(12)

Special Report 84-1

A DESIGN GUIDE FOR NONSPEECH AUDITORY DISPLAYS

B. E. Mulligan*, D. K. McBride and L. S. Goodman

September 1984

NAVAL AEROSPACE MEDICAL RESEARCH LABORATORY
PENSACOLA, FLORIDA

Approved for public release; distribution unlimited

A DESIGN GUIDE FOR NONSPEECH AUDITORY DISPLAYS

B. E. Mulligan*, D. K. McBride and L. S. Goodman

Naval Air Systems Command
61153N WR04210.WR0421001.6142

Reviewed by

F. E. Guedry, Ph.D.
Senior Scientist

Approved and Released by

Captain W. M. Houk, MC, USN
Commanding Officer

January 1985

Naval Aerospace Medical Research Laboratory
Naval Air Station, Pensacola, Florida 32508-5700

SUMMARY PAGE

THE PROBLEM

Vast attention has been devoted to the investigation of various sensory and perceptual characteristics of the human auditory system. It is not often obvious, however, how the aggregate findings provided by these efforts might effectively be utilized to design auditory displays of information. This report condenses and synthesizes critical research findings on the (1) detection , (2) loudness, and (3) distinctiveness of non-speech auditory displays. The format of this report provides a unique guide for the design of nonspeech auditory displays.

FINDINGS

Eight tables and two algorithms (in flow-chart form) were developed and are provided to assist the auditory display engineer in (1) increasing the detectability of signals presented in noise and (2) increasing the loudness of signals without increasing signal level. The algorithms are coded in the BASIC computer language and are enclosed as appendices.

RECOMMENDATIONS

The scope of this report and the algorithms provided are limited to three important areas of auditory display engineering. Similar attention should be devoted to other critical aspects of audition, such as, reaction time, stimulus-response compatibility, attention, recognition, and memory.

Acknowledgments

Thanks are due Mr. Glenn Davis, M.S., for coding the two auditory design algorithms in BASIC.

* Dr. B. E. Mulligan's address is Department of Psychology, University of Georgia, Athens, Georgia 30602.

TABLE OF CONTENTS

Page

TABLES

Subject

ALGORITHMS

Subject

1. DETECTION: A principal factor in evaluating the suitability of acoustic signals in human communication systems is the detectability of the signals which depends on both the physical characteristics of the signals and interfering noise as well as the sensitivity and frequency-selectivity of the auditory system. The primary parameters that limit detectability of tonal and complex signals by the human auditory system will be discussed in the following sub-sections for both externally noise-free (quiet) conditions and for noise-corrupting conditions. In addition to discussions of limiting parameters, wherever possible minimal conditions for reliable detection will be specified.

1.1. SIGNALS IN QUIET: If the term "quiet" is taken to mean the total absence of sound, then it must be found only in a vacuum. Obviously, the term does not imply either a vacuum or the total absence of sound. Nor does it indicate sound levels on the order of that resulting from the random motions of air molecules (Brownian motion). All that we intend "quiet" to mean is ambient sound pressure levels below those which mask pure tones presented at absolute threshold pressures (6).

1.1.1. TONAL SIGNALS: The detectability of sinusoidal signals under quiet conditions depends on signal duration, frequency, and sound pressure level. Although the precise shape of the relationship between detectability and sound pressure level (SPL) depends on the particular index that is chosen for detection measurement, it is generally found that signal detectability is an increasing function of SPL, usually somewhat ogival in shape. Some point on this "psychometric function" is taken as the absolute threshold (AB), typically the SPL that results in 75 percent correct detection performance (previously it was the 50 percent point). Because the auditory system is differentially sensitive to sound throughout the range of normal hearing (approximately 20 Hz to 20 kHz), the value of SPL at the absolute threshold varies as a function of signal frequency (see Table I). The smaller the value of SPL that is required for threshold detection, the greater is the sensitivity of the system. The region of greatest auditory sensitivity occurs at about 1 kHz and diminishes as signal frequency is either reduced below, or raised above, this region. Furthermore, as the duration of very brief (less than about 1 sec) signals increases, threshold decreases to some minimal value, e.g., that reported for the absolute threshold.

1.1.1.1. THRESHOLD INTERPRETATION: Several items pertinent to interpretation of absolute thresholds are worth noting. First, the absolute threshold is not a demarcation point between no detection and perfect detection. Rather, it is just one value on a psychometric function (e.g., the SPL corresponding to 75 percent correct detection) which extends over a range of SPL of about 10 decibles (dB) for the performance range of 0 percent to 100 percent detection. The particular point on the psychometric function that is selected for the absolute threshold is, in a sense, arbitrary but located in a region of the function where detection performance varies approximately linearly with SPL.

Thus some detectability of signals presented at SPLs below threshold should be expected. Second, standardized threshold values (see Table I) are averages over a number of individuals and may not hold exactly for any particular person, even if that person's hearing is within "normal" limits. Furthermore, standardized absolute thresholds have been established under relatively ideal listening conditions that probably would not be duplicated within "real-world" environments even if they were quiet (i.e., if noise levels were below absolute threshold). Consequently, setting SPLs of signals a few decibles above absolute threshold probably would not ensure 100 percent correct detection even in quiet environments. Familiarity with the signals, the probability of their occurrence, attentional demands on the listener, etc., may be expected to exert non-acoustic influences on signal detection and should be taken into account in selecting signal SPL, frequency, and duration.

TABLE I

Signal Frequency (Hz)	SPL (dB) ISO*	SPL (dB) ANSI**
80	--	61.0
125	45.5	45.5
250	24.5	28.0
500	11.0	12.5
1,000	6.5	5.5
1,500	6.5	8.5
2,000	8.5	10.5
3,000	7.5	7.0
4,000	9.0	9.5
6,000	8.0	10.5
8,000	9.5	9.0
10,000	--	17.0
12,000	--	20.5
15,000	--	39.0
18,000	--	74.0

*ISO 389-1975, "Standard Reference Zero for the Calibration of pure tone Audiometers".
**ANSI S3.6-1969, "ANS Specifications for Audiometers".

1.1.1.2. SPECIFICATION OF TONE SENSATION LEVEL: In specifying SPLs of signals to be presented in quiet, "real-world" situations, perhaps the most useful aspect of absolute thresholds is that they permit the determination of effectively equivalent SPLs for signals of different frequency. For example, if a 250 Hz signal and a 1000 HZ signal are both presented to a listener at the same SPL (e.g., 40 dB re 20 μ N/m^2), they will not be effectively equal in intensity and, therefore, not equally

detectable. Effective intensity, or sensation level (SL), of the 250 Hz signal will be about 16 dB while the SL of the 1000 Hz signal will be about 34 dB. To set different signals at equal SLs, the number of decibles above the absolute threshold to which the SPL of each signal is set should be the same. That is, $SL_1 = SL_2$, where $SL = SPL_1 - SPL(AB)_1 = SPL_2 - SPL(AB)_2$ and where SPL (AB) is sound pressure level at the absolute threshold and SPL is the sound pressure level to which the signal is set. Thus, as values of SPL(AB) vary with signal frequency (see Table I), so must signal SPLs in order that SLs remain equal. It is recommended that, when tonal signals are to be presented to listeners under quiet conditions, the SLs should be set between about 40 dB and 70dB, depending on the presence of non-acoustic sources of interference with detection. However, in any situation where unacceptable risk is contingent upon failure to detect a signal occurrence, the SL of the signal that will be required to produce 100 percent detection performance should be determined through empirical testing.

1.1.1.3. TONE DURATION: The minimally detectable SPL required for tones under quiet conditions depends not only on signal frequency, but also the duration of brief signals. Threshold SPL decreases as a function of signal duration up to times between about 0.05 and 1.0 sec due to temporal integration of signal energy by the auditory system. The threshold for tones may be reduced by more than 25 dB (depending on frequency) by increasing signal duration from about 1 msec to 1 sec. The threshold SPLs listed in Table I are for signals of durations greater than 1 sec. Since the calculation of SL requires values of SPL(AB) (as described in section 1.1.1.2.), and because duration and frequency interact in determining thresholds of very brief signals, it is recommended that durations of at least 1 sec be specified for tonal signals.

1.1.1.4. TONE RISE-DECAY TIMES: If the onsets or offsets of tonal signals are too rapid, wide-spectrum transients will be produced. Essentially, these transients are bursts of noise. If the tonal quality (frequency integrity) of the signal is important, the onsets and offsets of the signals should be gradual. The rate at which the signal amplitude increases from zero to its peak or steady-state value (rise-time), and *vice versa* (decay-time), probably should not be less than about 5 to 10 msec, depending on signal frequency. Generally, slightly longer rise-decay times are required for low-frequency signals. However, in no case should rise-decay be less than about 1/6 of the total signal duration.

