COMBINING SENSORY INFORMATION FROM TWO SEPARATE CRITICAL BANDS

MARTIN ERHARD RICKERT

COMBINING SENSORY INFORMATION
FROM TWO SEPARATE
CRITICAL BANDS

by

© Martin Erhard Rickert, *B.Sc.*

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

Department of Psychology
Memorial University of Newfoundland
August 1987

St. John's *Newfoundland*

ISBN 0-315-45077-0

ABSTRACT

This study investigated how subjects combine auditory signals and noise from two separate frequency-selective channels, or *critical bands*. In the first experiment, listeners were trained to detect a 0.5 *kHz* sinusoidal tone in a continuous background of noise. Then, without informing listeners, a 1.3 *kHz* tone was added to the signal. Initially the detectability of the two-tone complex was no better than the detectability of the 0.5 *kHz* signal, but it improved in subsequent sessions. Following this, the 1.3 *kHz* signal was removed, and performance dropped to, or below, the original level. These results indicated that human listeners have the ability to use either a single- *or* multiple-band listening strategy. In order to determine more about how information from widely separated critical bands is combined, two mathematical models were considered. The *information integration* model (Green, 1956) assumes that signal and noise energy from individual frequency-selective channels is statistically combined prior to decision. By contrast, the *decision threshold* model (Schafer and Gales, 1949) postulates that independent decisions are made about the information in each channel, and that the outcomes of the decisions are combined according to an overall decision rule. Data from the first experiment are consistent with the information integration model, but are contrary to predictions made using the decision threshold model. In order to further distinguish between these models; a second, more sensitive, experiment was designed. In this experiment the 0.5 *kHz*, 1.3 *kHz*, and combined signals were randomly mixed within the same block of trials. Therefore, subjects presumably used a similar listening strategy for each of the signals. Data from this experiment fit the neither the information integration model nor the decision threshold model. Furthermore, comparisons of detection performance between experiments argue against each of these models. In conclusion, while this study shows that neither the information integration model nor the decision threshold model explains how listeners combine information from two critical bands, it demonstrates that human listeners have the ability to switch from using one critical band to using two.

ACKNOWLEDGEMENTS

First, I would like to gratefully acknowledge the encouragement, and unhesitating assistance provided by my supervisor Michael A. Zagorski. Throughout the course of this study, his "sound" advice, which has often extended beyond the domain of science, has proven invaluable. I sincerely hope that he will remain my mentor and trusted friend in years to come. I'd like also to thank the members of my thesis committee, Michael Bruce-Lockhart and Russ Adams for critically reading, and commenting on, an earlier version of this thesis. To MPBL, I extend a further and most deserved thanks. His guidance and support has been, and will continue to be, greatly appreciated.

This thesis is dedicated to the "gang" at home who have remained a constant source of inspiration. The completion of this project would not have been possible without their unconditional support.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The main advantage introduced by signal detection theory (Peterson, Birdsall, and Fox, 1954) to auditory psychophysics, is that it provides a framework for distinguishing between sensory and decision processes. This framework includes empirical and analytic methods that permit listeners' sensitivity to be estimated independent of non-sensory variables. Consequently, signal detection methods have been used to address many substantive issues about auditory signal processing.

One issue that has received considerable attention, is how listeners utilize information from different parts of the auditory spectrum when detecting signals in noise. Evidence suggests that listeners can adopt a listening "strategy" that is suited to the requirements of a specific task (Swets, 1963, 1984). Thus if information is restricted to a relatively narrow part of the spectrum, listeners have the ability to process that information while ignoring the rest. If, however, information arrives in several spectral regions, there are a number of ways in which it can be processed.

Presented here are two experiments. The first investigates a subjects' ability to make transitions in listening strategy when the signal changes unexpectedly; the second investigates performance when the same strategy must be used for different signals. In addition, these experiments test various assumptions about how listeners combine auditory information in order to make a decision concerning the presence or absence of a brief tonal signal embedded in noise.

A. Auditory filtering: *The critical band*

Unlike the visual system, which integrates energy at different spectral frequencies to provide the sensation of a single color, the auditory system allows us to resolve, or separate, individual spectral components of a complex sound (Green and Swets, 1974). This ability stems primarily from the fact that the auditory stimulus is *filtered* at various levels of processing. For example, physiological experiments conducted by von Bekesy (1961b) show that for a pure sound of a given frequency, there is a region of maximum vibratory response along the basilar membrane (BM). The mechanical displacement of the BM is then converted to a distribution of neural activity across the cochlear fibers (Sachs and Kiang, 1968). However, electrophysiological measurements indicate that this excitation pattern is not perfectly correlated with mechanical activity in the cochlea. The fact that displacement patterns along the BM do not match the frequency selectivity exhibited by the discharge characteristics of the VIIIth nerve has led to the hypothesis that a "second" filter exists between the basilar membrane and the initial segment of the auditory nerves (Evans and Wilson, 1973).

Psychophysical studies have also provided strong evidence for the existence of auditory filtering, and, have been used to determine some important filter characteristics, including bandwidth and shape. The width of the assumed auditory filter was first investigated by Fletcher in 1940. In this study Fletcher systematically varied the bandwidth of a white noise (i.e., constant power per unit bandwidth) that was centered on a sinusoidal target tone. At first, the signal level of the tone was set so that it was just detectable in a wideband noise masker. As the bandwidth of the noise was progressively narrowed, detectability of the tone remained unaffected until the noise band became smaller than a certain value, called the *critical bandwidth*. For noise bandwidths smaller than this, performance improved for the same signal level.

On the basis of results obtained by using tonal signals of different frequencies, Fletcher concluded that only a narrow band of noise around the signal frequency is effective as a masking stimulus. Stated otherwise, masking energy outside a critical band does not contribute to the masking process but only adds to the loudness of the noise. It was further argued that when a tone was just detectable in the noise band, signal and noise energy within the critical band were equivalent.

Later, in a refinement of the critical band study, Shafer, Gales, Shewmaker, and Thompson (1950) attempted to estimate the actual shape of the auditory filter. They used noise with a rectangular spectrum which was constructed by adding a number of closely spaced sinusoids, and, in comparison to Fletcher, took many more measurements of detectability as the bandwidth was narrowed. Their results failed to indicate a sharp discontinuity at the expected critical value, which implied that the auditory filter was not rectangular. Rather, Shafer, et al (1950) inferred that the shape more closely resembled that of a "single-tuned," or "universal-resonance" filter.

Another procedure, the probe-signal technique, has been used to obtain measurements of the frequency-response characteristics of the auditory filter (Greenberg and Larkin, 1968; Greenberg, 1969a, 1969b). This method assumes that observers are extensively trained to detect a near-threshold signal of a single, specified frequency. This signal is presented with high *a priori* probability of occurrence. Then, on some small proportion of trials various other "unexpected" frequencies, probe-signals, are presented in lieu of the main signal and, in the vicinity of the assumed filter. Any change in the level of detectability for probe-signals relative to the main signal is interpreted as reflecting both the filter shape and bandwidth. Note that the analysis of detection performance is based on the sensitivity index d' and assumes that observers center a critical band on the main signal. The underlying rational is that if observers attenuate masking noise outside the critical band, then signals should also be attenuated.

One general conclusion made on the basis of data obtained in "simultaneous" masking conditions (i.e., signal and masker presented at the same time), is that detectability depends upon the ratio of the signal energy to noise energy within a single critical band (Weber, 1978; Green, 1983). These energies are related to both the power spectrum of the masking noise and the weighting, or transfer function of the auditory filter (Patterson and

Nimmo-Smith, 1980). Nevertheless, other masking conditions yield results which indicate that the filtering process is somewhat more complicated.

