INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

University Microfilms International
300 North Zeeb Road
Ann Arbor, Michigan 48106 USA
St. John's Road, Tyler's Green
High Wycombe, Bucks, England HP10 8HR

7900802

MOSCHETTO, CHARLES FRANCIS
ON THE MEASUREMENT OF NOISE-INTENSITY
DISCRIMINATION BY UP-DOWN PSYCHOPHYSICAL
METHODS.

CITY UNIVERSITY OF NEW YORK, PH.D., 1978

University
Microfilms
International 300 N. ZEEB ROAD, ANN ARBOR, MI 48106

ON THE MEASUREMENT OF NOISE-INTENSITY DISCRIMINATION BY UP-DOWN PSYCHOPHYSICAL METHODS

by Charles F. Moschetto

A dissertation submitted to the Graduate Faculty
in Psychology in partial fulfillment of the
requirements for the degree of Doctor of
Philosophy, The City University of New York

1978

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy

7-24-78
date

David H. Raab
Chairman of the Examining Committee

7/24/78
date

Florence L. Denmark
Executive Officer

I. Abramov

H. Levitt

N. Macmillan

E. Osman

D. Raab
Supervisory Committee

The City University of New York

Abstract

ON THE MEASUREMENT OF NOISE-INTENSITY DISCRIMINATION BY UP-DOWN PSYCHOPHYSICAL METHODS

by

Charles F. Moschetto

Advisor: Professor David H. Raab

Noise-intensity discrimination was investigated in a two-interval forced-choice task, using the Up-Down Transformed-Response method (UDTR). It was found that the level at which UDTR runs are initiated could have profound effects on performance. When UDTR is initiated with a signal that is well below threshold, the results approach the predictions of Green's ideal-receiver model for the case of noise-masked-by-noise. On the other hand, when adaptive testing begins with a suprathreshold signal, the pattern of results is inconsistent with that predicted for an ideal energy detector. It is proposed that practice at suprathreshold signal levels could be harmful in that the observer might adopt a nonoptimal detection strategy based on a stimulus attribute that is salient at high signal levels, but which becomes unavailable as the signal level is decreased.

In addition, it was found that UDTR, as employed here and elsewhere, may yield estimates of performance that are biased for purely statistical reasons. The magnitude of this bias was assessed by employing Monte Carlo computer simulations for an ideal observer. Recommendations are made for data collection and reduction in order that methodologically bias-free estimates of performance may be obtained.

ACKNOWLEDGEMENTS

I would like to express sincere appreciation to my advisor, Dr. David H. Raab, for his guidance in both this thesis project, and my intellectual development in general.

I am also grateful to Drs. Neil Macmillan and Eli Osman for serving on my committee and offering valuable criticism on earlier versions of this paper.

A special note of thanks goes to Richard Bernstein, not only for many tedious subject hours, but also for his friendship and moral support as a fellow graduate student.

This research was supported by faculty grants (to Dr. Raab) from the City University of New York. Many hours of computer time were also provided by the City University. A special debt of gratitude is owed to Dr. James Wang and his staff at the Brooklyn College Computer Center for their valuable software support.

Finally, I dedicate this dissertation to all my family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES, Continued

LIST OF TABLES, Continued

Table Page

LIST OF FIGURES

Chapter I

INTRODUCTION

In an early investigation of noise-masked-by-noise, Miller (1947) concluded that when the masking and masked stimuli are identical, the detection of a masked signal is equivalent to the discrimination of signal-plus-masker from masker alone. This observation, that intensity discrimination is a special form of masking, has served as a foundation for subsequent models of the detection process.

For the case of noise-masked-by-noise, ideal detector models have often looked to stimulus characteristics to provide a limit on discrimination. Peterson, Birdsall, and Fox (1954) described the optimal receiver for each of a number of signals masked by Gaussian noise as one that bases each of its decisions on a likelihood ratio — i.e., the probability that an observation resulted from the presentation of the signal added to noise, relative to the probability that it resulted from the presentation of noise alone. For the case of noise-masked-by-noise, it was demonstrated that the likelihood-ratio receiver is an energy detector.

This analysis was then extended by Green (1960) to the two-interval forced-choice procedure (2IFC). The development of such a model allows for specific predictions to be made concerning the effects of stimulus characteristics on the performance of an optimal receiver.

Its applicability as a model of the human detection process may then be directly ascertained by comparing the predicted effects of stimulus characteristics to those obtained for a human observer.

In this dissertation, Green's model for noise-intensity discrimination will be presented along with its predictions for the performance of the optimal receiver. We will then review the psychoacoustic literature pertaining to these predictions, and show that Green's model alone is insufficient to account for the performance of a human observer.

A discussion of the psychophysical procedures employed to obtain these data indicates that the practice of providing clearly audible signals in the early stages of testing is common to all methods. It is possible that early in testing, the observer adopts a nonoptimal detection strategy based on some stimulus characteristic which is salient at high signal levels. This dissertation, therefore, investigates noise-intensity discrimination with a procedure that deprives the observer of such suprathreshold information. An adaptive testing method (UDTR; see Levitt, 1971) is intiated from a level where the signal is clearly well below threshold. Data collected in this manner indicate that the performance of human observers closely approximates that of Green's ideal receiver, whereas data obtained with the more traditional application of UDTR (starting "high" with audible signals) replicate the differences between real and ideal observers noted in the literature.

A. Green's Energy Detection Model for Noise-Intensity Discrimination

Before describing Green's model, its assumptions, and predictions, it would be instructive to define the 2IFC paradigm. A sample of noise (which we shall call the signal) is added to a background noise (simply referred to as noise) with equal probability in one or the other of two temporal intervals. The observer must designate the interval which contains the signal. Let the finite duration of the observation intervals be called T, and the bandwidth of the signal and masker waveforms W_s and W_n, respectively.

The noise waveform is defined as Fourier-series bandlimited white Gaussian noise. That is, the distribution of instantaneous amplitudes is Gaussian with zero mean, and variance equal to the average power, N. In addition, the waveform has the same average power in each of its Fourier components. The waveform resulting from the addition of a signal is likewise Gaussian since the sum of two Gaussian random variables is itself Gaussian. The distribution of amplitudes for the signal-plus-noise waveform has zero mean, and variance equal to the average power, S+N. Inasmuch as the average power of the signal-plus-noise waveform will be greater than that of the noise alone, the observer need measure only the average power (or energy) in each observation interval and report the larger of the two. Performance will be less than perfect, however, due to the random nature of the stimuli. It is possible that on some trials the interval containing noise alone will have greater energy than the interval with signal-plus-noise. As a result, an optimal receiver will exhibit less than perfect performance in this situation. The probability of a correct response will depend on the distributions of noise and signal-plus-noise. Green's model, therefore, begins with the derivation of the distributions of energy for the two waveforms.

Green makes two assumptions in the derivation of his model: (1) The observer has exact knowledge of the bandwidth and center frequency of the signal, and employs a rectangular filter matched to W_s. (It should be noted that a rectangular filter is appropriate since Green's model deals solely with rectangular noise spectra.) If the masker bandwidth is wider than the signal, energy outside this internal filter plays no part in masking. Since W_n is now bandlimited to match Ws, let bandwidth simply be called W. (2) The observer has exact knowledge of the starting time and duration of the signal, and samples the waveforms only during the observation intervals. Only events during time T, therefore, have a bearing on his decision.

Shannon (1949) showed that a noise waveform of infinite duration may be reconstructed by sampling amplitudes every 1/2W seconds. A noise sample of finite duration, T, may be discretely approximated by making 2WT observations. Consequently, the energy of the noise burst may be approximated by obtaining 2WT estimates of power.

Let us denote (X_i) as the amplitude sampled at the ith point in time. The energy of the noise burst (E_n) is then

$$E_n = \sum_{i=1}^{2WT} X_i^2 .$$

The amplitude distribution of the noise is Gaussian, with zero mean, and variance N. If we normalize this distribution so that the mean equals zero and the variance equals 1, then

$$z_i = \frac{X_i - 0}{\sqrt{N}} .$$

Or,

$$X_i = z_i\sqrt{N} .$$

Power, then, is

$$X_i^2 = z_i^2 N .$$

Since the noise waveform may be approximated by constructing amplitudes spaced at intervals of 1/2W, energy will be the sum of 2WT products of power and interval width (1/2W). That is,

$$E_n = \sum_{i=1}^{2WT} z_i^2 N(1/2W) .$$

Or simply,

$$E_n = N(1/2W) \sum_{i=1}^{2WT} z_i^2 .$$

Power divided by N (which is z^2) is distributed as chi-square with 2WT degrees of freedom. The mean of a chi-square distribution is equal to its degrees of freedom (df), and the variance is simply 2df. In this case, the mean equals 2WT and the variance is 4WT.

Green considers the chi-square distribution to be normal when

$2WT$ is greater than 30. The mean power of the noise distribution is $2WTN$ (recall that $X_i^2 = z_i^2 N$), and the variance is $4WTN^2$.

The mean energy of the noise distribution is therefore $2WTN(1/2W)$, and the variance is $4WTN^2(1/2W)^2$. The mean reduces to WTN_0, and the variance to WTN_0^2, where $N = N_0 W$, and N_0 is the spectral power density of the noise waveform.

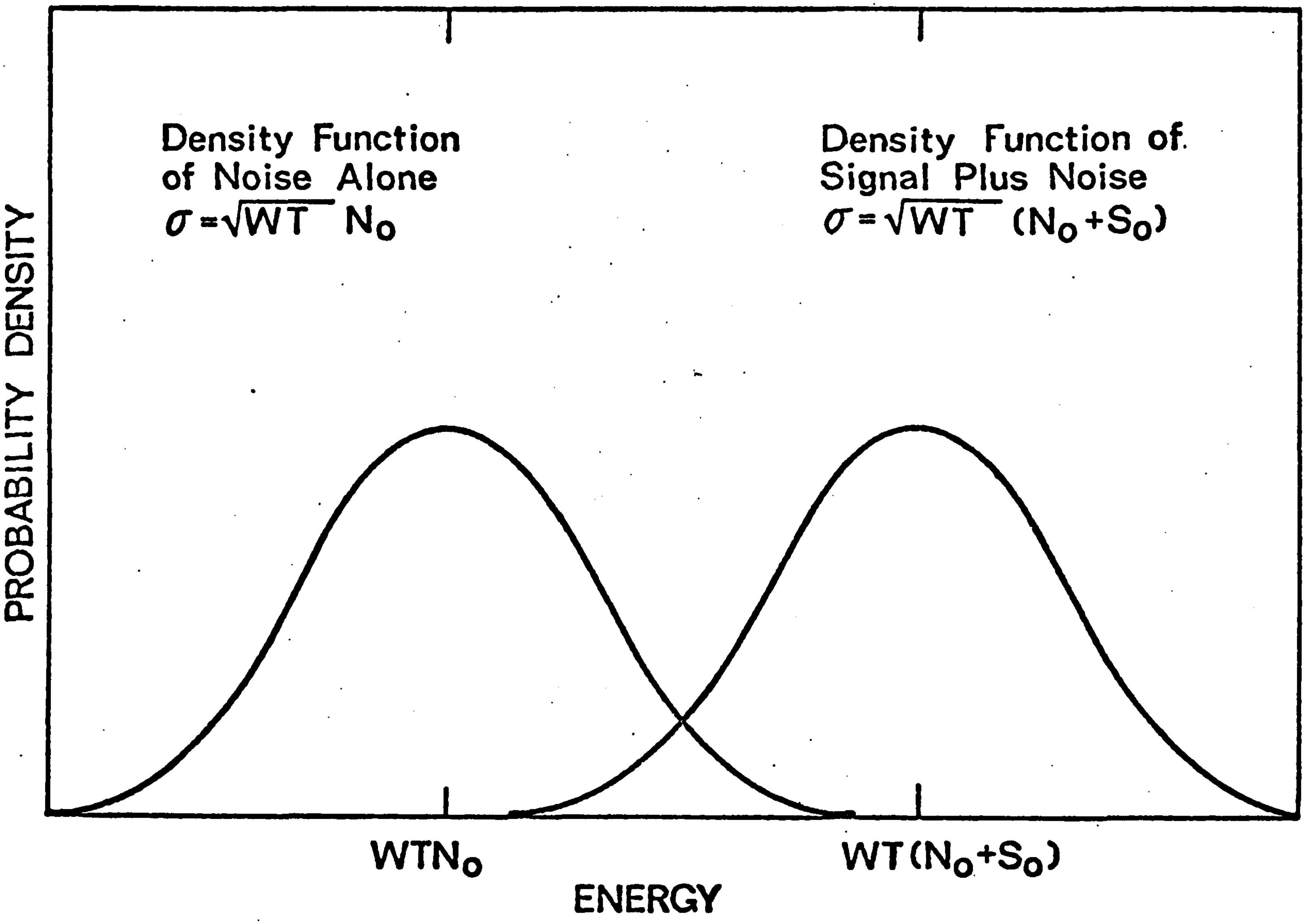
A similar analysis may be extended to the case of signal-plus-noise, by merely substituting $S+N$ for the variance of the amplitude distribution.

The energy distributions for noise and signal-plus-noise are depicted in Figure 1-1: they are of the form Normal(WTN_0, $\sqrt{WT}\, N_0$) and Normal($WT(S_0+N_0)$, $\sqrt{WT}\,(S_0+N_0)$), respectively.

It has been stated that for noise-intensity discrimination, the optimal decision rule is to select the interval with the greater energy. The probability of a correct response is, therefore, the probability that the energy of the sample drawn from the signal-plus-noise distribution (E_{s+n}) is greater than that drawn from the distribution of noise alone (E_n). This is simply the probability that the statistic ($E_{s+n} - E_n$) is greater than zero. This probability is calculable only if the form of the distribution from which this statistic is drawn is known. Since E_{s+n} and E_n are assumed to be normally distributed, so too are their differences. The mean of the difference distribution is the difference of the means, and the variance of the difference distribution is the sum of the individual variances, since the distributions of noise and signal-plus-noise are independently sampled at two intervals in time. The resulting difference distribution is of the form

Figure 1-1

Theoretical distributions for signal-plus-noise and noise alone.


Density Function
of Noise Alone
$\sigma = \sqrt{WT}\, N_o$
Density Function of
Signal Plus Noise
$\sigma = \sqrt{WT}\,(N_o + S_o)$
PROBABILITY DENSITY
WTN_o
$WT(N_o + S_o)$
ENERGY

$\mathrm{Normal}(WTS_o, \sqrt{WTN_o^2 + WT(S_o+N_o)^2})$.

The normalized mean of this approximately Gaussian difference distribution is

$$M = \frac{WTS_o}{\sqrt{WTN_o^2 + WT(S_o+N_o)^2}} \quad . \qquad (1\text{-}1)$$

Or simply

$$M = (WT/2)^{1/2} \cdot S_o/N_o \cdot \frac{1}{\left[1 + S_o/N_o + \tfrac{1}{2}(S_o/N_o)^2\right]^{1/2}} \quad . \qquad (1\text{-}2)$$

Several aspects of the ideal receiver's performance are predicted by Equation 1-2: (1) Weber's Law obtains — i.e., for fixed W and T, S_o is proportional to N_o, in order to maintain a given level of detectability (M); (2) For a fixed masker level (N_o) a decade increase in T will yield a half log-unit (5 dB) decrease in the Weber fraction; (3) Since W and T are treated symmetrically by the model, a decade increase in W will also yield a 5-dB improvement in the Weber fraction; (4) The form of the psychometric function is specified as the upper half of a normal ogive (percent correct plotted as a function of S_o/N_o) with a "slope" such that, for example, a 7.1-dB increase in S_o/N_o is required for performance to improve from 60% ($M = 0.25$) to 90% ($M = 1.28$) correct detections. (Percent correct is merely the integral of the difference distribution from $E_{s+n} - E_n = 0$ to infinity, and may be obtained from a table of the normal probability integral. It should also be noted that the more commonly used measure of performance, d', is equal to $\sqrt{2}\,M$.)

In addition to the predictions resulting from Equation 1-2, two further predictions follow from the assumptions of this model: (5) Since the receiver is posited to have exact knowledge of the signal and noise waveforms, and employs a rectangular filter matched to the signal bandwidth, masker energy outside the signal passband will have no effect on performance. The ideal observer, therefore, performs equally well with noise maskers which are either matched in bandwidth to the signal ("homogeneous") or wider than the signal ("heterogeneous"). (6) The ideal receiver is equally selective in the time domain. Events occurring before and after the signal is presented will not influence the observer's decision process. Equivalent performance is therefore predicted for paradigms in which the noise masker is externally gated simultaneously with the signal, or is continuously present throughout an experimental session. These two masker conditions will henceforth be referred to as "gated" and "continuous."

Data from real listeners bearing on the above predictions and assumptions will be presented in the following sections. The psychoacoustic literature employing heterogeneous noise stimuli ($W_s < W_n$) will be reviewed first. It will be shown that prediction 5 fails to be supported: the observer appears unable to filter accurately to the bandwidth of the signal. If this assumption is satisfied by filtering the waveforms externally, confirmation of the remaining predictions of the model is anticipated. The literature employing homogeneously filtered bands of noise will, therefore, be reviewed next. It will be seen that many of the predictions of Green's model fail to be supported for human observers.

First, however, it would be instructive to describe some of the psychophysical methods most widely employed in the studies to be reviewed. Special consideration will be devoted to the Up-Down

Transformed-Response method, inasmuch as it has been used to collect the data presented in this dissertation.

B. Psychophysical Methods

Methods used by psychophysicists to measure thresholds may be divided into two classes — constant stimulus or fixed methods, and adaptive or tracking methods. In the former, a number of observations are placed at each of several predetermined stimulus levels. The order of presentation may be randomized, or organized into blocks of fixed intensity. The probability of detecting the signal may then be calculated for each intensity level, and if the stimulus levels were judiciously chosen to cover a reasonable portion of the psychometric function, desired levels of performance may be interpolated from the best-fitting psychometric function.

Adaptive, or staircase, methods differ from fixed methods in that rather than seeking to determine the percentage of correct responses for a given stimulus level, the stimulus level which would yield a specific proportion of correct responses is sought directly. In order to track a single point on the psychometric function, adaptive methods concentrate the trials in the region of interest by using some portion of the history of the run in determining which stimulus level to test next. A variety of adaptive procedures exist, differing basically in how level-changing decisions are made.

1. The Up-Down Method

The simplest adaptive procedure is merely an extension of the method of limits. Dixon and Mood (1948) first described the "Up-and-Down" method in relation to testing the sensitivity of explosives to shock. A weight is dropped on an explosive mixture from various heights in order to determine the height at which the mixture explodes 50% of the time. With this procedure, testing begins by dropping the weight from some arbitrarily chosen height. If the mixture fails to explode, the next weight is dropped from a higher level. Testing continues in this manner until the mixture explodes; the subsequent trial is made from a lower level, and so on. This procedure has the advantage that trials will automatically be concentrated near the 50% point.

The application to auditory psychophysics (intensity-discrimination in particular) is immediately apparent — if an observer fails to correctly detect a signal, the signal will be made more intense on the next trial; otherwise, it will be made less intense. This method will converge on the level at which the observer correctly detects the signal 50% of the time.

(a) The Up-Down Transformed-Response Method (UDTR)

Wetherill (1963) described a modified version of the up-down procedure, generalized to track performance levels other than 50%. (This procedure was actually first implemented by Heinemann (1961) to obtain luminance difference limens in a 2 AFC paradigm.) Rather than altering the stimulus intensity after each response, levels are changed only after a predetermined sequence of responses. This is perhaps best explained by example. A typical strategy involves decreasing the level only after two consecutive correct responses have been made, and increasing the level after each incorrect response.

It can be shown that this strategy converges on the stimulus level for which the probability of a correct response is .707. The stimulus level will be decreased under one condition only — the occurrence of two consecutive responses. An increase in level will follow either of two response sequences — a single incorrect response, or a correct response followed by an incorrect response. If the probability of obtaining a correct response equals P(X) and the probability of obtaining an incorrect response equals 1 - P(X), it follows that:

$$\text{P(decreasing the level)} = P((P(X))^2$$

$$\text{P(increasing the level)} = P(X)(1-P(X)) + (1-P(X)) .$$

Since the up-down procedure converges on a stimulus level where the probability of increasing the level equals the probability of decreasing the level,

$$(P(X))^2 = .5 \text{ and, therefore,}$$

$$P(X) = .707 .$$

This modified staircase has been termed by Wetherill (1963) the "Up-Down Transformed-Response" procedure, and will be referred to in this thesis as UDTR, followed by the percent correct that that particular version was designed to track (for example, UDTR(70.7%) above). This method has the advantage that any point on the psychometric function may be estimated by merely altering the strategy for level change. Levitt (1971) provides some rules for tracking points other than 50%; these and some others are presented in Table 1-1. Notice that in order to track points on the psychometric function appreciably higher than 75% correct, long blocks are necessary, and may be impractical to implement.

With UDTR, the experimenter makes several methodological decisions. These include: (1) the initial stimulus level, (2) the size of the step taken upon level change, and (3) the method by which the parameter is estimated from the history of the run.

(1) The choice of the initial stimulus level _seems_ straightforward. Trials will be maximally concentrated in the region of interest if the initial stimulus is chosen to equal the target level. A number of invesigators have consequently suggested that the run begin at the anticipated threshold (see for example, Dixon and Mood, 1948; Cornsweet, 1962; and Wetherill and Levitt, 1965). In practice, however, this is rarely the case. In addition to the experimenter's initial uncertainty as to the threshold level, it is currently in vogue to begin a staircase run with clearly discriminable stimuli in order to familiarize the observer with the task (see Campbell, 1963; Stuckey, Hutton, and Campbell, 1966; Taylor and

Table 1-1

Decision Rules for Tracking Various Levels of Detectability with UDTR

Percent Correct	Number Correct Responses	Number Incorrect Responses
50.0	1	1
57.9	4	3
63.6	5	3
70.7	2	1
73.6	5	2
80.0	7	2
89.1	6	1
93.3	10	1

Creelman, 1967; Campbell and Lasky, 1968; Pollack, 1968; and Raab and Goldberg, 1975). This practice continues despite the many recommendations made against it. Dixon and Mood (1948) first warned that the estimate of threshold will be biased appreciably toward the initial testing level if the initial testing level is poorly chosen (i.e., far from the anticipated threshold). This can be especially troublesome if the step size is small. This caution has since been reiterated by Wetherill and Levitt (1965), Pollack (1968), Creelman and Taylor (1969), and Levitt (1971).

Stuckey, Hutton, and Campbell (1966) used the results of a computer simulation to defend their practice of starting with audible signals. The initial stimulus level as well as other procedural parameters were varied in a Monte Carlo simulation of Campbell's adaptive method. Values of the initial stimulus level were chosen to be equal to the anticipated threshold (75% point on the psychometric function presented to the computer), close to, and well above this level. Simulations of Campbell's method, which includes only levels visited more than once in the parameter estimation, show a negligible bias resulting from starting far from the anticipated threshold. This conclusion, however, is misleading in that the results presented in Stuckey's Table III-I (p. 1175) were averaged across several procedures, one of which was the length of the run, which was allowed to vary from 96 to 240 trials. It is entirely possible (as was suggested by Dixon and Mood, 1948) that the bias arising from an inappropriate choice of starting level will be reduced by extending the length of the run. Campbell and Lasky (1968), in fact, found thresholds to decrease with number of trials for human observers. In order to avoid observer fatigue, runs as long as 240 are rarely employed. In a later chapter of this dissertation, Monte Carlo simulations of UDTR employing runs of 80 trials show an appreciable dependence of the estimated parameter on the choice of starting level. If the run is allowed to continue for as many as 240 trials, however, this bias becomes negligible.

It remains, that although it is optimal to begin adaptive testing as close as possible to the anticipated threshold level, this condition is rarely realized in practice.

(2) UDTR typically employs a fixed step size. Cornsweet (1960) suggests that stimuli should be spaced so as to approximate equal sensory steps. In audition, a logarithmic scale is well suited. Both Cornsweet, and Dixon and Mood (1948) state that the up-down procedure becomes maximally efficient when the steps are the size of the differential threshold. In noise-intensity discrimination, a step size of 1.5 dB is frequently employed, which allows approximately 6 levels to cover a range of detectability of 60% to 90% correct detections for Green's ideal observer.