1.1.1.5. AUDITORY FATIGUE: The detectability of a tonal signal may be reduced, i.e., its threshold may be elevated, due to previous exposure to sound within the same frequency region. The SPL of a tonal signal required for threshold detection increases as a function of the level and duration of pre-exposing sounds, and decreases as a function of (1) the time between termination of the exposing sound and onset of the signal, and (2) the difference in frequency between the exposing sound and the

signal. Pre-exposures no greater than about 30 dB SL produce little fatigue even after several minutes of exposure. However, the effect of pre-exposure to signals of greater SLs tends to increase as the duration of the pre-exposing signals increases, especially at frequencies above about 500 Hz. So long as the pre-exposure level is below about 80 dB SL, the elevated threshold can be expected to return to baseline within approximately 200 msec. Thus, if the signal to be detected follows the pre-exposing sound by times greater than about 200 msec, its detectability will not be affected. If, however, the level of the pre-exposing sound is much greater than about 80 dB SL, the resulting threshold elevation may endure for minutes, or even hours if the duration of pre-exposure has been long. Therefore, even though a signal may be presented during a quiet interval, its threshold may be elevated above its SPL(AB) due to previous exposure. If the time interval between offset of pre-exposure and onset of the signal cannot be increased to allow for recovery from fatigue, it is recommended that the signal frequency be shifted away from the frequency region of the pre-exposing sound by at least one octave. Since fatiguing effects are generally confined to a narrow band in the immediate vicinity of the exposing frequency (except for very intense sounds in which case the effect is maximal about 1/2 octave above the exposing frequency), thresholds for signals +1 octave away from the exposing frequency can be expected to be unaffected.

1.1.2. COMPLEX SIGNALS: Any signal with a non-sinusoidal wave form is regarded as complex. According to this definition, a pure tone is simple, but a mixture of pure tones is not. The signals produced by bells, buzzers, engines, and voices are all complex. These may be characterized by prominent periodicities, discontinuous spectra, and distinguishable frequency modulations, or they may be completely random (i.e., random with regard to amplitude and phase) with continuous spectra as in the case of "white" (wide band) or "pink" (narrow band) noise. The detectability of such signals in quiet is subject to the same considerations as in the case of tonal signals with the exception that the threshold for each particular signal must be determined, i.e., there is no standardized table of threshold values available for such signals. It is recommended that thresholds for such signals be determined in quiet following established psychophysical procedures. Once the threshold is known, then the SPL for the signal may be specified in terms of SL. This procedure is desirable because it yields signal specifications that are stated in terms of sensitivity of the auditory system to the signal in question. Detectability of signals not specified with respect to SL cannot be properly evaluated. It is essential to further specify the spectrum, duration, and rise-decay times of such signals since their threshold values are valid only if these signal parameters remain unchanged.

1.2. SIGNALS IN NOISE: The detectability of signals in noise depends not only on the frequency selectivity of the auditory system and its temporal integrating (and differentiating) capacity, but also on the physical characteristics of both

signals and masking noise. All of the signal parameters indicated in section 1.1. are important here, in addition to interaural parameters that may be present in the case of binaural signals. The relative importance of these various signal parameters depends on the characteristics of the noise, especially the temporal and spectral proximity of the noise to the signals. Research on the dectability of signals in noise has been primarily conducted using white and pink noise. In what follows, findings on the detection of signals in white and pink noise will be generalized to all situations where the detectability of signals is reduced by auditory masking.

1.2.1. TONAL SIGNALS: Precisely the same considerations raised in section 1.1.1. apply here. Again, by tonal we mean sinusoidal. However, here threshold refers to the masked threshold rather than the absolute threshold which is pertinent only under quiet conditions. In both cases, the threshold is determined for a particular performance value (e.g., 75 percent correct detection) from the psychometric function obtained over a range of signal-to-noise ratios (S/N in decibles), usually about 10 dB.

1.2.1.1. MONAURAL DETECTION: The detectability of tonal signals in noise under monaural conditions is a matter of practical interest only for signals presented to one ear from a single headphone which also transmits noise. This assumes that the input to the non-signal ear is of relatively low magnitude and uncorrelated with the noise presented through the headphone to the signal ear. The effect is a functional isolation of the two ears such that binaural interactions are rendered negligible. Signal detectability under these conditions is equivalent to that obtained under binaural diotic conditions (discussed in section 1.2.1.2.). However, if both ears are exposed to the same noise while the signal is presented to one ear alone, a condition of binaural imbalance occurs and detectability may exceed that obtained under the monaural condition (discussed in section 1.2.1.2.1.).

1.2.1.1.1. SIGNAL-TO-NOISE-RATIO: For a given signal frequency, the S/N ratio necessary to achieve a specified level of detection performance (e.g., the masked threshold, defined as 75 percent correct detection) remains approximately constant regardless of noise level. This means that, if the noise spectrum level changes, the signal level required to maintain constant detectability must also change by approximately the same amount. This is fortunate because, to achieve a desired level of detectability, it is necessary only to specify the required S/N ratio for the signal frequency in question. Consequently, the necessity of providing *a priori* specifications of signal levels for conditions where noise levels are either unknown, or subject to change, is avoided. Table II lists S/N ratios required to obtain 75 percent correct detection (masked thresholds) for a range of signal frequencies between 150 Hz and 6000 Hz (8). Since the tabulated values represent the performance of highly trained listeners under ideal conditions, it is recommended that

the values of S/N ratio given in the table be increased by at least 10 dB to take into account departures from ideal listening conditions. If noise levels vary or if distracting non-acoustic events occur, the tabulated S/N ratios should be further increased. Except in extraordinary circumstances, the signal SPL should not exceed 80 dB (re 20 μN/m^2). Signal levels of tones above 80 dB SPL not only may be aversive, but also may be unsafe if exposure is prolonged. In any case, tonal signals greater than 80 dB SPL may induce some degree of auditory fatigue (see section 1.1.1.5.) or forward masking (see section 1.2.1.1.4.) resulting in threshold shifts and consequent reductions in sensitivity to subsequent signals of the same frequency. (If the signal level cannot safely be increased enough to achieve a S/N ratio that will yield the required level of signal detectability, alternative steps should be considered (see Algorithm I in section 1.2.1.2.3.). The 80 dB SPL limit recommended here applies only to tonal signals. For signals of wider bandwidth, it is the spectrum level which should not exceed 80 dB. It will be apparent from the values listed in Table II that, as signal frequency increases, the magnitude of S/N ratio at the masked threshold also increases. This occurs because the width of the band of noise that is effective in masking the signal increases as signal frequency increases, i.e., auditory selectivity decreases (see section 1.2.1.1.2.). Thus even when the noise spectrum is flat across the frequency domain, i.e., of constant spectrum level, a greater effective level of masking noise affects high frequency signals than signals of lower frequency. This illustrates that it is the spectrum level, or average level, of the noise in the immediate vicinity of the signal frequency which must be known in order to determine the effective S/N ratio. This is especially important in the case of noise spectra that depart dramatically from uniformity across frequency. The S/N ratios listed in Table II are applicable only if the noise term (N_0) in the ratio represents the average power of the noise over the band of frequencies ranging about $\pm$200 Hz on each side of the signal.

TABLE II

SIGNAL-TO-NOISE RATIOS FOR TONES AT MASKED THRESHOLD

Signal Frequency (Hz)	S/N_O in dB*
150	17.7
250	17.9
500	18.3
800	18.8
1000	19.1
1300	19.6
1800	20.4
2000	20.7
2500	21.5
3000	22.2
3500	22.8
4000	23.5
4500	24.0
5000	24.6
6000	25.6

*The quantities reported here are 10 $\log(S/N_O)$ where S is signal power required for 75 percent correct detection against a noise power per unit bandwidth N_O.

1.2.1.1.2. AUDITORY FREQUENCY SELECTIVITY: In the previous section, it was stated that frequency selectivity of the auditory system decreases as frequency increases. This means that the bandwidth of the noise that is effective in masking a signal located at the center frequency of the band increases as a function of center frequency, i.e., the effective bandwidth widens as center frequency increases. This relationship is tabulated in Table III. For example, at center frequencies of 155 Hz, 503 hz, 1,060 Hz, and 2,130 Hz, the effective bandwidths are 90 Hz, 110 Hz, 175 Hz, and 320 Hz, respectively. These effective bandwidths are known as "critical bands" (W), and they represent the range of frequencies over which the auditory system sums (integrates) noise. The importance of W for estimating the magnitude of the S/N ratio needed to achieve the masked threshold level of detectability can be illustrated as follows. Recall that N_O is the average noise power over a range of frequencies inclusive of the critical band (W). The total effective noise power that is available to mask a signal at the center of W is, therefore, simply the product WN_O. WN_O is the integral of the noise power spectrum over the range W. In case the noise spectrum is so irregular that a simple average N_O is not meaningful, the noise spectrum will have to be integrated over the range W in order to obtain a quantity equivalent to WN_O. Since the signal power S needed to be detectable at the masked threshold is approximately equal to WN_O, i.e., $S = WN_O$, to ensure

that the signal power that is specified exceeds the masked threshold for a given noise masker, it is necessary that $S > WN_0$. Thus, it is the noise only in the immediate vicinity of the signal which contributes to its masking. It is not necessary that the signal level exceed the overall noise level to be detectable; rather, it must exceed that within the critical band. For example, the overall noise level may be 110 dB SPL while the level within the critical band may be only 45 dB SPL. So long as the signal level exceeds 45 dB SPL, in this case, its detectabiblity will exceed the masked threshold.

