 - P(T_\beta)]}} \end{aligned} \tag{A1.1}$$

where the probabilities refer to observed relative frequencies. It has been found that tau estimated in this way is often relatively unaffected by grouping of the variables. That is, if a pair of continuous variables are categorized into 5 or 10 equally probable values we may expect the three corresponding estimates of τ to be quite close in numerical value. Of course, it is impossible to obtain consistent estimates of tau using grouped data unless the underlying continuous bivariate distribution is known and the grouping procedure is under the control of the statistician.

The estimate of gamma used by Goodman and Kruskal (1964) can be viewed in exactly the same way as Kendall's estimate of tau. Since

$$\gamma = \mathrm{Corr}(a_{ij}, b_{ij} | a_{ij} \neq 0,\ \ b_{ij} \neq 0)$$

the estimate is given by

$$\text{est } \gamma = \text{est Corr}(a_{ij}, b_{ij} | a_{ij} \neq 0,\ \ b_{ij} \neq 0)$$

$$= \frac{P(S) - P(D)}{1 - P(T)}$$

where $P(T) = P(T_\alpha \cup T_\beta)$.

In connection with Theorem 4.1 these estimates suggest an appropriate estimator for the area under the ROC curve which attempts to correct for the grouping of an observer's decision variable. The theorem proved that in case the observer's decision variable is continuous

$$P_A = (\gamma + 1)/2.$$

Since z_α is continuous a_{ij} cannot equal 0. Thus, in this special case

$$\gamma = \mathrm{Corr}(a_{ij}, b_{ij} | b_{ij} \neq 0).$$

According to the preceding estimates we should take

$$\text{est } P_A = \tfrac{1}{2}[\text{est Corr}(a_{ij}, b_{ij} | b_{ij} \neq 0) + 1].$$

Now

$$\mathrm{Cov}(a_{ij}, b_{ij} | b_{ij} \neq 0) = \frac{P(S) - P(D)}{1 - P(T_\beta)} \quad ,$$

$$\mathrm{Var}(b_{ij} | b_{ij} \neq 0) = 1,$$

and

$$\mathrm{Var}(a_{ij} | b_{ij} \neq 0) = 1 - P(T_\alpha | b_{ij} \neq 0)$$

$$= 1 - P(T_\alpha | \overline{T}_\beta).$$

Therefore

$$\text{est } P_A = \frac{1}{2}\left[\frac{P(S) - P(D)}{\sqrt{1 - P(T_\alpha|\overline{T}_\beta)[1 - P(T_\beta)]}} + 1\right]. \tag{A1.2}$$

We can obtain a somewhat simpler expression to work with using the following identities.

$$P(T_\alpha|\overline{T}_\beta) = \frac{P(T_\alpha,\overline{T}_\beta)}{1 - P(T_\beta)}$$
$$= \frac{P(T_\alpha) - P(T_\alpha,T_\beta)}{1 - P(T_\beta)}.$$

But

$$P(T_\alpha,T_\beta) = P(T_\alpha) + P(T_\beta) - P(T)$$

so

$$P(T_\alpha|\overline{T}_\beta) = \frac{P(T) - P(T_\beta)}{1 - P(T_\beta)}$$

and

$$1 - P(T_\alpha|\overline{T}_\beta) = \frac{1 - P(T)}{1 - P(T_\beta)}.$$

Thus (A1.2) becomes

$$\text{est } P_A = \frac{1}{2}\left[\frac{P(S) - P(D)}{\{[1 - P(T)][1 - P(T_\beta)]\}^{1/2}} + 1\right]. \tag{A1.3}$$

A1.2 Computation

In this section I develop a common notation for the computational formulas*of est P_A, est τ and the trapezoidal estimate of the area.

Let the response categores for two observed decision variables R and R' be R_1, R_2, ..., R_m and R'_1, R'_2, ..., R'_n, respectively, where R_1 and R'_1 indicate the greatest confidence in H_1 and R_m and R'_n indicate the greatest confidence for the alternative hypothesis H_0. The observed joint frequency table is then in the following format:

---

* Kendall (1962) also gives a similar computational scheme for est τ.

	R'_1	R'_2	$\cdots$	R'_n	
R_1	f_{11}	f_{12}	$\cdots$	f_{1n}	$f_{1\cdot}$
R_2	f_{21}	f_{22}	$\cdots$	f_{2n}	$f_{2\cdot}$
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	
R_m	f_{m1}	f_{m2}	$\cdots$	f_{mn}	$f_{m\cdot}$
	$f_{\cdot 1}$	$f_{\cdot 2}$		$f_{\cdot n}$	$f_{\cdot\cdot}$

To compute $P(S) = P(a_{ij}b_{ij} > 0)$ we may take each joint observation and compare it with any other observation in the table. The comparisons which contribute to $P(S)$ from an observation in a cell (i,j) are all those observations in cells below and to the right of (i,j). The total number of such observations is

$$\sum_{h=i+1}^{m} \sum_{\ell=j+1}^{n} f_{k\ell}.$$

Now the comparison is made with each of the f_{ij} observations in cell (i,j). Therefore the total number of comparisons in the numerator of $P(S)$ is

$$S = \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} f_{ij} \sum_{k=1}^{m} \sum_{\ell=1}^{m} f_{k\ell} \tag{A1.4}$$

The denominator is simply the total number of comparisons

$$N = f_{\cdot\cdot}(f_{\cdot\cdot} - 1)/2. \tag{A1.5}$$

Thus

$$P(S) = S/N. \tag{A1.6}$$

The probability of disagreement in signs is computed similarly as

$$P(D) = D/N, \tag{A1.7}$$

where

$$D = \sum_{i=2}^{m} \sum_{j=2}^{n} f_{ij} \sum_{k=1}^{i-1} \sum_{\ell=1}^{j-1} f_{k\ell}. \tag{A1.8}$$

Now the number of ties T_α in the decision variable R is simply

$$T_\alpha = \sum_{i=1}^{m} f_{i\circ}(f_{i\circ} - 1)/2 \tag{A1.9}$$

so that

$$P(T_\alpha) = T_\alpha/N. \tag{A1.10}$$

Similarly

$$T_\beta = \sum_{j=1}^{n} f_{\circ j}(f_{\circ j} - 1)/2 \tag{A1.11}$$

and

$$P(T_\beta) = T_\beta/N. \tag{A1.12}$$

Thus from (A1.1), after slight rearrangement of the N's, we have

$$\text{est } \tau = \frac{S - D}{\sqrt{(N - T_\alpha)(N - T_\beta)}} \tag{A1.13}$$

as the computational formula for est τ.