In attempting to circumvent the possibility of compounding the bias produced by an inappropriate choice of starting level with too small a step size, it has often been suggested that UDTR start with relatively large steps, which would be made smaller as the run progressed (see Dixon and Mood, 1948; Cornsweet, 1962; Wetherill and Levitt, 1965; Levitt and Bock, 1967; Levitt and Treisman, 1969; and Levitt, 1971). Levitt (1971) suggests a strategy whereby the step size is halved after the first, third, seventh, fifteenth, etc. reversals, whereas Levitt and Bock (1967) suggest a simpler strategy of halving step size only once after a few initial trials. Such strategies, however, require additional stimulus programming. In noise-intensity discrimination, a fixed step size is frequently employed, combined with the exclusion of some portion of the run in which the region of interest is approached.

(3) In estimating threshold from the history of a staircase run, it is common practice to begin by discounting some portion of the initial trials in which the region of interest is approached. A typical strategy is to collect data only after the first reversal in the direction of level-change occurs. This should hopefully reduce the bias introduced by an inappropriate choice of starting level. It will be seen in a later chapter, however, that a sizeable bias remains despite this precautionary measure.

A number of alternatives exist for estimating threshold from the remaining history of the run. First, one may average the "peaks and valleys" of the staircase run, assuming these points of change in direction of level-change represent momentary thresholds. Second, threshold may be defined as the stimulus level that receives an equal number of correct and incorrect responses (treating the staircase as a modified method of constant stimuli). Third, the mean of levels revisited, weighted by the frequency of presentation may be calculated. Fourth, the median of levels revisited may be calculated.

Feeny, Kaiser, and Thomas (1966) tested the equivalence of the first three methods. UDTR(70.7%) was employed in a 2AFC task in which the observer had to report in which of two positions a blurred contour was presented. After discarding the first 20 trials of 60-trial runs, threshold was estimated by each method. It was concluded that the three methods yielded equivalent results. Rose, Teller, and Rendleman (1970) investigated the equivalence of using the mean or median as a threshold estimate. For a computer-simulated listener, a 2IFC task was employed with UDTR(70.7%); the number of trials per run was varied from 25 to 200. After excluding the first 10 trials, mean and median estimates did not differ, although the within-run variability for median estimates averaged 5% higher than that obtained for means. It may be concluded that all four measures yield equivalent estimates of threshold.

In sections to follow, the psychoacoustic literature bearing on the predictions of Green's model will be reviewed. When UDTR is used, the following conditions are typically employed: the initial stimulus level is high (stimuli are easily discriminable), each run is approximately 100 trials, step size is fixed, trials occurring before the first reversal are discarded, and the median of the remaining history is taken as the estimate of threshold.

(b) The Blocked Up-Down Two-Interval Forced-Choice Procedure (BUDTIF)

Campbell (1963) offers another staircase procedure designed to track points on the psychometric function other than 50%. Campbell's Blocked Up-Down Two-Interval Forced-choice method (BUDTIF) differs from UDTR in that the stimulus level need not be changed after the completion of a block of trials. For example, in tracking a level yielding 73.4% correct detections, a BUDTIF run will call for a level change if the number of correct responses is greater or less than 3 in a block of 4 trials; if exactly 3 trials are correct, the level remains unchanged for the next block. In addition, blocks must be completed. If the first two responses in a block are incorrect, it is clear that the next block will be tested at a higher level regardless of the responses on the next two trials. Whereas UDTR would change levels at this point without completing the block, BUDTIF blocks are run to completion.

In all other respects, BUDTIF is procedurally identical to UDTR: the initial stimulus level is high (stimuli are easily discriminable), each run is approximately 100 trials, step size is fixed, trials occurring before the first reversal are discarded, and the median of the remaining history is taken as the estimate of threshold. (For a discussion of these procedural variables with respect to the BUDTIF method, see Campbell, 1963; Campbell, Hutton, and Stuckey, 1966; Stuckey, Hutton, and Campbell, 1966; Campbell, 1967; and Campbell and Lasky, 1968.)

2. Parameter Estimation by Sequential Testing (PEST)

In an attempt to increase the efficiency of adaptive procedures, Taylor and Creelman (1967) developed a method differing substantially from UDTR and BUDTIF. Although this method has not been used to provide the data to be reviewed in the following sections, a brief review is nonetheless warranted.

Taylor and Creelman's version of PEST (Parameter Estimation by Sequential Testing) employs the following decision rules:

(1) A Wald (1947) sequential likelihood-ratio test is used to determine whether performance at the current level is above or below the target probability. In short, a running account of the entire history of the run is used to determine whether performance falls within some predetermined boundary surrounding the target percent correct. If performance falls below the lower bound, the next trial is presented at a higher level, if above the upper bound, the next trial is at a lower level. If performance falls within the bounds, no change is made and testing continues.

(2) Step size is not fixed with PEST. The initial stimulus level is chosen to be easily discriminable, and the step size is initially large. As testing continues, step size is allowed to change by the following rules: (a) Step size is halved for every reversal in step direction. (b) If a second step is called for in a given direction, the step size remains the same as the first. (c) The fourth and subsequent steps in a given direction each double their predecessor (within some predetermined upper bound on step size). (d) The size of the third sucessive step in a given direction is determined by the history of reversals. If the step immediately preceeding the most recent reversal resulted from a doubling, the third step is not doubled, while if that reversal was not the result of a doubling, the third step is doubled.

(3) A PEST run is terminated when the step size reaches a predetermined minimum. By computer simulation, Creelman and

Kaplan (1973) estimate that 60 trials are necessary to reach this criterion.

(4) In the original implemetation of PEST, the threshold is the current stimulus level at which this final change of level is called for.

Another version of PEST is offered by Pollack (1968), which differs from Creelman's in its rules for level-changing. Pollack employs the same decision rules as are used in BUDTIF: in tracking 73.4% correct detections, the stimulus level is increased if fewer than 3 correct responses are made in a block of 4 trials, decreased if 4 correct responses are made, and left unchanged if exactly 3 out of 4 responses are correct. In a Monte Carlo simulation, Rendleman, Rose, and Teller (1971) report that the between-runs variability of Pollack's PEST(73.4%) is equal to that of UDTR(70.7%). This finding is unexpected in light of PEST's more sophisticated stimulus programming. In addition, Rendleman et al. question the efficiency of Pollack's PEST. Although the number of trials required to achieve the minimum step size averaged 80, variability was considerable. Computer simulated runs ranged in length from 20 to 310 trials. This variability could become troublesome in practice, since neither the experimenter nor the observer would know beforehand how long the run would last.

In summary, adaptive methods are well-suited to track a given level of performance. Fixed methods, on the other hand, are inefficient in tracking a single point on a psychometric function since many observations are placed in regions of little interest. In keeping with a suggestion made by Smith (1961), adaptive methods have been widely used in psychoacoustics when a single level of performance is sought, while fixed methods have been largely reserved for instances where the form of the entire psychometric function is of interest. An exception to this generalization is Campbell's (1965) use of BUDTIF to track three levels of difficulty in a noise-intensity discrimination paradigm. Decision rules were designed to track 62%, 75%, and 88% correct detections by repeating a level if exactly 5, 6, or 7 correct responses were made in 8-trial blocks. The resulting three-point psychometric functions exhibited the same "steepness" characteristic of those more arduously obtained by Green (1960) with blocks of fixed intensity.

C. Studies Employing Heterogeneous Noise

1. Intensity-Discrimination Functions - Weber's Law

It has long been known to psychophysicists that for many sensory continua, the Weber fraction (expressed as the ratio of the just-noticeable-increment in intensity, ΔI, to the standard comparison intensity, I) is constant over a moderate range of intensities. That is,

$$\Delta I/I = k \quad (1\text{-}3)$$

The extent to which Weber's Law is obeyed may be assessed by plotting the Weber fraction as a function of intensity. A function well fit by a horizontal line is predicted. Alternatively, the intensity-discrimination function may be displayed by plotting the just-noticeable-increment as a function of intensity. When plotted on log-log coordinates, a straight line with a slope of 1.0 is predicted. Consider a more general form of Equation 1-3:

$$\Delta I = kI^{m} . \qquad (1\text{-}4)$$

Taking logarithms of both sides of Equation 1-4, we obtain:

$$\log \Delta I = m \log I + \log k . \qquad (1\text{-}5)$$

The slope of the resulting intensity-discrimination function is m.

The few studies that have used heterogeneous bands of noise in their investigations of Weber's Law yield conflicting results. Campbell (1964) obtained a departure from Weber's Law for a number of heterogeneous masker conditions. Campbell's subjects were required to detect the presence of a one second burst of narrowband noise added to a continuous wideband masker (W_n = 4940 Hz). The signal was white Gaussian noise with a bandwidth of either 300 or 700 Hz, each centered at a number of frequencies; a wideband signal (W_s = 3200 Hz) was also employed. In each condition, the overall masker level was varied from 0 to 90 dB SL, in 10-dB steps. Thresholds tracked with BUDTIF yielded Weber's Law for masker levels ≥20 dB SL with the exception of a single point. For all conditions, Weber fractions were 2 dB larger when the masker level equalled 60 dB SL. This small but consistent departure from Weber's Law is evident in Campbell's homogeneous data as well, but curiously enough, is not evidenced in the data of any other researcher.

Moore (1975) presents data for 300-msec bursts of 1/3-octave bands of noise gated simultaneously with 2-octave maskers. Masker levels of 10, 20, 30, 40, and 50 dB SL were investigated with a two-alternative forced-choice procedure (UDTR(73.6%)). Weber's Law obtained for signals centered at 1000, 3150, and 4000 Hz (W_s = 232, 729, and 926 Hz, respectively). For signals centered at 5000 and 6300 Hz (W_s = 1158 and 1459 Hz), performance worsened with level, thereby violating Weber's Law.

Schacknow and Raab (1976) replicated many of the conditions of the Campbell (1964) and Moore (1975) studies. UDTR(73.6%) provided thresholds for 100-msec signals with both gated and continuous maskers. Signal bandwidths of 100, 316, 1000, and 3160 Hz were presented with maskers of broader bandwidth. Masker levels of 5, 25, and 45 dB SPL were tested: Weber's Law obtained in every condition. These results are in accord with the findings of Moore (1975) for narrowband signals, but differ from Moore's findings, where performance deteriorated with increasing masker level for signal bandwidths of 1158 and 1459 Hz. A direct comparison cannot be made between the findings of Campbell (1964) and Schacknow and Raab (1976) for continuously masked signals, due to the latter's choice of a limited range of masker levels.

The effects of masker level on heterogeneous noise-intensity discrimination are summarized in Table 1-2. Weber's Law obtains for continuously masked signals (with the possible exception of a small decrement in performance at approximately 60 dB SL). For narrowband signals ($W_s \leq 1000$ Hz) with gated maskers, Weber's Law obtains in both

Table 1-2

Studies Investigating Weber's Law for Heterogeneous Masker Conditions

Study	Method	Mode of Masker	Bandwith*	Findings
Campbell (1964)	BUDTIF (75%)	Cont.	S = 300, 700, 3200 M = 4940	Weber's Law
Schacknow and Raab (1976)	UDTR (73.6%)	Cont.	S = 100, 316, 1000, 3160 M = 1000, 3160, 10000	Weber's Law
Schacknow and Raab (1976)	UDTR (73.6%)	Gated	S = 100, 316, 1000, 3160 M = 1000, 3160, 10000	Weber's Law
Moore (1975)	UDTR (73.6%)	Gated	S = 232, 729, 926 M = 1500, 4725, 6000	Weber's Law
Moore (1975)	UDTR (73.6%)	Gated	S = 1158, 1459 M = 7500, 9450	Performance worsens w/level

*S and M refer to signal and masker bandwidths, respectively.

the Moore (1975) and Schacknow and Raab (1976) studies. For signal bandwidths wider than 1000 Hz, results are conflicting: Schacknow and Raab report Weber's Law, while Moore reports a decline in performance with increasing masker level.

2. Slope of the Psychometric Function

For a 2IFC paradigm, the psychometric function is predicted to be the upper half of a normal ogive when percent correct is plotted as a function of S_0/N_0. Its "slope" can be calculated over a reasonably linear portion; an approximately 7-dB increase in S_0/N_0 is required for performance to improve from 60% to 90% correct detections. There is a paucity of data bearing on this prediction. Most studies employing heterogeneous maskers have used adaptive procedures to track a signal level yielding approximately 75% correct detections, thereby discarding the information necessary for the reconstruction of the observer's psychometric function.

Psychometric functions are provided by Green (1960), who collected data by presenting blocks of fixed intensity. In this manner, psychometric functions were constructed from 5 or 6 points spanning a range of approximately 50% to 100% correct detections. These are the only available psychometric functions for heterogeneous noise. Green's signals were 250-msec bursts of narrowband noise (Ws = 650 Hz), presented at several center frequencies. The masker bandwidth was always 5100 Hz, and was presented continuously throughout an experimental session. All psychometric functions were considerably steeper than predicted. Detectability grew rapidly: a change of only 3 dB was required for performance to increase from 60% to 90% correct detections.

Green (1960a) attributed this characteristic steepness to the observer's uncertainty concerning the signal onset and duration. Further discussion of the uncertainty model will be reserved for a later section, where it will be shown that the psychometric function is of the proper slope for gated masker conditions in which temporal uncertainty is reduced.

3. Effects of the Mode of Masker

A basic assumption of Green's model is that the observer filters in the time domain -- i.e., decisions are based on events occurring only during the observation intervals (which are assumed to be equal to the signal duration). Equivalent performance is therefore predicted in paradigms in which the masker is continuous, or gated simultaneously with the signal.

It is well documented that this equivalence is not found in the literature for tone-intensity discrimination (see for example, Green, 1964; Campbell, 1966; Campbell and Lasky, 1967; Green, 1969; Zwicker and Fastl, 1972; and Leshowitz and Cudahy, 1975), or tone-masked-by-noise (see Sherrick and Albernaz, 1961; Green, 1964; Dirks and Norris, 1966; Tucker, Williams and Jeffress, 1968; Campbell, 1969; and Wier, Green, Hafter, and Burkhardt, 1977). For tone-intensity discrimination, a gated masker is more effective than a continuous masker; thresholds differ by approximately 10 dB when the signal and masker are of the same frequency. For tone-masked-by-noise, again gated maskers are more effective, the difference in performance averaging 4 dB. It seems paradoxical that gated maskers should be more effective. If the observer's temporal filtering is inefficient, performance with a continuous masker should be worse due to the additional noise present before and after the signal is presented. Explanations for these obtained gated-continuous differences in performance are, therefore, not to be found in a straightforward analysis of the temporal filtering abilities of the observer.

Schacknow and Raab (1976) present the only within-subjects analysis of performance with gated and continuous maskers for the case of noise-masked-by-noise. For a wide range of signal and masker bandwidth combinations, 100-msec signals were easier to detect (by 4 dB) when masked continuously.

4. Effects of Increasing Masker Bandwidth

Green's ideal observer is assumed to have exact knowledge of the signal bandwidth, and to employ an appropriately matched rectangular filter. Masker energy outside the signal passband will, therefore, have no effect on performance. A direct test of this assumption is to increase the masker bandwidth as the signal bandwidth remains fixed. Performance should remain unaffected by additional masker energy outside the observer's internal filter.

Moore (1975) presents such data. For each of three center frequencies, a 1/3-octave noise signal was masked by either a 1/3-octave (homogeneous) or a 2-octave band of noise (heterogeneous). In each case, signal and masker waveforms were gated simultaneously. Increasing the masker bandwidth resulted in an average decrement in performance of 5.6 dB.

Schacknow and Raab (1976) present data that are in accord with Moore's findings. For each of three signals (W_s = 316, 1000, and 3160 Hz), the masker bandwidth was increased from homogeneous to 10000 Hz in half log-unit steps. Signals were 100-msec bursts embedded in either a gated or continuous background. For the two narrowband signals, performance declined dramatically with the addition of a small amount of masker bandwidth, and then remained constant as the masker bandwidth grew further. For continuously masked signals, the decrement in performance resulting from a half log-unit increase in W_n was approxiamtely 8 dB for the 316 Hz signal and 7 dB for the 1000 Hz signal. Performance with the 3160 Hz signal declined only 3.5 dB as a result of a half log-unit increase in W_n. The effect of increasing the masker bandwidth was similar for gated conditions, although the decrement in performance was somewhat less.

The above results pose serious problems for Green's detection scheme as applied to human observers. First, a basic assumption of the model is violated — the subject does not impose an internal filter matched to the signal bandwidth. Second, even if the model is modified by increasing the bandwidth of the internal filter, the observed decrement in performance due to increasing masker bandwidth is too large to be accounted for by additional stimulus fluctuation. That is, a 5-dB increase in W_n (a half log-unit) could not produce the 8-dB decrement in performance reported by Schacknow and Raab. Appropriate modification of the model, however, is difficult at this point, since the increases in masker bandwidth in the above studies were too coarsely spaced to accurately trace the decline in performance accompanying the increase in energy outside the signal passband.

5. Intensity-Bandwidth Reciprocity

Stimulus fluctuations are the sole source of error for Green's optimal receiver. A decrease in stimulus variability will, therefore, result in an improvement in performance. Consider the relative variability of a noise signal (sigma-to-mean ratio):

$$\frac{\sigma}{\mu} = \frac{\sqrt{WT}\,N_o}{WTN_o} \qquad (1\text{-}6)$$

This reduces to:

$$\frac{\sigma}{\mu} = \frac{1}{\sqrt{WT}} \qquad (1\text{-}7)$$

(Recall that for the ideal receiver, $W = W_s = W_n$ since an internal filter is set equal to the signal passband.)

Note from Equation 1-7, that as either the bandwidth (W) or duration (T) of the signal increases, the relative variability decreases by the square root of the corresponding increase in either W or T. The change in the Weber fraction (measured in dB) as a result of a log-unit increase in either W or T, is known as the "reciprocity factor." For the ideal receiver, a 5-dB decrement in performance is predicted for either W or T. To date, no experiments have been performed with heterogeneous noise to empirically determine the reciprocity factor for duration, whereas a number of studies confirm Green's prediction for increasing bandwidth.

Green (1960) presented his subjects with 250 msec bursts of narrowband noise (Ws = 650 Hz, at a number of center frequencies), masked by a continuous wideband noise (Wn = 5100 Hz). At an overall masker level of 40 dB SPL, performance was independent of center frequency; Weber fractions averaged -6.15 dB. The same subjects were then required to detect a 250 msec increment in the continuous wideband masker; Weber fractions averaged -9.80 dB. This approximately 8-fold change in signal bandwidth (from 650 to 5100 Hz) resulted in a 3.65-dB improvement in performance. Extrapolated to a full log-unit increase in W, the obtained reciprocity factor is -4.08 dB, somewhat smaller than predicted.

Campbell (1964) replicated Green's findings with an adaptive method of data collection (BUDTIF). Weber fractions were obtained for 1-second signals masked by a continuous wideband noise (W_n = 4940 Hz). Signal bandwidths were: (1) 300 Hz arithmetically centered at 250, 500, 1000, 1500, and 2000 Hz; (2) 700 Hz centered at 500, 1000, 2000, and 3000 Hz; and (3) 3200 Hz centered at 1750 Hz. For each condition, the masker level was varied from 0 to 90 dB SL in 10-dB steps. Center frequency was found to interact with masker level for both weak and intense maskers. At moderate intensities (40 dB SL) Weber fractions were independent of center frequencey. Averaged across subjects and center frequencies for 40-dB maskers, Weber fractions were -6.00, -7.75, and -11.00 dB for signal bandwidths of 300, 700, and 3200 Hz, respectively. The reciprocity factor is calculated to be -4.87 dB/decade increase in bandwidth.

Schacknow and Raab (1976) provide IW trades for both gated and continuous conditions. A wideband noise (W_n = 10000 Hz) was used to mask 100-msec signals whose bandwidth varied in half log-unit steps from 100 to 10000 Hz. Averaged across masker levels and subjects, IW reciprocities were -6.0 and -5.1 dB/decade increase in W for gated and continuous conditions, respectively.

Averaged across studies, the intensity-bandwidth reciprocity factor is -5 dB (see Table 1-3), a finding which is in accord with the prediction of Green's model.

6. Real-Ideal Differences in Performance

It has been stated that stimulus fluctuations are the sole source of error for Green's optimal receiver. The performance of a real observer can be no better than that which the inherent variability of the stimulus allows. It is then likely to be degraded, in that a human sensory system is not a perfect transducer of physical stimuli. This additional noise serves to make the detection process less efficient. The extent of this degradation is, however, not a priori known. Green (1960), in the original test of his model, reported Weber fractions to be approximately 5-dB worse than ideal for 250-msec narrowband signals (W_s = 650 Hz) masked by a continuous wideband noise (W_n = 5100 Hz). A number of other studies are reviewed below in which obtained Weber fractions for heterogeneous noise are

Table 1-3

Studies Investigating the Intensity-Bandwidth Reciprocity for Heterogeneous Masker Conditions

Study	Method	Mode of Masker	Reciprocity Factor (in dB)
Green (1960)	Fixed I	Cont.	-4.08
Campbell (1964)	BUDTIF (75%)	Cont.	-4.87
Schacknow and Raab (1976)	UDTR (73.6%)	Cont.	-5.10
Schacknow and Raab (1976)	UDTR (73.6%)	Gated	-6.00

compared to Green's theoretical ideal.

Campbell (1964), in a replication of Green (1960), reported Weber fractions for continuously masked signals that averaged 4.18-dB worse than ideal. (Although Campbell's signals were 1 second in duration, a signal duration of 300 msec was used in calculating the ideal Weber fraction since for signals longer than 300 msec, performance is independent of duration (Green, 1960). This rule will be used for all studies employing signal durations longer than 300 msec.) Schacknow and Raab (1976) provide data for continuously masked noise. For heterogeneous conditions, Weber fractions averaged across signal and masker bandwidths were 5.50-dB worse than ideal.

It has been shown that noise maskers are more effective when gated simultaneously with the signal, than when presented throughout an experimental session. It follows that real-ideal differences in performance will be greater for gated maskers. Schacknow and Raab (1976) show that for gated maskers, the real-ideal difference in performance is 10.22 dB (averaged across signal and masker bandwidths). Moore (1975), presenting 1/3-octave signals in 2-octave maskers, reports Weber fractions that are 12.29-dB worse than ideal (averaged across signal bandwidth, at an overall masker level of 40 dB SL).

7. Summary of Results Obtained with Heterogeneous Noise

The psychoacoustic literature for the detection of a Gaussian noise waveform added to a masker of greater bandwidth has been reviewed. To summarize the data bearing on the predictions of Green's model as listed in section A:

(1) Weber's Law is obtained for continuously masked signals, with the possible exception of a small decrement in performance at a masker level of approximately 60 dB SL. For gated maskers, Weber's Law is upheld for narrowband signals ($W_s \leq 1000$ Hz) embedded in wideband maskers, while the literature is inconsistent for signal bandwidths wider than 1000 Hz.

(2) There are no existing data tracing the course of detection as a function of signal duration.

(3) Intensity-bandwidth reciprocities are of the appropriate magnitude, averaging -5 dB/decade increase in W for both gated and continuous conditions.

(4) The psychometric function is steeper than predicted for continuously masked signals; no heterogeneous data have been reported employing gated maskers.

(5) With both gated and continuous maskers, the observer behaves as though unable to employ an internal filter matched to the signal bandwidth -- as a noise masker increases in bandwidth beyond the signal passband, performance progressively declines.

(6) The observer is unable to perform equivalently in gated and continuous masking paradigms — gated maskers are more effective than continuous maskers.

The two basic assumptions upon which the development of Green's model rests do not apply to human observers. First, the observer performs as though unable to limit the bandwidth of his internal filter to the signal passband — as a noise masker increases in bandwidth beyond the signal passband, performance progressively declines. Second, the observer appears unable to filter in the time domain — performance is not equivalent in gated and continuous masker paradigms. In neither case will a simple modification of Green's model reconcile the data and predictions.

It would not be unreasonable to assume that the observer is unable to limit the bandwidth of his internal filter to match that of the signal passband, and instead allows masker energy outside the signal passband to enter the decision process. This leads to the prediction that performance will decline as masker bandwidth increases beyond the signal passband. Although Schacknow and Raab (1976) present such data, the decrement in performance due to increasing masker bandwidth is too large to be accounted for by additional stimulus fluctuation. That is, a 5-dB increase in masker bandwidth (half log-unit) could not produce the obtained 8-dB decrement in performance. Before the human observer's internal filtering may be accurately modeled, however, the decline in performance due to increasing masker bandwidth must be traced with finer steps than those employed by Schacknow and Raab.

Similarly, it is not sufficient to attribute the gated-continuous difference in performance to merely the observer's inability to provide accurate temporal filtering of the stimulus waveforms. If this were the case, a gated masker should provide more accurate timing information that a continuous masker, and therefore lead to better performance in gated masker conditions. Performance is, in fact, worse with gated maskers. Further consideration of gated-continuous differences in performance will be reserved for a later chapter, where it will be seen that for homogeneous masker conditions, IT trades, IW trades, and the slope of the psychometric function are all dependent on the mode of the masker.