For example, Glasberg, Moore, and Smith (1984) masked a 1 *kHz* sinusoid of fixed signal level using (1) a noise with a spectral notch of variable width, (2) two tones with variable frequency separation, and (3) noise with a sinusoidally rippled spectrum with variable density. In each case, the masking spectrum was symmetrically placed with respect to signal. The level of the masking stimulus was varied in order to determine threshold. Glasberg, et al then compared the auditory filter shapes that were derived under both simultaneous and "forward" masking conditions (i.e., the masking stimulus is presented immediately AFTER the signal). For simultaneous maskers, the characteristics of the filters were found to be similar for notched and two-tone masking but slightly broader for rippled noise. In contrast, were the results from the forward masking condition which consistently revealed sharper bandwidths. This finding is compatible with an edge-enhancement mechanism of frequency selectivity produced by suppression (Houtgast, 1974; see Plomp, 1976).

Based on available psychophysical evidence, others have generated mathematical formulae which describe various properties of auditory filtering. Zwicker and Terhardt (1980) derived an analytic expression for the critical bandwidth as a function of center frequency. Their proposed formula assumes that the critical bandwidth is approximately constant at low frequencies (below 0.5 *kHz*), but increases with the logarithm of frequency at higher frequencies. Moore and Glasberg (1983) argue this point and provide data which show that the bandwidth does decrease at center frequencies below 0.5 *kHz*. In addition, these authors outline a method for calculating excitation patterns from a simplified filter shape (i.e., equivalent rectangular bandwidth). Likewise, Patterson (1974) reports an expression that describes the filter characteristics of bandwidth, center-frequency to bandwidth ratio (Q), shape and attenuation of the filter skirts (1974).

An important point to be made here is that the auditory system is designed in a way that it allows us to analyze a complex time-varying sound into its individual frequency components. Indeed, the process of *filtering* has been clearly demonstrated both at the physiological and psychophysical level. It is equally clear, however, that certain parameters of this process are not fixed. That is, depending on the requirements of the task, listeners can apparently adjust the bandwidth, frequency stability, spectral location and the number of bands used (Swets, 1963).

In summary, most current models of the auditory system that have been proposed to account for listeners' ability to detect signals in noise assume that initial filtering is a consequence of peripheral processing mechanisms, and that this provides the conceptual basis for the critical band. Because the critical bandwidth has been clearly established as an empirical phenomenon (see Sharf, 1971 for review), it will be used in the remaining discussions as an explanatory construct. Further, we will assume that the auditory stimulus is presented to a bank of critical band filters.

whose output the listener uses to perform the task.

B. Single- and multiple-band listening

In a typical auditory detection experiment listeners are asked to detect a signal of known frequency and duration in the presence of wideband Gaussian noise. During a well-defined observation interval $(0,T)$ one of two possible waveforms is presented with fixed *a priori* probability: either $x(t)=n(t)$, or signal-plus-noise $x(t)=s(t)+n(t)$, where $0 \leq t \leq T$. Using a critical band interpretation, the observer presumably centers a single filter on that part of the spectrum which best represents the signal, thereby maximizing the signal-to-noise ratio. In particular, according to an energy-detection model proposed by Pfafflin and Mathews (1962), the output of the filter $x(t)$ is squared and a running integral is performed on the squared output. The squared output is instantaneous power and the output of the integrator is an estimate of energy in the integration period (which is on the order of 100–200 ms for humans). In effect, this energy is represented by a single quantity $\Omega = \int_0^T x^2(t)\,dt$ which is the prime determinant of detectability within a critical band. An important question is how listeners process signals which contain spectral components that are separated by more than a single critical bandwidth. Consider the following models.

First, according to the *single-band* model of auditory processing (Tanner, et al, 1956) a decision concerning the presence or absence of a signal in background noise is governed by the information available in relatively narrow spectral bandwidth. Stated otherwise, regardless of the number of components and the bandwidth of the signal, it is assumed that detection is based on only one critical band. This model goes beyond Fletcher's earlier hypothesis by asserting that the center frequency (i.e., spectral location) of the band can be adjusted by the listener. Thus, for a wideband signal, the listener must scan the spectrum by continuously adjusting the center frequency of the filter.

By contrast, the *multiple-band* energy-detection model (Green, 1958) assumes that more than one listening band may be used at any given instant. According to this model, it is possible to select both the number and the center frequencies of the critical bands such that a separate critical band is positioned on each signal component. Furthermore, if components fall into different bands, the detection process common to each band would occur. Two crucial assumptions are:

(1) that the combination of the filter outputs is linear, and

(2) that an optimal weighting of the channels is used. That is, the channel with the more detectable signal is given more weight in the decision. This model is discussed further below.

Three modified versions of the basic detection paradigm, as described above, have been used to examine the main assumptions of both single- and multiple-band views. As Swets (1984) argues, the contradictory results obtained by these methods can be reconciled by assuming that the observer can select *either* mode of processing. Moreover, it is apparent that selection of a specific strategy depends heavily on the requirements of the task. Because comprehensive reviews are available elsewhere (see de Boer, 1966; Green and Swets, 1974; Swets, 1984), the following discussion is restricted to those experiments on which the methodology used in the present study is based.

1. Detection of unexpected, low-probability signals

This version of the basic method deals with the detection of an unexpected signal frequency and was first used in a study reported by Tanner, et al (1956). In the initial part of their experiment, unpracticed listeners were trained to detect a 1 *kHz* signal, presented during a well defined observation interval, in a continuous background of approximately Gaussian noise. Then, after several sessions, the signal was changed to a 1.3 *kHz* signal of the same energy as the original signal. They found that the probability of a correct response, $P(C)$, for this new signal dropped to a value which represented chance performance. However in subsequent sessions, after subjects were given a preview of the 1.3 *kHz* signal without background noise, performance was restored to previous levels.

In a related study, Karoly and Isaacson (unpublished, see Swets, 1984) obtained similar results using an expected 1 *kHz* signal and *two* unexpected signals of 0.5 *kHz* and 1.5 *kHz*. On randomly selected trials the unexpected frequencies were presented with low *a priori* probability of occurrence. Karoly, et al found that listeners' $P(C)$ for each of the unexpected signals was significantly lower than was $P(C)$ for the expected signal. Decrements in performance for an unexpected signal are logically consistent with both single- and multiple-band views because information in other spectral locations is presumably attenuated, or tuned-out.

2. Detection of multicomponent signals

A second version of the basic detection paradigm requires that listeners detect a multicomponent signal. In this condition, the target signal is a complex tone consisting of two or more sinusoidal components simultaneously presented in masking noise. In an early study, Schafer, et al (1949) obtained data on the detectability of two-, four-, and eight-component tones in noise. They found that, for two signals separated by more than a critical band, the energy of each of the components could be reduced by $0-2\,dB$ and detectability of the complex signal would be maintained. For stimuli consisting of 4 to 8 sinusoids, components could be reduced by $0-3\,dB$.

Green (1958) also examined the detection of single and multicomponent signals masked by noise. Four frequencies were used in this study: $0.5\,kHz$, $1\,kHz$, $1.823\,kHz$, and $2\,kHz$. Six two-tone complexes were then constructed from all possible pairings of the single components. Including the individual components, Green presented a total of ten signals to listeners. For each signal type, $P(C)$ was estimated from four 100 trial blocks. Each signal type was used for one complete block of trials before a second block was conducted for any signal. Green's data indicated that detection performance was better for the tonal complexes than for any of the individual components. Furthermore, the data were best explained by a model which assumed that auditory signals and noise were statistically combined prior to decision. The implication of these results is that, in some instances, complex signals are no more detectable than a single component, while in other instances detection performance is aided by the presence of signal energy in more than one critical band.