To estimate the area the ROC curve we use the data format given below

	R_1	R_2	$\cdots$	R_m	
H_1	f_{11}	f_{12}	$\cdots$	f_{1m}	$f_{1\circ}$
H_0	f_{01}	f_{02}	$\cdots$	f_{0m}	$f_{0\circ}$
	$f_{\circ 1}$	$f_{\circ 2}$	$\cdots$	$f_{\circ m}$	$f_{\circ\circ}$

Now S and D simplify to

$$S = \sum_{i=1}^{m-1} f_{1i} \sum_{j=i+1}^{m} f_{0i} \tag{A1.14}$$

$$D = \sum_{i=1}^{m-1} f_{0i} \sum_{j=i+1}^{m} f_{1j} \tag{A1.15}$$

The number of ties T in *either* variable is

$$T = N - (S + D). \tag{A1.16}$$

Now,

$$\begin{aligned} T_\beta &= [f_{1.}(f_{1.} - 1) + f_{0.}(f_{0.} - 1)]/2 \\ &= [f_{1.}^2 + f_{0.}^2 - (f_{1.} + f_{0.})]/2 \\ &= (f_{1.}^2 + f_{0.}^2 - f_{..})/2 \end{aligned}$$

Thus

$$\begin{aligned} 1 - P(T_\beta) &= \frac{N - T_\beta}{N} \\ &= \frac{f_{..}(f_{..} - 1) - (f_{1.}^2 + f_{0.}^2 - f_{..})}{2N} \\ &= \frac{f_{..}^2 - f_{1.}^2 - f_{0.}}{2N} \end{aligned}$$

The area estimate becomes, upon substitution and rearrangement,

$$\text{est } P_A = \frac{S - D}{\sqrt{2(S + D)(f_{..}^2 - f_{1.}^2 - f_{0.})}} + \frac{1}{2}\,. \tag{A1.17}$$

This estimate of the area is a distinct improvement over the one I proposed earlier (Wilcox, 1967) which was based on using the estimate of gamma given in Section A1.1.

For comparative purposes the computational formula for the area using the trapezoidal rule is given below

$$\text{trap } P_A = \frac{S + \frac{1}{2}\sum_{i=1}^{m} f_{1i}f_{0i}}{f_{1.}f_{0.}} \tag{A1.18}$$

For two response alternatives this reduces to the probability of a correct response

$$P(C) = P(H)P(H_1) + [1 - P(F)]P(H_0)$$

if and only if $P(H_1) = P(H_0) = 1/2$. This P(C), of course, is not to be confused with the probability of a correct response in a two interval-forced choice procedure which, under certain assumptions, is P_A.

For less than 5 response alternatives (A1.18) provides a rather gross underestimate of the area under the ROC curve. With as many as 10 response alternatives that are approximately equally spaced in probability trap P_A is almost indistinguishable from est P_A.

APPENDIX II

STATISTICS OF LINEAR UNCERTAIN DECISION VARIABLES

The marginal and joint moments for linear-uncertain observers reported in Chapter IV are derived here.

A2 1 Preliminaries

The decision variable of the linear-uncertain observer α is, from Equation 3.27,

$$z_\alpha = \frac{1}{N_0}[\lambda_\alpha y' D_\alpha y + \bar{\lambda}_\alpha \mu'_\alpha D_\alpha (2y - \mu_\alpha)].$$

To compute the moments of z_α, it is more convenient to express z_α in summation notation. Since D_α is a projection operator, there exists an orthogonal transformation of F in which the equivalent projection operator to D_α is a diagonal matrix with entries on the diagonal either 0 or 1. Since the trace of D_α is the dimensionality of the subspace F_α, the number of 1's in the equivalent matrix is $m_\alpha = 2W_\alpha T = 2\delta_\alpha$. We shall let $m_{\alpha\beta} = \min\{m_\alpha, m_\beta\}$ under the assumption that there exists a subspace $F_{\alpha\beta}$ in F for which $F_{\alpha\beta} = F_\alpha \cap F_\beta$ and $F_{\alpha\beta} = F_\alpha$ or $F_{\alpha\beta} = F_\beta$. Thus, the decision variable z_α (with no internal noise) in the equivalent representation (i.e., the frequency representation) becomes

$$z_\alpha = \frac{1}{N_0} \sum_{i=1}^{m\alpha} [\lambda_\alpha y_i^2 + \bar{\lambda}_\alpha \mu_{\alpha i}(2y_i - \mu_{\alpha i})]$$

Several further preliminaries are necessary to ease the burden of the derivations.

Definition. The *joint central moment* of type $(i,j,\ldots,k)$ for a sequence of random variables $(x,y,\ldots,z)$ is the expectation

$$\mu_{ij\ldots k}(x,y,\ldots,z) = E\{[x - E(x)]^i [y - E(y)]^j \ldots [z - E(z)]^k\}.$$

Thus, $\mu_1(x^2)$ is *not* the mean of x^2; $\mu_2(y) = \mu_{11}(y,y)$ is the variance of y; and $\mu_{11}(x,y^2)$ is the covariance between x and $z = y^2$.

(The mean vector μ_α of an observer's prior distribution will always contain a Greek subscript and so should cause no confusion with the notation for joint moments.)

Let a, b, and c be independent random variables. We observe that if $w = ax + by + c$ where a and b are independent of x, y, and z, then

$$\mu_{11}(w,z) = \mu_{11}(ax + by + c,\ z)$$
$$= E(a)\mu_{11}(x,z) + E(b)\mu_{11}(y,z) + \mu_{11}(c,z).$$

Further if c is a constant, or independent of z, then $\mu_{11}(c,z) = 0$. Of course, these relations may be generalized to joint moments of type $(1,1,\ldots,1)$. Finally, we recall that

$$\mu_{11}(x,y) = E(xy) - E(x)E(y).$$

Proposition. Let y have a normal distribution with mean zero and variance $\sigma^2 = \mu_2(y)$. Then, all odd moments of y are identically zero, and the even moments of y are given by

$$\mu_{(2r)}(y) = E(y^{2r}) = \sigma^{2r}(2r)!/2^r r!$$

for r a positive integer.

A proof may be found in Kendall and Stuart (1958, p. 60).

Corollary. Let y_i be a sample of the modified noise vector, i.e., $y_i = n_i - \mu_{n_i}$. Then,

$$E(y_i^{2r}) = N_0^r(2r)!/2^{2r}r!.$$

The corollary follows from the fact, found in Chapter II, that $\mu_2(y_i) = N_0/2$.

Three independent normally distributed internal noise sources are considered. e_1 is added to the input y. e_2 is added to the memory-specified signal vector μ_α. e_3 is added to the decision variable itself. The variances, with convenient scale factors are

$$\mu_2(e_{1i}) = N_0 V_1/2,$$
$$\mu_2(e_{2i}) = N_0 V_2/2,$$
$$\mu_2(e_3) = V_3.$$

The mean of each error source is zero.