D. Studies Employing Homogeneous Noise

In this section, the literature for homogeneous noise-intensity discrimination will be reviewed. It is anticipated that the performance of a real listener will more closely approximate that of Green's ideal observer if the assumptions of his model are satisfied externally (homogeneously filtered signals and maskers satisfy the assumption that the observer filters the masker to match the signal passband). It will be especially instructive to consider the homogeneous conditions in which external filtering is provided in the time domain as well; gated homogeneous masker conditions satisfy both basic assumptions of Green's model.

The literature for homogeneous noise-intensity discrimination is considerably more voluminous than that which has been reviewed for the heterogeneous condition. Data pertaining to each of the predictions of Green's model are available. In addition, data from several labs are remarkably consistent. In this section, it will be seen that satisfying the temporal filtering assumption of Green's model by gating the masker simultaneously with the signal allows for a closer approximation to some of the predictions of the model than is obtained with a continuous masker — Weber's Law is more consistently obtained, and the slope of the psychometric function is of the proper slope. In other ways, however, gated maskers lead to less efficient processing of the signal — intensity-duration trades are considerably more shallow (when $10 \log \Delta I/I$ is plotted as a function of T) than those obtained with continuous maskers, and for long-duration signals (longer than approximately 10 msec) the Weber fraction is larger by about 2.5 dB. In neither gated nor continuous masker conditions, however, are IW trades as predicted — IW trades obtained with homogeneously filtered noise are more shallow ($10 \log \Delta I/I$ as a function of W_s) than predicted.

1. Intensity-Discrimination Functions - Weber's Law

A large number of studies support the claim that the detectability of a noise increment is independent of masker intensity. Those employing continuous maskers will be reviewed first.

Miller (1947), in one of the earliest investigations of noise-intensity discrimination, required his listeners to detect a 1.5-second increment in wideband noise (W = 6850 Hz). The increments were periodically added to a continuous masker at intervals of 4.5 seconds; the listener had to report whether or not the signal was present on that trial. A number of trials were performed at each of 5 to 8 signal intensities, psychometric functions were constructed, and the signal intensities necessary for 50% correct detections were interpolated. For masker levels greater than 30 dB SL, the Weber fraction was constant over a range of 70 dB. Harris (1950), in a replication of Miller's experiment, obtained Weber's Law for intensities greater than 15 dB SL.

Two studies employing two-alternative forced-choice procedures also report a confirmation of Weber's Law for continuous masker conditions. Raab, Osman, and Rich (1963) used an adaptive procedure (UDTR(73.6%)) to determine the Weber fraction for 500-msec increments in a continuous wideband noise (W = 6900 Hz). Weber's Law obtained for masker levels greater than 20 dB SPL. Viemeister (1974b) presented increments in wideband noise which varied in duration from 0.8 to 1600 msec. For each duration, Weber's Law described the data for the continuously masked signals.

Weber's Law has also been obtained in paradigms in which the signal and masker waveforms are gated simultaneously.

Postman (1946) investigated the time-order error for differential sensitivity by varying the interstimulus interval (ISI) between two bursts of noise from 0 to 6 seconds. Eleven comparison stimuli were paired with each of three standard noise bursts ranging from 35 to 75 dB SL. The subject judged the comparison stimulus as "louder" or "softer" than the standard. The duration of the noise bursts (W = 2500 Hz) was 1 second. Psychometric functions (percent judged "louder" as a function of stimulus intensity) were constructed and difference limens interpolated. For ISI's of 1 and 2 seconds, differential sensitivity was greatest and was independent of the standard intensity — Weber's Law obtained.

Harris (1950, 1963) also employed the method of constant stimulus differences in an investigation of wideband noise-intensity discrimination (W = 6900 Hz). For 1-second bursts of standard and comparison stimuli, Weber's Law obtained for intensities greater than 15 dB SL.

Small, Bacon, and Fozard (1959), in a similar paradigm, had subjects adjust the intensity of a 1-second noise burst which was alternated with a 1-second standard burst until a continuous noise was heard. This alternation method corresponds to Postman's, in which ISI = 0. For octave bands of noise whose bandwidths equalled 128, 1040, 4080, and 8033 Hz, performance was described by Weber's Law, although differential sensitivity was relatively poor due to the contiguous nature of the standard and comparison stimuli. (Performance was worst in the Postman study when ISI = 0.)

Pollack (1951) employed the method of limits in a decrement detection paradigm, in which a 5-second presentation of continuous wideband noise was alternated with a 5-second presentation of periodically decremented noise. Decrements lasted 55 msec and were preceeded and followed by 45 msec of the standard intensity. The size of the decrement was either increased or decreased until the 2 bursts were indistinguishable. The Weber fraction was constant over a range of 35 to 95 dB SPL.

Four studies employing 2IFC procedures have also confirmed Weber's Law for gated masker conditions. Viemeister (1974a) employed an adaptive procedure (UDTR(70.7%)) to measure thresholds for 200-msec bursts of wideband noise. Weber's Law held over an 80-dB range of masker intensity. In the course of investigating the effects of signal duration on wideband noise-intensity discrimination, Viemeister (1974b) reported Weber's Law for conditions employing signal durations of .8 to 1600 msec. Moore (1975) also used a tracking procedure (UDTR(73.6%)) to assess the differential threshold for 300-msec signals embedded in gated maskers. For 1/3-octave bands of noise centered at 1000, 4000, and 6300 Hz, the Weber fraction was independent of masker intensities greater than 20 dB SL. In addition, Moore and Raab (1975) confirmed Weber's Law for wideband noise bursts of 10 and 250 msec, tracked by UDTR(73.6%).

Although Weber's Law generally obtains for noise-intensity discrimination, a few studies have reported departures from Weber's Law.

Campbell (1964) found a decline in performance of approximately 2 dB at a masker level of 65 dB SL, for 1-second increments in a continuous wideband noise (W = 5000 Hz). Aside from this decrement in performance, the intensity-discrimination function was essentially linear.

It is well established that the Weber fraction decreases with intensity for pure-tone intensity discrimination (see for example, Dimmick and Olson, 1941; Campbell, 1966; Campbell and Lasky, 1967; McGill and Goldberg, 1968a,b; Viemeister, 1972; Schacknow and Raab, 1976; Penner, Leshowitz, Cudahy, and Ricard, 1974; and Jesteadt, Wier, and Green, 1977). These intensity-discrimination functions have slopes between 0.80 and 1.00. This so-called "near miss" to Weber's Law is also evidenced in a number of studies employing homogeneous noise stimuli.

Green and Sewall (1962) collected thresholds with a 2IFC paradigm in which 100-msec increment signals of fixed intensity were masked by a continuous wideband noise (W = 4000 Hz). After collecting data at 6 stimulus levels, psychometric functions were constructed from which signal levels required for 75% correct detections were interpolated. A near miss to Weber's Law obtained - as the masker intensity increased from 55 to 95 dB SPL, performance improved by approximately 2 dB.

Bos and deBoer (1966) had subjects adjust the intensity of 125-msec increments in a continuous narrowband noise with a Bekesy audiometer. For bandwidths of 200 and 800 Hz, centered at 1000 Hz, Weber fractions were found to vary inversely with masker intensity.

Viemeister (1974a) obtained a near miss to Weber's Law for 200-msec increments in a gated, bandlimited masker (W = 1350 Hz). When a broadband noise floor was added, however, Weber's Law was reinstated.

Schacknow and Raab (1976) provide homogeneous data for both gated and continuous masker conditions. In all conditions, thresholds were determined by UDTR(73.6%) for 100-msec increments. For gated masker conditions, a near miss obtained for bandwidths of 100, 316, 1000, and 3160 Hz. With continuous maskers, a near miss obtained for 100, 316, and 1000-Hz stimuli. Detectability was found to be independent of masker level, however, for broadband noise (W = 10000 Hz, gated; 3160 and 10000 Hz, continuous).

Studies investigating Weber's Law for homogeneous noise-intensity discrimination are summarized in Table 1-4. Weber's Law is obtained almost without exception for wideband stimuli in both

Table 1-4

Studies Investigating Weber's Law
for Homogeneous Masker Conditions

Study	Method	Mode of Masker	Bandwidth	Finding
Miller (1947)	Fixed I	Cont.	6850	Weber's Law
Harris (1950)	Fixed I	Cont.	6900	Weber's Law
Raab, Osman, and Rich (1963)	UDTR (73.6%)	Cont.	6900	Weber's Law
Viemeister (1974b)	?	Cont.	Wideband	Weber's Law
Schacknow and Raab (1976)	UDTR (73.6%)	Cont.	3160, 10000	Weber's Law
Campbell (1964)	BUDTIF (75%)	Cont.	5000	Weber's Law
Green and Sewall (1962)	Fixed I	Cont.	4000	Near Miss
Bos and De Boer (1966)	Békésy Audiometry	Cont.	200, 800	Near Miss
Schacknow and Raab (1976)	UDTR (73.6%)	Cont.	100, 316, 1000	Near Miss
Postman (1946)	Fixed I	Gated	2500	Weber's Law
Harris (1950, 1963)	Fixed I	Gated	6900	Weber's Law

(Continued)

Table 1-4, Continued

Study	Method	Mode of Masker	Bandwidth	Finding
Small, Bacon, and Fozard (1959)	Adjustment	Gated	128, 1040, 4080, 8030	Weber's Law
Pollack (1951)	Method of Limits	Gated	Wideband	Weber's Law
Viemeister (1974a)	UDTR (70.7%)	Gated	Wideband	Weber's Law
Viemeister (1974b)	?	Gated	Wideband	Weber's Law
Moore (1975)	UDTR (73.6%)	Gated	232,729, 926	Weber's Law
Moore and Raab (1975)	UDTR (73.6%)	Gated	Wideband	Weber's Law
Schacknow and Raab (1976)	UDTR (73.6%)	Gated	10000	Weber's Law
Viemeister (1974a)	UDTR (70.7%)	Gated	1350	Near Miss
Schacknow and Raab (1976)	UDTR (73.6%)	Gated	100, 316, 1000, 3160	Near Miss

gated and continuous masker paradigms. For narrowband noise ($W \leq 1000$ Hz), a near miss to Weber's Law obtains with both gated and continuous maskers -- detectability improves slightly as the masker level is increased. Exceptions to this generalization include the near miss obtained by Green and Sewall (1962) for a bandwidth of 4000 Hz, and Moore's (1975) report of Weber's Law for narrowband noise.

2. Intensity-Duration Reciprocity

Increasing the sample size from which the observer makes his decisions, by lengthening the signal duration, results in a decrease in the relative variability of the sample (see Equation 1-7). Accordingly, the ideal observer's performance varies with signal duration so that a decade increase in T will result in a 5-dB improvement in the Weber fraction.

The psychoacoustic literature is replete with studies investigating the course of temporal integration for homogeneous noise stimuli. Those employing continuous maskers will be reviewed first.

Green (1960), in a test of his model, used a 2IFC task in which observers were required to detect increments in a continuous bandlimited noise ($W = 3862$ Hz). A number of trials were presented at each of 5 or 6 signal levels, and 75% points were interpolated from the resulting psychometric functions. For increments of 3 to 300 msec, the Weber fraction decreased at a rate of approximately 6 dB/decade increase in T. For signals longer than 300 msec, performance was independent of duration. Campbell (1963) replicated Green's experiment using an adaptive procedure (BUDTIF) to track a signal level yielding 75% correct detections. The detectability of increments in a broadband noise (W = 5100 Hz) improved at a rate of 5.5 dB/log-unit increase in duration, over a range of 1 to 100 msec. Raab, Osman, and Rich (1963) reported a duration reciprocity factor of -7.00 dB for increments in a broadband noise (W = 6900 Hz), ranging from 5 to 100 msec. Rochester (1971), using PEST, obtained a duration reciprocity factor of -8.00 dB for increments in wideband noise. Additional data for wideband noise-intensity discrimination (W = 5000 Hz) is provided by Macmillan (1973, 1974). In the former experiment, Macmillan employed signal durations ranging from 50 to 300 msec; in the latter, signal durations varied from 20 to 640 msec. Duration reciprocity factors were calculated individually for each subject, including only durations where intensity traded for duration, and were then averaged across subjects. Intensity traded with duration at a rate of -4.10 dB and -6.69 dB/decade increase in T for the 1973 and 1974 studies, respectively.

Two studies were performed using gated masker conditions exclusively. Moore and Raab (1975) reported an improvement in the Weber fraction of only 3.8 dB when the signal duration was increased from 10 to 250 msec. For a log-unit increase in T, the reciprocity factor is -2.70 dB. Raab and Goldberg (1975) obtained Weber fractions for both narrowband and wideband noise at each of two durations (T = 10 and 100 msec). For the narrowband stimuli, (W = 500 Hz), the reciprocity factor equalled -3.5 dB; for the wideband stimuli (W = 5000 Hz), a reciprocity factor of -3.7 dB obtained.

It is evident from the studies reviewed above that detectability improves more rapidly with increasing duration when signals are masked continuously than when the masker is gated simultaneously with the signal. A comparison of the effects of the mode of masker would be more appropriate if it were not made across studies, and therefore subjects. Three studies have been performed in which IT trades were obtained for both gated and continuous masker conditions in a within-subject design.

Wightman and Green (1966) employed a 2IFC up-down technique to determine the Weber fraction for wideband signals (W = 6000 Hz). For signal durations of 0.3 to 300 msec, intensity-duration reciprocity factors were -3.6 and -6.3 dB for gated and continuous conditions, respectively. Viemeister (1974b) reported similar IT trades for wideband noise signals which ranged in duration from 0.8 to 1600 msec. For signal durations less than 200 msec, intensity-duration reciprocity factors were -3.6 and -6.7 dB for gated and continuous conditions, respectively. Berner, in an unpublished study, obtained Weber fractions for wideband signals of 10, 20, and 100 msec (W = 5000 Hz). Performance improved at a rate of 2.95 and 6.35 dB/decade increase in duration for gated and continuous masker conditions, respectively.

In summary, continuously masked signals become more detectable with increasing duration than is predicted by Green's model: averaged across studies, a log-unit increase in duration results in a 6.29-dB improvement in the Weber fraction (see Table 1-5). For gated maskers, detectability grows less rapidly: an improvement of only 3.29 dB results from a decade increase in signal duration when reciprocity factors are averaged across studies.

Table 1-5

Studies Investigating the Intensity-Duration Reciprocity for Homogeneous Masker Conditions

Study	Method	Mode of Masker	Reciprocity Factor (in dB)
Green (1960)	Fixed I	Cont.	-6.00
Campbell (1963)	BUDTIF (75%)	Cont.	-5.50
Raab, Osman and Rich (1963)	UDTR (73.6%)	Cont.	-7.00
Rochester (1971)	PEST	Cont.	-8.00
Macmillan (1973)	UDTR (73.6%)	Cont.	-4.10
Macmillan (1974)	UDTR (73.6%)	Cont.	-6.69
Wightman and Green (1966)	"up-down"	Cont.	-6.30
Viemeister (1974b)	?	Cont.	-6.70
Berner (unpublished)	UDTR (73.6%)	Cont.	-6.35
Moore and Raab (1975)	UDTR (73.6%)	Gated	-2.70
Raab and Goldberg (1975)	UDTR (73.6%)	Gated	-3.60
Wightman and Green (1966)	"up-down"	Gated	-3.60
Viemeister (1974b)	?	Gated	-3.60
Berner (unpublished)	UDTR (73.6%)	Gated	-2.95

3. Intensity-Bandwidth Reciprocity

The increase in sample size and resulting decrease in relative variability discussed in the previous section may also be achieved by increasing the bandwidth of the signal. A 5-dB improvement in the Weber fraction is, therefore, predicted for a tenfold increase in signal bandwidth.

Bos and deBoer (1966) investigated the effects of signal bandwidth on intensity discrimination for 250-msec increments in a continuous masker. Signal bandwidths ranged from 5 to 10000 Hz, each centered at 500, 1000, 2000, 4000, and 8000 Hz. Signal and masker waveforms were identical in bandwidth, and had a uniform spectral-power density. A secondary masker was also employed, with an intensity adjusted to be 20 dB below that of the primary masker, as measured in a 1/3-octave band situated around the center frequency of the signal passband. This noise floor was filtered to have uniform masking properties; i.e., above 500 Hz, masker components were attenuated at a rate of 3 dB per octave. Difference limens, determined with a Békésy audiometer, indicate two interesting results. First, detectability is independent of center frequency. Second, for homogeneously filtered noise stimuli, the effect of increasing bandwidth is far less than predicted by Green's model. The intensity-bandwidth reciprocity factor may be calculated for the data presented in Bos and deBoer's figures 6 and 7 (pp. 712, 713). Averaged across subjects and center frequencies, the plot of 10 log S/N as a function of log W appears to be reasonably linear for bandwidths ranging from 300 to 10000 Hz. Over this range, the intensity-bandwidth reciprocity factor is calculated to be -2.29 dB.

This small effect of bandwidth has been replicated by Schacknow and Raab (1976). Using an adaptive procedure (UDTR(73.6%)) to track the detectability of 100-msec increments in a continuous masker, a decade increase in bandwidth afforded only a 1.85-dB improvement in the Weber fraction. Signal bandwidths ranged from 100 to 10000 Hz, in half log-unit steps.

Shallow IW trades (plotted as 10 log ΔI/I as a function of Ws) have also been reported when the signal and masker are gated simultaneously.

Small, Bacon, and Fozard (1959) employed a method of adjustment in which an incremented noise burst was alternated with a standard intensity burst. Each burst lasted 1 second. The observer adjusted the size of the increment until the noise appeared to be continuous. Four bandwidths were employed (W = 128, 1040, 4080, and 8030 Hz), each at a number of sensation levels. The intensity-bandwidth reciprocity factor has been calculated for Small's data at a sensation level of 20 dB; performance improved with signal bandwidth at a rate of 2.88 dB per decade.

DeBoer (1965) generated psychometric functions for 400-msec increments in noise bursts filtered homogeneously in both time and bandwidth. The noise stimuli were described by deBoer as "uniform masking noise" — i.e., spectral components were attenuated at a rate of 3 dB per octave ("pink noise"). Stimulus bandwidths were 63, 250, 1000, and 4000 Hz, centered at 1000 Hz. All four bandwidths were used with observer GR, whereas all but the widest were presented to observer HV. Reciprocity factors were calculated from deBoer's figures 1 and 2 (p. 4) by estimating the signal level required to obtain d' = 1. Averaged across subjects, the reciprocity factor was -3.97 dB.

Moore (1975), employing UDTR(73.6%), collected data for 300-msec increments in 1/3-octave bands of noise centered at 1000, 4000, and 6300 Hz. For a masker level of 40 dB SL, the resulting intensity-bandwidth reciprocity factor for Moore's observers averaged -2.55 dB.

The smallest effect of signal bandwidth may be found in the data of Raab and Goldberg (1975) for gated masker conditions. Averaged across subjects and masker levels, IW trades are -0.93 and -1.13 dB/decade increase in W, for 10 and 100-msec signals, respecitvely.

More recently, Schacknow and Raab (1976) provided IW trades for gated and continuous masker conditions in which the reciprocity factor equalled -1.35 dB for signal bandwidths which ranged from 100 to 10000 Hz.

A summary of the studies investigating the intensity-bandwidth reciprocity for homogeneous noise-intensity discrimination is presented in Table 1-6. Averaged across studies, the IW reciprocity factor equals -2 dB for both gated and continuous masker conditions: Stimulus bandwidth, therefore, has a much smaller effect on discrimination than is predicted by Green's model. The decrease in the relative variability of the noise sample due to a tenfold increase in signal bandwidth should result in a 5-dB improvement in the Weber fraction.

These results also differ from those reported for studies employing heterogeneous masker conditions, in which IW reciprocity

Table 1-6

Studies Investigating the Intensity-Bandwidth Reciprocity for Homogeneous Masker Conditions

Study	Method	Mode of Masker	Reciprocity Factor (in dB)
Bos and De Boer (1966)	Békésy Audiometry	Cont.	-2.29
Schacknow and Raab (1976)	UDTR (73.6%)	Cont.	-1.85
Schacknow and Raab (1976)	UDTR (73.6%)	Gated	-1.35
Small, Bacon and Fozard (1959)	Adjustment	Gated	-2.88
De Boer (1965)	Fixed I	Gated	-3.97
Moore (1975)	UDTR (73.6%)	Gated	-2.55
Raab and Goldberg (1975)	UDTR (73.6%)	Gated	-1.03

factors are as predicted (-5 dB). This discrepancy is particularly troublesome in that homogeneously filtered signal and masker waveforms setify Green's internal filtering assumption artificially and should consequently allow the human observer to more closely approximate the performance of the ideal receiver.

4. Slope of the Psychometric Function

Evidence has been presented indicating that the psychometric function for heterogeneous noise stimuli is too steep for continuously masked signals -- i.e., detectability grows more rapidly with signal intensity than is predicted by Green's model. So too is this the case for homogeneously filtered noise.

Green (1960) presented his subjects with increments in continuous wideband noise. Signal durations varying from 3 to 1000 msec were employed with white Gaussian noise whose passband equalled 3862 Hz. A wideband noise (W = 5100 Hz) was also employed; signal duration in this case was 250 msec. Psychometric functions were constructed, spanning a range of detectability from approximately 50% to 100% correct detections. For all conditions, psychometric functions were considerably steeper than predicted. An increase in S_o/N_o of only 3 dB was sufficient for performance to improve from 60% to 90% correct detections, as opposed to the 7 dB required of Green's optimal receiver.

Campbell (1965) obtained three-point psychometric functions for noise-intensity discrimination by adaptively tracking each point separately with BUDTIF. Signal levels required for nominally 62%, 75%, and 88% correct detections were obtained. For 1-second increments in a continuous, wideband noise (W = 4950 Hz), psychometric functions were similar in shape to those of Green (1960). A least squares fit was obtained for Campbell's data -- the psychometric function is predicted to be linear when S_O/N_O is plotted as a function of the normalized percent correct. (Recall that the statistic, M, is the normalized percent correct, and Equation 1-2 reduces to the form y = mx + b.) Extrapolated to performance levels of 60% and 90% correct, the Weber fraction increased by 2.78 dB for experienced listeners.

Green and Sewall (1962) provide psychometric functions collected under both continuous and gated masker conditions. Their psychophsyical procedure was the same as that employed by Green (1960). Blocks of fixed-intensity increments (T = 100 msec) were presented in a continuous wideband noise (W = 4000 Hz). Psychometric functions were constructed by repeating this procedure at a number of signal levels. An increase of only 3 dB was required for performance to improve from 60% to 90% correct detections. Using different subjects, Green and Sewall collected data for increments in a gated masker. The psychophysical procedure and stimulus parameters were the same as above. Increments were presented in a gated masker whose overall level was 75 dB SPL. In addition to the gated masker, a continuous noise floor of identical frequency spectrum was added at levels ranging from 55 to 90 dB SPL, including one condition in which the noise floor was absent. For all conditions, psychometric functions were indistinguishable from Green's theoretical ideal.

DeBoer (1965) replicated Green and Sewall's findings for gated maskers. Using the same psychophysical procedure, psychometric functions were constructed for 400-msec bursts of "pink noise". Four bandwidths were used: 63, 250, 1000, and 4000 Hz, all centered at 1000 Hz. In order for performance to improve from 60% to 90% correct detections, a 7-dB increase in S_O/N_O was required.

The literature for homogeneous noise-intensity discrimination is consistent in its findings that the psychometric function is as predicted when signal and masker waveforms are gated simultaneously, whereas for continuously masked signals, detectability grows more rapidly with signal intensity than is predicted by Green's model (see Table 1-7).

An explanation for the gated-continuous difference in the slope of the psychometric function has been offered by Green and Sewall (1962). Temporal uncertainty concerning signal onset and duration in continuous masker paradigms is positted to adversely affect performance more for weak signals than for easily detectable signals. Performance will be relatively worse at low levels of detectability, which will result in a psychometric function that is steeper than predicted by Green's model. Gating the masker simultaneously with the signal should provide the observer with more accurate timing information and therefore result in a psychometric function of the predicted slope. The literature reviewed above indicates that the slope of the psychometric function is as predicted when gated maskers are employed.