TABLE III

CRITICAL BANDWIDTHS AND CENTER FREQUENCIES*

Center Frequencies (Hz)	Critical Bandwidth (Hz)
155	90
250	95
503	110
755	140
1060	175
1580	240
2130	320
2480	380
3120	500
4020	680
5200	920
6200	1150

* From Zwicker, E., Flottorp, G. and Stevens, S. S. (15).

1.2.1.1.3. SIGNAL DURATION: Just as in the case of quiet conditions (see section 1.1.1.3.), under noise conditions the masked thresholds of brief tonal signals decrease as their duration increases to some limit, usually reported as falling between about 200 msec and 1 sec. It is therefore recommended that minimum signal duration be specified at 1 sec to obtain the lowest possible S/N ratios at the masked threshold.

1.2.1.1.4. FORWARD AND BACKWARD MASKING: In these two forms of masking, signals and noise are not simultaneously present at the ear. Forward masking is similar to auditory fatigue (see section 1.1.1.5.), in that the threshold for a given signal is elevated by previous exposure to noise. The magnitude of threshold elevation increases as the time interval between noise offset and signal onset decreases. The threshold of a very brief signal (e.g., 5 msec) that follows offset of a 90 dB SPL noise by no more than about 2 msec may be elevated by more than 50 dB. If the interval is lengthed to 15 msec, the threshold elevation will diminish to about 10 dB. Only marginal forward masking seems to occur for intervals greater than about 50 msec. Approximately

equivalent results are obtained in backward masking where the onset of the noise masker occurs after offset of the signal. Obviously, both forward and backward masking can be avoided by separating signals and noise in time. It appears that intervals of separation greater than 50 msec are sufficient. If temporal separation of signals and noise by intervals greater than 50 msec cannot be achieved, it may be possible (if the noise is confined to a narrow band) to move the signal away from the noise in the frequency domain. As is the case in auditory fatigue (section 1.1.1.5.), if signal and masker are separated by 1 to 1.5 octave, little or no threshold elevation should occur.

1.2.1.2 BINAURAL DETECTION: The detectability of signals under binaural conditions involving interaural dichotic imbalances is superior to the detection that may be achieved with the same signals under either monaural or binaural diotic conditions, i.e., conditions involving no interaural imbalances. The binaural advantage may be as great as 20 dB. The interaural imbalances responsible for this superiority of binaural over monaural detection are interaural time (or phase) and intensity differences which serve as potent cues for signal detection. Only amplitude increments are available as detection cues under monaural conditions (listener familiarity with the signal may aid detection under either monaural or binaural conditions). The relative difference in the detectability of signals in noise under binaural and monaural conditions is designated as the "masking level difference", or MLD. The MLD is, simply, the difference in decibles between the signal-to-noise ratios required to achieve a given level of detection (e.g., 75 percent correct detection) under binaural and monaural (or binaural diotic) conditions. The pragmatic importance of the MLD is that it represents an improvement in the detectability of signals in noise which may be achieved without increasing S/N ratio. Creation of the interaural imbalances necessary to produce MLDs may be accomplished most readily when signals and noise are presented to the two ears through a pair of headphones.

1.2.1.2.1. INTERAURAL IMBALANCES: If a signal presented in noise to one, or both ears results in an interaural imbalance, then that signal will be more readily detectable than if no imbalance occurs. For example, assume that a S/N ratio of 18 dB is necessary to attain 75 percent correct detection performance when a 500 Hz signal is briefly added to noise at one ear alone. Now, if a duplicate of the noise is also presented to the other ear such that the interaural correlation of the two noises is +1, then a S/N ratio of only 10 dB would be needed in order to achieve the same level of detectability as when the same signal and noise were presented to one ear only. In this case the MLD would be 8 dB. Merely the addition of +1 correlated noise at the non-signal ear reduced the S/N ratio required for 75 percent correct detection by 8 dB from what it was in the purely monaural condition. This amounts to more than a 6-fold reduction in signal power. The explanation is this: With +1 correlated noise at both headphones, the acoustical waveforms at the two ears are in near-perfect syncrony. When the signal is then presented to

one ear, an interaural phase shift as well as an amplitude increment occurs. Under the monaural condition, only the amplitude increment is contingent upon signal occurrence. The interaural phase shift obtained with +1 noise at both ears thus contributes powerfully to detectability of the signal. This cue may be eliminated by presenting the signal in-phase at the two ears. In this case, both noise and signals are in near-perfect interaural syncrony (diotic condition) and no interaural imbalance is contingent upon occurrence of the signal. In this example, the S/N ratio required for 75 percent correct detection would be 18 dB, just what it was in the purely monaural condition. The MLD would be 0 dB. If, however, the same signal is added 180 degrees out of phase at the two ears to the +1 correlated noise, a very large interaural phase shift would occur and the necessary S/N ratio would be about 0 dB yielding an MLD of 18 dB. This is the equivalent of a 63-fold reduction in signal power from that required for equally detectable monotic or diotic signals. In this example, only three of the many possible interaural conditions were discussed: They were NOSm (noise diotic, signal monotic), NOSO (noise diotic, signal diotic), and NOSπ (noise diotic, signal dichotic by 180 degrees). Table IV summarizes the various interaural temporal relationships of signals and noise known to yield MLDs. The magnitude of the MLD that can be obtained by manipulating interaural temporal relations ranges between the zero MLD conditions (NmSm, NπSπ, NOSO) and the extreme antiphasic conditions (NπSO and NOSπ), a range of 14-20 dB for signals below about 800 Hz.

TABLE IV

INTERAURAL TEMPORAL RELATIONS BETWEEN SIGNALS AND NOISE

Nm: Noise monotic
NO: Noise diotic ($\alpha=1$; $\theta=0^{\circ}$) or dichotic ($\alpha\neq1$; $\theta=0^{\circ}$).
Nπ: Noise dichotic ($\alpha=1$ or $\alpha\neq1$; $\theta=180^{\circ}$).
Nτ: Noise dichotic ($\alpha=1$ or $\alpha\neq1$; $\tau>0$) or diotic ($\alpha=1$; $\tau=0$).
Np: Noise dichotic ($\alpha=1$ or $\alpha\neq1$; p<+1) or diotic ($\alpha=1$; p=+1).
Nu: Noise dichotic ($\alpha=1$ or $\alpha\neq1$; p=0).

Sm: Signal monotic.
SO: Signal diotic ($\alpha=1$; $\theta=0^{\circ}$) or dichotic ($\alpha\neq1$; $\theta=0^{\circ}$).
Sθ: Signal dichotic ($\alpha=1$ or $\alpha\neq1$; $\theta>0^{\circ}$).
Sπ: Signal dichotic ($\alpha=1$ or $\alpha\neq1$; $\theta=180^{\circ}$).
Sp: Signal dichotic ($\alpha=1$ or $\alpha\neq1$; p<+1) or diotic ($\alpha=1$; p=+1) noise.

α: Interaural intensity ratio.
θ: Interaural phase difference.
τ: Interaural time delay.
p: Normalized interaural correlation coefficient.

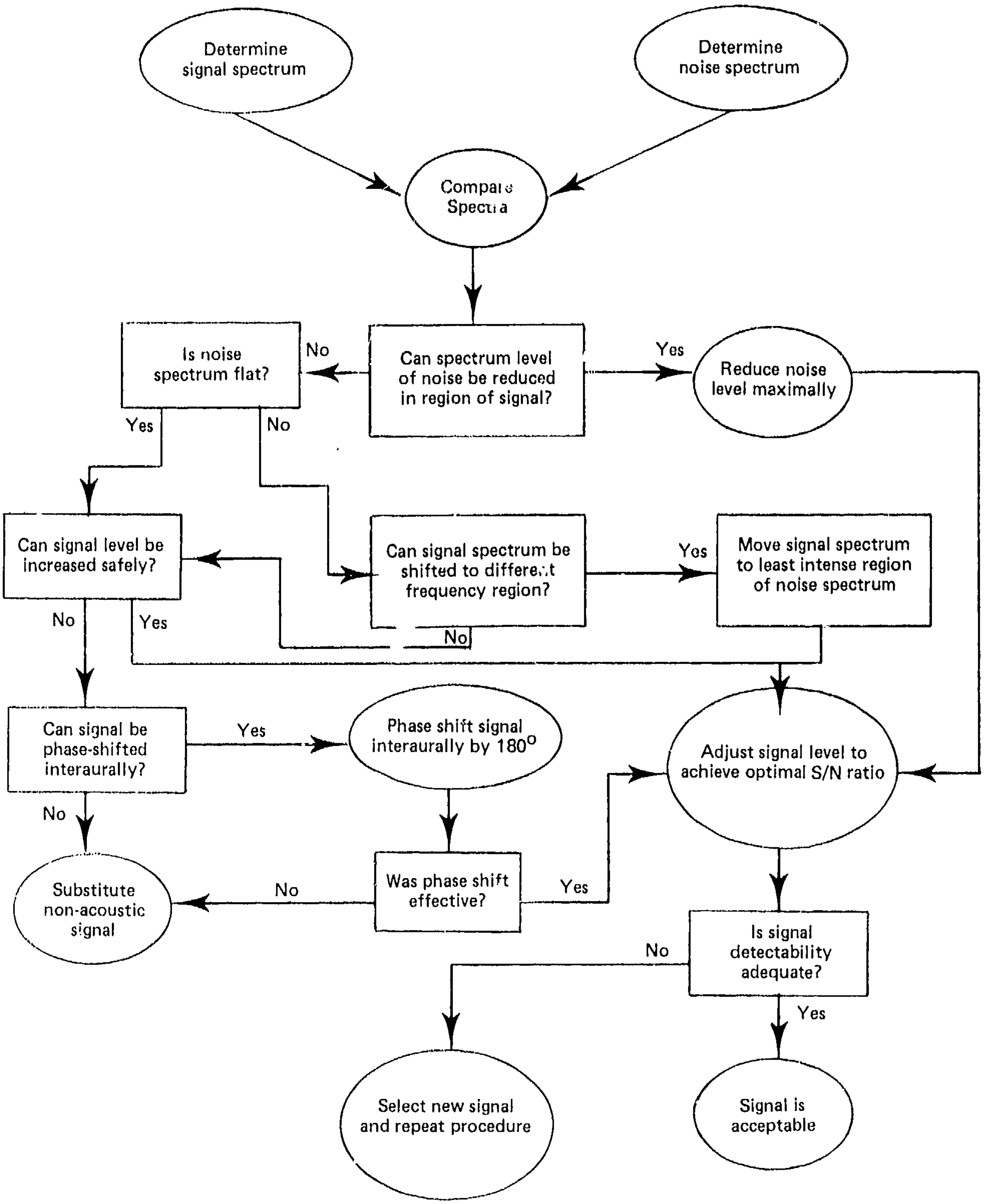
Typical combinations of the above include: NmSm, NOSm, NOSO, NOSπ, NOSθ, NpSO, etc.