3. Uncertain frequency detection

In this type of detection experiment listeners are asked to detect a signal that is equally likely to be one-of-M orthogonal frequencies (Nolte and Jaarsma, 1967). That is, only one frequency, selected at random from a specified set of frequencies, occurs on any given trial. It is assumed that the signal waveforms $s(t)$ have equal energy and are orthogonal over the observation interval. Thus,

$$\int_0^T s_i(t)s_j(t)\,dt = \begin{cases} E, & i=j \\ 0, & i \neq j, \quad i = 1, 2, ..., M \end{cases} \tag{1}$$

where E is the energy of each signal in $(0, T)$. Consider the simplest case, for which the set of potential signals consists of only two potential frequencies. So, the uncertainty parameter, M, is equal to 2. The problem faced by the observer is to detect the presence or absence of either frequency regardless of which has been presented. Clearly, if a listener uses a single-band strategy, then on any given trial that band can be focused on only one part of the spectrum. Thus, a signal presented in a different, remote part of the spectrum, would not be detected. By contrast, if a multiple-band strategy is used, then a listener presumably centers a separate critical band in the vicinity of each of the two specified frequencies. However, detectability for a single frequency would be lower relative to detectability for that same component when only one band is used owing to the fact there is an increase in total variance in the decision process due to noise power contained in the additional band.

Although there is sufficient experimental evidence to show that listeners often utilize the output of a single critical band, there is also evidence to show that listeners can employ more than one critical band. Assuming that both listening strategies are available, the experiments described below represent an attempt to learn more about how listeners switch from using one strategy to another, and also about the manner by which information from two widely separated critical bands is combined. The first experiment examines detection performance in a condition that allows subjects to adopt different listening strategies.

II. EXPERIMENT 1. *Blocks using the same signal throughout*

Greenberg, et al (1968) showed that when the frequency separation between an expected and unexpected signal is sufficient ($150-200\,Hz$), listeners are unable to distinguish the novel signal from background

Gaussian noise. Likewise, MacMillan and Schwartz (1975) found that observers could listen for two widely spaced signals while ignoring frequencies between the primary signals. The argument put forth by these authors is that subjects position an auditory listening band at high-probability frequencies and, in effect, attenuate signals and noise occurring in other parts of the spectrum.

The goal in the initial part of Experiment 1 is to determine whether the observer is listening to only a narrow band of the spectrum. This would support critical band theory and, more specifically, the single-band model. The question addressed is what happens to detection performance when the signal is shifted from a single sinusoidal in isolation to a combined signal which includes both the old, expected component and the new, unexpected component. In particular, two frequencies ($0.5\,kHz$ and $1.3\,kHz$) are used that, according to current estimates of the critical bandwidth, excite two non-overlapping filters.

Consider the following. If listeners use the output of a single filter centered on the expected component to form a decision variable, then our *a priori* expectation is that a novel component should not affect detection performance. As a result, the detectability index d' for the combined signal, d'_{com}, should be equal to the d' obtained for the single component. That is, dependent on which signal is expected

$$d'_{com} = \{ d'_l \text{ or } d'_h \} \qquad (2)$$

where d'_l is the d' for the 0.5 kHz signal and d'_h is the d' for the 1.3 kHz signal. As discussed earlier, this is Marill's (1956) conclusion: the detectability of a two-component signal is no more detectable than the most detectable member of the pair. However, even though listeners might continue to focus on only one critical band and not pick up the second, unexpected component, does not preclude the possibility that, given subsequent experience on the combined signal, they can begin to utilize the additional signal component. Therefore, the signal used in the next phase of this experiment consists of both the 0.5 kHz component as well as the 1.3 kHz component. This is done to allow listeners to adopt a multiple-band strategy, and also to see what happens to the detectability of the two-tone complex as listeners make the transition.

Let us suppose that listeners begin to utilize both signal components. If the outputs of the integrators associated with the critical band filters are combined with the weights, ω_l, and ω_h, respectively, then according to Green and Swets (1974, pp. 273), d'_{com} equals

$$d'_{com} = \frac{\omega_l \cdot \Delta x_l + \omega_h \cdot \Delta x_h}{\sqrt{(\omega_l^2 \cdot \sigma_l^2 + \omega_h^2 \cdot \sigma_h^2)}} \qquad (3)$$

where $\Delta x_l = (\mu_{sn_l} - \mu_{n_l})$ and $\Delta x_h = (\mu_{sn_h} - \mu_{n_h})$, and where $var\, x_i = \sigma_i^2$.

Next we come to a critical point in this first experiment. In order to determine whether both components contribute to the detectability of the combined signal, we can eliminate one of the signal components without informing the listener. In terms of a multiple-band model, performance should be impaired if the observer is using two critical bands because signal energy is reduced while noise energy is not. Thus, the d' for the remaining component is proportional to the signal energy of the remaining component divided by the total variance due to noise power in both bands. That is,

$$d'_l = \frac{\omega \cdot \Delta x_l}{\sqrt{(\omega_l^2 \cdot \sigma_l^2 + \omega_h^2 \cdot \sigma_h^2)}} \tag{4}$$

Conversely, removing the 0.5 *kHz* signal component unexpectedly should result in a similar decrement of d'_h. Thus, by dropping out one component from the combined signal after a listener is using both we have a method for obtaining an estimate of the detectability for either signal component, in isolation, with noise from two bands. The integration model further predicts that the d' for the isolated signal should drop *below* previous levels.

Let us now consider an alternative processing model for combining information from multiple bands: the *decision threshold* model (see Green and Swets, 1974 for formal discussion). In contrast to the integration model, which assumes that sensory information is statistically combined prior to decision, the underlying assumption of this model, is that listeners make an independent decision for each critical band and then combine *decisions* to arrive at an overall response. More specifically, it is assumed that the listener establishes a criterion for making a decision about signal presence in each critical band. There is *no* guessing in the sensory process; "guessing" is part of the decision process. If a signal produces sufficient excitation, the criterion is exceeded and a positive detect is made. At this point, the listener uses some rule to combine the independent binary (yes-no) decisions. One relatively straightforward rule for making an overall positive response is to say "yes" if the outcome of any one band exceeds threshold (Shafer and Gales, 1949; Zagorski, 1975; Zagorski, 1984; Buus, 1986). Assuming that detection is based on n bands and that the individual probabilities associated with each band are equal, the model is,

$$p_n = 1-(1-p_i)^n \tag{5}$$

For the signals used in the current study, an overall positive response should be made when either, or both critical bands yield a positive detect. Eq. (5) then reduces to

$$P(C) = p_l + p_h - (p_l \cdot p_h) \tag{6}$$

where $P(C)$ is the probability of correct response and p_i represents a positive detect on each of the individual components (ie., hit rate), and where $i = h, l$.

To summarize, this experiment looks at how observers utilize information from two critical bands and how detection performance is affected when changes are made to the signal. In particular, it is designed to test the assumption that either a single- or multiple-band strategy can be used, and that the listener can shift from using one strategy to using the other. The experiment consists of four basic signal detection tasks carried out in the following sequence:

(1) the signal is a single specified frequency (0.5 *kHz*),

(2) the signal contains a second unexpected frequency component (1.3 *kHz*) in addition to the 0.5 *kHz* component,

(3) the signal contains both signal components, but the observer is made aware of them both, and

(4) each of the separate components is removed unexpectedly.