Further discussion of the uncertainty model and the slope of the psychometric function will be presented in Chapter III, where psychometric functions are reported for gated and continuous masker conditions obtained in a within-subjects design. It will be seen that

Table 1-7

Studies Investigating the Slope* of the Psychometric Function for Homogeneous Masker Conditions

Study	Method	Mode of Masker	Slope
Green (1960)	Fixed I	Cont.	3.00
Campbell (1965)	BUDTIF (62%, 75%, 88%)	Cont.	2.78
Green and Sewall (1962)	Fixed I	Cont.	3.00
Green and Sewall (1962)	Fixed I	Gated	7.00
De Boer (1965)	Fixed I	Gated	7.00

*The "slope" of the psychometric function will be given as the increase in signal power (in dB) necessary for performance to increase from 60% to 90% correct detections.

the uncertainty argument is insufficient to reconcile the gated-continuous difference in the slope of the psychometric function.

5. Effects of the Mode of Masker

The failure of gated and continuous maskers to yield equal detectabilities has been reviewed for a number of signal and masker waveforms. Four studies allow a direct comparison of performance with gated and continuous maskers for homogeneously filtered noise.

Schacknow and Raab (1976) collected data for 100-msec increments in both gated and continuous maskers. An average gated-continuous difference in performance of 1.08 dB obtained for signal bandwidths ranging from 100 to 10000 Hz: continuously masked signals were more easily detected.

A number of studies indicate that the magnitude as well as the direction of the gated-continuous difference is dependent on signal duration. As indicated in section D-2, intensity-duration trades differ in magnitude for gated and continuous masker paradigms. If plotted on a single set of coordinates, the intensity-duration functions must somewhere intersect. For durations longer than that at the point of intersection, continuously masked signals are easier to detect; for shorter durations, the reverse is true.

Wightman and Green (1966) employed a 2IFC, up-down procedure to determine the detectability of increments in wideband noise (W = 6000 Hz). Signal duration varied from 0.3 to 300 msec in approximately half log-unit steps. The size of the gated-continuous difference may be seen in Table 1-8, where Weber fractions obtained with continuous maskers are subtracted from those obtained with gated maskers for each signal duration. A positive difference indicates performance that is better with continuously masked signals. Equivalent performance for gated and continuous conditions appears to obtain for a signal duration between 3 and 10 msec. For shorter duration signals, a continuous background provides more effective masking; at longer durations the reverse is true.

Berner, in an unpublished study, replicated the findings of Wightman and Green. Using UDTR(73.6%) with wideband noise (W = 5000 Hz), Berner obtained thresholds for increments of 10, 20, and 100 msec. An orderly effect of duration on gated-continuous differences may be seen in Table 1-8. The gated-continuous difference grows from +0.85 to +4.20 dB as signal duration increases from 10 to 100 msec. These results suggest that had Berner employed even shorter duration signals, the gated-continuous difference would have changed in direction, affording an advantage to signals embedded in gated maskers. The point of equality for gated and continuous maskers may be extrapolated to be slightly less than 1 msec.

Viemeister (1974b) employed both gated and continuous wideband maskers; signal duration was varied from 0.8 to 1600 msec. The following conclusions were made for data presented in abstract form: (1) Performance was better with gated than continuous maskers for durations shorter than 70 msec; this relationship reversed for longer duration signals. (2) Performance with both maskers was independent of duration for durations longer than 200 msec. (3) Intensity traded with duration at a rate of -3.6 dB/decade increase in T with gated maskers,

Table 1-8

Gated-Continuous Differences in Performance (a)

Study	W (Hz)	T(msec)	Gated-Continuous Difference
Schacknow and Raab (1976)	(b)	100.	+1.08
Wightman and Green (1966)	6000	0.3	-3.50
		1.	-2.00
		3.	-2.00
		10.	+2.00
		30.	+3.00
		100.	+3.00
		300.	+4.00
Berner (unpublished)	5000	10.	+0.85
		20.	+1.65
		100.	+4.20
Viemeister (1974b)	Wideband	0.8	-6.02
		7.	-3.10
		70.	0.00
		≥ 200.	+1.40

(a) All values are Weber fractions obtained with gated maskers minus those obtained with continuous maskers (in dB). A positive difference indicates better performance with a continuous masker.

(b) Gated-continuous differences have been averaged over bandwidths ranging from 100 to 10,000 Hz.

and -6.7 dB/decade with continuous maskers. From the above information, gated-continuous differences were reconstructed. An orderly effect of duration on the gated-continuous difference is indicated inasmuch as the intensity-duration functions differ in slope, but are both linear. For the shortest duration (0.8 msec), the gated-continuous difference was calculated to be -6 dB. For durations of 200 msec and longer, a gated-continuous difference of +1.4 dB obtains. Gated-continuous differences for some representative durations are presented in Table 1-8.

In summary, it has been demonstrated that long-duration signals are more easily detected if the masker is present continuously throughout an experimental session. The reverse is true for short-duration signals. The point at which this relationship reverses, however, appears to differ among studies.

6. Real+Ideal Differences in Performance

Differences between obtained and theoretically ideal Weber fractions were reviewed for heterogeneous noise stimuli in section C-6. With continuous maskers, observers are 5-dB worse than the ideal receiver. With gated maskers, the real-ideal difference is 11 dB. The human observer is simply not as efficient a signal processor as Green's optimal receiver.

Raw data have been reported for a large number of studies employing homogeneous noise. These Weber fractions will be compared to those predicted by Equation 1-2 of Green's model. In studies where the signal duration exceeds the limit of temporal integration, ideal Weber fractions will be calculated for a duration of 300 msec, the value exhibited by Green's (1960) listeners for this upper limit. Since all studies to be included in this analysis have been fully reviewed in previous sections, real-ideal differences in performance will simply be presented in tabular form. Table 1-9 has been divided into two sections — one presents data for wideband noise (W $\geq$1000 Hz), the other for narrowband noise (W $<$1000 Hz). The rationale for such a division comes from the shallow IW reciprocities reported for homogeneous noise. Intensity-bandwidth functions for real and ideal observers are not parallel. As a result, real-ideal differences in performance will increase with bandwidth. This is reflected in the data of Moore (1975), Raab and Goldberg (1975), and Schacknow and Raab (1976).

To summarize the data presented in Table 1-9, when the masker is continuous, results for homogeneous and heterogeneous wideband noise differ little — performance is degraded by approximately 6 dB. For wideband gated maskers, discrimination is relatively better with homogeneous noise stimuli — a real-ideal difference of approximately 7 dB obtains (as compared to the 11-dB difference obtained with heterogeneous noise). The few studies employing narrowband homogeneous noise yield performance that is closer to ideal than that obtained with wideband noise. Schacknow and Raab (1976), in particular, report Weber fractions that are close to, and in some cases slightly better than, those predicted for the ideal receiver.

Table 1-9

Difference Between Real and Ideal Performance (a)

Study	Mode of Masker	Bandwidth	Real-Ideal Difference
	W ≥1000 Hz		
Miller (1947	Cont.	6850	+6.56
Green (1960)	Cont.	3862	+6.05
Campbell (1963)	Cont.	5000	+7.50
Campbell (1964)	Cont.	5000	+5.88
Wightman and Green (1966)	Cont.	6000	+7.67 (b)
Macmillan (1973)	Cont.	5000	+5.25 (b)
Macmillan (1974)	Cont.	5000	+4.89 (b)
Schacknow and Raab (1976)	Cont.	1000	+3.40
		3160	+5.35
		10000	+5.68
Berner (unpublished)	Cont.	5000	+6.12 (b)
Wightman and Green (1966)	Gated	6000	+7.10 (b)
Viemeister (1974a)	Gated	1350	+7.00
Moore (1975)	Gated	1459	+7.09
Raab and Goldberg (1975)	Gated	5000	+7.14 (b)
Schacknow and Raab (1976)	Gated	1000	+4.12
		3160	+5.68
		10000	+7.93
Berner (unpublished)	Gated	5000	+8.35 (b)

(Continued)

Table 1-9, Continued

Study	Mode of Masker	Bandwidth	Real-Ideal Difference
	W <1000 Hz		
Schacknow and Raab (1976)	Cont.	100 316	-1.48 +1.42
Moore (1975)	Gated	232 926	+5.19 +6.82
Raab and Goldberg (1975)	Gated	500	+3.14 (b)
Schacknow and Raab (1976)	Gated	100 316	-0.22 +2.08

(a) All values are real minus ideal Weber fractions (in dB). A positive difference indicates performance that is worse than ideal.

(b) Real-ideal differences are averaged across signal duration.

7. Summary of Results Obtained With Homogeneous Noise

By externally filtering both signal and masker waveforms to have identical frequency spectra, one assumption of Green's model is satisfied by the experimenter -- the observer need not employ an internal filter matched to the signal bandwidth. If in addition, the masker is gated simulaneously with the signal, still another assumption is satisfied - the observer need not filter in the time domain. It is therefore anticipated that homogeneously filtered noise stimuli should provide data that more closely satisfy the predictions of Green's model than those which have been reviewed for heterogeneous noise. Futhermore, gated homogeneous maskers should provide data that fit Green's model best of all, inasmuch as both assumptions have been satisfied by the experimenter.

The literature for homogeneous noise-intensity discrimination employing continuous maskers will be summarized first:

(1) Support for Weber's Law appears limited to wideband stimuli: Five studies employing wideband noise obtained Weber's Law (Miller, 1947; Harris, 1950; Raab, Osman, and Rich, 1963; Viemeister, 1974b; and Schacknow and Raab, 1976, for W = 3160 and 10000 Hz). A near miss to Weber's Law was obtained for narrowband noise (Bos and de Boer, 1966; and Schacknow and Raab, 1976) and in one instance for wideband noise (Green and Sewall, 1962). Campbell (1964) obatined a slight decrement in performance at a masker level of 65 dB SL with wideband noise.

(2) Several studies report that intensity trades with duration at a rate (averaged across studies) of -6.3 dB/decade increase in T.

(3) The intensity-bandwidth reciprocity factor is considerably less than predicted by Green's model. Averaged across studies, the IW reciprocity factor is -2.1 dB/decade increase in W.

(4) Psychometric functions for continuously masked signals are too steep — an increase of only 3 dB is necessary for performance to improve from 60% to 90% correct detections.

(5) For wideband noise stimuli, obtained Weber fractions average 6.1-dB worse than predicted on the basis of stimulus variability alone; for narrowband noise, no appreciable difference between real and ideal performance has been reported.

Data obtained with gated maskers present a somewhat different picture of the detection process:

(1) Except for two studies, support for Weber's Law is widespread. A near miss was obtained by Viemeister (1974a) for narrowband noise (W = 1350 Hz), and by Schacknow and Raab (1976) for bandwidths of 100, 316, 1000, and 3160 Hz.

(2) Intensity-duration trades are more shallow than predicted — a tenfold increase in T will effect only a 3.3-dB improvement in the Weber fraction.

(3) Intensity-bandwidth trades are similarly shallow — the average reciprocity factor equals -2.4 dB.

(4) Psychometric functions, however, are as predicted — a 7-dB increase in S_o/N_o is required for performance to improve from 60% to 90% correct detections.

(5) Obtained Weber fractions average 7.1-dB worse than ideal for wideband noise; for narrowband stimuli, an average real-ideal difference of 3.4 dB obtains.

Satisfying the temporal filtering assumption of Green's model by gating the masker simultaneously with the signal allows for a closer approximation to some of the predictions of Green's model — Weber's Law is more consistently obtained, and the psychometric function is of the proper slope. In other ways, however, gated maskers lead to less efficient processing of the signal — intensity-duration trades are considerably shallower than those obtained with continuous maskers, and for long duration signals the Weber fraction is larger (when evaluated within subjects, this gated-continuous difference is approximately 2.5 dB). In neither gated nor continuous masker conditions, however, do proper IW reciprocity factors obtain. Only when the signal and masker bandwidths are heterogeneous and the observer must provide his own filtering mechanism, are IW trades as predicted.

E. Plan of Research

Real listeners depart from Green's ideal mechanism in a number of ways. These differences include: IW trades are too shallow; IT trades are too shallow with gated maskers, and too steep with continuous maskers; the slope of the psychometric function is too steep when maskers are continuous; Weber fractions are dependent on the mode of masker.

Examination of Tables 1-2 through 1-9 reveals that the literature is remarkably consistent regardless of the method used to collect the data. Common to all these methods, is the practice of providing clearly audible signals in the early stages of testing —

adaptive procedures start "high", fixed methods provide suprathreshold practice trials. It is possible, that with such methods the observer adopts a nonoptimal detection strategy that is based on some stimulus characteristic which is salient at high signal levels. If this is the case, altering the psychophysical procedure in such a way as to deprive the observer of potentially "useful" information about the signal at the beginning of the experimental session might result in the choice of alternate detection strategies.

It is therefore proposed, that noise-intensity discrimination be investigated with an adaptive procedure which is initiated from a level where the signal is clearly below threshold. In addition, no suprathreshold practice trials are to be given. This condition will henceforth be referred to as an "ascending" run, as opposed to the usual practice of initiating a staircase run with a clearly audible signal (a "descending" run).

This dissertation was consequently designed to test the applicability of Green's model of noise-masked-by-noise to the human detection process, by employing both ascending and descending versions of UDTR. The predictions tested include: (1) the intensity-duration reciprocity, (2) the intensity-bandwidth reciprocity, (3) the slope of the psychometric function, (4) the equivalence of gated and continuous maskers, and (5) the equivalence of ascending and descending runs.

The assumption that observers filter the masker waveform to match that of the signal was satisfied externally by employing homogeneously filtered bands of noise in all conditions. Weber functions were not obtained inasmuch as the literature largely confirms Weber's Law for homogeneous bands of noise (see Table 1-4).

It will be seen in the following chapters that the performance of real listeners closely approximates that of Green's ideal receiver for ascending conditions, whereas the differences between real and ideal observers noted in the literature are replicated with descending staircase runs.

Chapter II

METHOD

Three experiments were conducted in an attempt to discover whether listening strategies depend on the information made available to the observer. Two experiments investigated the effects of duration and bandwidth on noise-intensity discrimination. A third experiment examined the slope of the psychometric function. In all experiments, difference thresholds were determined by starting with the signal either clearly audible ("descending runs") or well below detectability ("ascending runs"); the noise masker was either presented continuously throughout an experimental session or gated with each observation interval.

A. Subjects

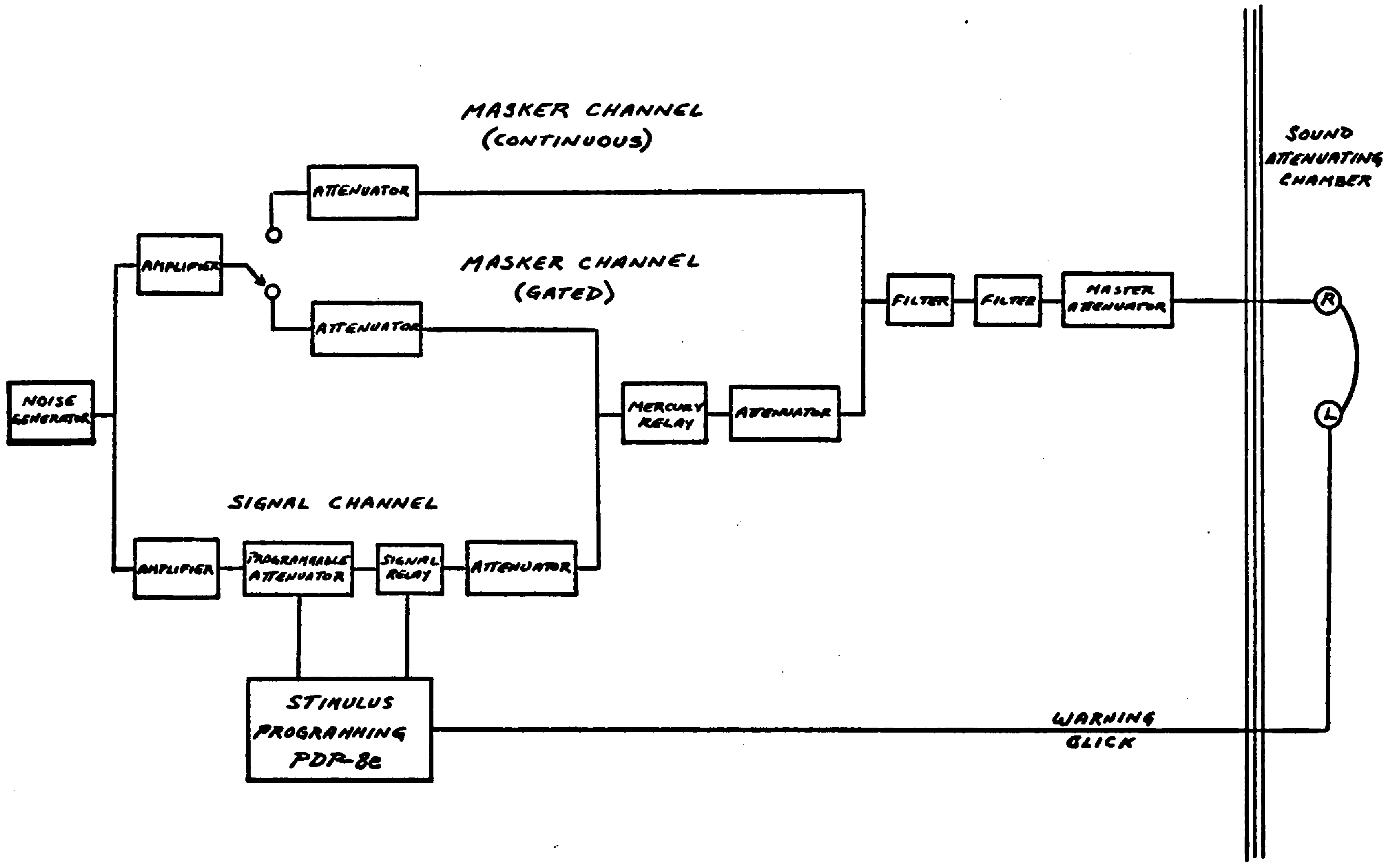
Two male graduate students (RB and the author, CM) served as subjects. Both were experienced listeners and received considerable training before any data were collected.

B. Apparatus

A block diagram for the apparatus used in Experiment 1 is presented in Figure 2-1. Experiments 2 and 3 required minor modifications which are described later.

Figure 2-1

Block diagram of the apparatus used in Experiment 1.


MASKER CHANNEL
(CONTINUOUS)
MASKER CHANNEL
(GATED)
SIGNAL CHANNEL
NOISE GENERATOR
AMPLIFIER
ATTENUATOR
ATTENUATOR
AMPLIFIER
PROGRAMMABLE ATTENUATOR
SIGNAL RELAY
ATTENUATOR
MERCURY RELAY
ATTENUATOR
FILTER
FILTER
MASTER ATTENUATOR
STIMULUS
PROGRAMMING
PDP-8e
SOUND
ATTENUATING
CHAMBER
R
L
WARNING
CLICK

In all conditions, the signal and masker bandwidths were identical. The output of a General Radio (1390-B) noise generator was divided and led to two channels — one for the increment signal (ΔV), the other for the masker (V). Each channel contained an amplifier (Calex 175) and attenuators for level adjustment. The magnitude of ΔV was changed between blocks of trials by a Grason-Stadler (1284) programmable attenuator. The activation of a "signal" relay placed ΔV in one of two observation intervals with a probability of 0.5, which was determined by a Coulbourn (S35-20) probability generator.

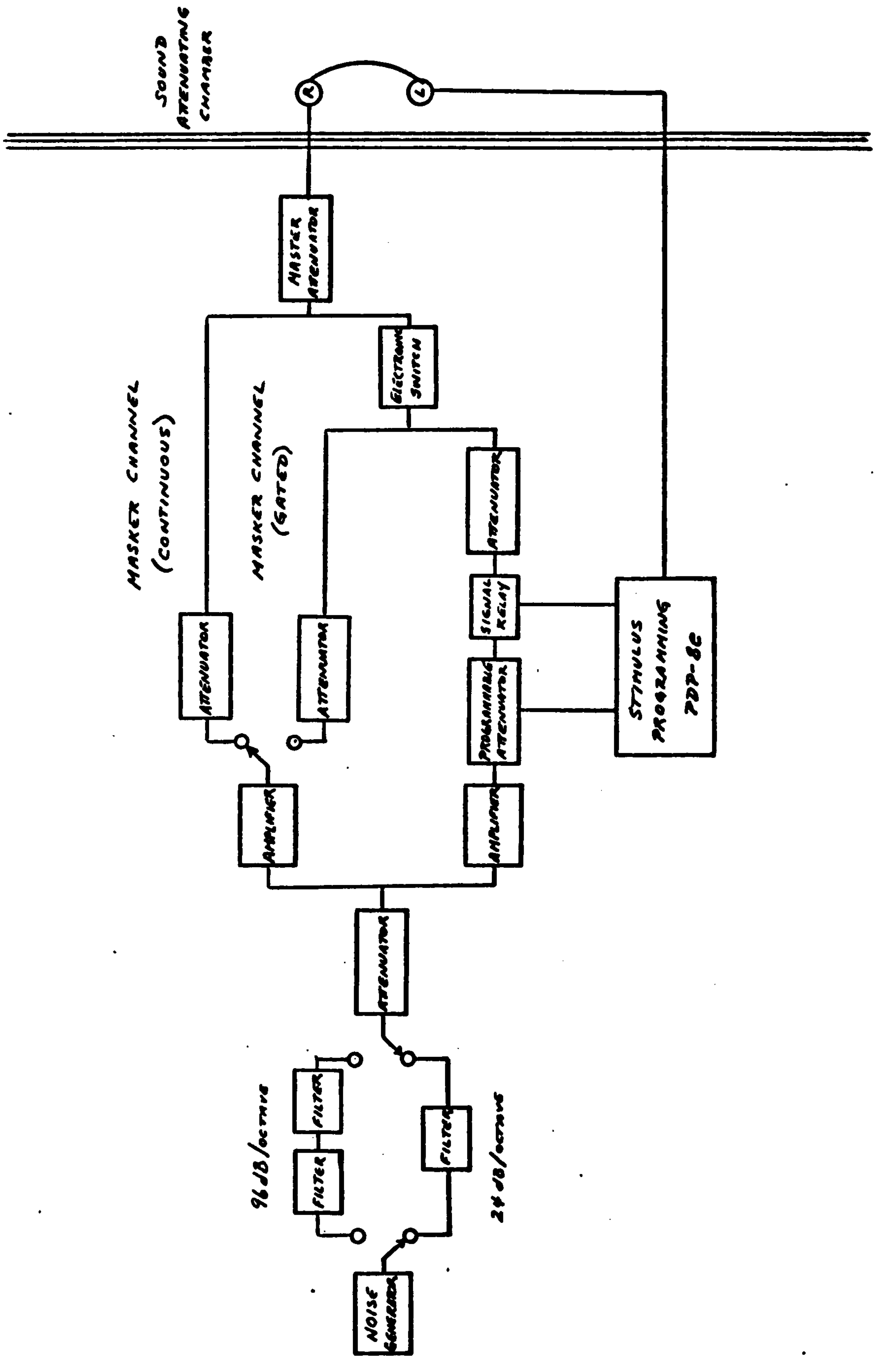
For the gated masker condition, the signal and masker were added coherently by means of a resistive mixing network and then gated by means of a mercury-wetted relay. When the masker was continuous, the increment signal alone was gated. The two waveforms were then mixed and the sum was bandlimited by two Rockland (1042-F) electronic filters connected in cascade. The resulting spectrum skirts had an attenuation rate of 96 dB/octave. The noise was then passed through a master attenuator, patched into a sound-proof room (IAC 1200-ACT), and finally delivered to the subject's right earphone (Telephonics, TDH-39), mounted in an MX/41-AR cushion.

The apparatus for Experiment 2 differed from the above only in that a single Rockland (1042-F) filter was used to provide 48 dB/octave skirts.

A block diagram for the apparatus used in Experiment 3 is presented in Figure 2-2. The stimulus waveforms were first filtered and then divided into signal and masker channels. Either a single Rockland (432) filter was used to provide spectrum skirts with an attenuation rate of 24 dB/octave, or two Rockland (1042-F) filters were connected in cascade to provide 96 dB/octave skirts. After filtering, a level-setting attenuator was adjusted to compensate for the difference in gain between the two filters. A Grason-Stadler (829-C) electronic switch, set to give a

Figure 2-2

Block diagram of the apparatus used in Experiment 3.


SOUND ATTENUATING CHAMBER
R
L
MASTER ATTENUATOR
ELECTRONIC SWITCH
MASKER CHANNEL (CONTINUOUS)
MASKER CHANNEL (GATED)
ATTENUATOR
ATTENUATOR
ATTENUATOR
SIGNAL RELAY
PROGRAMMABLE ATTENUATOR
STIMULUS PROGRAMMING PDP-8E
AMPLIFIER
AMPLIFIER
ATTENUATOR
96 dB/OCTAVE
FILTER
FILTER
FILTER
24 dB/OCTAVE
NOISE GENERATOR

rise/fall time of 10 msec, replaced the mercury-wetted relay. Finally, the resulting waveform was delivered to a Sennheiser HD-414 earphone, which provided a wider frequency response than the TDH-39 phones used in Experiments 1 and 2.