1.2.1.2.2. DICHOTIC SIGNAL FREQUENCY: The size of the MLD that can be obtained with any combination of signal and noise temporal relations depends on signal frequency. The maximum MLD seems to occur in the region of 250 Hz, dropping off rapidly as signal frequency is reduced. At frequencies above 250 Hz, the drop in MLD is less rapid. For example, between 250 Hz and 500 Hz, the MLD declines by about 3 dB. Between 250 Hz and 1000 Hz, the MLD declines about 8 dB. Therefore, if phase-shifting the signal is used to improve the detectability of signals in noise, best results can be expected if the signals are low-frequency, i.e., near 250 Hz. Little improvement can be obtained by this method for dichotic signals below about 150 Hz or above 1500 Hz.

1.2.1.2.3. BINAURAL FREQUENCY SELECTIVITY: The role of the critical band in monaural detection of signals in noise was discussed in the section on auditory frequency selectivity (section 1.2.1.1.2.). There it was shown that the critical band increases as a function of center frequency (see Table III) which accounts for the increase with frequency in S/N ratio needed to reach masked threshold (see Table II). Recall that it is only the noise within the critical band centered on the monaural signal frequency that is responsible for masking that signal. Likewise, for binaural signals, it is only the noise within corresponding critical bands at the two ears that influences detection. Under ordinary headphone listening conditions, it is likely that the noise spectra at the two ears will be nearly the same and, consequently, corresponding critical bands will receive essentially +1 correlated noise. Under free-field listening conditions, turning of the head relative to the noise source may alter the correlation of the noise at the two ears (time-delays may be translated into correlations), but the spectral distributions of noise energy within corresponding critical bands will remain unchanged with head movements. It is only in the unlikely event that corresponding bands receive very different noise spectra (e.g., through headphones) that a problem might arise. In this case, the relations itemized in Table IV are not applicable. In any case, it is only the spectrum of noise in the immediate vicinity of the signal frequency that need be of concern. If the noise spectrum is narrow, the best strategy may be to move the signal frequency away from the noise thereby improving the effective S/N ratio. This is but one of several strategies that may be used to enhance signal detectability as is illustrated in Algorithm I.

ALGORIGHM I

PROCEDURE TO ENHANCE DETECTABILITY OF SIGNALS IN NOISE



1.2.1.2.4. BINAURAL SIGNAL DURATION: The same considerations apply in the case of binaural signal durations as in the case of monaural signal durations (see section 1.2.1.2.3.). Phase-shifting interaurally may, however, be employed as an alternative strategy if signal duration cannot be increased. If signal frequency is low and duration brief (e.g., 500 Hz signal of 150 msec duration), a simple phase-reversal across the headphones may be as effective in improving detectability as a 10-fold increase in duration.

1.2.1.2.5. FORWARD AND BACKWARD MASKING: Precisely the same constraints apply for binaural signals and maskers as for their monaural counterparts (see section 1.2.1.1.4.). Since phase-shifting is effective in improving signal detectability only when both signals and noise are simultaneously present, the conditions where maskers either precede (forward masking) or follow (backward masking) the signal are equivalent for binaural and monaural signals. In both cases, signal detectability may be improved by moving the signal away from the noise in time.

2.2. COMPLEX SIGNALS: As was indicated in section 1.1.2., any signal consisting of more than a single frequency is considered complex. Most natural and machine-produced sounds are of a complex nature. The detectability of such signals in noise depends not only on their spectra at any moment in time, but also on their time-varying properties (e.g., amplitude and/or frequency modulation). Fortunately, the same principles apply for the detection of complex signals as for tonal signals (i.e., psychometric functions relating detection performance to S/N ratio are of the same shape; interaural temporal imbalances result in improved detectability, etc.). However, because such signals may occur in nearly an infinite variety, it is not possible to provide *a priori* specifications. Rather, taking into account the principles that govern the detectability of tonal signals in noise, parameters appropriate for the particular signals in question may be determined empirically. With respect to speech signals, some standards have been developed for evaluating speech interference (3) for measuring word intelligibility (2) and for determining an articulation index (4). As these standards suggest, the interest in speech signals is not limited merely to their detectability but extends to reception of their informational content. Obviously, an acoustic signal that is recognizable must also be detectable. The reverse does not apply. Signals may be detectable at levels below those needed for the more complete processing involved in recognition. If the listener's task is to identify one among several signals that may occur against a noise background, a higher S/N will be required than if the task simply is to determine the occurrence of a signal. In any case, the S/N ratio that will be necessary to achieve the desired performance will depend not only on the parameters of the signal, but also those of the noise, and these must be known before an effective S/N ratio can be specified. In the case of non-speech acoustic signals, it is essential that their power spectra be given. If spectra undergo any changes as a function of time (as in modulated waveforms), the defining

parameters of these changes should be stated. Essentially the same requirements exist for specifying the background noise against which signals are to be presented. Wherever possible, signal spectra should be positioned in regions of the noise spectrum containing least energy to permit the choice of detectable signal intensities that do not exceed comfortable listening levels. In any case, both signal and noise levels should be specified in terms of sound pressure levels present at the listener's ears.

2. LOUDNESS: Although loudness of sound increases monotonically as a function of sound intensity, loudness is also influenced by parameters of sound other than intensity. Sounds of different frequencies may be perceived as being of different loudness even though their intensities are the same. The loudness of brief sound may increase as its duration increases, whereas the loudness of a prolonged sound may decrease as its duration increases. Furthermore, the presence of a masking noise may reduce the loudness of a signal. Consequently, loudness may not be considered as bound invariantly to a single physical dimension of sound, i.e., intensity. This is particularly important when loudness needs to be increased without increasing intensity (as illustrated in Algorithm II). One should also be aware that the form of the relationship between loudness and intensity (for a given signal frequency) depends upon how loudness is measured. The most widely accepted scale for loudness is the sone scale. Unit loudness, one sone, is defined as the loudness of a 1 kHz tone at 40 dB above absolute threshold. The function relating loudness in sones of a 1 kHz tone to sound pressure level (SPL) in decibles (plotted on log-log coordinates) is negatively accelerating, becoming approximately linear for SPLs greater than about 30 dB above absolute threshold (see Table V). The significance of this function is that it serves as a standard yardstick against which the loudness of any sound may be measured. If, for example, in order to match the loudness of some sound against the loudness of a 1 kHz tone, the latter has to be set at 50 dB SPL, then the loudness in question will be 2 sones. This procedure is analogous to matching the lengths of various objects to the scale values of an ordinary ruler.