A. Methods

1. *Observers*

Three paid subjects with normal hearing participated in this study. Two females and one male, ages 18 to 21, were tested during a six-week period. Although none had had any previous experience in psychoacoustical tasks, each listener received a minimum of 5,000 trials before any of the data reported here were collected.

2. *Apparatus*

Signal synthesis and presentation, trial sequence and timing, and response recording were controlled by an Intel 8080-based microcomputer. Sinusoidal signals were produced through an 8 bit digital-to-analog converter. A Grason-Stadler Model 901B noise generator provided a continuous background of approximately Gaussian noise. Both signal and noise stimuli were amplified, analogically mixed, and low-pass filtered with -3 dB cutoff set at 3.2 *kHz*. Presentation levels were controlled by a Hewlett-Packard 350D attenuator and monitored on a Hewlett-Packard 3400A RMS voltmeter. Subjects listened diotically (ie., identical waveform delivered to both the right and left ear) through calibrated Grason-Stadler TDH-39 300Z headphones.

3. *Stimuli*

A modified table look-up method was used to generate the sinusoidal stimuli on each trial. The stimuli generated were two sinusoidal tones of 0.5 *kHz* and 1.3 *kHz*, and one combined signal consisting of both of these frequencies. Components for the combined signal were phase-locked, in-phase, with a repetition rate of 0.1 *kHz*. A 25 *kHz* sampling rate provided high temporal resolution of the signals. Total signal duration, including 12 *ms* linear rise and decays, was 134 *ms*. Signal levels (in *dB*) were calculated using the relationship $10\log(E_s/N_0)$, where E_s represents signal energy, and N_0 is the spectral density of the noise (ie., noise power per unit bandwidth). All data were collected with signal levels adjusted as follows: 7.5 for both the 0.5 *kHz* and the 1.3 *kHz* components; 10.5 for the complex signal. The output level of the noise was held constant in the 3.2 *kHz* bandwidth at 40 *dB SPL*.

4. *Trial, block, and session design*

All data reported here were collected with a single-interval, yes-no procedure. Each block consisted of a random order of noise-alone (N) and signal-plus-noise trials (SN). The *a priori* probabilities of N and SN trials were set equal to 0.5. A brief visual warning, approximately 500 *ms* in duration, was used to indicate the beginning of a trial. It consisted of a white bar, approximately 4 degrees of visual angle vertically, across a video monitor. Except for a white cursor, the screen was blank. When a signal was presented, its onset coincided with the offset of the warning light. Listeners were instructed to press one of two keys to indicate whether or not a signal was detected. A new trial began 1.7 sec after the subject entered a response. At the end of each block the hit rate and correct rejection rate were displayed on the observer's terminal.

Experimental sessions consisted on average of ten 120-trial blocks, and lasted between 1.5 and 2 hours. Each listener was given a brief rest between blocks and one longer rest period midway through the session.

5. *Procedure*

In the first experiment the same signal was used *throughout* a block of trials. Thus, when a signal was changed, the change occurred either between successive blocks, or between sessions. With the exception of session 2 and session 8, during which unexpected shifts occurred, each listener was given a preview of the signal, both with and without background noise, prior to running any experimental blocks. To obtain baseline estimates of d' for subsequent comparison, all listeners were trained using the $0.5\,kHz$ sinusoid as the signal. After training was judged to be complete, performance was measured for an entire session (session 1).

Session 2 began with a preview of the $0.5\,kHz$ component. However, prior to the first experimental block the $1.3\,kHz$ component was added to the signal. Note that this component was added without informing the observer.

Beginning with session 3 and continuing through to session 7, only the combined signal was used. This was done in order to provide a sufficient number of trials to establish a stable performance level on the combined signal. Again, the listeners' task was simply to indicate whether or not a signal was detected.

Session 8 was designed to measure performance on the individual components of the combined signal. An important aspect of the procedure used in this session (which involved removing and reintroducing signal components between blocks) was that subjects were not informed about the manipulation of signal types. The changes from one signal type to another were the same for all subjects and occurred in the following sequence. After receiving a preview of the combined signal, observers were tested for two blocks during which only the $0.5\,kHz$ tone was presented. Next, two blocks were run using the combined signal in order to restore performance on both components. Two blocks were then run using only the $1.3\,kHz$

tone and these were followed by two blocks using the complex signal.

Final estimates for performance on the single component signals were obtained from one complete session (session 9) which was conducted using the 0.5 *kHz* signal, and two sessions (session 10 and 11) during which the 1.3 *kHz* signal was used. This ended the experiment for which the same signal was presented throughout a block of trials.

B. Results

The values of d' obtained in this experiment are listed in Table 1. To facilitate comparison, these values are displayed in Figures 1 to 3, to the left of the dotted line. Note that for session 8 the values in Table 1 are given sequentially, on a block-by-block basis. However, the plotted $d's$ are based on two blocks for both single component signals, and on four blocks for the combined signal. All $d's$ shown are based on the assumption that the noise and signal-plus-noise distributions are normal, and are of equal variance. They were computed in the following way. From the four possible stimulus-response conjunctions, the frequencies for "hits" and "false alarms" were converted to conditional probabilities, $P(yes/SN)$ and $P(yes/N)$, respectively. Using these two probabilities, $d's$ were then found directly from the tables published by Elliot (1958).

Session 1 to session 2.

Recall that in session 1 the signal was a 0.5 *kHz* sinusoid whereas in session 2 it was the combined signal. Comparison of the d' from session 1 to the d' from session 2 shows there was no initial change in detectability in shifting from a single sinusoid to the two-component signal.

Session 3 to session 7.

For the complex signal, d' increases to a maximum of 1.69 for AC and a maximum of 1.84 for SH but does not increase for DG.

Session 8.

From Table 1, we can see that removing the 1.3 *kHz* component results in a drop in d' for all listeners. For both SH and DG, performance on the 0.5 *kHz* signal falls below performance on this signal in session 1. During the second block, d' is higher for each of the listeners. When the 1.3 *kHz* component is reintroduced, performance improves for AC but not for DG. For SH, performance on the complex signal is initially similar to that on the 0.5 *kHz* signal, but on the second block it improves substantially. Drops in d' also occur when the 0.5 *kHz* component is removed. When this component is reintroduced, recovery is immediate for AC and DG, but again occurs in the second block for SH.

Figure 1 shows that, for this listener (AC), the combined signal is more detectable than either of the individual signal components in isolation. By contrast, Figures 2 and 3 (SH and DG, respectively) show that one of the individual frequencies is as detectable as the combined signal.

Session 9.

When compared to the d'_l obtained in session 1, the detectability of the 0.5 *kHz* component has remained stable over the course of the experiment for observer SH. For AC however, detectability has improved to a level where d' for this component nearly matches some of the $d's$ for the complex signal that were obtained during sessions 3 to 7.

Session 10 and session 11.

For both AC and DG, d' for the 1.3 *kHz* is lower than d' for either of the other two signal types. However, for SH, the level d'_h measured in session 11 exceeds the maximum d'_{com} on the complex signal found in session 6.

C. Discussion of Experiment 1

The results indicate that detection performance, as measured by the sensitivity index d', did not change when the high component was unexpectedly added to the signal. If, indeed, listeners were using a single critical band centered on the low component, then we might argue that spectral energy in the region of the high component was attenuated. This finding is consistent with previous evidence concerning the detectability of unexpected signals (Tanner, et al, 1956), and supports a sensory filtering hypothesis (Greenberg and Larkin, 1968).