The decision variable with e_1 and e_2 is

$$z_\alpha = \frac{1}{N_0} [\lambda_\alpha (y_\alpha + e_{\alpha 1})' D_\alpha (y_\alpha + e_{\alpha 1}) + \overline{\lambda}_\alpha (\mu_\alpha + e_{\alpha 2})' D_\alpha (2(y_\alpha + e_{\alpha 1}) - (\mu_\alpha + e_{\alpha 2})], \tag{A2.1}$$

where the input y has been given a subscript to allow for the possibility that two observers may have different inputs. The noisy decision variable with criterion variability is

$$z^*_\alpha = z_\alpha + e_{\alpha 3}. \tag{A2.2}$$

The decision variable for a second observer β is obtained from z^*_α by replacing each α with the subscript β.

A2.2 Derivation of the Mean

The mean z^*_α may be found directly. Since $E(e_3) = 0$, $E(e_{\alpha 1}) = E(e_{\alpha 2}) = 0$, and $e_{\alpha 1}$, $e_{\alpha 2}$, and y_α are mutually independent, we have

$$\begin{aligned} E(z^*_\alpha) &= E(z_\alpha) \\ &= \frac{1}{N_0} \{\lambda_\alpha [E(y'_\alpha D_\alpha y_\alpha) + E(e'_{\alpha 1} D_\alpha e_{\alpha 1})] \\ &\qquad + \overline{\lambda}_\alpha [2\mu'_\alpha D_\alpha E(y_\alpha) - \mu'_\alpha D_\alpha \mu_\alpha - E(e'_{\alpha 2} D_\alpha e_{\alpha 2})]\} \\ &= \frac{1}{N_0} \{\lambda_\alpha [E(y'_\alpha D_\alpha y_\alpha) + \delta_\alpha N_0 V_{\alpha 1}] \\ &\qquad + \overline{\lambda}_\alpha [2\mu'_\alpha D_\alpha E(y_\alpha) - E_\alpha - \delta_\alpha N_0 V_{\alpha 2}]\}. \end{aligned}$$

When a signal s_α is present (H_1), $y_\alpha = y + s_\alpha$. Then, defining, as in Chapter IV,

$$\mu_{s_\alpha} = s_\alpha$$

$$R_{\alpha\beta} = \mu_{\alpha}^{\prime}D_{\alpha\beta}\mu_{\beta}, \quad r_{\alpha\beta} = 2R_{\alpha\beta}/N_0,$$
$$E_{\alpha} = R_{\alpha\alpha}, \quad d_{\alpha} = 2E_{\alpha}/N_0$$

we have

$$E(z_{\alpha}^{*}) = \frac{1}{N_0}\ [\lambda_{\alpha}(\delta_{\alpha}N_0 + E_{\ \alpha} + \delta_{\alpha}N_0V_{\alpha 1}) + \bar{\lambda}_{\alpha}(2R_{\alpha\ \alpha} - E_{\alpha} + \delta_{\alpha}N_0V_{\alpha 2})],$$

and collecting terms,

$$E(z_{\alpha}^{*}) = \lambda_{\alpha}[\delta_{\alpha}(1 + V_{\alpha 1}) + d_{s_{\alpha}}/2] + \bar{\lambda}_{\alpha}[r_{\alpha s_{\alpha}} - d_{\alpha}/2 - \delta_{\alpha}V_{\alpha 2}]. \qquad \text{(A2.3)}$$

A2.3 Derivation of the Variance and Covariance

The covariance between two observers is

$$\begin{aligned}\mu_{11}(z_{\alpha}^{*}, z_{\beta}^{*}) &= \mu_{11}(z_{\alpha} + e_{\alpha 3}, z_{\beta} + e_{\beta 3}) \\ &= \mu_{11}(z_{\alpha}, z_{\beta}). \qquad \text{(A2.4)}\end{aligned}$$

Let $y_{\alpha}^{*} = y_{\alpha} + e_{\alpha 1}$ and $\mu_{\alpha}^{*} = \mu_{\alpha} + e_{\alpha 2}$, and similarly for observer β. Then the covariance becomes

$$\begin{aligned}\mu_{11}(z_{\alpha}, z_{\beta}) = \frac{1}{N_0^2}\,\mu_{11}[&\lambda_{\alpha}y_{\alpha}^{*\prime}D_{\alpha}y_{\alpha}^{*} + \bar{\lambda}_{\alpha}\mu_{\alpha}^{*\prime}D_{\alpha}(2y_{\alpha}^{*} - \mu_{\alpha}^{*}), \\ &\lambda_{\beta}y_{\beta}^{*\prime}D_{\beta}y_{\beta}^{*} + \bar{\lambda}_{\beta}\mu_{\beta}^{*\prime}D_{\beta}(2y_{\beta}^{*} - \mu_{\beta}^{*})].\end{aligned}$$

This may be expanded into four terms:

$$\begin{aligned}N_0^2\mu_{11}(z_{\alpha}, z_{\beta}) &= \lambda_{\alpha}\lambda_{\beta}\mu_{11}[y_{\alpha}^{*\prime}D_{\alpha}y_{\alpha}^{*}, y_{\beta}^{*\prime}D_{\beta}y_{\beta}^{*}] \\ &+ \bar{\lambda}_{\alpha}\lambda_{\beta}\mu_{11}[\mu_{\alpha}^{*\prime}D_{\alpha}(2y_{\alpha}^{*} - \mu_{\alpha}^{*}), y_{\beta}^{*\prime}D_{\beta}y_{\beta}^{*}] \\ &+ \lambda_{\alpha}\bar{\lambda}_{\beta}\mu_{11}[y_{\alpha}^{*\prime}D_{\alpha}y_{\alpha}^{*}, \mu_{\beta}^{*\prime}D_{\beta}(2y_{\beta}^{*\prime} - \mu_{\beta}^{*}] \\ &+ \bar{\lambda}_{\alpha}\bar{\lambda}_{\beta}\mu_{11}[\mu_{\alpha}^{*\prime}D_{\alpha}(2y_{\alpha}^{*} - \mu_{\alpha}^{*}), \mu_{\beta}^{*\prime}D_{\beta}(2y_{\beta}^{*} - \mu_{\beta}^{*})] \\ &= \lambda_{\alpha}\lambda_{\beta}T_1 + \bar{\lambda}_{\alpha}\lambda_{\beta}T_2 + \lambda_{\alpha}\bar{\lambda}_{\beta}T_3 + \bar{\lambda}_{\alpha}\bar{\lambda}_{\beta}\,{}_4. \qquad \text{(A2.5)}\end{aligned}$$

Each of the four terms will be evaluated for two special cases, namely where the observers are different ($\alpha \neq \beta$) and where they are the same ($\alpha = \beta$). The latter, of course, is the variance. Further, we let $y_{\alpha} = y + s_{\alpha}$, and $y_{\beta} = y + s_{\beta}$, and so consider the possibility that the observers may be presented different signals. Notice that here y denotes the modified noise vector $n - \mu_n$. The first term, T_1, will be evaluated in some detail; for the other terms the discussion will be brief.