ALGORITHM II

PROCEDURE TO INCREASE LOUDNESS WITHOUT INCREASING LEVEL

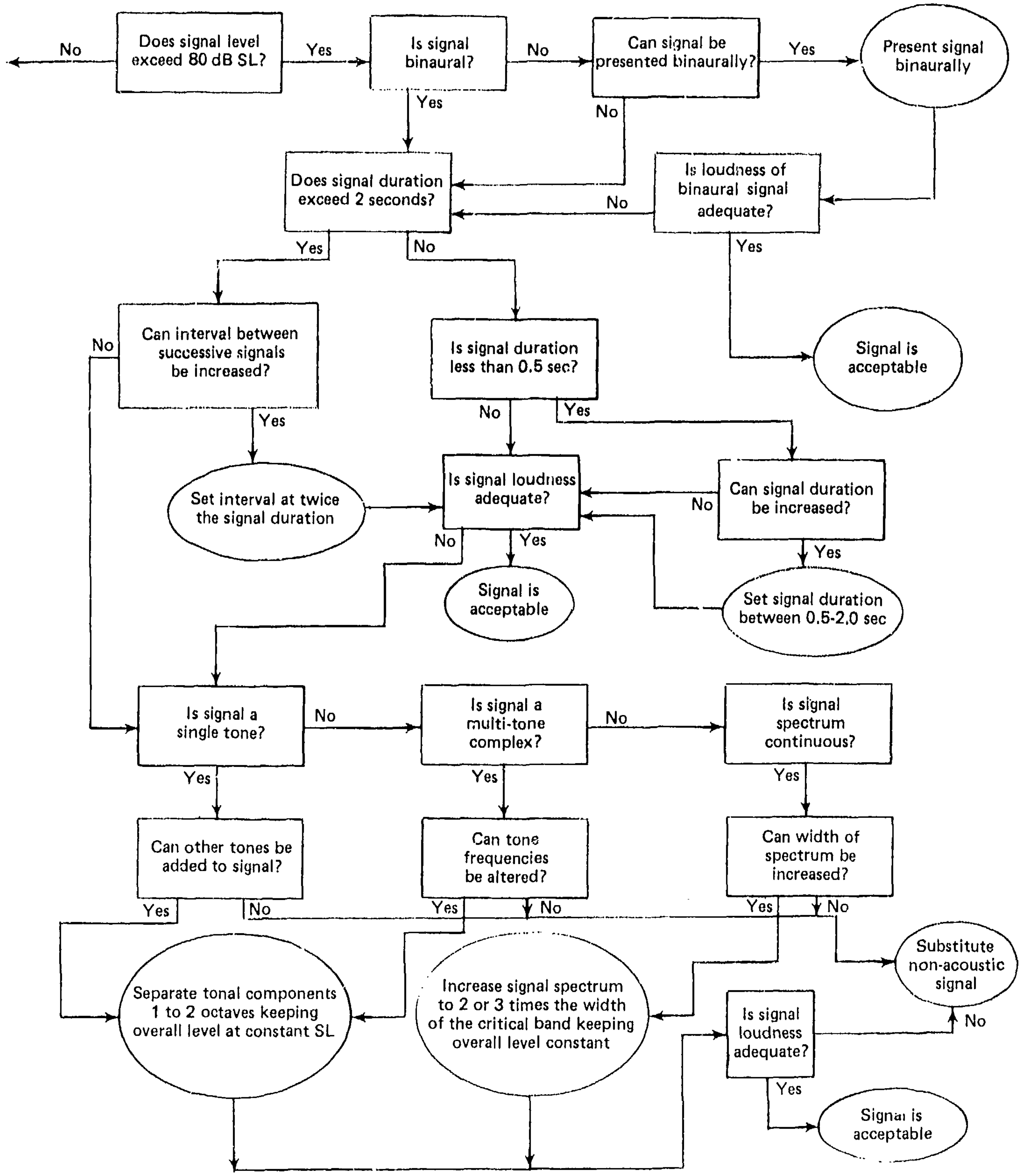


TABLE V

SONE VALUES FOR 1000-Hz TONE AND FOR WHITE NOISE *

SPL dB	Tone	Noise	SPL dB	Tone	Noise	SPL dB	Tone	Noise
10	.052	-	50	2.00	3.85	90	32.0	46.0
12	.072	-	52	2.30	4.45	92	36.8	50.5
14	.095	-	54	2.64	5.20	94	42.2	57.5
15	.110	-	55	2.83	5.60	95	45.3	61.0
16	.125	-	56	3.03	6.00	96	48.5	65.0
18	.155	-	58	3.48	7.00	98	55.7	72.0
20	.190	-	60	4.00	7.85	100	64.0	80.0
22	.230	-	62	4.59	8.90	102	73.5	91.0
24	.280	-	64	5.28	10.20	104	84.4	102.0
25	.103	-	65	5.66	10.90	105	90.5	108.0
26	.330	-	66	6.06	11.50	106	97.0	114.0
28	.395	.450	68	6.96	13.00	108	111.0	128.0
30	.460	.580	70	8.00	14.70	110	128.0	-
32	.550	.720	72	9.19	16.40	112	147.0	-
34	.640	.900	74	10.60	18.50	114	169.0	-
35	.700	1.000	75	11.30	19.50	115	181.0	-
36	.750	1.100	76	12.10	20.60	116	194.0	-
38	.860	1.360	78	13.90	23.20	118	223.0	-
40	1.000	1.650	80	16.00	26.00	120	256.0	-
42	1.115	2.000	82	18.40	29.00			
44	1.320	2.400	84	21.10	32.50			
45	1.410	2.600	85	22.60	34.80			
46	1.520	2.800	86	24.30	36.50			
48	1.740	3.280	88	27.90	41.00			

* From Scharf, B.(12).

2.1. MONAURAL LOUDNESS: Most parameters of sound that exert any differential influence on loudness are equally effective under monaural and binaural conditions. These influences thus may be regarded as monaural parameters. They include signal frequency (or spectra), duration, and masking of both tonal and complex signals.

2.1.1. TONAL SIGNALS: Parametric studies of loudness typically have utilized tonal signals which permit precise mapping of loudness relations throughout the range of human hearing. Both the rate at which the loudness of tones grows with increasing intensity and the intensity required to maintain constant loudness change systematically as tonal frequency is changed. The pragmatic importance of this is that it enables determinations of the dynamic (loudness) ranges available at certain frequencies and permits determinations of SPLs required to make tonal signals equally or differentially loud.

2.1.1.1. SIGNAL FREQUENCY: Since the loudness of tones, especially tones below about 200 Hz and above about 4,000 Hz, changes as frequency is changed even though SPL remains constant. Therefore, in order to specify the levels at which two or more tones will be judged equally loud, equal loudness contours should be consulted. Two sets of such contours are available; those provided originally by Fletcher and Munson (5) for tones presented through headphones, and those provided by Robinson and Dadson (11) for tones presented in a free field. From these contours may be obtained the sound pressure levels required for tones to be heard at specified loudness levels ranging between absolute threshold and about 120 phones. The unit of loudness level is the phone, where the number of phones is equal to the number of decibles above absolute threshold to which the SPL of a 1,000 Hz tone must be set to match the loudness of another sound. As in the case of the sone, the 1,000 Hz tone is used as the standard in terms of which loudness levels of other sounds are measured. Phones may be converted to sones by regarding the SPLs for the 1,000 Hz tone in Table V as phones. This is a valid procedure since both SPLs and phones are expressed in decibles relative to the same standard reference. For phones the reference is the sound pressure at absolute threshold for a 1,000 Hz tone, i.e., 20 $\mu N/m^2$, which is the same reference for SPLs. Thus, for a signal frequency of 1,000 Hz, the number of phones is its SPL. Corresponding sone values are listed in the column adjacent to SPLs in Table V. As an example, assume that we want to set the loudness of 500 Hz and 4.000 Hz tones equal to 2 sones. From Table V we find that the loudness of a 1,000 Hz tone at 50 dB SPL is equal to 2 sones, i.e., its loudness level is 50 phones. Turning to the equal loudness contours of Robinson and Dadson (11), we find that the SPLs of 500 Hz and 4,000 Hz tones must be about 47 dB and 33 dB, respectively, to attain a loudness level of 50 phones (or a loudness of 2 sones). Thus, if 500 Hz, 1,000 Hz, and 4,000 Hz tones are all to be presented in a free field (say, through a loudspeaker) at 2 sones loudness, then their respective intensities must be set to produce sound pressure levels at the listeners ears of 47, 50, 33 dB. Use of Table V in conjunction with the equal loudness contours may also be extended to assessments of the loudness of tones of known SPLs. For example, if a 500 Hz tone is presented at 37 dB SPL and a 4,000 Hz tone is presented at 52 dB SPL, the loudness levels of these two tones will be 40 phones and 60 phones, respectively, and their loudnesses will be 1 sone and 4 sones. The 4,000 Hz tone will thus be 4 times louder than the 500 Hz tone. If both tones are presented at the same sound pressure level (say, 37 dB), the loudness of the 4,000 Hz tone would be only about 1.3 times greater than that of the 500 Hz tone. It should be clear from the above examples that equating the intensities of tones of different frequencies does not result in equal loudness.

2.1.1.2. SIGNAL DURATION: The loudness of brief signals tends to increase as duration increases up to some limit between about 50 and 200 msec. Put another way, the sound pressure level required to maintain a constant loudness decreases as signal

duration increases. Although the loudness of tones of different frequencies may grow as a function of duration up to different limits, research has not yet provided a systematic relationship between such temporal limits and signal frequency. At present, it seems safe to recommend that signal durations should be at least 0.5 sec in order that loudness be stable. The decrease in SPL required for constant loudness as duration increases from about 1 to 70 msec may be as much as 20 dB. Hence, an increase in the duration of brief signals may be an effective way of increasing loudness under conditions where signal level cannot be safely increased, as illustrated in Algorithm II.