Note, however, that for listeners AC and SH detection performance improves with subsequent experience on the combined signal, indicating that these listeners began to use the high component. It is possible that they were making a shift in strategy from using a single critical band to using two separate bands. In order to determine whether these observers utilized the information in both bands we can examine the results obtained during session 8.

The data from the two blocks in which only the low component was used, and the two blocks in which only the high component was used, enable us to use each of the models to predict performance on the four blocks of combined signal. Specifically, the information integration model for two uncorrelated components is $pred\, d'_{com} = \sqrt{(d_l')^2 + (d_h')^2}$. By substituting the obtained value for the individual $d's$ the model predicts a d'_{com} of 1.68 for AC whereas the obtained d'_{com} is equal to 1.69. For SH the predicted d'_{com} is 2.08 whereas the obtained is 2.10. Both of these predicted values are within a standard error of the obtained $d's$. By contrast, the decision-threshold model predicts a probability correct, $P_{com}(C)$, of 0.85 for AC whereas the obtained is 0.80. For SH the predicted value is 0.90 and the obtained is 0.82. These discrepancies are each several standard errors.

In qualitative agreement with the information integration model are the decrements in performance that occurred when either of the individual components was removed from the combined signal. For example, in the first block of session 8, SH's performance on the 0.5 *kHz* signal dropped *below* performance on this same component as measured in session 1. This is consistent with the multiple-band energy-detection view because it is assumed that the subject listens to noise power from both filters but suddenly receives a signal in only one of them (see Eq. 4).

The fact that performance increased to previous levels during the second block suggests that SH detected a change in the signal and then, in effect, attenuated the output associated with the critical band containing only noise. Similarly, the data for this observer (session 8) show that detectability did not initially increase when either of the single components was reintroduced (see Table 1). Once again, the sharp increase in d' indicates that a shift in listening strategy occurred during the second block.

Taken together, the results from this experiment suggest that subjects can use a different listening strategy in response to changes in the spectral composition of a signal. In particular, a listener can improve detection performance either by adopting a single-band strategy when the signal is restricted to a single frequency, or by adopting a multiple-band strategy when the signal consists of two, widely separated frequencies. The next experiment is designed as an attempt to determine more about how auditory information from two critical bands is combined. A task is considered in which listeners cannot adopt an optimal strategy for each of the different signals.

III. EXPERIMENT 2. *Blocks with a mixture of signal types*

It was found that when signals were changed between blocks, subjects apparently detected properties of the signal and then adjusted their listening strategy accordingly. So, for example, when a component was eliminated from the combined signal, performance was initially impaired but it improved in subsequent blocks. In this experiment the three signals used in Experiment 1 are randomized, on a trial-by-trial basis, within the *same* block of trials. As a result, listeners cannot select a strategy that is best suited to each signal. Rather, they must select a strategy that deals with all signals simultaneously and maintain it throughout a block of trials.

Swets (1984) has described various models for how observers "aggregate," or combine information which excites separate critical bands. First let us reconsider the integration model. This model asserts that observations from n statistically independent critical bands, $(x_1, x_2, \ldots, x_n)$, are combined either via likelihood ratio or by an optimally weighted linear combination. Because of an assumed post-detection integration, the detection process common to each band would occur, resulting in an overall decision that is made on the basis of the combined information. Hence, when n observations are combined, the model is

$$d'_n = \left[\sum_{i=1}^{n} (d_i')^2 \right]^{1/2} \qquad (7a)$$

In words, the combined value of d', denoted d_n', is equal to the square root of the sum of the squares of the individual values of d'. For a signal consisting of two equal energy, orthogonal components the integration model is

$$d'_{com} = \sqrt{(d_l')^2 + (d_h')^2} \qquad (7b)$$

Recall that the results obtained in Experiment 1 demonstrate that listeners can attenuate both signal and noise input to one of the bands. Thus, in some cases, the value of d' for a given signal changed gradually over many blocks, whereas in other cases, it changed dramatically from one block to the next. In contrast, because listeners in Experiment 2 are presented with signals that excite either filter separately, or both filters simultaneously, the output of both critical bands must be used. This means that, independent of which signal is presented, noise from each critical band is delivered to the decision process. Therefore, in a mixed block condition, the total noise that "contributes" to variance in the decision process is equal to a weighted linear combination of the variance in each band,

$$\sigma_{tot} = \sqrt{(\omega^2_l \cdot \sigma^2_l + \omega^2_h \cdot \sigma^2_h)} \qquad (8)$$

Assuming that the filter outputs are uncorrelated, then a signal which excites either critical band would produce an increase of signal energy in that band alone. In the current study, this will be true if either of the single component signals ($0.5\,kHz$ *or* $1.3\,kHz$) occurs. It is important to note, however, that the total variance due to the noise power, σ_{tot} is still the combined noise from both bands. Thus, the appropriate d' for each individual component is

$$d'_l = \frac{(\omega_l \cdot \Delta x_l)}{\sigma_{tot}} \tag{9a}$$

and,

$$d'_h = \frac{(\omega_h \cdot \Delta x_h)}{\sigma_{tot}} \tag{9b}$$

where $\Delta x_l = (\mu_{sn_l} - \mu_{n_l})$.

But on trials when the complex signal occurs, the signal energy in both critical bands increases, and therefore

$$d'_{com} = \frac{(\omega_l \cdot \Delta x_l + \omega_h \cdot \Delta x_h)}{\sigma_{tot}} \tag{10}$$

By substituting Eq. 9a,b into Eq. 10 we can see that the relationship between the detectability for the individual components and the detectability for the complex signal is given by

$$d'_{com} = d'_h + d'_l \tag{11}$$

Thus, by randomly mixing the single component signals and the combined signal within the same block of trials, the obtained d' for the combined signal should be the simple sum of the individual $d's$.

In this experiment the d' summation model (Eq. 11) and the decision threshold model (Eq. 6) will be used to quantitatively test how listeners combine information from two separate critical bands. While the integration model assumes that the outputs of the critical band filters are statistically combined to form the basis for decision, the decision threshold model assumes that only the decisions made on each band are retained for further processing.

A. Procedure

All signals and signal levels were the same as those used in Experiment 1. In addition, the overall *a priori* probability of SN and N trials was set equal to 0.5. Out of the 60 SN trials, each signal type was equally likely to be selected for presentation and thus, the $0.5\,kHz$, $1.3\,kHz$, and combined signal each were presented 20 times within the same block. The listeners' task was to indicate whether or not a signal was detected; identification was not required. Subjects AC, SH, and DG were tested for 30, 60, and 50 mixed blocks respectively.

B. Results

In order to facilitate comparison of the detectability of signal types across experiments, the obtained values of d' listed in Table 2 are displayed in Figures 1 to 3 above the label "mixed blocks." Note that the $d's$ reported here were calculated using the overall false alarm rate. Table 2 shows the experimentally obtained hit rate and d' for each of the signal types, as well as the estimated hit rate and d' found by iteratively fitting each of the models to the data. The maximum likelihood estimates were found by adjusting the estimated hit and false alarm rate to minimize the chi square discrepancy between the data and the estimates subject to the constraint of the models. The final iteration involved step sizes that correspond to changes in the estimates of 0.001.

The results of the analysis show that the models tested here do not provide an accurate description of the data. For the d' summation model, the values of χ^2 are all significant at the 1% level (with $3\,df$). In addition, the form of the decision threshold model analyzed here does not fit the data for observers SH and DG, but does fit AC's data. Although the experimentally obtained $d's$ for the combined signal are less than those predicted by the models, it is apparent that for AC and SH, the combined signal is more detectable than either of the single components (see Figures 1 and 2). For DG, the $0.5\,kHz$ signal and combined signal are equally detectable.