Evaluation of Term 1.

We first rewrite the moment in summation notation for the appropriate equivalent representation space:

$$\begin{aligned} T_1 &= \mu_{11}[y_\alpha^{*\prime} D_\alpha y_\alpha^*, y_\beta^{*\prime} D_\beta y_\beta^*] \\ &= \mu_{11}[\sum_{i=1}^{m_\alpha}(y_{\alpha i}^*)^2, \sum_{j=1}^{m_\beta}(y_{\beta j}^*)^2] \\ &= \sum_{i=1}^{m_\alpha}\sum_{j=1}^{m_\beta}\mu_{11}[(y_{\alpha i}^*)^2, (y_{\beta j}^*)^2]. \end{aligned}$$

Now, when $i \neq j$ the random variables in the brackets are independent regardless of the case being considered, so that such μ_{11} terms are zero. Moreover, by assumption, either F_α is a subspace of F_β or the reverse. Thus, in the equivalent representation space the summation extends only to $m_{\alpha\beta} = \min\{m_\alpha, m_\beta\}$. These comments imply that T_1 may be simplified to

$$\begin{aligned} T_1 &= \sum_{i=1}^{m_{\alpha\beta}} \mu_{11}[(y_{\alpha i}^*)^2, (y_{\beta i}^*)^2] \\ &= \sum_{i=1}^{m_{\alpha\beta}} T_{1i}, \end{aligned}$$

where the definition of T_{1i} is implicit.

For $\alpha \neq \beta$, substitution for $y_{\alpha i}^*$ and $y_{\beta i}^*$ yields

$$\begin{aligned} T_{1i} &= \mu_{11}[(y_{\alpha i} + e_{\alpha 1i})^2, (y_{\beta i} + e_{\beta 1i})^2] \\ &= \mu_{11}[(y_{\alpha i}^2 + 2y_{\alpha i}e_{\alpha 1i} + e_{\alpha 1i}^2, y_{\beta i}^2 + 2y_{\beta i}e_{\beta 1i} + e_{\beta 1i}^2] \\ &= \mu_{11}[y_{\alpha i}^2, y_{\beta i}^2]. \end{aligned}$$

However, if $\alpha = \beta$, we have

$$T_{1i} = \mu_{11}(y_{\alpha i}^2, y_{\alpha i}^2) + 4\mu_{11}(y_{\alpha i}e_{\alpha 1i}, y_{\alpha i}e_{\alpha 1i}) + \mu_{11}(e_{\alpha 1i}^2, e_{\alpha 1i}^2),$$

where the obviously zero terms have been dropped. For the second term on the right of the latter expression we find

$$\begin{aligned}\mu_{11}(y_{\alpha i}e_{\alpha 1i}, y_{\alpha i}e_{\alpha 1i}) &= E(y_{\alpha i}^2 e_{\alpha 1i}^2) - [E(y_{\alpha i}e_{\alpha i})]^2 \\ &= E(y_{\alpha i}^2)E(e_{\alpha 1i}^2) - [E(y_{\alpha i})E(e_{\alpha i})]^2 \\ &= E(y_{\alpha i}^2)N_0V_{\alpha 1}/2,\end{aligned}$$

since $E(e_{\alpha i}) = 0$. For the last term,

$$\begin{aligned}\mu_{11}(e_{\alpha 1i}^2, e_{\alpha 1i}^2) &= \mu_2(e_{\alpha 1i}^2) \\ &= E(e_{\alpha 1i}^4) - [E(e_{\alpha 1i}^2)] \\ &= 3\left(\frac{N_0V_{\alpha 1}}{2}\right)^2 - \left(\frac{N_0V_{\alpha 1}}{2}\right)^2 \\ &= N_0^2V_{\alpha 1}^2/2,\end{aligned}$$

where the proposition was used to obtain $E(e_{\alpha 1i}^4)$.

Now collecting terms again, for $\alpha = \beta$,

$$T_{1i} = \mu_{11}(y_{\alpha i}^2, y_{\alpha i}) + 2E(y_{\alpha i}^2)N_0V_{\alpha 1} + N_0^2V_{\alpha 1}^2/2.$$

T_1, Covariance. Since $y_{\alpha i} = y_i + s_{\alpha i}$ and $y_{\beta i} = y_i + s_{\beta i}$ we have

$$\begin{aligned}T_{1i} &= \mu_{11}[y_{\alpha i}, y_{\beta i}] \\ &= \mu_{11}[(y_i + s_{\alpha i})^2, (y_i + s_{\beta i})^2] \\ &= \mu_{11}[y_i^2, y_i^2] + 4s_{\alpha i}s_{\beta i}\mu_{11}[y_i, y_i] \\ &= 2(N_0/2)^2 + 4s_{\alpha i}s_{\beta i}(N_0/2)^2 \\ &= N_0^2/2 + 2N_0s_{\alpha i}s_{\beta i}.\end{aligned}$$

Therefore,

$$\begin{aligned}T_1 &= \sum_{i=1}^{m_{\alpha\beta}} T_{1i} \\ &= m_{\alpha\beta}N_0^2/2 + 2N_0 \sum_{i=1}^{m_{\alpha\beta}} s_{\alpha i}s_{\beta i} \\ &= \delta_{\alpha\beta}N_0^2 + 2N_0R_{s_\alpha s_\beta}\end{aligned}$$

T_1, Variance.

$$\begin{aligned}T_{1i} = {} & \mu_{11}[(y_i + s_{\alpha i})^2, (y_i + s_{\alpha i})^2] \\ & + 2E[(y_i + s_{\alpha i})^2]N_0V_{\alpha 1} + N_0^2V_{\alpha 1}/2\end{aligned}$$

$$= \mu_{11}(y_i^2,y_i^2) + 4s_{\alpha i}^2\mu_{11}(y_i,y_i)$$

$$+[E(y_i^2) + s_{\alpha i}^2]N_0V_{\alpha 1} + N_0^2V_{\alpha 1}^2/2$$

$$= N_0^2[1 + V_{\alpha 1}^2 + 2V_{\alpha 1}]/2 + 2N_0s_{\alpha i}^2(1 + V_{\alpha 1})$$