Stimulus sequencing and data acquisition were controlled on-line by a DEC PDP-8e computer. At the end of an 80-trial run, the computer provided a frequency histogram of signal levels visited.

C. Procedure

1. Psychophysical Method

Intensity-discrimination thresholds were determined using a two-interval, forced-choice, Up-Down Transformed-Response method (cf. Levitt, 1971). A faint warning click, presented to the subject's left earphone, marked the beginning of each trial. Two observation intervals, whose onsets were spaced 800 msec apart, were presented in succession one second after the warning click. The increment signal (ΔV) was placed with equal probability in one of the two intervals. The subject indicated the interval containing the signal by pressing one of two microswitch buttons. Indicator lights on his response panel marked the observation intervals and provided immediate feedback. Trials were self-paced, which yielded one trial approximately every five seconds.

Each threshold determination began with the signal either clearly audible, or well below detectability. When tracking 73.6% correct detections, the following decision rule was employed: five correct responses before two errors at a given signal level resulted in a 3-dB attenuation of the increment signal (ΔV). Otherwise, ΔV was increased.

This "up-down" procedure allows the observer to cross and recross his threshold an average of 15 times in the course of an 80-trial run. Upon completion of a run, the median value of attenuation in the signal channel, ΔV, was computed and then converted to increment power, ΔI, taking into account the in-phase addition of the signal and masker waveforms. (A 3-dB change in ΔV, therefore led to a 1.5-dB change in ΔI over most of the range tracked by the staircase.) Only levels that were revisited were included in the analysis. The ratio of ΔI to I, expressed in decibels, is used as an index of intensity discrimination. Five such determinations were made for each experimental condition.

2. Experimental Design

Experiment 1 investigated the effects of duration on wideband noise-intensity discrimination. The signal and masker had identical passbands; the half-power bandwidth was 5000 Hz, centered at 1500 Hz. Signal durations were 10 or 100 msec. The masker, presented in either the gated or continuous mode, had an overall level of 45 or 75 dB SPL. The signal increment yielding 73.6% correct detections was tracked by descending runs beginning with $\Delta I/I = +4.77$ dB. Ascending runs began with $\Delta I/I = -19.48$ dB for 100-msec signals, and -14.96 dB for 10-msec signals.

Experiment 2 examined the slope of the psychometric function as affected by gated and continuous maskers. The wideband noise employed in this study had a half-power bandwidth of 6040 Hz, centered at 775 Hz. Signal duration was 100 msec. The noise masker, either gated or continuous, had an overall level of 45 dB SPL. Three-point psychometric functions were obtained for descending runs by tracking the signal increment yielding 57.9%, 73.6%, and 89.1% correct detections. These runs began with $\Delta I/I$ = +4.77 dB for gated, and -0.78 dB for continuous conditions.

Ascending runs tracked 63.6% and 73.6% correct detections, starting with $\Delta I/I$ = -19.48 dB. Decision rules which track these detectabilities are given in Table 2-1.

Experiment 3 addressed the question of how the observer makes use of bandwidth information in ascending and descending runs. The noise was bandlimited to produce a half-power bandwidth of 469 Hz, centered at 1060 Hz, with attenuation rates of 24 or 96 dB/octave. In each case, the spectrum level at the center of the passband was 46.4 dB SPL (measured by a General Radio Wave Analyzer Type 1568-A). As in Experiment 2, signal duration was 100 msec, and the masker was presented either continuously or gated with the signal. Starting levels were +4.77 dB for descending runs, and -14.96 dB for ascending runs.

In all experiments, ascending and descending runs were presented in blocks in order to minimize such transfer effects as might result from differences in listening strategies. The order of ascending and descending runs was counterbalanced across experiments, while other stimulus parameters were randomized.

Table 2-1

Decision Rules for Tracking
Various Levels of Detectability*

Percent Correct	Number of responses to effect a level change	
	Correct	Incorrect
57.9	4	3
63.6	5	3
73.6	5	2
89.1	6	1

*See text for details.

Chapter III

RESULTS AND DISCUSSION

The general pattern of results will be presented first. Detailed data will be found in Appendix A, Tables A-1 through A-6.

Results obtained with ascending and descending runs differ substantially. Data collected with descending staircase runs indicate that the human detection process is not independent of the mode of masker. Furthermore, for neither masking paradigm can Green's ideal receiver model completely describe the performance of the human observer. Ascending runs, on the other hand, yield data that are in accord with the predictions of Green's model for both gated and continuous masker paradigms.

In addition, examination of Tables A-1 through A-6 reveals a profound effect of starting level on detectability in each of the experiments. For all conditions, performance is approximately 5-dB better for ascending runs. It is possible that this effect is purely statistical. The random-walk nature of the staircase procedure is such that the signal level tracked may not be properly sampled in a finite run of 80 trials. If this is so, Weber fractions might be spuriously large for descending runs and small for ascending runs.

Analysis of all other parameter manipulations shall, therefore, be postponed until after the magnitude of this statistical bias is assessed. This was done for an ideal observer, by varying the initial stimulus level in computer-simulated staircase runs.

A. Monte Carlo Computer Simulations

In adaptive procedures, such as BUDTIF, PEST, and the Up-Down Transformed Response method (UDTR), a target percent correct is tracked by varying the signal level in accordance with a running account of the subject's performance. The decision rules are chosen such that the staircase will eventually converge on the signal level required to achieve the target percent correct. It has been traditional in psychoacoustics to choose an initial level where the signal is easily detectable (see for example, Stuckey, Hutton, and Campbell, 1966; Taylor and Creelman, 1967; Pollack, 1968; and Raab and Goldberg, 1975). In this manner, the subject is allowed to "familiarize" himself with the signal to be detected. For adaptive procedures utilizing a fixed step size (such as BUDTIF and UDTR), when the initial signal level is high relative to threshold, a number of trials are wasted in the descent, diminishing the usable portion of the run. Ascending runs are subject to the same pitfall; a number of trials are wasted in the ascent to the threshold signal level.

It is conceivable that the large ascending-descending differences observed in our data may be attributed wholly to such statistical considerations. That is, thresholds might be spuriously high for descending runs and low for ascending runs. A computerized Monte Carlo procedure was therefore utilized in an effort to quantify this statistical bias potentially inherent in UDTR.

Of prime importance in this simulation are all parameters which affect the degree of convergence on the target percent correct. Such factors include: (1) the initial stimulus level, (2) the slope of the presumed psychometric function, (3) the target percent correct tracked, (4) the block size used to make a decision concerning level change, (5) the step size between levels, and (6) the number of trials in a run. Since the computer simulations were performed to provide an *a posteriori* analysis of the methods used to obtain the data reported herein, manipulation of the above factors was limited to the values employed in the data collection (except for initial starting level which was thoroughly investigated).

1. Description of the Computer Program

A computer program, written in Fortran IV for an IBM-370/168 computer, provided a Monte Carlo solution to UDTR threshold determination. The method of collecting and analyzing the simulated data was identical to that employed with real listeners. By keeping the listener invariant, the simulated staircase runs could assess the exact magnitude of the statistical bias arising from varying the starting level.

The computer program was provided with a psychometric function for our invariant listener, consisting of a set of probabilities of a correct response for each level tracked by the staircase. A representative psychometric function (WT = 500) is presented in Table 3-1. The listener was ideal in two senses. First, the psychometric function was determined by evaluating Green's equation for M — the performance of an optimal energy detector (see Introduction, Equation 1-2). The normalized mean of the approximately Gaussian difference distribution (derived from the energy distributions for noise

Table 3-1

Probability of a Correct Response as a Function of Signal Level for WT = 500

10 log $\Delta V/V$	10 log $\Delta I/I$	P(C)
0	+4.77	1.000000
-3	+2.83	1.000000
-6	+0.98	1.000000
-9	-0.78	1.000000
-12	-2.48	1.000000
-15	-4.12	1.000000
-18	-5.72	0.999895
-21	-7.30	0.996358
-24	-8.85	0.973122
-27	-10.39	0.916161
-30	-11.92	0.837369
-33	-13.44	0.757990
-36	-14.96	0.690460
-39	-16.47	0.637881
-42	-17.97	0.598784
-45	-19.48	0.570374

and signal plus noise), was then converted to percent correct accurate to six decimal places (National Bureau of Standards, 1953). Second, since the sole determinant of a response is this psychometric function, the simulated observer exhibits perfect attention and motivation, and responds free from any biases or sequential dependencies.

The program generates runs for an ideal observer in the following manner:

1. An initial stimulus value is chosen, equal to that employed in the collection of real-listener data for a descending run (e.g., $\Delta I = +4.77$ dB when WT = 500).
2. A seven digit random number (in the range 0 to 1.0 inclusive) is generated from a rectangular distribution provided by the IMSL (International Mathematical and Statistical Library) subroutine, GGUB, which employs a multiplicative congruential method of generation (see Appendix B). This random number is then compared to the percent correct for the level being tested. If the random number is less than or equal to the percent correct, a correct response is stored in memory. Otherwise, an incorrect response is stored.
3. Step 2 is repeated until a criterion is reached for a level change. When tracking 73.6% correct detections, for example, the occurrence of 5 correct responses before 2 incorrect responses results in a 3-dB attenuation of ΔV, otherwise ΔV is incremented by 3 dB.

4. Upon changing levels, a count is entered into memory indicating that the staircase run had visited that level. (As was true in the collection of real-listener data, the staircase is allowed to span 16 stimulus levels. If the decision rule calls for a change in level that would exceed this limit, the run is aborted and excluded from the data analysis.)
5. Steps 2-4 are repeated until 80 statistical trials are completed. The median of levels revisited is then computed.
6. This cycle is repeated until 1000 runs are obtained for a given starting level. The mean of the medians is then computed.
7. The initial stimulus value (ΔV) is decreased by 3 dB and 1000 runs are obtained for this new starting level. This procedure is repeated until 1000 staircase runs are simulated for each of the 16 levels covered by the staircase.
8. Computer simulated runs were obtained for each of the stimulus conditions employed in this dissertation by repeating this procedure with the appropriate psychometric function for each condition.

2. Results of Computer Simulations

The effects of starting level on the signal level tracked by UDTR are depicted in Tables 3-2 through 3-4. Each Weber fraction is the mean of 1000 simulated runs. Included in each table is the performance of an ideal energy detector calculated from Equation 1-2. The effect of starting level on the estimated threshold can be assessed by comparing the simulated signal level tracked with ideal performance.

Table 3-2 presents the effects of starting level for each of the 16 starting levels used in the simulations for Experiment 1. The orderly effect of starting level on the estimated threshold is readily apparent. Inasmuch as the trends in the simulations were the same for the conditions of Experiments 2 and 3, the effects of only the starting levels used in real-listener data collection are included in Tables 3-3 and 3-4.

From Table 3-2, it may be seen that starting UDTR at the signal level it should ideally be tracking will result in no error of estimation. For example, when WT = 50, ideal performance is -8.71 dB. Beginning UDTR(73.6%) at -8.85 dB results in a simulated signal level of -8.93 dB. The error in tracking, therefore, will be minimal when UDTR is started at the level it is designed to track. This is true for all conditions simulated.

With <u>real</u> listeners, performance is less than optimal. Consequently, the signal level to be tracked by UDTR is initially unspecified. This uncertainty, combined with a desire to provide the observer with suprathreshold practice trials, can result in faulty tracking. The simulations presented in Table 3-2 indicate that starting a

Table 3-2

Simulated Performance (10 log $\Delta I/I$) for an "Ideal" Listener in Experiment 1 (a)

Starting Level	Signal Level Tracked (WT = 50)	Signal Level Tracked (WT = 500)
+4.77 (b),(c)	-7.71	-12.42
+2.83	-7.82	-12.76
+0.98	-7.84	-12.82
-0.78	-7.96	-13.00
-2.48	-8.08	-13.01
-4.12	-8.12	-13.15
-5.72	-8.26	-13.22
-7.30	-8.56	-13.25
-8.85	-8.93	-13.33
-10.39	-9.30	-13.42
-11.92	-9.74	-13.67
-13.44	-10.05	-13.90
-14.96 (b)	-10.45	-14.26
-16.47	-10.77	-14.65
-17.97	-11.03	-14.92
-19.48 (c)	-10.97	-15.08
Ideal	-8.71	-14.08

(a) UDTR tracks 73.6% correct detections.

(b) Starting levels employed in Experiment 1 when WT = 50.

(c) Starting levels employed in Experiment 1 when WT = 500.

Table 3-3

Simulated Performance (10 logΔI/I) for an "Ideal" Listener in Experiment 2 (WT = 604)*

Starting Level		Signal Level Tracked			
	Percent Correct =	57.9	63.6	73.6	89.1
+4.77		-15.27	-13.50	-12.72	-10.39
-0.78		-15.59	-14.81	-13.35	-10.80
-19.48		-17.04	-16.78	-15.30	-12.11
Ideal		-19.43	-17.03	-14.47	-11.64

*Only starting levels used in real-listener data collection are included. The effects of all 16 starting levels are presented in Appendix A, Table A-7.

Table 3-4

Simulated Performance (10 log $\Delta I/I$) for an "Ideal" Listener in Experiment 3 (WT = 46.9)*

Starting Level	Signal Level Tracked
+4.77	-7.54
-14.96	-10.33
Ideal	-8.55

*Only starting levels used in real-listener data collection are included. The effects of all 16 starting levels are presented in Appendix A (see Table A-8).

run with an easily detectable signal will result in a biased estimate of performance. When WT = 50, beginning a run with $\Delta I = +4.77$ dB will result in an estimated threshold of -7.71 dB rather than the ideal value of -8.71 dB that this version of UDTR is designed to track. All descending runs tend to underestimate performance.

Ascending runs overestimate performance; simulated Weber fractions are too small. For WT = 50, a run starting at -14.96 dB will track a signal level of -10.45 dB rather than the predicted -8.71 dB. Such an error in parameter estimation appears in all of the ascending runs.

Since descending runs yield inflated Weber fractions while ascending runs yield deflated Weber fractions, an ascending-descending difference is to be expected even with an ideal listener. This bias, however, is not of sufficient magnitude to account for the ascending-descending differences presented at the beginning of this chapter. Tables 3-5 through 3-7 present the statistical bias for each experimental condition, calculated by subtracting the simulated Weber fraction (given in Tables 3-2 through 3-4) from the ideal Weber fraction obtained from Equation 1-2. If this bias is then added to the real listeners' Weber fractions, a measure of performance, corrected for methodological bias, is obtained. The remaining differences between ascending and descending thresholds may be presumed to be of psychological origin.

Table 3-5

Statistical Bias Inherent in UDTR for the Starting Levels Used in Experiment 1*

Starting Level	Statistical Bias	
	WT = 50	WT = 500
+4.77	-1.00	-1.66
-14.96	+1.74	+0.18
-19.48	+2.26	+1.00

*All values are ideal Weber fractions minus simulated ideal Weber fractions (in dB).

Table 3-6

Statistical Bias Inherent in UDTR for the Starting Levels Used in Experiment 2 *

Starting Level	Statistical Bias Percent Correct = 57.9	63.6	73.6	89.1
+4.77	-4.16	-3.53	-1.75	-1.25
-0.78	-3.84	-2.22	-1.12	-0.84
-14.96	-2.39	-0.25	+0.83	+0.47

*All values are ideal Weber fractions minus simulated ideal Weber fractions (in dB).

Table 3-7

Statistical Bias Inherent in UDTR for the Starting Levels Used in Experiment 3*

Starting Level	Statistical Bias
+4.77	-1.01
-14.96	+1.78

*All values are ideal Weber fractions minus simulated ideal Weber fractions (in dB).

3. Implications for Data Collection and Analysis

That UDTR will not properly track a target percent correct in a finite run of 80 trials, raises serious questions about data collected with this method. Runs of 100 and 120 statistical trials were also simulated. The effects of starting level were essentially the same as reported for 80-trial runs. A further discussion of data collection and reduction schemes will be presented in Chapter IV.

In practice, it is convenient to begin staircase runs at one comfortable signal level even under two (or more) experimental conditions which are known to affect performance. Tables 3-5 and 3-6 show that for descending runs starting with ΔI = +4.77 dB, the amount of bias introduced by UDTR increases with the difference between the starting level and the signal level tracked. For example (see Table 3-5), UDTR(73.6%) ideally tracks signal levels of -8.71 and -14.08 dB when WT = 50 and 500, respectively. It may be seen that the bias is larger when the signal level tracked is further removed from the starting level. The same is true for the case illustrated in Table 3-6. UDTR(57.9%) will track a lower signal level than will UDTR(89.1%). Starting both runs at +4.77 dB will, therefore, result in a larger bias when tracking 57.9% correct detections. Since the tracking error differs between conditions, intensity-duration reciprocity factors (Experiment 1) and slopes of psychometric functions (Experiment 2) will be contaminated.

A similar argument was presented by Creelman and Taylor (1969) in reference to IT trades obtained with the BUDYEN (Blocked Up-Down Yes-No) method. Campbell and Counter (1969) employed BUDYEN (which is merely

BUDTIF without the second interval) to collect thresholds for pure tones which varied in duration from 4 to 100 msec. Intensity-duration reciprocities were dependent on frequency for frequencies lower than 250 Hz. Creelman and Taylor argued against the use of BUDYEN on the grounds that computer-simulated thresholds tracked with BUDYEN are biased low. Furthermore, the magnitude of the bias is dependent on step size (in percent correct). Creelman and Taylor then argued that if the slope of the psychometric function for pure-tone intensity discrimination is dependent on signal duration (Creelman, 1963), the bias introduced by BUDYEN will vary with signal duration if the step size is constant (in dB) across durations. Intensity-duration reciprocity factors will therefore be contaminated.

If data are to be compared across experimental conditions which are believed to affect performance, the magnitude of the bias introduced by the psychophysical method must be first assessed, and then compensated for. In Section A2 of this chapter, the bias introduced by UDTR was assessed for an ideal receiver. The magnitude of the statistical bias varied across experimental conditions. It is therefore proposed that real-listener data obtained with UDTR be corrected by adding the statistical bias to the obtained Weber fractions. In this manner, thresholds obtained with descending runs, which underestimate performance, will be lowered. Similarly, thresholds obtained with ascending runs, which overestimate performance, will be increased.

Appropriate correction factors have been added to the Weber fractions from Tables A-1 through A-6, and are reported in Tables A-9 through A-14. The remainder of this chapter will consider these data. In addition to reporting performance free from statistical bias, uncorrected descending runs will be considered in order to make direct comparisons to the literature. It will be seen that IT and IW reciprocity factors as well as the slope of the psychometric function all change in magnitude when the data are corrected for methodological bias.

B. Results Corrected for Statistical Bias

Even after correction for methodological bias, data obtained with ascending and descending runs differ substantially.

The basic results found with descending runs are as follows:
(1) Intensity-duration reciprocity factors are smaller than predicted with gated maskers, and larger than predicted with continuous maskers.
(2) Intensity-bandwidth trades are too shallow with both modes of masker.
(3) The slope of the psychometric function is as predicted when the masker is gated simultaneously with the signal. With continuous maskers, the psychometric function is steeper than predicted. (4) Long-duration signals (T = 100 msec) are more easily detected if the masker is present continuously throughout an experimental session. The reverse is true for short-duration signals (T = 10 msec).

The pattern of results is different when staircase runs begin with a signal that is well below detectability: Weber fractions are consistently smaller than those obtained with descending runs; IT and IW trades are as predicted; gated-continuous differences with respect to thresholds, IT trades, IW trades, and the slopes of the psychometric functions all disappear.

In addition, after correcting for statistical bias, sizeable differences in the Weber fractions remain, reflecting a psychological difference in performance on ascending and descending runs. Table 3-8 presents the difference between corrected ascending and descending Weber fractions; a negative value indicates better preformance on ascending runs.

Results appear different for narrowband and wideband noises. For wideband noise (W = 5000 Hz and W = 6040 Hz), averaged across experiments, the ascending-descending difference is -1.51 dB with continuous maskers, and -3.60 dB with gated maskers. For narrowband noise, the effect of starting level is smaller: +0.47 and -1.38 dB for continuous and gated maskers, respectively.

1. Real-Ideal Differences in Performance

The effect of starting level on average detectability may also be assessed by considering the discrepancy between real and ideal behavior. Real-ideal differences in performance are presented in Table 3-9. Each difference score is computed by subtracting the ideal Weber fraction (from Equation 1-2) from the obtained Weber fraction; a positive difference indicates performance that is worse than the limit imposed by stimulus statistics. Data are presented for ascending and descending runs corrected for statistical bias as well as for uncorrected descending runs.

(a) Descending Runs

Uncorrected, descending runs employing UDTR(73.6%) yield Weber

Table 3-8

Differences Between Ascending and Descending Runs (a)

T	Attenuation Rate	Gated			Continuous		
(msec)	(dB/octave)	W = 469	5000	6040(b)	469	5000	6040(b)
10	96		-1.75			-2.25	
100	96	-0.62	-3.73		+0.15	-1.38	
100	48			-5.33			-0.91
100	24	-2.14			+0.79		

(a) All values are ascending Weber fractions minus descending Weber fractions (in dB). A negative difference indicates better ascending performance.

(b) Ascending-Descending differences for W = 6040 Hz are for 73.6% correct detections only.

Table 3-9

Difference Between Obtained and Ideal Weber Fractions*

T (msec)	Attenuation Rate (dB/octave)	Percent Correct	Gated W = 469	Gated 5000	Gated 6040	Continuous 469	Continuous 5000	Continuous 6040
				Descending (uncorrected)				
10	96	73.6		6.10			6.85	
100	96	73.6	3.64	8.30		3.12	5.83	
100	48	89.1			10.47			4.54
100	48	73.6			9.86			5.04
100	48	57.9			9.67			7.67
100	24	73.6	3.04			1.75		
				Descending (corrected)				
10	96	73.6		5.10			5.85	
100	96	73.6	2.63	6.44		0.74	4.17	
100	48	89.1			8.42			3.70
100	48	73.6			8.11			3.92
100	48	57.9			6.31			3.83
100	24	73.6	2.03			2.11		

(Continued)

Table 3-9, Continued

T (msec)	Attenuation Rate (dB/octave)	Percent Correct	Gated			Continuous		
			W = 469	5000	6040	469	5000	6040
				Ascending (corrected)				
10	96	73.6		3.35			3.58	
100	96	73.6	2.01	2.91		2.26	2.79	
100	48	73.6			2.78			2.90
100	48	63.6			2.69			3.01
100	24	73.6	-0.11			1.52		

***All values are Obtained Weber fractions minus Ideal Weber fractions (in dB).**

fractions that average 8.08 and 5.91-dB worse than ideal for gated and continuous wideband maskers, respectively. These data agree with those found in the literature (see Table 1-9), where real-ideal differences are presented for performance levels of approximately 75%.

For gated conditions, real-ideal differences obtained with UDTR(89.1%) and UDTR(57.9%) equal those obtained with UDTR(73.6%). Since the difference in sensitivity between the real and ideal observer is constant across performance levels, the obtained three-point psychometric function is equal in slope to the ideal psychometric function calculated from Equation 1-2. For continuous conditions, real-ideal differences obtained with UDTR(57.9%) are greater than, and those obtained with UDTR(89.1%) are less than those obtained with UDTR(73.6%). The obtained psychometric function is, therefore, steeper than ideal. This result replicates the findings of a gated-continuous difference in the slope of the psychometric function reviewed in Chapter I, Section D-4, and will be discussed more fully in a later section.

For narrowband noise, descending runs yield uncorrected Weber fractions that are closer to ideal than are those obtained with wideband stimuli. Performance is 3.34-dB worse than ideal with gated maskers, and 2.44-dB worse with continuous maskers. These values compare favorably with those reported in the literature (see Table 1-9). Averaged across studies (excluding the 100-Hz condition employed by Schacknow and Raab), the real-ideal difference in performance is approximately 4 dB for gated masker conditions. For continuous masker conditions (W = 316 Hz), Schacknow and Raab (1976) report Weber fractions which average 1.42-dB worse than ideal. (The exclusion of Schacknow and Raab's 100-Hz condition is discussed in Appendix C.)

After correction for statistical bias, Weber fractions obtained with descending runs are reduced, and therefore closer to ideal. For wideband noise, average performance is 6.88 and 4.29-dB worse than ideal for gated and continuous maskers, respectively. For narrowband noise, performance is 2.33-dB worse than ideal with gated maskers, and 1.43-dB worse than ideal with continuous maskers.