C. Discussion of Experiment 2

The results from this experiment show that neither the information integration nor the decision threshold model can adequately describe the relationship between performance on the single component signals and the combined signal. Nevertheless, additional comparisons of performance can be made between the experiments.

According to both models the detectability of the two-tone complex should be the same in both experiments. For example, the information integration model asserts that the listener combines sensory information from both critical bands. Regardless of whether this signal is presented throughout a block of trials or, is interspersed randomly with the isolated components in the same block, detectability would be unaffected. That is, for the complex signal his decisions are based on the same signal strengths and the same noise (compare Eq. 4 and Eq. 9). Likewise, for the decision threshold model, the observer is combining decisions resulting from the same processes in both experiments. So, it too predicts identical performance on the combined signal in both experiments. By contrast, these models do not make the same predictions for the single component signals. In particular, the decision threshold model predicts that performance should be the same for both experiments because the probability of the variable in a channel exceeding some fixed value (threshold) does not depend on whether the signal is the same in every trial or varies from trial to trial (Buus, et al, 1986). The multiple-band model, on the other hand, does predict lower performance in the second experiment. The reason for decreased performance is that the listener must simultaneously use both listening bands, which causes an increase in the variance of the decision variable relative to a single-band strategy. Put simply, the noise that degrades performance in the first experiment comes from only a single critical band while in the second experiment it must come from both critical bands.

For the combined signal SH, AC, and DG show performance decrements equivalent to signal decreases of 0.6, 0.6, and 1.8 *dB* respectively, where $d's$ have been converted to *dB* using the relationship $dB = 10 \log (d'_{mixed} / d'_{fixed})$. Each of these values represent changes of at least two standard errors. The decreases for the low components are 0.2, 0.9, and 1.7 *dB* respectively. Finally, decreases for the high component are 2.1 and 5.5 *dB* for SH and AC. The decreases for the single components represent changes of at least two standard errors for all but the smallest decrease. We can argue, then, that these decreases for the single component signals are contrary to the predictions made according to the decision threshold model while the decreases on the combined signal are contrary to both models. Thus not only do the models *not* fit the results of this experiment, but they do not predict the differences in performance between experiments.

IV. CONCLUSION

Experiment 1 shows that subjects can adjust their listening strategy to adapt to changing signals and also that this transition in strategy can occur between successive blocks, or over many blocks. In addition, the quantitative analysis in experiment 2 shows that subjects do not combine information according to the information integration or decision threshold models. This conclusion is further supported by the comparisons of performance between experiments.

Therefore, while the results from this study demonstrate that subjects have the ability to switch from a single- to a multiple-band listening strategy, presumably to optimize performance when the signal changes, neither the information integration nor the decision threshold models account for the details of their performance.

Table 1. *Obtained values of d′ from Experiment 1.*

Condition		Observers					
		AC		SH		DG	
Session	*Signal*	*d′*	*trials*	*d′*	*trials*	*d′*	*trials*
1	0.5 kHz	1.27	1200	1.35	1200	1.21	1200
2	combined	1.29	"	1.39	"	1.28	"
3	"	1.58	"	1.56	"	1.11	"
4	"	1.49	"	1.71	"	1.19	"
5	"	1.57	"	1.80	"	1.17	"
6	"	1.51	960	1.84	960	1.18	"
7	"	1.69	1200	1.74	1080	1.23	"
8	0.5 kHz	1.26	120	1.19	120	1.04	120
	"	1.56	"	1.28	"	1.35	"
	combined	1.74	"	1.35	"	1.34	"
	"	1.88	"	2.04	"	1.30	"
	1.3 kHz	1.05	"	1.78	"	0.00	"
	"	0.88	"	1.61	"	0.15	"
	combined	1.54	"	1.62	"	1.07	"
	"	1.65	"	2.58	"	1.31	"
9	0.5 kHz	1.48	1200	1.38	960	1.34	1200
10	1.3 kHz	1.12	"	1.81	1200	0.29	"
11	"	1.06	"	1.92	"	0.11	"

Table 2. *Summary of the analysis of Experiment 2.*

		Data		Maximum Likelihood Estimates		
				d′ *Summation model*		*Decision Threshold model*
Observer	Signal	*d′*	Hit rate	*d′*	Hit rate	Hit rate
AC	1.3 *kHz*	0.29	0.44	0.44	0.50	0.43
	Combined	1.44	0.85	1.55	0.86	0.87
	0.5 *kHz*	1.21	0.78	1.12	0.75	0.76
		(n = 3240)		$\chi^2 = 13.20$		$\chi^2 = 3.34$
SH	1.3 *kHz*	1.19	0.70	0.92	0.61	0.65
	Combined	1.65	0.84	1.98	0.91	0.90
	0.5 *kHz*	1.33	0.75	1.20	0.71	0.70
		(n = 7200)		$\chi^2 = 142.27$		$\chi^2 = 56.41$
DG	1.3 *kHz*	0.20	0.47	0.28	0.50	0.43
	Combined	0.88	0.73	0.96	0.75	0.80
	0.5 *kHz*	0.90	0.74	0.69	0.66	0.72
		(n = 6000)		$\chi^2 = 30.60$		$\chi^2 = 4.33$

Note: A χ^2 value of 11.30 (with 3 *df*) is significant at the .001 level.

Figure 1. Obtained values of d' for AC.