Therefore,

$$T_1 = \delta_\alpha N_0^2(1 + V_{\alpha 1})^2 + 2N_0E_{s_\alpha}(1 + V_{\alpha 1}).$$

Evaluation of Term 2

$$T_2 = \sum_{i=1}^{m_{\alpha\beta}} \mu_{11}[\mu_{\alpha i}^*(2y_{\alpha i}^* - \mu_{\alpha i}^*),(y_{\beta i}^*)]$$

$$= \sum_{i=1}^{m_{\alpha\beta}} \{2\mu_{11}[\mu_{\alpha i}^*y_{\alpha i}^*,(y_{\beta i}^*)^2] - \mu_{11}[(\mu_{\alpha i}^*)^2,(y_{\beta i}^*)^2]\}$$

$$= \sum_{i=1}^{m_{\alpha\beta}} [2T_{21i} - T_{22i}]$$

$$T_{21i} = \mu_{11}[(\mu_{\alpha i} + e_{\alpha 2i})(y_{\alpha i} + e_{\alpha 1i}),(y_{\beta i} + e_{\beta 1i})^2]$$

$$= \mu_{\alpha i}\mu_{11}(y_{\alpha i},y_{\beta i}^2) + 2\mu_{\alpha i}\mu_{11}[e_{\alpha 1i},y_{\beta i}e_{\beta 1i}]$$

$$T_{22i} = \mu_{11}[(\mu_{\alpha i} + e_{\alpha 2i})^2,(y_{\beta i} + e_{\beta 1i})^2]$$

$$= 0.$$

T_2, Covariance.

$$T_{21i} = \mu_{\alpha i}\mu_{11}[y_i + s_{\alpha i},(y_i + s_{\beta i})^2]$$

$$= \mu_{\alpha i}[2s_{\beta i}\mu_{11}(y_i,y_i)]$$

$$= N_0\mu_{\alpha i}s_{\beta i}.$$

Therefore,

$$T_2 = 2N_0 \sum_{i=1}^{m_{\alpha\beta}} \mu_{\alpha i}s_{\beta i}$$

$$= 2N_0R_{\alpha s_\beta}$$

T_2, Variance

$$
\begin{aligned}
T_{21i} &= \mu_{\alpha i}\mu_{11}[y_i + s_{\alpha i}, (y_i + s_{\alpha i})^2] \\
&\quad + 2\mu_{\alpha i}\mu_{11}[e_{\alpha 1i}, (y_i + s_{\alpha i})e_{\alpha 1i}] \\
&= N_0\mu_{\alpha i}s_{\alpha i} + 2\mu_{\alpha i}s_{\alpha i}\mu_{11}[e_{\alpha i}, e_{\alpha i}] \\
&= N_0\mu_{\alpha i}s_{\alpha i}(1 + V_{\alpha 1})
\end{aligned}
$$

Therefore,

$$T_2 = 2N_0R_{\alpha s_\alpha}(1 + V_{\alpha 1}).$$

Evaluation of Term 3.

Because of symmetry T_3 is T_2 with β replacing α.

T_3, Covariance.

$$T_3 = 2N_0R_{\beta s_\alpha}$$

T_3, Variance.

$$T_3 = 2N_0R_{\beta s_\beta}(1 + V_{\beta 1})$$

Evaluation of Term 4

$$T_4 = \mu_{11}[2\mu_\alpha^{*\prime}D_\alpha y_\alpha^* - \mu_\alpha^{*\prime}D_\alpha\mu_\alpha^*, 2\mu_\beta^{*\prime}D_\beta y_\beta^{*\prime} - \mu_\beta^{*\prime}D_\beta\mu_\beta^*]$$

$$= \sum_{i=1}^{m_{\alpha\beta}} \{4\mu_{11}[\mu_{\alpha i}^* y_{\alpha i}^*, \mu_{\beta i}^* y_{\beta i}^*] - 2\mu_{11}[\mu_{\alpha i}^* y_{\alpha i}^*, (\mu_{\beta i}^*)^2]$$

$$- 2\mu_{11}[(\mu_{\alpha i}^*)^2, \mu_{\beta i}^* y_{\beta i}^*] + \mu_{11}[(\mu_{\alpha i}^*)^2, (\mu_{\beta i}^*)^2]$$

$$= \sum_{i=1}^{m_{\alpha\beta}} [4T_{41} - 2T_{42} - 2T_{43} + T_{44}].$$

$$T_{41} = \mu_{11}[(\mu_{\alpha i} + e_{\alpha 2i})(y_{\alpha i} + e_{\alpha 1i}), (\mu_{\beta i} + e_{\beta 2i})(y_{\beta i} + e_{\beta 1i})]$$

$$= \mu_{\alpha i}\mu_{\beta i}\{\mu_{11}[y_{\alpha i}, y_{\beta i}] + \mu_{11}[e_{\alpha 1i}, e_{\beta 1i}]\}$$

$$+ \mu_{11}[y_{\alpha i}e_{\alpha 2i}, y_{\beta i}e_{\beta 2i}] + \mu_{11}[e_{\alpha 2i}e_{\alpha 1i}, e_{\beta 2i}e_{\beta 1i}]$$

$$T_{42} = \mu_{11}[(\mu_{\alpha i} + e_{\alpha 2i})(y_{\alpha i} + e_{\alpha 1i}), (\mu_{\beta i} + e_{\beta 2i})^2]$$

$$= \mu_{11}(y_{\alpha i}e_{\alpha 2i}, 2\mu_{\beta i}e_{\beta 2i}).$$

By symmetry to T_{42}

$$T_{43} = \mu_{11}(2\mu_{\alpha i}e_{\alpha 2i}, y_{\beta i}e_{\beta 2i})$$

$$T_{44} = \mu_{11}[(\mu_{\alpha i} + e_{\alpha 2i})^2, (\mu_{\beta i} + e_{\beta 2i})^2]$$

$$= 4\mu_{11}[\mu_{\alpha i}e_{\alpha 2i}, \mu_{\beta i}e_{\beta 2i}] + \mu_{11}[e_{\alpha 2i}^2, e_{\beta 2i}^2]$$

T_4, Covariance.

$$T_{41} = \mu_{\alpha i}\mu_{\beta i}\mu_{11}[y_i + s_{\alpha i}, y_i + s_{\beta i}]$$

$$= N_0\mu_{\alpha i}\mu_{\beta i}/2.$$

$$T_{42} = T_{43} = T_{44} = 0$$

Therefore,

$$T_4 = 2N_0R_{\alpha\beta}$$

T_4, Variance.