(b) Ascending Runs

Even after correction, Weber fractions are smaller for ascending than descending conditions. It follows that real-ideal differences for ascending conditions will be smaller than those reported above for descending runs. For wideband noise, gated and continuous maskers yield average real-ideal differences of 2.93 and 3.07 dB, respectively. For narrowband noise, the discrepancy is 0.95 dB with gated maskers, and 1.89 dB with continuous maskers.

In summary, descending runs yield performance that is consistent with that found in the literature. With wideband noise, performance is approximately 6-dB worse than if stimulus fluctuations alone governed detection. For narrowband noise, performance is only 2.5-dB worse than ideal. Performance on ascending runs, on the other hand, is considerably better. Wideband maskers yield average Weber fractions that are only 3-dB worse than ideal. Narrowband maskers yield Weber fractions that are approximately 1.5-dB worse than ideal.

2. Effects of the Mode of Masker

In the Introduction, evidence was marshalled which indicates that, for noise-intensity discrimination, a gated masker is more effective than a continuous one. Table 3-10 presents the difference in performance between gated and continuous presentations for all three experiments, averaged across both observers and levels. A positive difference indicates better performance with a continuous masker. It should be noted that there need not be a correction for the statistical bias in assessing a gated-continuous difference if runs under both conditions are initiated from the same signal level. The simple energy detector, and therefore, the simulated ideal observer, makes no distinction between modes of the masker; the receiver provides its own temporal gating to correspond to the duration of the signal interval. With the exception of Experiment 2, gated-continuous differences on descending runs are identical before and after correction for the statistical bias. In Experiment 2, gated and continuous runs were initiated from different signal levels, which results in a slight difference between the gated-continuous differences reported before and after correction.

(a) Average Detectability

(1) Descending Runs

Descending runs confirm the findings reported in the literature (see Table 1-8). For long-duration signals, an increment is more easily detected when masked by a continuous background. Averaged across conditions, this difference is +3.46 dB for wideband noise. For

Table 3-10

Gated-Continuous Differences in Performance*

T (msec)	Attenuation Rate (dB octave)	Percent Correct	Gated-Continuous Difference		
			W = 469	5000	6040
		Descending (uncorrected)			
10	96	73.6		-0.75	
100	96	73.6	+0.51	+2.47	
100	48	89.1			+5.13
100	48	73.6			+4.82
100	48	57.9			+2.80
100	24	73.6	+1.29		
		Descending (corrected)			
10	96	73.6		-0.75	
100	96	73.6	+0.51	+2.47	
100	48	89.1			+4.72
100	48	73.6			+4.19
100	48	57.9			+2.48
100	24	73.6	+1.29		
		Ascending (corrected)			
10	96	73.6		-0.23	
100	96	73.6	-0.24	+0.12	
100	48	73.6			-0.21
100	48	63.6			-0.21
100	24	73.6	-1.63		

*All values are Weber fractions obtained with gated maskers minus those obtained with continuous maskers (in dB). A positive Gated-Continuous difference indicates better performance with a continuous masker.

narrowband noise, a smaller gated-continuous difference of +0.90 dB obtains. The effect of the mode of masker reverses for short duration signals: Weber fractions are 0.75-dB smaller with gated maskers. This reversal is consistent with the findings of Wightman and Green (1966), Viemeister (1974b), and Berner (unpublished data).

A similar situation is evidenced in the study of visual increment detection. Cornsweet and Pinsker (1965) obtained a gated-continuous difference in performance for luminance increment detection when a spatial 2AFC task was employed with a descending version of UDTR(70.7%). A luminance increment was added to one of two background fields which were either on continuously, or pulsed simultaneously with each other and the signal. With 4.5-msec signals, better performance was obtained in the pulsed (or gated) condition. These results with short-duration signals were replicated by Matin and Kornheiser (1976), who, in addition, obtained thresholds for 200-msec increments. With the longer duration signals, the gated-continuous difference reversed in direction: continuously masked signals were easier to detect. This reversal is consistent with that found in audition, and replicates the findings of Leshowitz, Taub, and Raab (1968) for vision. In the Leshowitz et al. study, continuously masked signals were easier to detect when the duration was equal to or greater than 10 msec. For durations equal to or less than 3 msec, performance was better when the masker was gated. In both studies, blocks of fixed-intensity signals were employed, preceded by suprathreshold practice trials. This method presumably provides information at suprathreshold signal levels similar to that provided by a descending staircase.

(2) Ascending Runs

Data obtained with ascending runs differ markedly from those reported in the literature; gated and continuous maskers have equivalent effects on discriminability. It should be noted that the lack of a gated-continuous difference is the result of greatly improved performance with gated maskers rather than impaired performance with continuous maskers. An exception to this finding is that continuous maskers are more effective for narrowband noise with 24 dB/octave skirts.

(b) Slope of the Psychometric Function

The results of Experiment 2 are presented in Figures 3-1 through 3-3, where Weber fractions (10 log $\Delta I/I$) are plotted for each level of detectability tracked. The leftmost function in each figure is the performance of an ideal energy detector, calculated from Equation 1-2. For an ideal observer, a change in performance from 57.9% to 89.1% correct detections requires a 7.79-dB increase in $\Delta I/I$. For real listeners (in descending conditions), psychometric functions were fit by the method of least-squares since the energy detection model predicts a linear function when $\Delta I/I$ is plotted as a function of M (or d') — see Equation 1-2. The increase in $\Delta I/I$ (in dB) required for real-listener performance to change from 57.9% to 89.1% correct detections was then derived from these functions. These values are reported below.

Figure 3-1

Percent correct as a function of 10 log $\Delta I/I$ (descending runs, uncorrected for bias). The open triangles represent continuous masker conditions, closed triangles - gated (upright for observer CM, inverted for observer RB). The leftmost function represents the performance of an ideal energy detector calculated from Equation 1-2, with $WT = 604$.

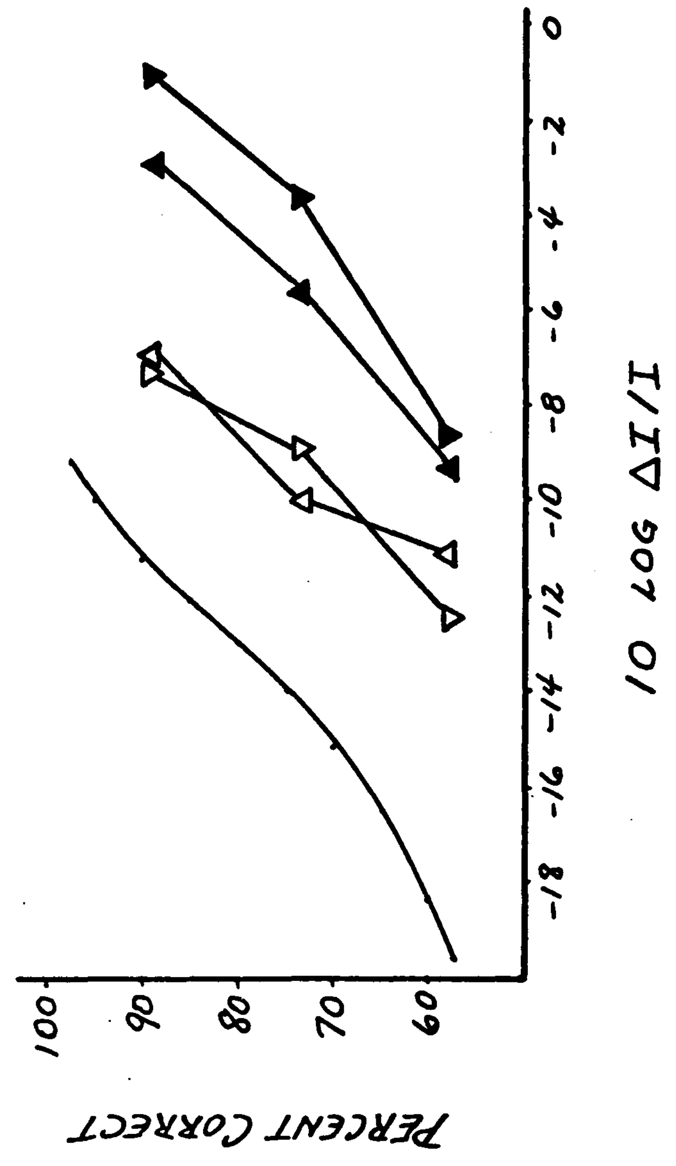

PERCENT CORRECT
100
90
80
70
60
-18
-16
-14
-12
-10
-8
-6
-4
-2
0
10 LOG ΔI/I

Figure 3-2

Percent correct as a function of 10 log $\Delta I/I$ (descending runs, corrected for bias). The open triangles represent continuous masker conditions, closed triangles - gated (upright for observer CM, inverted for observer RB). The leftmost function represents the performance of an ideal energy detector calculated from Equation 1-2, with WT = 604.

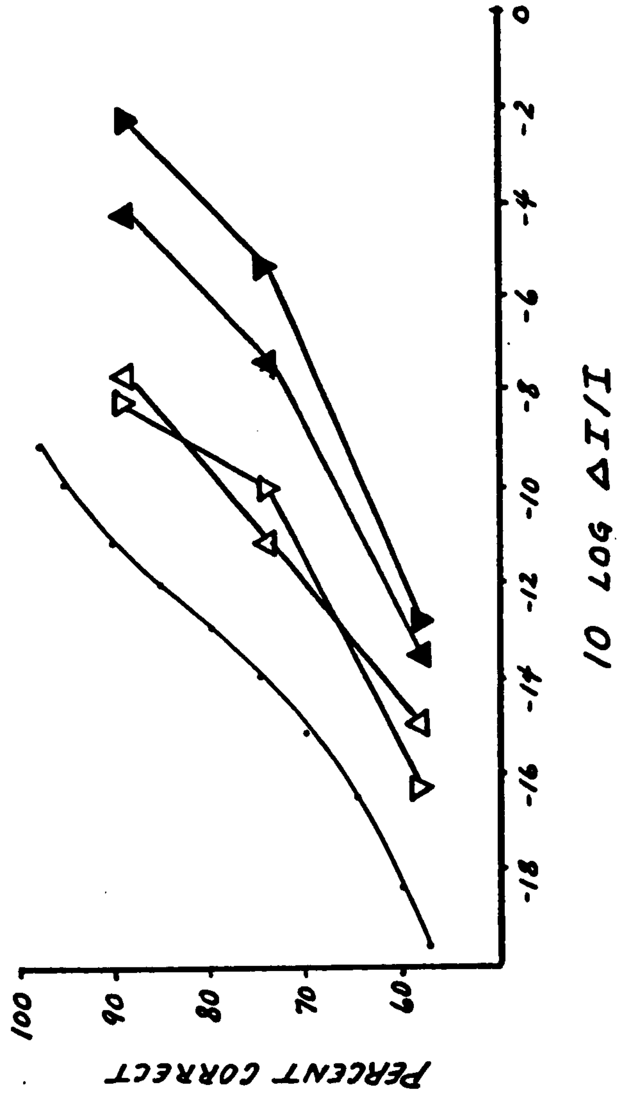
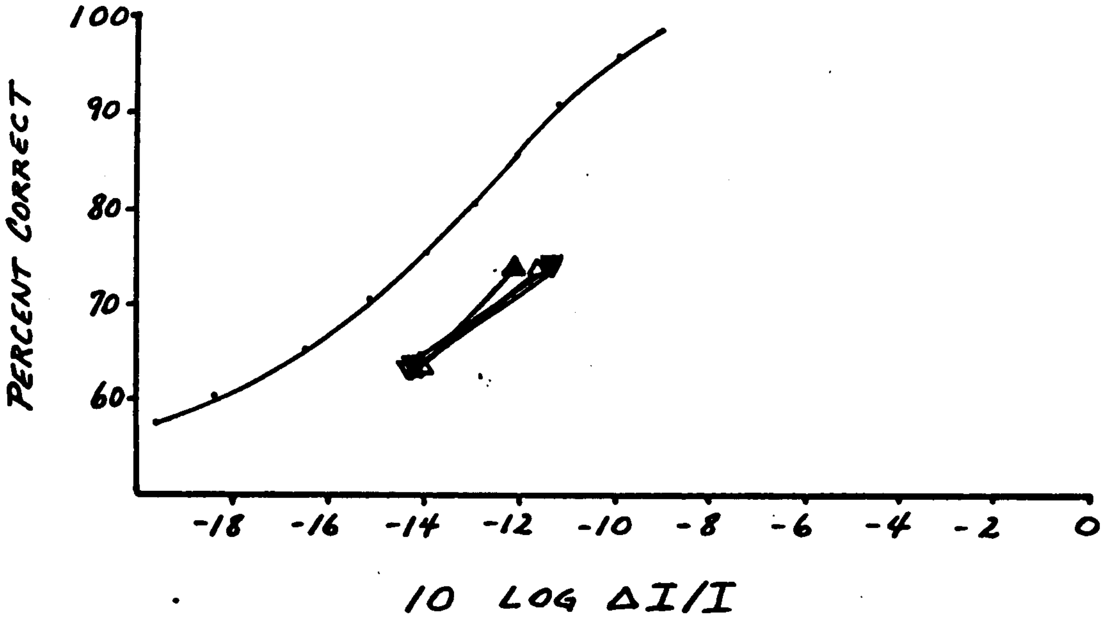

PERCENT CORRECT
100
90
80
70
60
-18
-16
-14
-12
-10
-8
-6
-4
-2
0
10 LOG ΔI/I

Figure 3-3

Percent correct as a function of 10 log $\Delta I/I$ (ascending runs, corrected for bias). The open triangles represent continuous masker conditions, closed triangles - gated (upright for observer CM, inverted for observer RB). The leftmost function represents the performance of an ideal energy detector calculated from Equation 1-2, with $WT = 604$.


PERCENT CORRECT
100
90
80
70
60
-18
-16
-14
-12
-10
-8
-6
-4
-2
0
10 LOG ΔI/I

(1) Descending Runs

A gated-continuous difference at all levels of detectability is evident in the descending runs depicted in Figures 3-1 and 3-2. The magnitude of this effect appears to be dependent on the level of detectability; the gated-continuous difference grows with the percent correct tracked (see Table 3-10). The psychometric function, therefore, is steeper for continuous conditions (detectability grows more rapidly with signal intensity for continuously masked signals). Uncorrected for simulations, our observers require an average increase of 6.91 dB for gated and 4.78 dB for continuous conditions in order to improve performance from 57.9% to 89.1% correct detections. Corrected for simulations, this difference in slope is unchanged, although both functions become more shallow.

These data replicate the findings of Green (1960) and Campbell (1965) for continuously masked noise, DeBoer (1965) for gated maskers, and Green and Sewall (1962) who employed both modes of the masker. Green and Sewall attributed this gated-continuous difference in the slope of the psychometric function to the reduction of temporal uncertainty afforded by gating the signal and masker simultaneously.

The concept of uncertainty affecting the slope of the psychometric function has its origin in a paper by Peterson, Birdsall and Fox (1954). For the case of tone-masked-by-noise, they showed that a cross-correlation receiver will exhibit a variety of psychometric functions, each of whose slope depends on the degree of uncertainty about signal parameters. Each function is generated by assuming that the receiver knows that the signal will be randomly selected from a population of M orthogonal signals. For pure tones, the nature of this uncertainty may be the frequency, phase, or time of occurrence of the signal. As the uncertainty parameter, M, increases, the psychometric function both increases in slope and is displaced to the right along the abscissa. In other words, performance declines with increasing uncertainty for all detectabilities, but more so at lower levels of detectability.

The uncertainty model has been supported by data obtained for human observers. The psychometric function displayed by a human observer for a pure tone signal masked by continuous noise is similar to that predicted for an ideal receiver with some uncertainty (see for example, Green, McKey, and Licklider, 1959). Assuming that the steepness of the human observer's psychometric function (even when the signal frequency is known) is due to uncertainty, several experimental techniques have been employed in an attempt to change its slope.

Frequency uncertainty might be reduced if a continuous pure tone were added to the noise masker. The observer would then have to detect an increment in the pure tone. The psychometric function is less steep with this procedure, and the signal becomes more detectable (Green, 1960a).

Although frequency information is provided by a continuous sinusoid in the above example, temporal uncertainty remains. This uncertainty may be minimized by gating the tonal background simultaneously with the signal. This so-called pedestal condition is equivalent to amplitude discrimination for two bursts of a pure tone, masked by a continuous noise. It is well documented (see Green and Swets, 1966) that the presence of a pedestal influences performance sufficiently so that the resulting psychometric function is parallel to that predicted for the case of the signal-known-exactly.

For noise-intensity discrimination, the ideal detector is one that measures the energy in each observation interval and reports the larger of the two. Uncertainty about signal parameters can include the frequency region, starting time, and duration of the signal. It has been shown for homogeneous noise-intensity discrimination, that the psychometric function for continuously masked signals is steeper than predicted. When the masker is gated, however, the resulting psychometric function is parallel to that predicted for the ideal receiver. Since there is no frequency uncertainty with a continuous, homogeneous background, Green and Sewall (1962) have attributed the steepness of the psychometric function to temporal uncertainty. The observed decrease in slope in the gated masker condition has therefore been attributed to temporal uncertainty reduction, although it is not perfectly clear how Green and Sewall make the transition from the uncertainty arising from a choice of M orthogonal tonal signals to temporal uncertainty in noise-intensity discrimination.

Uncertainty reduction should also result in improved discriminability — the gated psychometric function should be displaced to the left of the continuous function. Such is not the case for the functions reported by Green and Sewall. Performance in the continuous masker condition is better at all detectabilities. Uncertainty reduction alone is therefore an inadequate explanation for the difference in the slope of the psychometric functions, although it might be argued that for their data, the relative placement of the functions cannot be assessed since different subjects were employed in gated and continuous masker conditions. The data reported in this dissertation are for a within-subjects design and consequently allow this comparison.

Examination of Figure 3-1 shows that the psychometric function for the gated mode is more shallow than the continuous function, and is equal in slope to that predicted for Green's ideal receiver. It is, however, displaced to the right of the continuous function. That is, performance is worse at all levels of detectability in a condition that is assumed to reduce uncertainty and improve performance. The uncertainty argument alone, therefore, cannot account for the gated-continuous differences reported in this dissertation.

(2) Ascending Runs

Psychometric functions for ascending conditions must be inferred from Weber fractions obtained over a narrower range of detectability than that used with descending runs. Pilot data as well as computer simulation indicated that tracking a level of detectability as low as 57.9% with ascending runs would be impractical.

In all ascending conditions employed in this dissertation, the starting level was chosen to be approximately 6-dB below the performance level predicted for Green's ideal receiver. In this manner, the signal to be detected would remain "inaudible" for a number of trials while the observer's behavior is shaped by feedback. That is, a detection strategy would be adopted to maximize the occurrence of a green (correct) feedback light. The same approach was used in collecting pilot data where the signal level which yields 57.9% correct detections was tracked with ascending runs. For WT = 604, Green's ideal receiver would achieve this level of performance when 10 log $\Delta I/I = -19.5$ dB. Ascending runs were therefore started at -25.5 dB, where the probability of a correct response for the ideal receiver equals 52%.

Averaged across subjects, the Weber fraction tracked with UDTR(57.9%) equalled -24.91 dB. This suggested that in a run of 80 trials, the ascending version of UDTR(57.9%) was not moving from the starting level. It is conceivable that shaping a response strategy is extremely difficult when the probability of a correct response is approximately equal to the probability of an incorrect response (recall that the run was started at a level where P(C) = 52%). This situation may lead to a considerable amount of guessing behavior. If this is the case, 80 trials may be insufficient for UDTR(57.9%) to move appreciably from the starting level if responding is governed by chance.

Computer simulation of chance responding confirmed these findings. When UDTR(57.9%) was started at -25.5 dB, a signal level of -25.05 dB was tracked in 80 trials.

It is likely, therefore, that when an ascending version of UDTR(57.9%) is employed with real listeners, initial responding is so close to chance that the feedback received is insufficient to shape the observer's behavior in a run of 80 trials. Responding appears to remain at the chance level.

Tracking high levels of detectability with ascending runs poses a conceptual problem. Beginning a run 6-dB below the anticipated target level can easily result in what is effectively a descending run. When WT = 604, the signal level which yields 89.1% correct detections (for the ideal receiver) is approximately -11 dB. A starting level of approximately -17 dB would be chosen for an ascending run, where the probability of a correct response equals 65%. Since the decision rule for UDTR(89.1%) calls for an increase in signal level if one error is made before 6 correct responses, only a few errors during the initial shaping will provide the observer with a signal that is easily detectable (UDTR(89.1%), by definition, tracks an easily detectable signal). At this point, the observer is exposed to whatever cue and resulting alternate detection strategy that he is thought to use in a descending run.

In general, UDTR decision rules for tracking detectabilities greater than 75% call for an increase in signal level after a single incorrect response. Otherwise, lengthy sequences of responding are necessary to effect a level change, which would lead to few reversals in a run of 80 trials. In any case, since a high target level implies an easily detectable signal, it is difficult to determine whether performance is the result of the detection strategy employed in an ascending or a descending run.

Psychometric functions obtained with the particular ascending strategy used in the other experiments in this dissertation were therefore constructed from two points only -- 63.6% and 73.6% correct detections.

It is evident from Figure 3-3, that for ascending runs, performance is equivalent for gated and continuous conditions at both levels of detectability tracked. In this region the psychometric functions may be presumed to be identical. In addition, the slopes of the obtained psychometric functions parallel that of the ideal observer: the real-ideal difference in performance is the same at both levels of detectability (see Table 3-9).

(c) Summary

Data collected with descending staircase runs replicate the gated-continuous differences found in the literature. Noise increments are more easily detected if the masker is presented continuously, than if it is gated simultaneously with the signal. In addition, the slope of the psychometric function is steeper for continuously masked signals. It has been shown that these results are not adequately explained by an uncertainty model. These results, taken together with the gated-continuous differences in IT and IW trades to be presented in the following section, suggest that perhaps different discriminative processes underlie the detection of noise signals in the presence of gated and continuous maskers.

When threshold is tracked with ascending runs, performance is independent of the mode of masker — Weber fractions are the same for gated and continuous masker paradigms. Consequently, the slope of the psychometric function and the magnitude of IT and IW trades are also independent of the mode of masker. Furthermore, absolute detectabilities more closely approximate those of an ideal energy detector. This result is consistent with the idea that for ascending runs, a single discriminative process is employed which resembles Green's ideal energy detector.

3. Intensity-Duration Reciprocity

Experiment 1 examined the effects of signal duration on noise-intensity discrimination. The performance of an ideal energy detector, given by Equation 1-2, is such that, for a 5000-Hz band of noise, an increase in signal duration from 10 to 100 msec will yield a 5.37-dB improvement in the Weber fraction. Table 3-11 presents reciprocity factors (defined as the change in the Weber fraction resulting from a log-unit increase in duration) for the results obtained in Experiment 1. Data are averaged across masker levels and observers.

In all instances, discriminability is improved by increasing the signal duration by one log-unit.

Table 3-11

Intensity-Duration Reciprocities
Obtained in Experiment 1

	Descending	Ascending
	Uncorrected	
Gated	-3.17	-5.07
Continuous	-6.39	-5.42
	Corrected	
Gated	-3.83	-5.81
Continuous	-7.05	-6.16
	Ideal*	
	-5.37	-5.37

*Ideal reciprocity factor computed from Equation 1-2.

(a) Descending Runs

Descending runs will be considered first. For the gated masker condition, a decade increase in duration results in a 3.17-dB improvement in the uncorrected Weber fraction. With continuous maskers, a 6.39-dB improvement obtains. Corrected for bias, reciprocity factors become -3.83 and -7.05 dB for gated and continuous maskers, respectively. These results strongly confirm the findings of other investigators in which average IT reciprocities are reported to be -3.3 dB with gated maskers, and -6.3 dB with continuous maskers (see Table 1-5).

This gated-continuous difference in the course of temporal integration is also found in vision. Leshowitz, Taub, and Raab (1968), as well as Matin and Kornheiser (1976), report perfect temporal summation for the case of continuously masked luminance increments. When the background was pulsed, however, performance was affected little by increasing duration.

(b) Ascending Runs

A consequence of the lack of gated-continuous differences for all ascending conditions is that IT reciprocity factors will not differ as a function of the mode of masker. A log-unit increase in signal duration yields an improvement of 5.81 and 6.16 dB for gated and continuous conditions, respectively. Both values differ little from each other and from the IT reciprocity of -5.37 dB predicted by Green's model.

Once again, observers perform in a more optimal fashion when staircase runs are initiated with a signal that is well below detectability. As signal duration increases, the observer appears to take full advantage of the resulting decrease in relative variability, as would an ideal receiver. When descending staircase runs are employed, however, the human observer performs differently in gated and continuous masker conditions. When the masker is gated, the improvement in performance accompanying an increase in signal duration is less than that predicted from the consideration of stimulus statistics. For continuous masker conditions, the improvement is greater than predicted.