2.1.1.3. AUDITORY ADAPTATION: At the opposite extreme from the increase in loudness which occurs when the durations of very brief signals increase, there may occur a decrease in loudness during prolonged exposure to a signal due to adaptation of the auditory system. Although loudness adaptation seems to represent essentially the same process as that involved in auditory fatigue (see section 1.1.1.5.), adaptation refers to decreases in loudness that occur during stimulation where fatigue refers to decreases in sensitivity that are evident after cessation of stimulation. The same parameters are important in both cases, i.e., frequency, intensity, and duration of the exposing sound and the frequency and temporal proximity of the test sound to the exposing sound. The most rapid adaptation occurs within the first 30 seconds of continuous stimulation but loudness may continue to decrease for as long as several minutes in the case of intense stimulation. As much as 40 dB adaptation has been obtained with stimulation at 80 dB above threshold. About 70 percent recovery occurs after about 1 min of quiet; complete recovery is realized within several minutes. The region of adaptation appears to be confined within 1 to 1.5 octave of the adapting stimulus. Given these data, if it is important that the loudness of a signal remain constant, it is recommended that signal durations not exceed more than 1 or 2 seconds and that signal intensities be set below 80 dB above absolute threshold. If signal duration and intensities must exceed these limits, it is recommended that the signal be composed of frequencies separated by about 1 or 2 octaves and presented alternately for durations of no more than about 1 second. These considerations have been incorporated into Algorithm II.

2.1.1.4. LOUDNESS MASKING: The loudness of tones presented against a noise background will be less than the loudnesses of the same tones presented at the same SPLs in quiet. The noise effectively raises the tonal threshold, and loudness becomes approximately proportional to signal-to-noise ratio. For example, a 1,000 Hz tone presented at 80 dB SPL against a white noise, the overall level of which is 90 dB, would be matched in loudness by the same tone at about 50 dB SPL presented without the noise, a reduction in loudness level by about 30 phones. In addition to requiring that signal levels be increased to achieve a given loudness, the presence of noise also increases the slope of the loudness function, i.e., it increases the rate at which loudness grows as a function of intensity, ultimately reducing

the dynamic range of the loudness function. Because systematic data on these effects are not available for signals over a range of frequencies and S/N ratios, it is not possible to stipulate with any accuracy a specific correction which might be employed to offset the presence of noise. A best guess would be as follows: for every 10 dB increment in noise above threshold, there should be an increment of 10-20 dB in the signal. A precise assessment of the reduction in tonal loudness due to noise could be made by presenting the tone-plus-noise through one headphone and matching this tone's loudness with that of a duplicate tone (or 1,000 Hz tone) presented in quiet through another headphone to the opposite ear. The difference in tone SPLs required to achieve a loudness match would indicate the reduction in loudness level due to the noise in question.

2.1.2. <u>COMPLEX SIGNALS</u>: As in other sections of this report (see sections 1.1.2. and 1.2.2.), the term complex is applied to any signal consisting of multiple frequency components. While the relationships discussed in previous sections for tonal signals generally hold for complex signals, there is one aspect of the loudness of complex signals that is unique to them, i.e., monaural summation of loudness with increments in signal bandwidth.

2.1.2.1. <u>SUMMATION WITHIN CRITICAL BANDS</u>: The relationship of critical bandwidth to center frequency is given in Table III. Due to the frequency selectivity of auditory processing (see section 1.2.1.1.2.), all of the signal energy that falls within a critical bandwidth is summed (integrated). This means that each frequency component of the signal within a critical band will contribute to its loudness. Even if the energy contribution of all components are equal (flat spectrum), as the signal's bandwidth is increased by adding components on each side of the center frequency (note that adding components on just one side would shift the center frequency toward the side of the addition), the overall power within the band increases as does its loudness. For example, if the signal is centered at 1,000 Hz, the critical band there will be about 175 Hz wide. A signal that ranges ± 25 Hz on either side of 1,000 Hz will be less loud than one that ranges ± 50 Hz about the center frequency even if the components outside ± 25 Hz are less intense. The ear simply sums all the energy present within the critical band.

2.1.2.2. <u>SUMMATION OUTSIDE CRITICAL BANDS</u>: From the preceeding section, it is clear that the greater the energy within a critical band, the greater is the signal's loudness due to simple energy summation. However, loudness summation may result in louder signals even if the energy level within one critical band is reduced. This occurs when the signal bandwidth is greater than the critical band on which it is centered. For example, if the overall level of a signal centered on 1,000 Hz is held constant, as its bandwidth is increased up to about ± 87 Hz (the width of the critical band), the loudness will remain constant because the average energy in each component has to be reduced as additional components are added in order to keep the overall

level constant; i.e., signal loudness remains constant as signal bandwidth increases because overall level within the critical band remains unchanged. If this procedure of adding components on each side of the signal while reducing the average level of components to keep their overall level constant is continued beyond the width of the critical band, loudness begins to increase and continues as signal band increases beyond the critical band. This illustrates an interesting finding, viz., that loudness increases as a function of signal spectrum width even if the spectrum level of the signal is reduced. One practical consequence of this is that loudness adaptation may be avoided in the case of signals of long durations. Since loudness adaptation appears to be mainly confined to effects within critical bands, adaptation may be prevented by using wider band signals without sacrificing loudness. The lower spectrum levels of wide band signals adjusted to yield the required loudness would induce less adaptation than if all the signal's energy were concentrated within one critical band. This would be especially important in the case of signals of durations longer than several seconds of continuous presentation. In such cases, it is recommended that the signal bandwidth be set several times that of the critical band at its center frequency. The loudness in sones of a wide band noise has been listed in Table V as a function of overall SPL and a simplified procedure for calculating the loudness in sones of various noises has been developed (1).

2.2. BINAURAL LOUDNESS: If a signal is presented simultaneously to both ears, its loudness will be approximately twice that of the monaural signal alone. Consequently, less intense binaural signals would be preferred over purely monaural signals if the presence of masking noise requires that monaural signals be presented at uncomfortable or adapting intensities. Table V should be consulted to determine the change in SPL that a doubling in loudness represents. For example, a binaural tone of 1,000 Hz at 40 dB SPL would be equal in loudness to a monaural tone of the same frequency at 50 dB SPL. Obviously, this amounts to a reduction of 10 dB in the level of the binaural signal required to match the loudness of its more intense monaural duplicate. This 10 dB saving per doubling of loudness holds for tones as intense as 120 dB SPL, but not for noise. In the case of noise signals, Table V should be consulted. Binaural loudness summation may be particularly useful as a means of increasing the loudness of signals under conditions where signal levels cannot be increased, as illustrated in Algorithm II.

3. DISTINCTIVENESS: This section considers the primary parameters responsible for discrimination between acoustic signals and the organization of their components into perceptual patterns. These parameters are monaural intensity and frequency differences, interaural time and intensity differences, angle of origin (or directional) differences, and organizational factors simultaneously and/or sequentially present in complex sound arrays.

3.1. MONAURAL DISCRIMINATION: With the exception of interaural time and intensity differences which determine the localizability of sound sources in auditory space, it appears that all discriminable parameters of sound may be resolved by one ear. This includes spectral and sequential sound patterns which are discussed in a separate section (see section 3.3.). In a practical sense, the data describing monaural discrimination is of interest only for the condition where one ear is stimulated, e.g., by means of a single headphone. However, in most instances interaural resolution of the acoustic differences discussed in this section are no greater than monaural resolution and, consequently, limits on the latter may be taken as applicable under either monaural or binaural conditions.

3.1.1. FREQUENCY RESOLUTION OF TONES: The discriminability of frequency, or pitch, differences between tones has been shown to depend not only on frequency, but also on the level of tones above absolute threshold (sensation level, SL), on signal-to-noise (S/N) ratio, and on tone duration. Unlike the pitch, or frequency difference limen (Δ f) which may change substantially due to variation in SL or S/N ratio, the pitch of individual tones remains somewhat more stable (at least in the region 1-3 Hz) over a considerable range of intensities. Before examining pitch discrimination, it will be useful to become familiar with the relationship between pitch and tone frequency.

3.1.1.1. PITCH OF TONES: The unit of pitch is the mel which is defined as follows: 1,000 mels is the pitch of a 1,000 Hz tone 40 dB above absolute threshold. This unit is more a scaling convenience than a measurement device, i.e., it is not possible to change the pitch of a 1,000 Hz tone to match that of a tone of very different frequency. However, the mel scale may be taken as a rough index of the relationship of pitch to frequency of tones. This is given in Table VI. The mel scale may be useful in estimating approximately how much "higher" the pitches of tones in one frequency region are as compared with the pitches of tones in a lower frequency region. For example, a 4,000 Hz tone is a little more than twice the pitch of a 1,000 Hz tone (2,250 mels *vs* 1,000 mels) while a 1,000 Hz tone is just 1/3 as high in pitch as a 9,000 Hz tone (1,000 mels *vs* 3,000 mels). Note that above 1,000 Hz pitch changes vary gradually, although approximately linearly, with changes in frequency. The most dramatic pitch changes occur in the low frequencies. This relationship between pitch and frequency should be kept in mind especially if the pitches of two or more sounds must be readily recognized. Here the problem is not one of merely ensuring that the signals are discriminably different, but rather it is a matter of ensuring that the signals are of sufficiently different pitches that they will not be confused. Usually a pitch ratio of 2 to 1 would be more than adequate. For example, the pitches of 400 Hz and 1,900 Hz tones are approximately 2:1, as are the pitches of 700 Hz and 2,000 Hz tones. The point is that, if pitch differences are to be used to make signals individually recognizable, they must be considerably larger than frequency difference limens (Δ f).