$$T_{41} = \mu_{\alpha i}^2\{\mu_{11}[y_i + s_{\alpha i}, y_i + s_{\alpha i}] + \mu_{11}[e_{\alpha 1i}, e_{\alpha 1i}]$$
$$+ \mu_{11}[(y_i + s_{\alpha i})e_{\alpha 2i}, (y_i + s_{\alpha i})e_{\alpha 2i}]$$
$$+ \mu_{11}[e_{\alpha 2i}e_{\alpha 1i}, e_{\alpha 2i}e_{\alpha 1i}]$$
$$= N_0\mu_{\alpha i}^2(1 + V_{\alpha 1})/2 + \mu_{11}[y_ie_{\alpha 2i}, y_ie_{\alpha 2i}] + s_{\alpha i}\mu_{11}[e_{\alpha 2i}, e_{\alpha 2i}]$$
$$+ E(e_{\alpha 2i}^2, e_{\alpha 1i}^2)$$
$$= N_0[\mu_{\alpha i}^2(1 + V_{\alpha 1}) + s_{\alpha i}^2V_{\alpha 2}]/2 + N_0^2V_{\alpha 2}(1 + V_{\alpha 1})/4.$$

$$T_{42} = 2\mu_{11}[(y_i + s_{\alpha i})e_{\alpha 2i}, \mu_{\alpha i}e_{\alpha 2i}]$$
$$= 2\mu_{\alpha i}s_{\alpha i}\mu_{11}[e_{\alpha 2i}, e_{\alpha 2i}]$$
$$= N_0\mu_{\alpha i}s_{\alpha i}V_{\alpha 2}$$

$$T_{43} = N_0\mu_{\alpha i}s_{\alpha i}V_{\alpha 2}$$

$$T_{44} = 4\mu_{\alpha i}^2\mu_{11}[e_{\alpha 2i}, e_{\alpha 2i}] + \mu_{11}[e_{\alpha 2i}, e_{\alpha 2i}]$$
$$= 2N_0\mu_{\alpha i}^2V_{\alpha 2} + N_0^2V_{\alpha 2}^2/2$$

Therefore,

$$T_4 = 2N_0[E_\alpha(1 + V_{\alpha 1}) + E_{s_\alpha}V_{\alpha 2}] + 2\delta_\alpha N_0^2(1 + V_{\alpha 1})V_{\alpha 2}$$
$$- 2N_0R_{\alpha s_\alpha}V_{\alpha 2} + 2N_0E_\alpha V_{\alpha 2} + \delta_\alpha N_0^2V_{\alpha 2}^2$$

Evaluation of $\mu_{11}(z_\alpha^*, z_\beta^*)$

Having completed the derivation of the individual terms of $\mu_{11}(z_\alpha^*, z_\beta^*)$ in the previous sections, we may collect the terms as defined in Equation A2.5.

Covariance. $y_\alpha = y + s_\alpha$, $y_\beta = y + s_\beta$

$$\mu_{11}(z_\alpha^*, z_\beta^*) = \lambda_\alpha\lambda_\beta[\delta_{\alpha\beta} + r_{s_\alpha s_\beta}] + \overline{\lambda}_\alpha\lambda_\beta r_{\alpha s_\beta} + \lambda_\alpha\overline{\lambda}_\beta r_{\beta s_\alpha} + \overline{\lambda}_\alpha\overline{\lambda}_\beta r_{\alpha\beta} \tag{A2.6}$$

Variance. $y_\alpha = y + s_\alpha$

$$\begin{aligned}\mu_2(z_\alpha^*) = {} & \lambda_\alpha^2[\delta_\alpha(1 + V_{\alpha 1})^2 + d_{s_\alpha}(1 + V_{\alpha 1})] + 2\overline{\lambda}_\alpha\lambda_\alpha r_{\alpha s_\alpha}(1 + V_{\alpha 1}) \\ & + \overline{\lambda}_\alpha^2\{\delta_\alpha V_{\alpha 2}[2(1 + V_{\alpha 1}) + V_{\alpha 2}] + d_\alpha(1 + V_{\alpha 1} + V_{\alpha 2}) \\ & + (d_{s_\alpha} - r_{\alpha s_\alpha})V_{\alpha 2}\} + V_{\alpha 3}\end{aligned} \tag{A2.7}$$

APPENDIX III

GLOSSARY OF SYMBOLS AND TERMS

a_i	the coefficient of $\cos \frac{2\pi it}{T}$ in the finite Fourier series approximation to $x(t)$.
a_{ij}, b_{ij}	signum of the difference $z_{\alpha i} - z_{\alpha j}$ and $z_{\beta i} - z_{\beta j}$, respectively.
A	the matrix of a filter A.
b_i	the coefficient of $\sin \frac{2\pi it}{T}$ in the finite Fourier series approximation to $x(t)$.
B	a matrix such that $A = B'B$.
$c_i(t)$	short hand for $\sqrt{\frac{2}{T}} \cos \frac{2\pi it}{T}$.
Corr(,)	population correlation of the arguments.
Cov(,)	population covariance of the arguments.
C	the matrix of an orthogonal linear transformation.
d	standardized energy: $2E/N_0$.
d_s	standardized energy of the signal vector s.
d_{s_α}	standardized energy of the signal vector s_α presented to observer α. If s_α is the null signal 0, $d_{s_\alpha} = 0$.
d_α	standardized energy of the memory-specified reference signal μ_α.
d'	the square-root of the standardized energy d.
d'_s	$\sqrt{d_s}$
d'_z	sensitivity index of the decision variable z.
d'_{z_α}	d'_z for observer α.

det	determinant.
D_α	a projection operator matrix for observer α.
$D_{\alpha\beta}$	the product $D_\alpha D_\beta$.
Double-probability paper	graph paper linear in both coordinates with standard normal deviation scores.
e_3	an error term added to the decision variable z.
e	a random error vector.
$e_{\alpha 1}$	a random error vector added to the modified input y for observer α.
$e_{\alpha 2}$	a random error vector added to the memory-specified reference signal μ_α.
E	the energy of a constant signal waveform of duration T.
E_s	the energy of s.
E_α	the energy of μ_α.
E_{z_α}	the signal energy necessary for the ideal observer to performance at the level as observer α who uses the decision variable z_α.
$E(\ )$	the expectation operator.
f( )	a density function of the argument.
$F_0(z)$	the cumulative distribution function of z conditional on H_0.
$F_1(z)$	the cumulative distribution function of z conditional on H_1.
F	the frequency representation space.
F_α	a subspace of F generated by the projection operator D_α.
False-alarm rate	same as P(F).
g(s)	an *a priori* density function of the signal s.