4. The Effects of Bandwidth on Intensity Discrimination

(a) Intensity-Bandwidth Reciprocity

The improvement in performance resulting from an increase in bandwidth from 469 Hz (Experiment 3) to 5000 Hz (Experiment 1) may be seen in the intensity-bandwidth trades presented in Table 3-12. Data are averaged across observers and masker levels. Only data obtained with 96 dB/octave spectrum skirts are included in this analysis in an attempt to make more direct comparisons to theoretical IW trades which assume rectangular bands of noise. For an ideal energy detector, this increase in bandwidth will result in a 5.53-dB improvement in the Weber fraction.

(1) Descending Runs

Data obtained with descending runs replicate the findings summarized in Table 1-6 for homogeneous bands of noise. Stimulus bandwidth has a much smaller effect on discrimination than is predicted by Green's model.

Table 3-12

Intensity-Bandwidth Reciprocity

.	Descending	Ascending
	Uncorrected	
Gated	-0.86	-3.85
Continuous	-2.82	-4.22
	Corrected	
Gated	-1.51	-4.63
Continuous	-3.47	-4.99
	Ideal*	
	-5.53	-5.53

*Ideal IW trade computed from Equation 1-2.

Uncorrected for bias, improvements of 0.86 and 2.82 dB obtain for gated and continuous conditions, respectively. When corrected, IW trades are 1.51 and 3.47 dB. The bandwidth effect, therefore, is larger for continuous than for gated masker conditions.

(2) Ascending Runs

Observers appear to make more optimal use of bandwidth information on ascending runs. An increase in signal bandwidth from 469 to 5000 Hz results in an improvement in the corrected Weber fraction of 4.63 and 4.99 dB for gated and continuous conditions, respectively. These trades are closer to the predictions of Green's model than any heretofore reported for homogeneous bands of noise.

(b) Attenuation Rate of the Filter

Green's model for noise-intensity discrimination considers the case of "idealized" rectangular bands of noise. The most efficient means of processing such signals is for the observer to employ a rectangular input filter matched to the bandwidth of the signal. Schacknow and Raab (1976) have noted that for a noise-signal whose spectrum has less than infinitely steep skirts the bandwidth of this internal filter is critical: If the receiver's input filter extends beyond the half-power bandwidth of the signal, the relative variability (sigma-to-mean ratio) of the noise is somewhat reduced. A limit on this improvement is imposed by the residual noise of the filter as well as that of other system components. This system noise will mask the filter's output at frequencies removed from the cut-off frequencies, where signal energy is minimal. The "equivalent statistical bandwidth" of a non-rectangular band of noise may now be defined as the width of the rectangular band of noise having an equivalent sigma-to-mean ratio, provided that all spectral components of the stimulus have greater power than system noise. Since the noise samples used in the experiments reported here had finite attenuation rates, the equivalent statistical bandwidths were always greater than their corresponding half-power bandwidths. Such increases in "statistical" bandwidth and the resultant decreases in relative variability lead to the prediction of lower thresholds than if half-power bandwidths are considered (see Schacknow and Raab's Table IV, p. 900).

An even more efficient strategy for processing such "tinted" noises is for the observer to employ a transfer function which is the inverse of the stimulus spectrum (cf. Schacknow and Raab, 1976, p. 900). This effectively provides the listener with white Gaussian noise, whose relative variability is less than that of Gaussian noise having any other spectrum. This transformation has been called "prewhitening" (Blackman and Tukey, 1958). The accompanying decrease in relative variability should serve to increase the detectability of a noise increment.

If a listener is able to utilize either of the above strategies, the particular shape of the filtered waveform should have an effect on performance. Specifically, the attenuation rate of the filter for non-rectangular bands of noise should influence the effective passband, such that wider effective passbands will be realized for filtered waveforms with shallow skirts. It follows that if the observer prewhitens the input, or simply filters wider than the half-power bandwidth, performance should improve as spectrum skirts are made more shallow.

Table 3-13 presents the effects of varying the attenuation rate of the filter on intensity discrimination. A negative difference indicates

Table 3-13

Effects of Attenuation
Rate on Discrimination*

	Descending	Ascending
Gated	-0.58 -0.62	-1.89 -2.35
Continuous	-0.70 -2.05	-1.80 +0.33

*Values are differences between corrected Weber fractions obtained with 24 and 96 dB/octave attenuation rates (in dB).

The upper values refer to observer CM, the lower to observer RB.

better performance for noise filtered with 24 dB/octave skirts. Since there is no existing literature relevant to the comparisons made in this section, uncorrected results will not be considered. On the whole, performance is somewhat better for the 24 dB/octave conditions. For gated masker conditions, the advantage of decreasing the slope of the skirts is approximately 2 dB for ascending runs, whereas the advantage afforded on descending runs is only 0.60 dB. For continuously masked signals, results for observers CM and RB differ. Observer CM's data with continuous maskers for both ascending and descending runs are equivalent to those reported for gated maskers. For RB, performance on descending runs is improved by 2 dB as a result of decreasing the slope of the spectrum skirts, whereas on ascending runs no difference in performance obtains.

Evidence has been presented that supports Schacknow and Raab's contention that the observer does not simply employ a rectangular internal filter matched to the half-power bandwidth of the signal. As more signal bandwidth is made available beyond the half-power points of the noise spectrum (resulting in a reduction in the sigma-to-mean ratio) by decreasing the attenuation rate of the filter, performance improves. Although the exact shape and bandwidth of the internal filter remain unspecified, these results suggest that the observer's input transfer function extends beyond the half-power bandwidth of the stimulus. Furthermore, the effective passband appears to be wider on ascending runs than on descending runs.

That obtained IW trades are more shallow than predicted has been explained by Schacknow and Raab (1976) as partly due to the fact that theoretical reciprocity factors have been calculated for rectangular noise stimuli, rather than for the nonrectangular waveforms actually employed. Furthermore, if one considers the possibility that the observer employs an internal filter which is both wider than the half-power bandwidth of the signal, and an inverse transfer function of the stimulus spectrum (prewhitening), then new theoretical reciprocity factors may be calculated. Schacknow and Raab provide such an analysis for nonrectangular bands of noise which have attenuation rates of 48 dB/octave (see Table VII, p. 903). There, it is seen that theoretical reciprocity factors are considerably reduced from -5 dB. This results from the fact that the advantage afforded by prewhitening is bandwidth dependent. That is, the improvement in the performance of an ideal receiver resulting from such a consideration is far greater for narrowband noise.

This analysis, however, will not reconcile the shallow IW trades reported in this dissertation for descending staircase conditions. First, the noise waveforms were filtered to have attenuation rates of 96 dB/octave, thereby minimizing the difference between half-power and equivalent statistical bandwidths. Second, the results of Experiment 3 suggest that the human observer employs a wider internal filter on ascending runs than on descending runs. Bandwidth reciprocity factors for ascending runs should therefore be even more shallow than those for descending runs. Such is not the case.

5. Summary of Results

In summary, the main finding of this dissertation is that the starting level employed in adaptive testing can have profound effects on performance, possibly by altering the information made available to the observer.

Providing the observer with a suprathreshold signal at the beginning of a staircase run (descending runs) often yields results that are not in accord with the predictions of an energy-detection model — intensity-duration reciprocity factors are smaller than those predicted with gated maskers, and larger than those predicted with continuous maskers; intensity-bandwidth trades (with homogeneous noise) are too shallow with both modes of masker; the psychometric function is steeper than that predicted with continuously masked signals, although it is of the proper slope with gated maskers; gated maskers are more effective with long-duration signals, whereas the reverse is true for short-duration signals.

Beginning with a signal that is well below detectability, on the other hand, yields results that are consistent with the predictions of Green's model: Weber fractions are consistently smaller than those obtained with descending runs; IT and IW trades are as predicted; gated-continuous differences with respect to thresholds, IT trades, IW trades, and the slopes of the psychometric functions all disappear.

Chapter IV

RECOMMENDATIONS FOR DATA COLLECTION AND REDUCTION

In Chapter III, the magnitude of the statistical bias inherent in UDTR was assessed for an ideal listener. In the present chaper, the nature of this bias will be investigated by computer-simulation, and recommendations will be made for implementing relatively unbiased versions of the Up-Down Transformed-Response method.

It is probable that ascending and descending adaptive procedures tap different abilities in the human observer. Which procedure is more desirable in practice will, however, depend on the practical and theoretical interests of the individual investigator. If, for example, ultimate sensitivity is sought for noise-intensity discrimination, the solution is clear — an ascending version of UDTR should be employed. On the other hand, the human observer behaves differently when suprathreshold signals are presented at the beginning of a run. This is certainly worthy of consideration, especially since most of the literature has been obtained with such methods. Descending adaptive techniques, therefore, still have their place in psychophysics.

It is important to examine more closely the statistical bias introduced by UDTR so that recommendations may be made for the development of an unbiased adaptive method. Inasmuch as the descending version of UDTR has been widely used in the literature, this section will be directed to that bias. It should be kept in mind, however, that since ascending runs are biased toward the starting level by approximately the same amount as are descending runs, arguments similar to those to be presented below may be made for ascending runs. In addition, this discussion will be mainly concerned with UDTR(70.7%) and UDTR(73.6%) since these methods are commonly employed.

Monte Carlo computer simulations were used to assess the nature of the bias inherent in UDTR. Initial consideration was given to the particular method of data tabulation and reduction used in these studies. Recall that in both ascending and descending conditions, the target is approached from a remote signal level. As the target level is approached, a count is dropped in a separate bin for each signal level visited; as levels are revisited, additional counts are deposited. The threshold estimate is then the median of levels <u>revisited</u>. That is, levels visited only once are excluded from the data analysis. Consequently, those levels visited as a result of a poor placement of initial observations will not bias the threshold estimate toward the choice of starting level.

For example, the estimate of threshold for the hypothetical run in Table 4-1 would be half-way between levels 6 and 7 (or simply 6.5). Note that the first reversal occurred at level 7. In the initial descent from the starting level, counts are dropped in bins 1 through 7. These initial counts, however, are only ignored for levels 1 through 4. According to the above rationale, all counts deposited as the region of interest is approached should be ignored. If only the counts deposited after the first

Table 4-1

A Hypothetical Staircase Run

Signal Level	Reversal 1	2	3	4	5	Total	Total After 1st Reversal
1	1					1	
2	1					1	
3	1					1	
4	1					1	
5	1	1				2	1
6	1	1	1	1		4	3
7	1		1	1	1	4	3
8			1		1	2	2
9					1	1	1

reversal are considered, the estimate of threshold will be lower. In this example, the median becomes 6.83. The inclusion of these initial or underlying "1's" in the data analysis results in a threshold estimate that is biased in the direction of the starting level.

If these initial counts are deposited in an ascending version of UDTR, and the underlying "1's" are counted, the threshold estimate will be biased in the opposite direction. This could contribute to the ascending-descending difference in simulated staircase runs presented in Chapter III-B.

In order to test this hypothesis, 80-trial runs were simulated for an ideal observer. These simulations are identical to those presented in Chapter III-B, except for the following data reduction rule: counts are not deposited until after the first reversal (although all trials are counted to determine the run length). In effect, a number of initial trials are excluded from the data analysis. For comparison, simulations are presented in which counts are deposited for all levels, but only those levels revisited are included in the data analysis (these runs replicate those in Chapter III-B).

In Table 4-2 it can be seen that the orderly effect of starting level on the estimated threshold is present with both methods of data reduction, although the magnitude of the bias is somewhat less when only trials after the first reversal are considered. The inclusion of underlying "1's" in the data analysis is therefore an insufficient explanation of the methodological bias.

Let us now consider the number of trials that are actually presented in the region of interest. Column 4 of Table 4-2 presents the number of trials which remain after the first reversal for 80-trial runs. For all starting levels, the first reversal is made at approximately the level where

Table 4-2

Simulated Performance (10 log $\Delta I/I$) with UDTR(73.6%) for an "Ideal" Listener (WT = 500)

Starting Level	Signal Level Tracked		Number of Trials After 1st Reversal
	Method I(a)	Method II(b)	
+4.77	-12.43	-12.30	22
+2.83	-12.75	-12.84	26
+0.98	-12.82	-13.24	30
-0.78	-12.93	-13.34	34
-2.48	-13.01	-13.46	39
-4.12	-13.08	-13.49	44
-5.72	-13.11	-13.54	49
-7.03	-13.19	-13.57	54
-8.85	-13.31	-13.62	59
-10.39	-13.41	-13.66	64
-11.92	-13.57	-13.69	68
-13.44	-13.87	-13.77	71
Ideal	-14.08	-14.08	

(a) In method I, counts are deposited for all levels visited. The median is then calculated for all levels revisited.

(b) In method II, counts are deposited only after the first reversal. The median is then calculated for all levels revisited.

P(C) = .76 ($\Delta I/I$ = -13.44 dB). The number of trials remaining once this point is reached will depend solely on the choice of starting level. In the most extreme case, where the starting level is +4.77 dB, only 22 tials remain once the region of interest is reached. As UDTR is initiated closer to the target level, the number of trials in the region of interest increases and the magnitude of the bias decreases. Since the level at which the first reversal occurs is independent of starting level (although in all cases it is a bit high), these simulations suggest that the bias may be reduced by increasing the number of trials in the region of interest.

If the methodological bias is the result of too few trials in the region of interest, a number of solutions exist:

1) Begin UDTR at the target level. This approach, however, would have little practical value if the target level is initially unknown. In addition, if UDTR is initiated at the target level, suprathreshold practice trials would be sacrificed. Although it has been shown that such practice can interfere with performance in a noise-intensity discrimination task, it could be beneficial in other psychophysical tasks.

2) Increase the length of the run. A run beginning at +4.77 dB will show a small bias (0.38 dB) if the total number of trials is 200. Although this approach is of theoretical interest, 200 trials are more than most experimenters are willing to present given that observers might fatigue.

It is possible, however, to divide a 200-trial run in half by starting the second 100 trials (after a brief rest) from the last level tested. Although this would eliminate the bias in the final estimate, it is not clear that such a procedure would be equivalent to a single run for a human observer.

3) Step size may be varied so that the target level is rapidly approached in large steps, followed by smaller steps for more precise parameter estimation. Such a strategy is employed in PEST as well as some versions of UDTR.

4) A version of UDTR may be employed which changes level more rapidly than UDTR(73.6%) and tracks a similar level of detectability. The UDTR(70.7%) method satisfies this requirement since a level change will be called for after one or two trials.

5) If one prefers to use UDTR(73.6%) rather than UDTR(70.7%), a hybrid version of UDTR may be employed with multiple decision rules. More specifically, UDTR(70.7%) could be used to locate the region of interest. This would allow 2 practice trials at each stimulus level until a single incorrect response is made. A fixed number of trials could then be run with UDTR(73.6%). Since the first incorrect response will be in the region where P(C) = .707, the bias arising from starting at a level other than the target will be small.

In order to assess the relative merits of the last two suggestions, each was simulated for an ideal observer.

Suggestion 4 was examined by exploring the effect of starting level on UDTR(70.7%). Total run-length was 80 trials, and counts were not deposited in bins until after the first reversal. It is evident from Table 4-3 that the bias introduced by UDTR(70.7%) is considerably smaller that that introduced by UDTR(73.6%) — see Table 4-2, method II for comparison. In the most extreme case tested, starting UDTR(70.7%) at +4.77 dB results in a tracking error of only 0.23 dB, when WT = 500 and the target level is -14.08 dB. One can, therefore, safely initiate UDTR(70.7%) from a signal level far removed from the target and still obtain bias-free estimates of threshold.

Table 4-3

Simulated Performance (10 log Δ I/I) with UDTR(70.7%) for an "Ideal" Listener (WT = 500)

Starting Level	Signal Level Tracked	Number of Trials After 1st Reversal
+4.77	-13.86	57
+2.83	-13.84	59
+0.98	-13.92	61
-0.78	-13.91	63
-2.48	-13.89	65
-4.12	-13.93	67
-5.72	-13.93	69
-7.03	-13.90	71
-8.85	-13.88	73
-10.39	-13.93	75
-11.92	-13.96	76
-13.44	-14.03	77
Ideal	-14.08	

In lieu of UDTR(70.7%), some investigators have preferred to employ UDTR(73.6%), which allows an observer to make a single error without causing a change in level (recall that UDTR(73.6%) calls for a level change following either 5 correct or 2 incorrect responses). As a result of such a large block size, this method moves slowly, and many trials might be wasted in an attempt to locate the target level. It has been shown that this reduction in the number of trials in the region of interest will result in a biased estimate of threshold. A hybrid method, as suggested above, would allow for a rapid location of the region of interest with UDTR(70.7%), followed by the implementation of UDTR(73.6%).

In order to assess its feasability, this method was simulated for an ideal observer. All runs were initiated from +4.77 dB and theoretically tracked -14.08 dB (WT = 500). The region of interest was located with UDTR(70.7%). Data were not collected until after the first reversal, whereupon UDTR(73.6%) was implemented. From this point, run length was varied from 40 to 160 trials.

The results of the simulation are presented in Table 4-4. Since approximately 23 trials are presented prior to the first reversal, UDTR(73.6%) is initiated from the 11th signal level. (The first reversal is at level 12.) Again it is seen that when UDTR(73.6%) is initiated from a particular level, the magnitude of the bias depends on the number of trials. If only 40 trials are presented after the first reversal, the bias is sizable (0.89 dB), even though the starting level is close to the target. As the run-length is increased, the magnitude of the bias decreases.

This increase in accuracy is accompanied by improved between-runs variability. As sample size increases, the between runs standard deviation decreases by approximately $\sqrt{N}$. It would be desirable to obtain both a small

Table 4-4

Simulated Performance (10 log $\Delta I/I$) for an "Ideal" Listener (WT = 500) when Run-Length is Varied*

Number of Trials After 1st Reversal	Signal Level Tracked	Between-Runs Standard Deviation
40	-13.19	1.52
60	-13.44	1.23
70	-13.59	1.18
80	-13.62	1.11
100	-13.72	1.00
120	-13.83	0.91
140	-13.80	0.82
160	-13.86	0.74
Standard*	-13.87	0.91
Ideal	-14.08	

*See text for details.

bias and between-runs variance in a run of reasonable length. For an 80-trial run, this condition is met by initiating UDTR from the target level. This method is therefore included in Table 4-4 (the "standard" condition) for comparison with the proposed hybrid method. It can be seen that 160 trials are necessary for the hybrid method to yield as small a bias as when UDTR(73.6%) is initiated from the target level. For runs of more practical length (70 or 80 trials), however, there is only a small difference in either bias or between-runs variability when compared to this standard. This difference would be even smaller if the starting level is chosen to be closer to the target.

In summary, if a moderate level of detectability is to be tracked with UDTR, a methodologically bias-free estimate of performance may be obtained if UDTR(70.7%) is employed throughout an 80-trial run, or if it is merely used to locate the region of interest, after which UDTR(73.6%) is implemented for approximately 80 trials.

If other levels of detectability are to be tracked with a descending version of UDTR, it remains important to locate the region of interest without wasting a large number of trials. If a low P(C) is desired, a hybrid version of UDTR could be employed in which the region of interest is located with a decision rule that calls for a level change following each correct or incorrect response (UDTR(50%)). Other decision rules, such as UDTR(57.9%) could then be implemented. When tracking a high P(C), the methodological bias is less since the target level is closer to the starting level (see Table 3-6). A hybrid method may still be employed to minimize this bias. For example, if a performance level of approximately 90 percent correct is desired, UDTR(84.1%) could be used to locate the region of interest (the level is changed following either 4 consecutive correct responses, or a single incorrect response). After the first reversal, UDTR(89.1%) could then be employed.

If ascending runs are used to obtain a psychometric function (i.e., various levels of detectability are tracked), several problems arise which were discussed in Chapter III-B2. These arguments will be reviewed only briefly in this section. Low levels of detectability are difficult to track. If UDTR(57.9%), for instance, is initiated from below the level to be tracked, the initial signal level will be so weak that observers might engage in a considerable amount of guessing. Since initial responding is close to chance, the feedback received might be insufficient to shape the observer's behavior in a run of 80 trials. Tracking high levels of detectability poses a conceptual problem. Since a high target level implies an easily detectable signal, it is difficult to interpret whether performance is the result of the detection strategy used in ascending or descending runs. Performance in ascending conditions may therefore be obtained at only intermediate levels of detectability.

Chapter V

SUMMARY AND CONCLUSIONS

It has been shown that the starting level employed in adaptive testing can greatly influence performance in a noise-intensity discrimination task. In summary, the major findings of this dissertation are:

1) The Up-Down Transformed-Response method, as implemented here and elsewhere, may yield estimates of performance that are biased for purely statistical reasons. Computer simulations indicate that UDTR(73.6%) does not properly track the target percent correct in an 80-trial run.

2) Ascending runs, after correction for the aforementioned bias, yield data that are in accord with the predictions of Green's model.

3) Descending runs yield results which are at variance with the predictions of Green's model. For neither gated nor continuous masker paradigms can a simple energy detection model completely describe the performance of the human observer.

Examination of Table 5-1 reveals that the patterns of results differ for ascending and descending runs.

When ascending runs are employed, the data are in accord with the predictions of Green's model. It may be concluded, therefore, that a

Table 5-1

A Comparison of Results Obtained in These Experiments with Those Reported in the Literature

Predictions of Green's Model		Ideal	Ascending	Descending	Literature(a)
Weber's Law		predicted	no data	no data	generally obtained
IT Reciprocity Factor(b)	Gated	-5	-5.81	-3.17	-3.3
	Cont.	-5	-6.16	-6.39	-6.3
IW Reciprocity Factor(b)	Gated	-5	-4.63	-0.86	-2.4
	Cont.	-5	-4.99	-2.82	-2.1
Range of the Psychometric Function(c)	Gated	7.1	7.36	6.33	7
	Cont.	7.1	7.39	4.51	3
Gated-Continuous Difference(d)		0	0	+3.46	+2.5

(a) These figures represent averages across studies as presented in Tables 1-4 through 1-8.

(b) IT and IW reciprocity factors are given as the change in performance (10 log $\Delta I/I$) as a result of a 10-fold change in duration and bandwidth, respectively.

(c) The range of the psychometric function is reported as the change in $\Delta I/I$ (in dB) necessary for performance to improve from 60% to 90% correct detections.

(d) A positive gated-continuous difference indicates better performance in continuous masker conditions. Gated-continuous differences are reported for wideband, long-duration signals only ($T \geq 30$ msec).

simple energy detection scheme is sufficient to account for performance in this condition.

On the other hand, when staircase runs are initiated with a suprathreshold signal, the pattern of results replicates that found in the literature. Performance is dependent on the mode of the masker, and is often inconsistent with that predicted for an ideal energy detector.

In addition, for all conditions, performance with descending runs is worse than that obtained with ascending runs. This result is counterintuitive in that descending adaptive procedures allow for many practice trials at suprathreshold signal levels, whereas ascending procedures do not. It is generally found that stimulus previewing improves performance in single-interval Yes-No tasks. A number of studies have demonstrated that both pure-tone intensity discrimination and frequency discrimination are aided if a standard is presented on each trial prior to the presentation of either that standard or a comparison stimulus. (See Creelman and Macmillan, in press; Jesteadt and Bilger, 1974; Jesteadt and Sims, 1975; Long, 1973; and Sorkin, 1962; for a discussion of fixed-standard designs.) While a fixed standard may provide a memory aid in a single-interval paradigm, the psychophysical task employed in the studies reported here was two-interval forced-choice intensity discrimination. This paradigm allows for direct stimulus comparisons on every trial. Signal previewing in descending runs, therefore, might not offer any additional advantage in defining the stimulus.

Exposure to suprathreshold signals could, however, interfere with performance if a detection strategy is adopted on the basis of some stimulus attribute that is salient at high signal levels, but which becomes unavailable as the signal level is decreased. Such a nonoptimal detection strategy could account for the failure of descending runs to yield results consistent with the predictions of an energy detection model. It could also account for the fact that non-adaptive procedures have yielded results that are similar to those reported here for descending runs (see for example Green, 1960; and Green and Sewall, 1962). When blocks of fixed-intensity signals are presented, it is common practice to precede each block by a number of practice trials at a suprathreshold signal level.

While it would be interesting to speculate about the nature of a detection strategy that would produce the pattern of results obtained with descending runs, this dissertation was not designed as a test of such matters. The question of why these data are not in accord with the predictions of a simple energy detection model remains unresolved.