TABLE VI

RELATIONSHIP OF PITCH IN MELS TO FREQUENCY IN Hz*

Frequency (Hz)	Mels	Frequency (Hz)	Mels
100	161	4,000	2,250
200	301	5,000	2,478
400	508	6,000	2,657
700	775	7.000	2,800
1,000	1,000	8,000	2,911
1,500	1,296	9,000	3,000
2,000	2,545	10,000	3,075
3,000	1,962		

*From Stevens, S. S. (14).

3.1.1.2. PITCH DISCRIMINATION IN QUIET: Where signals are presented in close temporal contiguity, it may be useful to know the minimal frequency difference (Δ f) that can just be discriminated under quiet conditions by practiced listeners. These values of Δ f may be used as minimal, or ideal, frequency differences. Certainly, it should not be expected that listeners would resolve pitch differences between signals separated by frequencies closer than Δ f. Values of Δf for a range of signal frequencies are given in Table VII for a constant SL of 5 dB, and for a single frequency of 250 Hz over a range of sensation levels. As the values listed in Table VII for 250 Hz illustrate, the size of Δf decreases as tonal intensity increases. Little change in Δ f occurs above 60 dB SL. Also, the size of Δ f remains roughly constant, for a given SL, for frequencies below 1,000 Hz, but Δf increases as f increases above 1,000 Hz. Thus, discrimination of frequency, or pitch, differences is best at low frequencies and at moderate to high SIGNAL levels. The values of Δ f given in Table VII apply only for tonal signals of different frequencies that are alternately presented in rapid succession. Pitch memory is not sufficiently acute to permit such fine discriminations if the time between successive tone presentations is much greater than about 20 msec. It should be noted also that the values of Δ f in Table VII are valid only under quiet conditions. Larger values of Δf are required in the presence of noise.

TABLE VII

FREQUENCY DIFFERENCE LIMENS (Δf) FOR TONES*

5 dB SL		250 Hz	
f in Hz	Δf in Hz	SL in dB	Δf in Hz
125	7.8	5	9.0
250	9.0	10	5.5
500	8.5	20	3.3
1,000	9.5	40	2.8
2,000	16.0	60	2.4
4,000	26.0		

*From Shower, E. G. and Biddulph, R. (13).

3.1.1.3. PITCH DISCRIMINATION IN NOISE: It appears that the size of the difference limen Δf increases as tone levels decrease relative to the level of noise in their immediate vicinity. This increase in Δf with decreasing S/N ratio is roughly equivalent to what happens under quiet conditions when SL is decreased toward the absolute threshold. The changes in Δf at 250 Hz listed in Table VII illustrate the magnitude of reduction in discriminability of pitch differences that may occur as SL, or, by extension, S/N ratio is reduced. Since, under actual operating conditions, factors other than just noise (both acoustic and non-acoustic distractors) are likely also to diminish the acuity of pitch resolution, it seems necessary that the values of Δf required for reliable discrimination be determined under actual conditions. This is especially important in the case of signals involving sequences of necessarily distinguishable pitches, i.e., pitch patterns (see section 3.3.2.). Two or more such patterns may be clearly distinguishable in quiet, but may be confused in noise due to failure to resolve the successive pitch changes peculiar to each pattern. In such cases it may be possible to solve the problem simply by increasing tone levels. If this is not feasible, then tone frequency differences will have to be increased to make the pitch changes reliably discernable. However, this may alter the pitch pattern unacceptably if frequency ratios of signal components are changed significantly. It should also be kept in mind that, even under quiet conditions, if the components of multi-tone complexes (see section 3.1.2.) are to be individually identifiable, they should be separated in frequency by no less than one bandwidth (see Table III), and the number of components should be no more than 5 to 7. The problem of component identification (e.g., identification of the harmonics of complex sounds) is thus a more informationally demanding task than simple pitch discrimination and frequency differences must be several times larger than Δf whether or not noise is present in the signal channel.

3.1.1.4. TONE DURATION: Values of Δf are affected by tone duration only in the case of very brief tones, e.g., less than about 50 msec for tones below 1,000 Hz and less than about 25 msec for tones in the vicinity of 4,000 Hz. In such cases the sizes of Δfs will be larger than the Δfs achievable with longer duration signals. If, as in the case of signals with rapidly alternating components, durations of components cannot be increased as a means of reducing Δf, then to ensure discriminability larger frequency separations will be required. These should be empirically established under conditions that approximate the noise existing within the operational environment.

3.1.2. FREQUENCY RESOLUTION IN COMPLEX SOUNDS: Unlike simple pitch discrimination among tones of slightly different frequencies, identification of the pitches of the components of complex sounds is more difficult and requires that the components be spaced at greater frequency intervals. In part this is due to the fact that the individual components of a complex sound are all present simultaneously so that the listener's task is not just one of distinguishing between the pitches of alternately presented tones, but rather it is one of filtering out individually identifiable pitches on-going within the complex. Of interest here are the minimal frequency separations between components of tonal complexes that are necessary for their individual resolution. The answer seems to depend on the number of components contained in the complex, i.e., more components require larger frequency separations. In the case of a two-tone complex, the minimum separation necessary for individual pitches to be discerned is about one-fifth the width of the critical band (see Table III). For a three-tone complex the minimum separation is about one-third of the critical bandwidth. For five- to seven-tone complexes, the minimum separation is one critical bandwidth. It appears that no more than seven tonal components can be identified in a complex. It should be pointed out that resolution of individual components within complex sounds is not necessary in order for different complexes to be distinguishable (see section 3.3.1.). The problem of individual component resolution is of practical interest when, for example, one component in a complex serves as the signal for some event, or when some relationship among several components serves as the signal. As an illustration, presentation of a tone higher in pitch than that of an on-going tone may mean "to the right of" while presentation of a tone lower in pitch may mean "to the left of." Since the listener must be able to hear both pitches simultaneously, the two certainly must be resolvable. It is noteworthty that spatial relationships can be represented by utilizing the relative properties of pitch (and pitch changes) within multi-tone complexes.

3.1.3. INTENSITY RESOLUTION OF TONES: Whether the question of interest is how large an intensity fluctuation can be tolerated in a signal for its intensity to be regarded as acceptably constant, or how large an intensity increment must be in order for it to be detectable, the best answer available is the

intensity difference limen (Δ I). In decibles, the intensity difference limen is usually expressed in one of two forms, either as an absolute difference limen (ADL),

$$ADL = 10 \text{ Log } \Delta I/Io,$$

or as a relative difference limen (RDL),

$$RDL = 10 \log (\Delta I + I)/I,$$

where ΔI is the magnitude of the just detectable intensity increment,

I is the magnitude of the basic intensity from which the increment is made,

and Io is the intensity of the signal at absolute threshold.

It may be helpful to recall that the sensation level (SL) of a signal is the number of decibles that its intensity (I) is above its threshold (Io), i.e.,

$$SL = 10 \log I/Io.$$

Since both ADL and SL are expressed relative to the same reference term, viz., Io, precisely the same relationship exists between ADL and SL that exists between Δ I and I. The general nature of this relationship is such that ADL increases as a function of SL, i.e., greater intensity differences are required for discrimination at higher intensities. Stated more accurately, as SL increases from low to moderate levels (e.g., 30 to 40 dB SL), ADL increases with a positive acceleration. At higher levels, ADL increases approximately linearly as a function of SL. This does not mean that the auditory system is less efficient in resolving intensity differences at higher intensities, even though larger intensity increments are required to be discriminable. In fact, relative to the magnitudes of the higher intensities, the sizes of the Δ I are smaller, i.e., Δ I increases more slowly than I over the moderate to high range of intensities. Consequently, the ratio Δ I/I decreases as signal level increases, as shown in Table VIII. Likewise, RDL decreases as a function of SL, from about 1.5 dB at 5 dB SL to about 0.5 dB at 80 dB SL. The values given in Table VIII represent the relative magnitudes of just detectable increments in the intensities of tones ranging between 200 Hz and 8 kHz. Intensity fluctuations in tones that are smaller than the tabled values probably will be imperceptible even under quiet conditions and such tones may be regarded as effectively constant. Intensity increments equal to, or just slightly above the tabled values may be detectable to a careful observer listening for such increments under quiet conditions. If detection of intensity increments in tonal signals is critical, the size of the increment should

exceed RDL by a factor of at least 2 for quiet conditions. Under noisy or distracting conditions, the size of the increment required will have to be even larger, but it should be determined empirically.

TABLE VIII

INTENSITY DIFFERENCE LIMENS FOR TONES*

SL	10 log $(\Delta I + I)/I$**	$\Delta I/I$
5	1.57	0.44
10	1.50	0.41
15	1.43	0.39
20	1.36	0.37
25	1.29	0.35
30	1.22	0.32
35	1.15	0.30
40	1.08	0.28
45	1.01	0.26
50	0.94	0.24
55	0.87	0.22
60	0.80	0.20
65	0.73	0.18
70	0.66	0.16
75	0.59	0.14
60	0.52	0.13

*From Jesteadt, W., Wier, C. G. and Green, D. M. (7).
**Tabled values determined from equations used by Jesteadt *et al.* (7) to fit their data: $10 \log (\Delta I + I)/U = 1.644 - 0.0141 \times \overline{10} \log (I/Io)$; and $\Delta I/I = 0.463(I/Io) - 0.072$. $SL = 10 \log (I/Io)$.