$G(s)$	the cumulative prior distribution of s.
$h_1(x), h_0(x)$	the distribution densities of the input known to an observer conditional upon H_1 and H_0, respectively.
H_1, H_0	the experimenter-specified hypotheses of signal-plus-noise and noise-alone, respectively.
Hit-rate	same as P(H).
I	an identity matrix.
Iso-bias curve	jargon for the curve which cuts across a family of ROC curves at points of constant slope.
ln ( )	the natural logarithm of the argument.
$\ell(x)$	the likelihood ratio of x.
$M_0/2$	the imprecision (variance) of an observer's specification of a single component of the signal in the frequency representation space.
n_j	the $j^{\underline{th}}$ entry in a noise vector sample n.
n	a sample vector of the noise random vector $\tilde{n}$.
$\tilde{n}$	a random noise vector.
n_f	a sample noise vector explicitly represented in F.
$n(t,\tilde{n})$	a stochastic noise process specified by the random vector $\tilde{n}$ as a function of time.
N	the total average power of a bandlimited noise process.
N_0	the noise power per unit bandwidth N/W.
$N(\mu,\Sigma)$	a multivariate normal density function with mean vector μ and dispersion (variance-covariance) matrix Σ.

Observer	a subject or device in a detection task; when the stimulus is acoustic subjects are also called "listeners" and devices "receivers".
P_A	the area under an ROC curve.
$P(\cdot)$	the probability of an event.
$P(C)$	the probability of a correct decision. The meaning of "correct" depends upon the context.
$P(D)$	the probability that jointly observed differences have Different signs. D is the event $a_{ij}b_{ij} < 0$.
$P(F)$	the probability of a "false-alarm", i.e., $P(\text{say } H_1 \mid H_0 \text{ true})$. $P(F)$ is also referred to as the "false-alarm rate" or the "incorrect detection rate".
$P(H)$	the probability of a "hit", i.e., $P(\text{say } H_1 \mid H_1 \text{ true})$. $P(H)$ is also referred to as the "hit rate" or the "probability of (correct) detection".
$P(S)$	the probability that jointly observed differences have the same sign. S is the event $a_{ij}b_{ij} > 0$.
$P(T)$	the probability that either one or both of jointly observed differences are zero. T is the event $a_{ij}b_{ij} = 0$.
$P(T_\alpha), P(T_\beta)$	the probability of a tie in two independent samples of a random variable. T_α is the event $a_{ij} = 0$, T_β is the event $b_{ij} = 0$.
Psychometric function	a graph of the relationship between a performance index and a physical measure of detectability.
Q_m	the precision matrix $(\Sigma_n + \Sigma_s)^{-1}$
Q_n	the precision matrix Σ_n^{-1}.
$r_{\alpha s}$	the standardized inner product (or cross-correlation) of the vectors μ_α and s, i.e., $2\mu_\alpha' D_\alpha s/N_0$.

$r_{\alpha\beta}$	the standardized inner product of the vectors μ_α and μ_β, i.e., $2\mu_\alpha D_{\alpha\beta}\mu_\beta/N_0$.
$r_{\alpha s_\beta}$	the standardized inner product of the V vectors μ_α and s_β, i.e., $2\mu_\alpha' D_\alpha s_\beta$. When $s_\beta = 0$, $r_{\alpha s_\beta} = 0$. When $s_\beta = s$, $r_{\alpha s_\beta} = r_{\alpha s}$.
r	a response vector.
R_1, R_0	acceptance regions in the representation space for the hypotheses H_1 and H_0, respectively. R_1 is also called the criterion region.
$R_{\alpha\beta}$	the inner-product (cross-correlation) $\mu_\alpha' D_{\alpha\beta}\mu_\beta$.
ROC curve	the theoretical curve (the receiver operating characteristic) [P(F),P(H)] generated by a decision variable z for discrimination between the hypotheses H_1 and H_0.
$s_i(t)$	short hand for $\sqrt{\frac{2}{T}} \sin \frac{2\pi i t}{T}$.
$s(t)$	a constant signal waveform defined for $0 \leq t \leq T$.
$s_\perp(t)$	the Hibert transform of $s(t)$.
s	a signal vector representing $s(t)$.
$s_\perp$	the representation vector of $s_\perp(t)$.
s_α	the signal vector presented to observer α. Usually on H_1 trials $s_\alpha = s$ and on H_0 trials $s_\alpha = 0$.
T	the time interval over which x(t) is represented. Also, the duration of the presentation interval.
T	the temporal representation space.
$V_1, V_{\alpha 1}$	the variance, relative to $N_0/2$, of a component of e_1 or e_α, if the observer is specified.
$V_2, V_{\alpha 2}$	the variance, relative to $N_0/2$, of a component of e_2 or $e_{\alpha 2}$.

V_3, $V_{\alpha 3}$	the variance of e_3 or $e_{\alpha 3}$.
Var( )	the population variance of the argument.
W	the bandwidth of the noise process.
W_α	the equivalent square bandwidth of the (hypothetical) interval filter of observer α.
$W_{\alpha\beta}$	the minimum of W_α and W_β, i.e., the overlap in bandwidth of the interval filters for observers α and β.
x_i	the $i^{\underline{th}}$ entry in the vector x.
x(t)	a waveform; the input waveform sample presented to an observer.
$\hat{x}(t)$	the finite Fourier series approximation to x(t).
x(j/2 W)	the value of x(t) at t = j/2 W.
x	a representation vector for x(t).
x_f	the representation vector of x(t) explicitly with coordinates in the frequency representation space F.
x'	the transpose of x.
$\|\|x\|\|$	the length of x, i.e., $\sqrt{x'x}$.
X	the sample space of vectors x. X may be interpreted as either T or F.
y	the modified input vector $x - \mu_n$.
y^*	$y + e_1$.
z	a decision variable.
z_c	a particular value of z.
z_p	the "perfect" decision variable.
z_α	the decision variable of observer α.

$z_{\alpha 0}$	a value of z_α conditional on H_0.
$z_{\alpha 1}$	a value of z_α conditional on H_1.
z_α^*	the decision variable of observer α augmented by including various interval sources of error.
z_1	the decision variable of the ideal observer.
α	a label for an observer.
β	(1) a label for an observer; (2) a cut-off value of a decision variable z which determines an acceptance region R_1.
γ	the Goodman-Kruskal coefficient gamma.
δ_α	short-hand for the bandwidth-time product $W_\alpha T$.
$\delta_{\alpha\beta}$	short-hand for $W_{\alpha\beta}T$.
λ_α	the relative uncertainty parameter $= M_0/(M_0 + N_0)$.
$\overline{\lambda}_\alpha$	short hand for $1 - \lambda_\alpha$.
η, η_α	the efficiency of an observer.
$\psi_j(t)$	the $j^{\underline{th}}$ temporal interpolation function.
$\rho, \rho_{\alpha\beta}$	the linear correlation between the decision variables z_α and z_β.
ρ_1, ρ_0	the correlation conditional upon H_1 or H_0, respectively.
σ^2	variance.
Σ_m	the dispersion matrix $\Sigma_n + \Sigma_s$.
Σ_n	the dispersion matrix of the noise process.
Σ_n^{-1}	the inverse matrix of Σ_n.
Σ_s	the dispersion matrix of an observer's prior specification of the signal s.