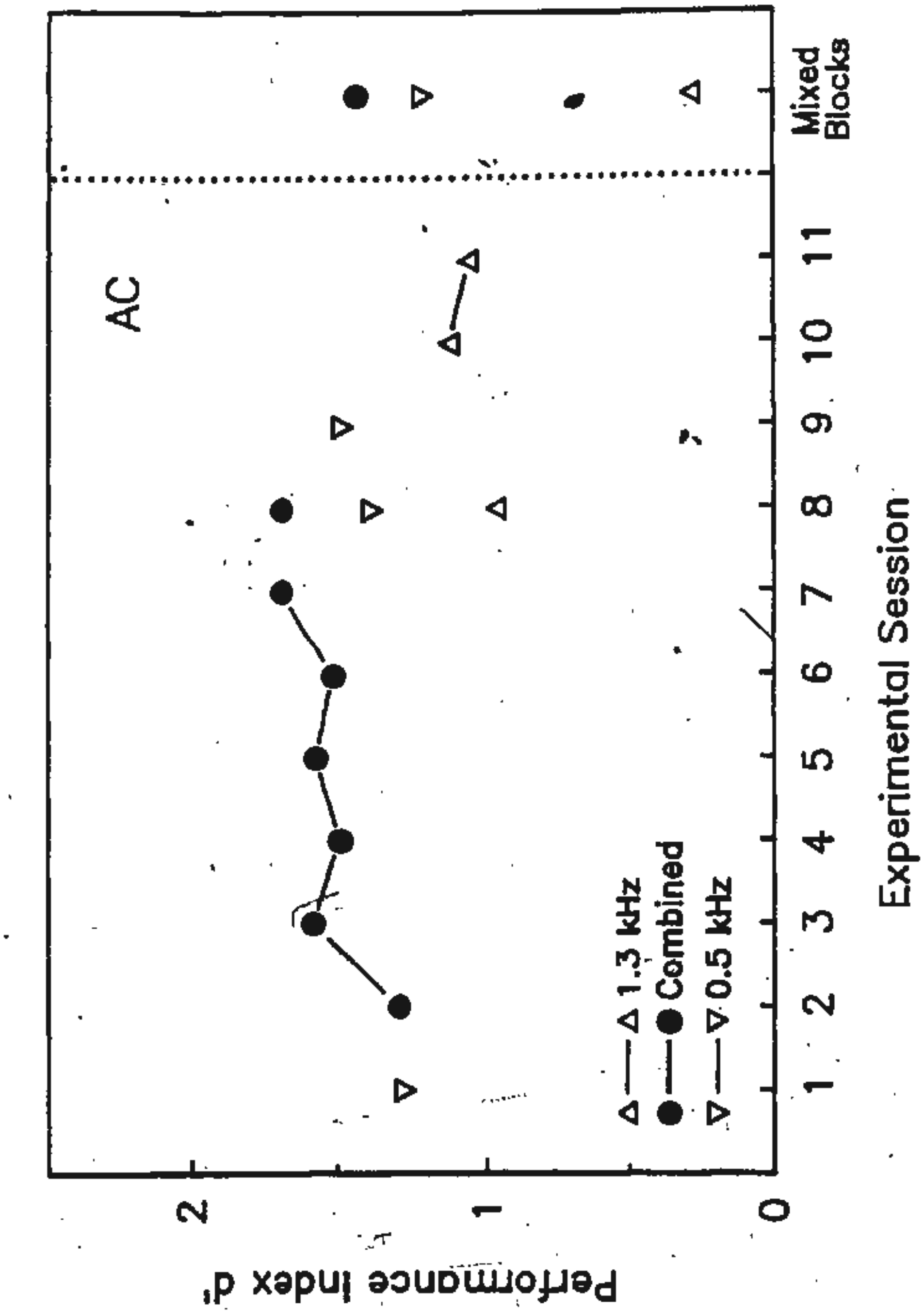


Figure 2. Obtained values of d' for SH.

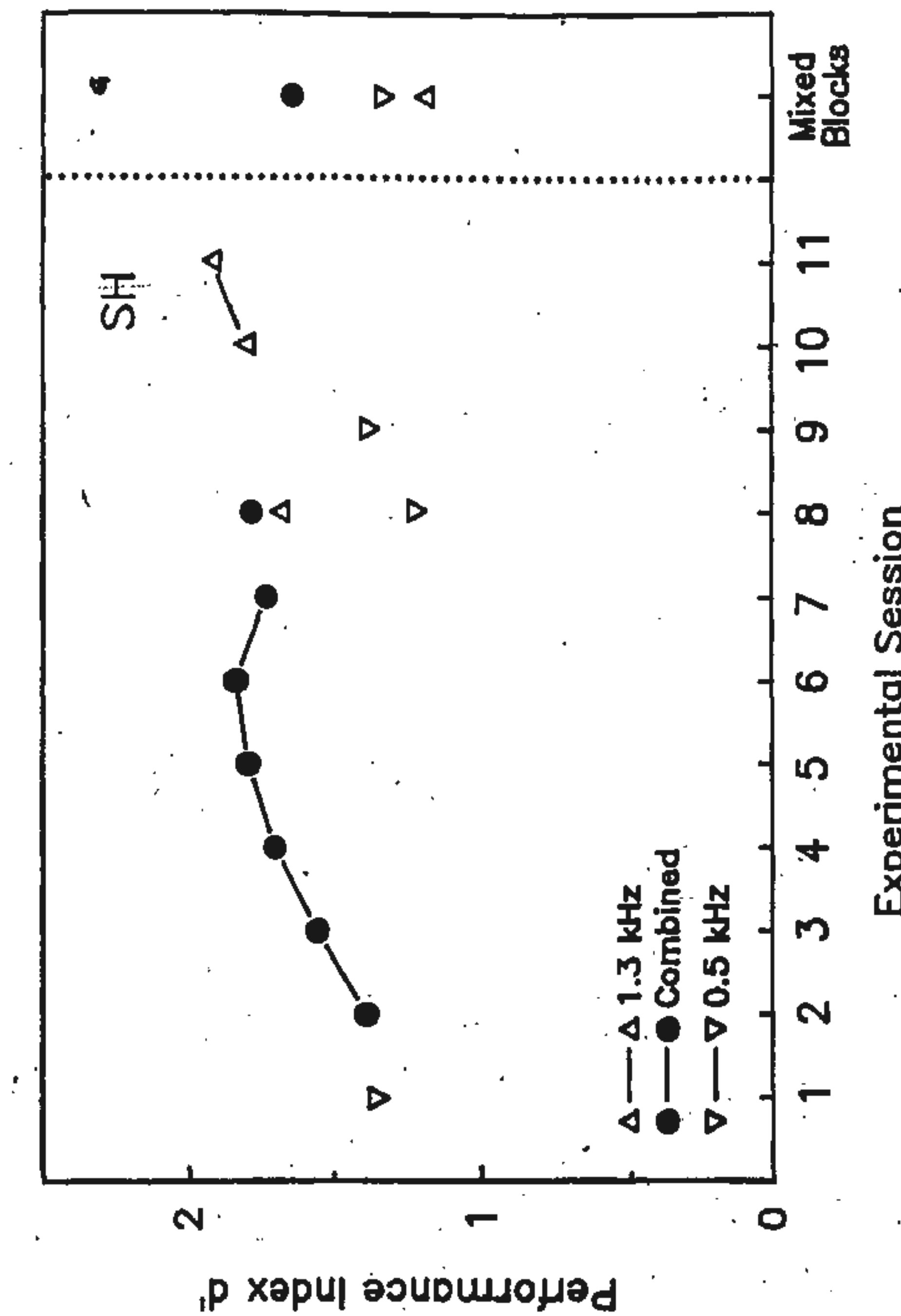
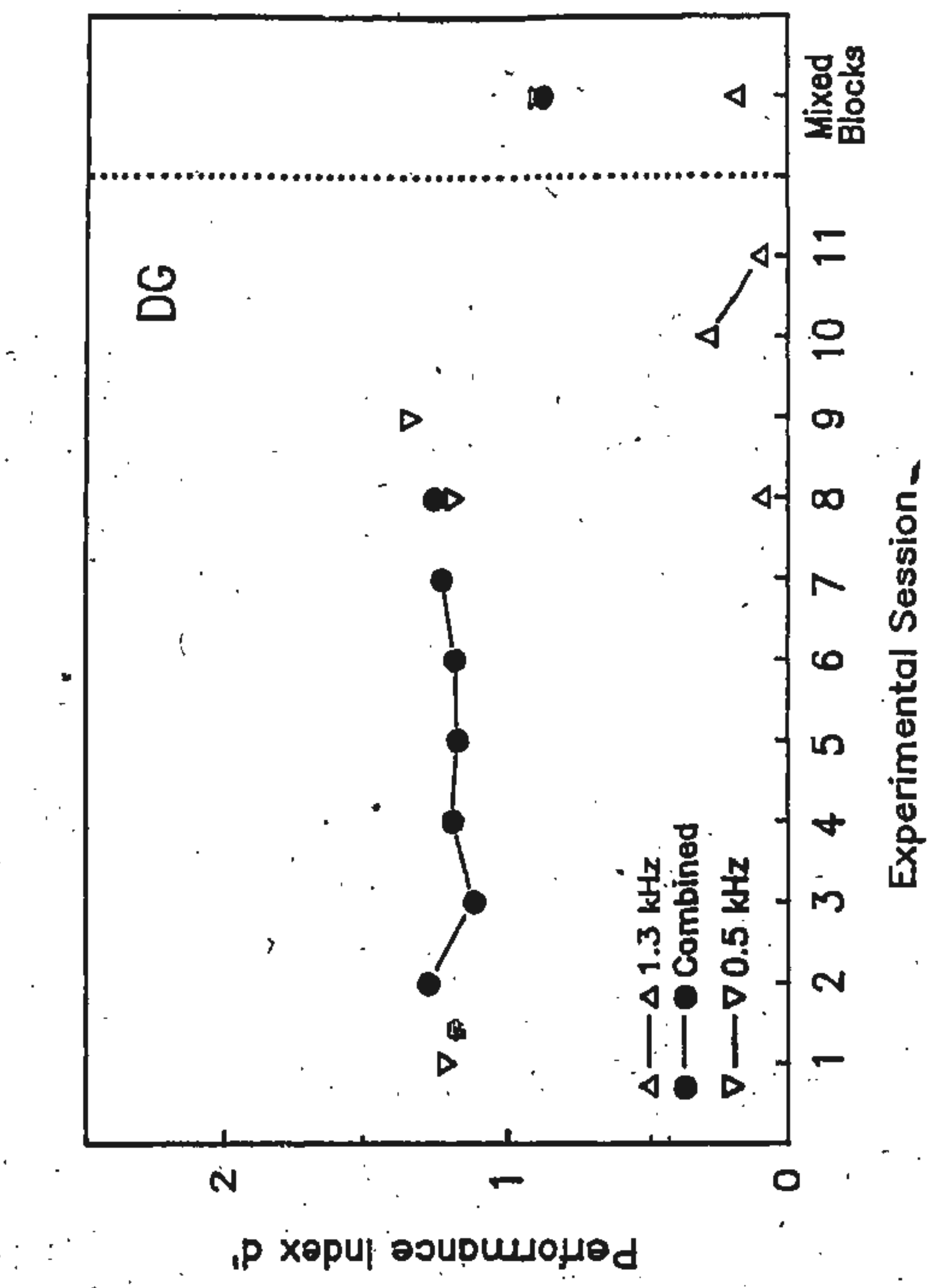