Appendix A

TABLES

Table A-1

Weber Fractions (10 log ΔI/I) for Experiment 1 (Descending Runs)

I*	Duration (msec)	Gated	Continuous
		Observer CM	
45	10	-2.94	-2.18
	100	-5.71	-8.02
75	10	-2.58	-1.54
	100	-5.72	-8.01
		Observer RB	
45	10	-2.54	-2.13
	100	-6.17	-8.97
75	10	-2.38	-1.60
	100	-5.52	-8.01
		Means Across Observers and Levels	
	10	-2.61	-1.86
	100	-5.78	-8.25

*Overall level of the masker (dB SPL).

Table A-2

Weber Fractions (10 log $\Delta I/I$) for Experiment 1 (Ascending Runs)

I*	Duration (msec)	Gated	Continuous
		Observer CM	
45	10	-8.22	-7.67
	100	-12.69	-12.89
75	10	-7.41	-7.61
	100	-12.42	-12.62
		Observer RB	
45	10	-7.11	-5.73
	100	-12.24	-11.73
75	10	-5.67	-6.48
	100	-11.31	-11.95
		Means Across Observers and Levels	
	10	-7.10	-6.87
	100	-12.17	-12.29

*Overall level of the masker (dB SPL).

Table A-3

Weber Fractions (10 log $\Delta I/I$) for Experiment 2 (Descending Runs)

Percent Correct	Observer CM		Observer RB	
	Gated	Continuous	Gated	Continuous
57.9	-9.35	-11.12	-8.57	-12.40
73.6	-5.62	-10.00	-3.59	-8.85
89.1	-2.97	-6.89	-0.97	-7.30
	Means Across Observers			
57.9	-8.96	-11.76		
73.6	-4.61	-9.43		
89.1	-1.97	-7.10		

Table A-4

Weber Fractions (10 log $\Delta I/I$ for Experiment 2 (Ascending Runs)

Percent Correct	Observer CM		Observer RB	
	Gated	Continuous	Gated	Continuous
63.6	-14.10	-13.81	-14.08	-13.94
73.6	-12.89	-12.43	-12.15	-12.15
		Means Across Observers		
63.6	-14.09	-13.88		
73.6	-12.52	-12.29		

Table A-5

Weber Fractions (10 log $\Delta I/I$) for Experiment 3 (Descending Runs)

Attenuation Rate	Gated	Continuous
	Observer CM	
24 dB/octave	-5.32	-6.43
96 dB/octave	-4.74	-5.73
	Observer RB	
24 dB/octave	-5.71	-7.18
96 dB/octave	-5.09	-5.13

Table A-6

Weber Fractions (10 log $\Delta I/I$) for Experiment 3 (Ascending Runs)

Attenuation Rate	Gated	Continuous
Observer CM		
24 dB/octave	-10.58	-9.58
96 dB/octave	-8.69	-7.78
Observer RB		
24 dB/octave	-10.30	-8.04
96 dB/octave	-7.95	-8.37

Table A-7

Simulated Performance (10 log $\Delta I/I$) for an "Ideal" Listener in Experiment 2 (WT = 604)*

Starting Level	Percent Correct =	Signal Level Tracked			
		57.9	63.6	73.6	89.1
+4.77		-15.27	-13.50	-12.72	-10.39
+2.83		-15.44	-13.95	-13.12	-10.64
+0.98		-15.56	-14.42	-13.28	-10.72
-0.78		-15.59	-14.81	-13.35	-10.80
-2.48		-15.63	-14.86	-13.47	-10.86
-4.12		-15.72	-14.92	-13.57	-10.92
-5.72		-15.82	-14.99	-13.58	-10.93
-7.30		-15.77	-15.19	-13.64	-10.95
-8.85		-15.85	-15.16	-13.65	-11.09
-10.39		-15.85	-15.15	-13.82	-11.32
-11.92		-15.92	-15.29	-14.01	-11.54
-13.44		-16.03	-15.52	-14.20	-11.68
-14.96		-16.34	-15.89	-14.58	-11.91
-16.47		-16.56	-16.26	-14.96	-12.03
-17.97		-16.85	-16.57	-15.18	-12.10
-19.48		-17.04	-16.78	-15.30	-12.11
Ideal		-19.43	-17.03	-14.47	-11.64

*UDTR tracks 73.6% correct detections.

Table A-8

Simulated Performance (10 log $\Delta I/I$) for an "Ideal" Listener in Experiment 3 (WT = 46.9)*

Starting Level	Signal Level Tracked
+4.77	-7.54
+2.83	-7.68
+0.98	-7.65
-0.78	-7.82
-2.48	-7.92
-4.12	-7.94
-5.72	-8.11
-7.30	-8.44
-8.85	-8.81
-10.39	-9.18
-11.92	-9.62
-13.44	-9.88
-14.96	-10.33
-16.47	-10.51
-17.97	-10.59
-19.48	-10.85
Ideal	-8.55

Table A-9

Corrected Weber Fractions (10 log $\Delta I/I$)
for Experiment 1 (Descending Runs)

I(a)	Duration (msec)	Gated	Continuous
	Observer CM		
45	10	-3.94	-3.18
	100	-7.37	-9.68
75	10	-3.58	-2.54
	100	-7.38	-9.67
	Observer RB		
45	10	-3.54	-3.13
	100	-7.83	-10.63
75	10	-3.38	-2.60
	100	-7.18	-9.67
	Means Across Observers and Levels		
	10	-3.61	-2.86
	100	-7.44	-9.91
	Ideal(b)		
	10	-8.71	-8.71
	100	-14.08	-14.08

(a) Overall level of the masker (dB SPL).

(b) Ideal performance computed from Equation 1-2.

Table A-10

Corrected Weber Fractions (10 log $\Delta I/I$) for Experiment 1 (Ascending Runs)

I(a)	Duration (msec)	Gated	Continuous
	Observer CM		
45	10	-6.48	-5.93
	100	-11.69	-11.89
75	10	-5.67	-5.87
	100	-11.42	-11.62
	Observer RB		
45	10	-5.37	-3.99
	100	-11.24	-10.73
75	10	-3.93	-4.74
	100	-10.31	-10.95
	Means Across Observers and Levels		
	10	-5.36	-5.13
	100	-11.17	-11.29
	Ideal(b)		
	10	-8.71	-8.71
	100	-14.08	-14.08

(a) Overall level of the masker (dB SPL).

(b) Ideal performance computed from Equation 1-2.

Table A-11

Corrected Weber Fractions (10 log ΔI/I)
for Experiment 2 (Descending Runs)

Percent Correct	Observer CM		Observer RB	
	Gated	Continuous	Gated	Continuous
57.9	-13.51	-14.96	-12.73	-16.24
73.6	-7.37	-11.12	-5.34	-9.97
89.1	-4.22	-7.73	-2.22	-8.14
	Means Across Observers			
57.9	-13.12	-15.60		
73.6	-6.36	-10.55		
89.1	-3.22	-7.94		
	Ideal*			
57.9	-19.43	-19.43		
73.6	-14.47	-14.47		
89.1	-11.64	-11.64		

*Ideal performance computed from Equation 1-2.

Table A-12

Corrected Weber Fractions (10 log $\Delta I/I$)
for Experiment 2 (Ascending Runs)

Percent Correct	Observer CM Gated	Observer CM Continuous	Observer RB Gated	Observer RB Continuous
63.6	-14.35	-14.06	-14.33	-14.19
73.6	-12.06	-11.60	-11.32	-11.32
	Means Across Observers			
63.6	-14.34	-14.13		
73.6	-11.69	-11.46		
	Ideal*			
63.6	-17.03	-17.03		
73.6	-14.47	-14.47		

*Ideal performance computed from Equation 1-2.

Table A-13

Corrected Weber Fractions (10 log $\Delta I/I$) for Experiment 3 (Descending Runs)

Attenuation Rate	Gated	Continuous
Observer CM		
24 dB/octave	-6.33	-7.44
96 dB/octave	-5.75	-6.74
Observer RB		
24 dB/octave	-6.72	-8.19
96 dB/octave	-6.10	-6.14
Ideal*	-8.55	-8.55

*Ideal performance computed from Equation 1-2.

Table A-14

Corrected Weber Fractions (10 log $\Delta I/I$)
For Experiment 3 (Ascending Runs)

Attenuation Rate	Gated	Continuous
	Observer CM	
24 dB/octave	-8.80	-7.80
96 dB/octave	-6.91	-6.00
	Observer RB	
24 dB/octave	-8.52	-6.26
96 dB/octave	-6.17	-6.59
Ideal*	-8.55	-8.55

*Ideal performance computed from Equation 1-2.

Appendix B

RANDOM NUMBER GENERATION

Random numbers (in the range 0 to 1.0 inclusive) were generated from a rectangular distribution provided by the IMSL (International Mathematical and Statistical Library) subroutine, GGUB, which employs the following multiplicative congruential method:

$$\text{Random number} = 7^5 \cdot \text{Seed} \pmod{2^{31} - 1} .$$

The seed is an integer in the exclusive range (1, 2147483647), and is replaced by the program for each new call.

__Example__

Seed = 123457

Random number = $7^5 \cdot 123457 \pmod{2^{31} - 1}$

Random number = .96622

Further discussion of the multiplicative congruential method for random number generation is given by Taussky and Todd (1954), and Schreider (1966).

Appendix C

DETECTABILITY OF NARROWBAND NOISE

The Weber fractions reported by Schacknow and Raab (1976) for 100-Hz bandwidth conditions are excluded from the discussion in Chapter III-B1 since they are difficult to interpret. For masker spectrum levels of 25 and 45 dB SPL, Weber fractions with gated maskers are about 1-dB better than predicted for an ideal receiver; with continuous maskers, performance is approximately 2.5-dB better than ideal.

For Green's energy-detection model, performance is limited solely by the stimulus fluctuations inherent in the noise stimuli. The derivation of this model, and its resulting predictions, are for Gaussian noise signals having rectangular power spectra. A direct comparison of real and ideal performance is, therefore, not straightforward if the noise stimuli do not have rectangular power spectra.

In order to compare real and ideal performance, the equivalent statistical bandwidth (ESB) of a nonrectangular noise stimulus must be calculated. The ESB is the bandwidth of the rectangular filter which would pass a white noise signal with the same relative variability as would the actual filter. Schacknow and Raab (1976) calculated the ESB for each of the noise stimuli employed in their study. In all cases, the ESB is larger than the corresponding half-power bandwidth (HPB), although this discrepancy diminishes with increasing bandwidth. For their narrowest stimulus (HPB = 100 Hz), the ESB equals 136 Hz. This increase in effective bandwidth results in an approximately 1.4-dB improvement in the performance of the ideal detector, over that calculated on the basis of the HPB. For wider bandwidths, the difference between the HPB and ESB is smaller. For example, the difference in theoretical thresholds for the ideal receiver when HPB = 3160 Hz and ESB = 3399 Hz is only about 0.1 dB. It follows, therefore, that if data from real observers were compared to ideal Weber fractions calculated using the half-power bandwidth, real-ideal differences will be artificially smaller for narrowband stimuli.

There is less need for these computations in analyzing our data, since the narrowest half-power bandwidth employed in this thesis was 469 Hz. From Schacknow and Raab's Table IV (p. 900), it can be seen that the difference between ideal Weber fractions calculated using the HPB and those using the ESB is about 0.5 dB for a bandwidth of 469 Hz and about 0.1 dB for a 5000-Hz bandwidth. Furthermore, ESB's calculated by Schacknow and Raab were for noise stimuli filtered with attenuation rates of 48 dB/octave. In one condition, our narrowband noise stimuli had attenuation rates of 96 dB/octave. These stimuli are more nearly rectangular than those of Schacknow and Raab, and will consequently be less affected by the HPB-ESB conversion. The real-ideal differences in performance reported in Table 3-9 for 100 msec stimuli with 96 dB/octave skirts show that observers perform better, relative to ideal, with narrowband stimuli than with wideband stimuli. The magnitude of this difference is too large (about 3 dB) to be accounted for by the use of the HPB in calculating the ideal Weber fraction from Equation 1-2. Since the distinction between half-power bandwidth and equivalent statistical bandwidth is small for the stimuli employed in this dissertation (in the most extreme case, less than 0.5 dB), all references to bandwidth will be to the half-power bandwidth.

BIBILIOGRAPHY

Berner, C.N. Unpublished data provided by D.H. Raab.

Blackman, R.B., & Tukey, J.W. *The Meaurement of Power Spectra*, New York: Dover, 1958.

de Boer, E. *Detection function for noise bursts*. Paper F26 presented at 5th International Congress on Acoustics, Liege, 1965.

Bos, C.E., & de Boer, E. Masking and discrimination. *Journal of the Acoustical Society of America*, 1966, 39, 708-715.

Campbell, R.A. Detection of a noise signal of varying duration. *Journal of the Acoustical Society of America*, 1963, 35, 1732-1737.

Campbell, R.A. Masker level and noise-signal detection. *Journal of the Acoustical Society of America*, 1964, 36, 570-575.

Campbell, R.A. Feedback and noise-signal detection at three performance levels. *Journal of the Acoustical Society of America*, 1965, 37, 434-438.

Campbell, R.A. Auditory intensity perception and neural coding, *Journal of the Acoustical Society of America*, 1966, 39, 1030-1033.

Campbell, R.A. BUDTIF thresholds: Variability and bias. *Journal of the Acoustical Society of America*, 1967, 42, 1149 (A).

Campbell, R.A. Thresholds re duration and levels of a continuous or gated masker. *Journal of the Acoustical Society of America*, 1969, 46, 895-897.

Campbell, R.A., & Counter, S.A. Temporal energy integration and periodicity pitch. *Journal of the Acoustical Society of America*, 1969, 45, 691-693.

Campbell, R.A., Hutton, C.L., & Stuckey, C.W. Computerized investigation of threshold-decision phenomenon. Journal of the Acoustical Society of America, 1966, 39, 1034-1036.

Campbell, R.A., & Lasky, E.Z. Masker level and sinusoidal-signal detection. Journal of the Acoustical Society of America, 1967, 42, 972-976.

Campbell, R.A., & Lasky, E.Z. Adaptive threshold procedures: BUDTIF. Journal of the Acoustical Society of America, 1968, 44, 537-541.

Cornsweet, T.N. The staircase method in psychophysics. American Journal of Psychology, 1962, 485-491.

Cornsweet, T.N., & Pinsker, H.M. Luminance discrimination of brief flashes under various conditions of adaptation. Journal of Physiology, 1965, 176, 294-310.

Creelman, C.D. Detection, discrimination, and the loudness of short tones. Journal of the Acoustical Society of America, 1963, 35, 1201-1203.

Creelman, C.D., & Macmillan, N.A. Auditory phase and frequency discrimination: A comparison of nine procedures. Journal of Experimental Psychology, in press.

Creelman, C.D., & Taylor, M.M. Some pitfalls in adaptive testing: Comments on "Temporal Integration and Periodicity Pitch". Journal of the Acoustical Society of America, 1969, 46, 1581-1582.

Dirks, D.D., & Norris, J.C. Shifts in auditory thresholds produced by ipsilateral and contralateral maskers at low-intensity levels. Journal of the Acoustical Society of America, 1966, 40, 12-19.

Dimmick, F.L., & Olson, R.M. The intensive difference limen in audition. Journal of the Acoustical Society of America, 1941, 12, 517-525.

Dixon, W.J., & Mood, A.M. A method for obtaining and analyzing sensitivity data. Journal of the American Statistical Association, 1948, 43, 109-126.

Feeny, S., Kaiser, P.K., & Thomas, J.P. An analysis of data gathered by the staircase method. American Journal of Psychology, 1966, 79, 652-654.

Green, D.M. Auditory detection of a noise signal. Journal of the Acoustical Society of America, 1960, 32, 121-131.

Green, D.M. Psychoacoustics and detection theory. Journal of the Acoustical Society of America, 1960(a), 32, 1189-1203.

Green, D.M. Continuous versus gated masking for noise and sinusoidal maskers. Journal of the Acoustical Society of America, 1964, 36, 2009 (A).

Green, D.M. Masking with continuous and pulsed sinusoids. Journal of the Acoustical Society of America, 1969, 46, 939-946.

Green, D.M., McKey, M.J., & Licklider, J.C.R. Detection of a pulsed sinusiod in noise as a function of frequency. Journal of the Acoustical Society of America, 1959, 31, 1446-1452.

Green, D.M., & Sewall, S.T. Effects of background noise on auditory detection of noise bursts. Journal of the Acoustical Society of America, 1962, 34, 1207-1216.

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

Harris, J.D. The effect of sensation-levels on intensive discrimination of noise. Journal of the Acoustical Society of America, 1950, 63, 409-421.

Harris, J.D. Loudness discrimination. Journal of Speech and Hearing Disorders, 1963, (Monograph Supplement Number 11).

Heinemann, E.G. The relation of apparent brightness to the threshold for differences in luminance. Journal of Experimental Psychology, 1961, 61, 389-399.

Jesteadt, W., & Bilger, R.C. Intensity and frequency discrimination in one- and two-interval paradigms. Journal of the Acoustical Society of America, 1974, 55, 1266-1276.

Jesteadt, W., & Sims, S.L. Decision processes in frequency discrimination. Journal of the Acoustical Society of America, 1975, 57, 1161-1168.

Jesteadt, W., Wier, C.C., & Green, D.M. Intensity discrimination as a function of frequency and sensation level. Journal of the Acoustical Society of America, 1977, 61, 169-177.

Leshowitz, B.F., & Cudahy, E. Masking patterns for continuous and gated sinusoids. Journal of the Acoustical Society of America, 1975, 58, 235-242.

Leshowitz, B., Taub, H.B., & Raab, D.H. Visual detection of signals in the presence of continuous and pulsed backgrounds. Perception and Psychophysics, 1968, 4, 207-213.

Levitt, H.A. Transformed up-down methods in psychophysics. Journal of the Acoustical Society of America, 1971, 49, 467-477.

Levitt, H., & Bock, D.E. Sequential programmer for psychophysical testing. Journal of the Acoustical Society of America, 1967, 42, 911-913.

Levitt, H., & Treisman, M. Control charts for sequential testing. Psychometrika, 1969, 34, 509-518.

Long, G.R. The role of the standard in an auditory amplitude discrimination experiment. Perception and Psychophysics, 1973, 13, 49-59.

Macmillan, N.A. Detection and recognition of intensity changes in tone and noise: The detection-recognition disparity. Perception and Psychophysics, 1973, 13, 65-75.

Macmillan, N.A. Intensity discrimination of continuous noise in five paradigms. Paper presented at the 87th meeting of the Acoustical Society of America, 1974.

Matin, L. & Kornheiser, A. Increment thresholds against pulsed and steady backgrounds. Perception and Psychophysics, 1976, 20, 250-266.

McGill, W.J., & Goldberg, J.P. Pure-tone intensity discrimination and energy detection. Journal of the Acoustical Society of America, 1968a, 44, 576-581.

McGill, W.J. & Goldberg, J.P. A study of the mear-miss involving Weber's Law and pure-tone intensity discrimination. Perception and Psychophysics, 1968b, 44, 105-109.

Miller, G.A. Sensitivity to changes in the intensity of white noise and its relation to masking and loudness. Journal of the Acoustical Society of America, 1947, 19, 609-619.

Moore, B.C.J. Mechanisms of masking. Journal of the Acoustical Society of America, 1975, 57, 391-399.

Moore, B.C.J., & Raab, D.H. Intensity discrimination for noise bursts in the presence of a continuous, bandstop background: Effects of level, width of the bandstop, and duration. Journal of the Acoustical Society of America, 1975, 57, 400-405.

National Bureau of Standards <u>Tables of Normal Probability Functions</u>, 1953, U.S. Government Printing Office, Wsshington, D.C.

Penner, M.J., Leshowitz, B., Cudahy, E., & Ricard, G. Intensity discrimination for pulsed sinusoids of various frequencies. <u>Perception and Psychophysics</u>, 1974, 15, 568-570.

Peterson, W.W., Birdsall, T.G., & Fox, W.C. The theory of signal detectability. In Luce, R., Bush, R., & Galanter, E. (Eds.), <u>Readings in Mathematical Psychology</u> (Vol. 1). New York: Wiley, 1963, 167-211.

Pollack, I. Sensitivity to differences in intensity between repeated bursts of noise. <u>Journal of the Acoustical Society of America</u>, 1951, 23, 650-653.

Pollack, I. Methodological examination of the PEST procedure. <u>Perception and Psychophysics</u>, 1968, 3, 283-289.

Postman, L. The time-error in auditory perception. <u>American Journal of Psychology</u>, 1946, 59, 193-219.

Raab, D.H., & Goldberg, I.A. Auditory intensity discrimination with bursts of reproducible noise. <u>Journal of the Acoustical Society of America</u>, 1975, 57, 437-447.

Raab, D.H., Osman, E., & Rich, E. Effects of waveform correlation and signal duration on detection of noise bursts in continuous noise. <u>Journal of the Acoustical Society of America</u>, 1963, 35, 1942-1946.

Rendleman, P., Rose, R.M., & Teller, D.Y. Statistical properties of Pollack's PEST procedure. <u>Perception and Psychophysics</u>, 1970, 9, 208-212.

Rochester, S. Detection and duration discrimination of noise increments. Journal of the Acoustical Society of America, 1971, 49, 1783-1789.

Rose, R.M., Teller, D.Y., & Rendleman, P. Statistical properties of staircase estimates. Perception and Psychophysics, 1970, 8, 199-204.

Schacknow, P.N., & Raab, D.H. Intensity discrimination of tone bursts and the form of the Weber function. Perception and Psychophysics, 1973, 14, 449-450.

Schacknow, P.N., & Raab, D.H. Noise-intensity discrimination: Effects of bandwidth conditions and mode of masker presentation. Journal of the Acoustical Society of America, 1976, 60, 893-905.

Shannon, C.E. Communication in the presence of noise. Proc. IRE, January, 1949, 37, 10-21.

Sherrick, C.E., & Albernaz, P.L.M. Auditory threshold shifts produced by simultaneously pulsed contralateral stimuli. Journal of the Acoustical Society of America, 1961, 33, 1381-1385.

Shreider, Y.A. The Monte Carlo Method. Oxford: Pergamon Press, 1966, pp. 260-263.

Small, A.M., Bacon, W.E., & Fozard, J.L. Intensive differential thresholds for octave-band noise. Journal of the Acoustical Society of America, 1959, 31 508-510.

Smith, J. Stimulus programming in psychophysics. Psychometrika, 1961, 26, 27-31.

Sorkin, R.D. Extension of the theory of signal detectability to matching procedures in psychoacoustics. Journal of the Acoustical Society of America, 1962, 34, 1745-1751.

Stuckey, C.W., Hutton, C.L., & Campbell, R.A. Decision rules in threshold determination. Journal of the Acoustical Society of America, 1966, 40, 1174-1179.

Taussky, O., & Todd, Generation of pseudo random numbers. In Symposium on Monte Carlo Methods, New York: Wiley, 1954, pp. 15-28.

Taylor, M.M., & Creelman, C.D. PEST: Efficient estimates on probability functions. Journal of the Acoustical Society of America, 1967, 41, 782-787.

Tucker, A., Williams, P.I., & Jeffress, L.A. Effects of signal duration on detection for gated and continuous noise. Journal of the Acoustical Society of America, 1965, 44, 813-816.

Viemeister, N.F. Intensity discrimination of pulsed sinusoids: The effects of filtered noise. Journal of the Acoustical Society of America, 1972, 51, 1265-1269.

Viemeister, N.F. Intensity discrimination of noise in the presence of band-reject noise. Journal of the Acoustical Society of America, 1974a, 56, 1594-1600.

Viemeister, N.F. Detection of a noise signal: the effects of duration and masker level. Paper presented at the 88th meeting of the Acoustical Society of America, 1974b.

Wald, A. Sequential Analysis, New York: John Wiley & Sons, 1947, esp. pp. 88-105 and 196-199.

Wetherill, G.B. Sequential estimation of quantal response curves. Journal of the Royal Statistical Society, 1963, Series B, 25, 1-48.

Wetherill, G.B., & Levitt, H. Sequential estimation of points on a psychometric function. British Journal of Mathematical and Statistical Psychology, 1965, 18, 1-10.

Wier, C.C., Green, D.M., Hafter, E.R., & Burkhardt, S. Detection of a tone burst in continuous- and gated-noise maskers; defects of signal frequency, duration, and masker level. Journal of the Acoustical Society of America, 1977, 61, 1298-1300.

Wightman, F.L., & Green, D.M. Effects of duration on detection of a noise signal. Journal of the Acoustical Society of America, 1966, 40, 1250 (A).

Zwicker, E. & Fastl, M. On the development of the critical band. Journal of the Acoustical Society of America, 1972, 52, 699-702.