$\tau, \tau_{\alpha\beta}$	Kendall's tau.
$\tau^{(1)}, \tau^{(0)}$	tau conditional upon H_1 or H_0, respectively.
μ_m	the vector sum $\mu_n + \mu_\alpha$.
μ_n	the mean vector of the noise process.
μ_{s_α}	same as s_α (for consistency in notation).
μ_α	the mean vector of the observer's prior specification of the signal.
μ^*_α	the vector sum $\mu_\alpha + e_{\alpha 2}$.
μ'_{ij}	a standardized moment of type (i,j).

BIBLIOGRAPHY

Ahumada, A., "Detection of Tones Masked by Noise: A Comparison of Human Observers with Digital-Computer Simulated Energy Detectors," University of California, Los Angeles: Human Communication Laboratory, Technical Report No. 29, 1967.

Birdsall, T. G., "Detection of a Signal Specified Exactly with a Noisy Stored Reference Signal," *J. Acoust. Soc. Am.*, *32*, 1038-1045 (1960).

Birdsall, T. G., *The Theory of Signal Detectability: ROC Curves and their Character*. Unpublished Ph.D. thesis, The University of Michigan, Ann Arbor, Michigan, 1966.

Goldman, S., *Information Theory*. Englewood Cliffs, New Jersey: Prentice-Hall, 1953.

Goodman, L. A., and Kruskal, W. H., "Measures of Association for Cross Classifications," *J. Amer. Statist. Assoc.*, *49*, 732-764 (1954).

Green, D. M., "Auditory Detection of a Noise Signal," *J. Acoust. Soc. Am.*, *32*, 121-131 (1960).

Green, D. M., "A Distribution Free Relation between Forced-choice and Yes-No." Paper read at the Psychonomic Society, Niagara Falls, Ontario, October, 1964 (a).

Green, D. M., "Consistency of Auditory Detection Judgments," *Psychol. Rev.*, *71*, 392-407 (1964b).

Green, D. M., Birdsall, T. G., and Tanner, W. P. Jr., "Signal Detection as a Function of Signal Intensity and Duration," *J. Acoust. Soc. Am.*, *29*, 523-531 (1957).

Green, D. M., and Moses, F. L., "On the Equivalence of Two Recognition Measures of Short-Term Memory," *Psychol. Bull.*, *66*, 228-234 (1966).

Green, D. M. and Swets, J. A., *Signal Detection Theory and Psychophysics*, New York: Wiley, 1966.

Greiner, R., "Über das Fehlersyster der Kollecktivmasslehre," *Keitschrift für Mathematik and Physik, 57*, 121, 225, and 337 (1909).

Hancock, J. C. and Wintz, P. A., *Signal Detection Theory*. New York: McGraw-Hill, 1966.

Henning, B. G., "Model for Auditory Discrimination and Detection," *J. Acoust. Soc. Am., 42*, 1325-1334 (1967).

Jeffress, L. A., "Stimulus-Oriented Approach to Detection," *J. Acoust. Soc. Am., 36*, 766-774 (1964).

Jeffress, L. A., "Stimulus-Oriented Approach to Detection Re-Examined," *J. Acoust. Soc. Am., 41*, 480-488 (1967).

Kendall, M. G., "Rank and Product-Moment Correlation", *Biometrika, 36*, 177-193 (1949).

Kendall, M. G., *Rank Correlation Methods*, Third Edition, New York: Hafner, 1962.

Kendall, M. G., and Stuart, A., *The Advanced Theory of Statistics, Vol. I.* London: Charles Griffin, 1958.

Kruskal, W. H., "Ordinal Measures of Association," *J. Amer. Statist. Assoc., 53*, 814-861 (1958).

Lamphiear, D. E., and Birdsall, T. G., "Approximations to Non-Central Chi-Square Distributions with Applications to Signal Detection Models." The University of Michigan: Department of Electrical Engineering, Electronic Defense Group, Technical Report No. 101, May, 1960.

Licklider, J. C. R., "The Phenomena of Localization." In Graham, A., (Ed.), *Sensorneural Hearing Processes and Disorders*. Boston: Little, Brown and Co., 1967.

Middleton, D., *An Introduction to Statistical Communication Theory*. New York: McGraw-Hill, 1960.

McGill, W. J., "Polynomial Psychometric Functions in Audition," Unpublished Manuscript. University of California, San Diego: Department of Psychology, 1967.

Peterson, W. W., Birdsall, T. G., and Fox, W. C., "The Theory of Signal Detectability," *Trans. IRE Professional Group on Information Theory*, PGIT-4, 171-212 (1954).

Pfafflin, S. M., and Mathews, M. V., "Energy-Detection Model for Monaural Auditory Detection," *J. Acoust. Soc. Am.*, *34*, 1842-1853 (1962).

Sherwin, C. W., Kodman, F. Jr., Kovaly, J. J., Prothe, W. C., and Melrose, J., "Detection of Signals in Noise: A Comparison between the Human Detector and an Electronic Detector," *J. Acoust. Soc. Am.*, *28*, 617-622 (1956).

Swets, J. A., (Ed.), *Signal Detection and Recognition by Human Observers: Contemporary Readings*. New York: Wiley, 1964.

Tanner, W. P. Jr., "Theory of Recognition," *J. Acoust. Soc. Am.*, *28*, 882-888 (1956).

Tanner, W. P. Jr., and Birdsall, T. G., "Definitions of d' and η as Psychophysical Measures," *J. Acoust. Soc. Am.*, *30*, 922-928 (1958).

Wainstein, L. A. and Zubakov, V. D., *Extraction of Signals from Noise*. Englewood Cliffs, New Jersey: Prentice Hall, 1962.

Watson, C. S., *Signal Detection and Certain Physical Characteristics of the Stimulus During the Observation Interval*. Unpublished Ph.D. thesis, Indiana University, Bloomington, Indiana, 1962.

Watson, C. S., Rilling, M. E., and Bourbon, W. T., "Receiver-Operating Characteristics Determined by a Mechanical Analog to the Rating Scale," *J. Acoust. Soc. Am.*, *36*, 283-288 (1964).

Wickelgren, W. A., "Unidimensional Strength Theory and Component Analysis of Noise in Absolute and Comparative Judgments," *J. Math. Psychol.*, *5*, 102-122 (1968).

Wilcox, G. W., "Measures of Performance Sensitivity." Paper read at the Acoustical Society of America, Miami Beach, November, 1967.