Figure 3. Obtained values of d' for DG.



REFERENCES

Bekesy, G. von (1961b). Pitch sensation and its relation to the periodicity of the stimulus. Hearing and skin vibrations. *Journal of the Acoustical Society of America*, 33, 341-348.

Buus, S., Schorer, E., Florentine, M., and Zwicker, E. (1986). Decision rules in detection of simple and complex tones. *Journal of the Acoustical Society of America*, 80, 1646-1667.

Evans, E.F. and Wilson, J.P. (1973). The frequency selectivity of the cochlea. In Moeller, A.R. (ed.), *Basic mechanisms in hearing*. pp. 519-554.

Fletcher, H. (1940). Auditory patterns. *Review of Modern Physics*, 12, 47-65.

Green, D.M. (1958). Detection of multiple component signals in noise. *Journal of the Acoustical Society of America*, 30, 904-911.

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

Green, D.M. (1961). Detection of auditory sinusoids of uncertain frequency. *Journal of the Acoustical Society of America*, 33, 897-903.

Green, D.M. (1983). Profile analysis. A different view of auditory intensity discrimination. *American Psychologist*, 38, 133-142.

Green, D.M., and Swets, J.A. (1974). *Signal detection theory and psychophysics*. New York: Krieger, Huntington.

Greenberg, G.Z. (1969a). Frequency-selective detection at three signal amplitudes. *Perception and Psychophysics*, 6, 297-301.

Greenberg, G.Z. (1969b). Frequency selectivity during amplitude discrimination of signals in noise. *Journal of the Acoustical Society of America*, 45, 1438-1442.

Greenberg, G.Z., and Larkin, W.D. (1968). Frequency response characteristic of auditory observers detecting signals of a single frequency in noise: The probe signal method. *Journal of the Acoustical Society of America*, 44, 1513-1523.

Greenwood, D.D. (1961a). Auditory masking and the critical band. *Journal of the Acoustical Society of America*, 33, 484-502.

Houtgast, T. (1974). Psychophysical evidence for lateral inhibition in hearing. *Journal of the Acoustical Society of America*, 51, 1885-1894.

Macmillan, N.A., and Schwartz, M. (1975). A probe-signal investigation of uncertain-frequency detection. *Journal of the Acoustical Society of America*, 58, 1051-1058.

Marill, T. (1956). Psychophysics and detection theory. Massachusetts Institute of Technology: Research laboratory of Electronics, Technical Report No. 319.

Moore, B.C.J., and Glasberg, B.R. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *Journal of the Acoustical Society of America*, 74, 750-759.

Nolte, L.W., and Jaarsma, D. (1967). More on the detection of one of *M* orthogonal signals. *Journal of the Acoustical Society of America*, 41, 497-505.

Patterson, R.D. (1974). Auditory filter shape. *Journal of the Acoustical Society of America*, 55, 802-809.

Patterson, R.D., and Nimmo-Smith, I. (1980). Off-frequency listening and auditory-filter asymmetry. *Journal of the Acoustical Society of America*, 67, 229-245.

Peterson, W.W., Birdsall, T.G., and Fox, W.C. (1954). The theory of signal detectability. *IRE Transactions on Information Theory*, 4, 171-212.

Pfafflin, S.M., and Mathews, M.V. (1962). Energy-detection model for monaural auditory detection. *Journal of the Acoustical Society of America*, 34, 1842-1853.

Plomp, R. (1976). *Aspects of tone sensation*. London: Academic Press.

Sachs, M.B., and Kiang, N.Y.S. (1968). Two-tone inhibition in auditory nerve fibers. *Journal of the Acoustical Society of America*, **43**, 1120-1128.

Shafer, T.H. and Gales, R.S. (1949). Auditory masking of multiple tones. *Journal of the Acoustical Society of America*, **21**, 392-398.

Shafer, T.H., Gales, R.S., Shewmaker, C.A., and Thompson, P.O. (1950). The frequency selectivity of the ear as determined by masking experiments. *Journal of the Acoustical Society of America*, **22**, 490-497.

Sharf, B. (1970). The critical band. In Tobias, J.. *Foundations of modern auditory theory*. pp. 159-202. New York: Academic Press.

Swets, J.A. (1963). Central factors in auditory frequency selectivity. *Psychological Bulletin*, **60**, 565-581.

Swets, J.A. (1984). Mathematical models of attention Parasuraman, R., and Davies, D.R. (eds.), *Varieties of attention*. pp. 183-234. New York: Academic Press.

Swets, J.A., and Birdsall, T.G. (1967). Deferred decisions in human signal detection: A preliminary experiment *Perception and Psychophysics*, **2**, 15-28.

Swets, J.A., and Birdsall, T.G. (1978). Repeated observations of an uncertain signal. *Perception and Psychophysics*, **23**, 269-274.

Tanner, W.P. Jr., Swets, J.A., and Green, D.M. (1956). Some properties of the hearing mechanism. University of Michigan: Electronic Defense Group, Technical report no. 86.

Weber, D.L. (1978). Suppression and critical band bandlimiting experiments. *Journal of the Acoustical Society of America*, **64**, 141-150.

Zagorski, M.A. (1975). Perceptual independence of pitch and loudness in a signal detection experiment: A model for 2ATFC (2IFC) experiments. *Perception and Psychophysics*, **17**, 525-531.

Zagorski, M.A. (1984). Vibratos and Tremelos: A case for communication among scientists and musicians. *Musicworks*, **26**.

Zwicker, E., and Terhardt, E. (1980). Analytic expressions for critical band rate and critical bandwidth as a function of frequency. *Journal of the Acoustical Society of America*, **68**, 1523-1525.