3.1.4. INTENSITY RESOLUTION OF COMPLEX SOUNDS: Perhaps the most elementary of complex sounds is obtained by adding together two tones of slightly different frequencies. In fact it was just such sounds that were first used to determine intensity difference limens. Differences in the intensities of two tones separated in frequency by only 3 Hz were gradually increased until the listener could detect the occurrences of "beats." The difference in intensities of the two tones at this point was taken as the value of ΔI. The main differences between limens determined in this fashion, as compared with those determined as just-detectable increments in tones (as given in Table VIII), is that the two-tone limens vary as a function of frequency and change over a greater range below 40 dB. The rate at which RDL decreases depends on the frequency of the primary tone. At any SL, the magnitude of the RDL is a function of frequency, decreasing in size as frequency increases from 35 Hz to 4 kHz,

and then reversing direction as frequency continues to increase. For example, the RDL for a tone-pair at 35 Hz decreases from about 5.5 dB at a SL of 15 dB to about 1.8 dB at a SL of 40 dB, while the RDL for a 4 kHz tone-pair decreases from about 1.4 dB to about 0.5 dB over the same range of SL. Little further decrease in RDL occurs above 40 dB at any frequency. At all frequencies above about 70 Hz, the RDL is less than 1 dB for SLs above 40 dB. Except for the differential frequency effect below 40 dB SL, the relationship of RDL to SL is the same as that given in Table VIII. In fact, the values of RDL given in the Table for SLs greater than 40 dB are good estimates of two-tone and white noise RDLs. In the case of signals of more complex spectra than tone combinations and flat bands of noise, intensity difference limens are difficult to measure. In such cases, loudness differences are more readily assessable (refer to section 2.1.2.).

3.2. <u>BINAURAL DISCRIMINATION</u>: This section is concerned with resolution of acoustic differences requiring both ears, *viz.*, differences in intensity and time between the two ears, and differences in the angular directions of sound sources relative to the orientation of the head.

3.2.1. <u>INTERAURAL INTENSITY DISCRIMINATION</u>: The question of concern here is, what is the smallest <u>change</u> in amplitude (Δa) between the two ears that can just be detected, and what are the parameters that influence it? Answers to this question come from experiments in which signals are presented to the two ears through headphones. The parameters that have been investigated include interaural time delay (τ) of the signal to one ear relative to that at the other ear; signal frequency; interaural amplitude imbalance ($a = A_1/A_2$, where A_1 and A_2 are the amplitudes at the two ears); and the overall signal amplitude (A). It appears that the just detectable amplitude change between the ears (Δa) is largely independent of variations in all the above parameters except overall amplitude (A). Between 250 Hz and 10 KHz, Δa ranges irregularly between about 1.0 and 0.4 dB, showing no obvious systematic relationship to frequency (but see section 3.2.3.). Furthermore, Δa remains approximately constant even if the presentation of the signal to one ear is delayed considerably. Interaural delays (τ) between 0 and 1,000 μ sec have been shown to result in no change in Δa (Δa = 0.9 dB for 500 Hz signals). In the case of interaural amplitude imbalances (a), Δa also remains remarkably constant. For example, it has been found that Δa holds at about 0.8 dB for variations in a from 0 to 55 dB, a very large difference between the amplitudes at the two ears. However, if the overall amplitude (A) of the two signals increases from 10 to 75 dB SL, Δa decreases linearly from about 1.5 dB to about 0.5 dB. This finding was obtained with balanced signals, i.e., no time delay ($\tau = 0$) and no amplitude imbalance $a = A_1/A_2 = 1$; $A - A_1 = A_2$). Thus it appears that interaural discrimination of amplitude changes is relatively insensitive to all parameters other than overall level, improving somewhat as level increases. In any case it seems that interaural resolution of amplitude differences

is no better than monaural resolution (see section 3.1.3.). The practical implications of this are: (a) either monaural or interaural changes in amplitude of about 1 dB will be discriminable; (b) this 1 dB change will be resolvable regardless of signal frequency or differences in the arrival times of signals at the two ears, but only for amplitudes above about 40 dB SL; and (c) signals at the two ears will not be noticeably different if fluctuations in amplitude either at one, or both, ears are much less than about 1 dB.

3.2.2. <u>INTERAURAL TEMPORAL DISCRIMINATION</u>: Unlike the discrimination of amplitude differences between the ears, the discrimination of changes in time delay ($\Delta\tau$) between signals presented through headphones to the two ears is influenced by interaural acoustic parameters. The discrimination of interaural time differences is remarkable. In the case of tones, interaural time differences of less than 20 μ sec can be resolved, and, in the case of long duration, low frequency noise, magnitudes of $\Delta\tau$ on the order of 6 μ sec can be discriminated. Just discriminable changes in interaural time delay ($\Delta\tau$) depend on signal frequency, interaural time delay (τ), interaural amplitude ratio ($a = A_1/A_2$), and overall amplitude (A). Under conditions where $\tau = 0$ and $a = 0$ dB, the relationship of $\Delta\tau$ to signal frequency is V-shaped (when frequency is plotted on log scale) and reaches a minimum $\Delta\tau$ of approximately 15 μ sec at 1 kHz. At 125 Hz, $\Delta\tau$ is about 57 μ sec while at 15 kHz, $\Delta\tau$ is indeterminately large. The relationship of $\Delta\tau$ to A (for the conditions $\tau = 0$ and $a = A_1/A_2 = 1$, where $A = A_1 = A_2$) is such that, as signal amplitude in both ears increases from 10 to 75 dB SL, $\Delta\tau$ decreases in a negatively accelerating function up to 40 dB SL and levels off at about 10 μ sec for further increases in A. The relationship of $\Delta\tau$ to τ (for the conditions $a = 1$, A = 50 dB SL), is such that, as τ increases from 0 to 400 μ sec, $\Delta\tau$ increases approximately linearly from 10 to 20 μ sec. The relationship of $\Delta\tau$ to a (for the conditions $\tau = 0$, A = 50 dB SL, A_2 variable) is such that, as the interaural amplitude imbalance (a) increases from 0 to 30 dB (by decreasing A_2), $\Delta\tau$ increases from about 10 μ sec to more than 100 μ sec. Thus it appears that sensitivity to changes in interaural time differences is best if signal amplitudes are greater than about 40 dB SL and differ between the ears by no more than about 10 dB, and when the change in τ to be detected (i.e., $\Delta\tau$) is made from $\tau = 0$ rather than from $\tau > 0$. These facts are especially pertinent under conditions where the localizability of signals in auditory space is important. Localization is heavily dependent on the resolution of changes in interaural temporal relations, and temporal resolution by the binaural system is, in turn, dependent on the degree of imbalance of signal amplitudes at the two ears.

3.2.3. <u>MINIMUM DISCRIMINABLE ANGLES OF AZIMUTH</u>: The smallest angular displacements of sound sources in the horizontal plane of the head are referred to as "minimum audible angles" ($\Delta\Phi$). These minimally discriminable differences in direction are determined by $\Delta\tau$ and Δa (see sections 3.2.1. and 3.2.2.) and the parameters on which they depend, <u>viz.</u>, τ, a, A, and signal

frequency. In a free sound field, the parameters τ and a are functions of the direction of the sound source relative to the orientation of the head (see section 4.1.1.). Likewise, $\Delta\tau$ and Δa also depend on the relative directions of sound sources. Sounds originating from directly in front of the head ($\Phi = 0^{o}$) permit better resolution of temporal and intensive differences (i.e., smaller values of $\Delta\tau$ and Δa) than sounds originating from one side ($\Phi > 0^{o}$), and, consequently, smaller angles ($\Delta\Phi$) can be discriminated in front. Signal frequency also is an important variable. Below 1.5 kHz, $\Delta\Phi$ is accounted for by $\Delta\tau$, and, at least between 1.5 kHz and about 5kHz, $\Delta\Phi$ is accounted for by Δa. In words, discrimination between the directions of low-frequency sounds depends primarily on interaural time differences while interaural amplitude differences are mainly responsible for discrimination between the directions of high-frequency sounds. Overall, the best discrimination of differences in angular directions occurs when sound sources are located in front of the head and the signals are low frequency. For example, if the sound source is located directly in front ($\Phi = 0^{o}$), angular displacements ($\Delta\Phi$) of about 1^{o} can be resolved for signal frequencies below about 1 kHz. As frequency increases from 1 kHz to 1.5 kHz, $\Delta\Phi$ increases from about 1^{o} to 3^{o}. From 1.8 kHz to 3 kHz, $\Delta\Phi$ decreases from about 3.2^{o} to about 1.7^{o} where it remains to about 6 kHz. The size of $\Delta\Phi$ over this frequency range becomes larger as the sound source(s) is moved to more lateral positions (i.e., $\Phi > 0^{o}$). For example, if the signal frequency is 500 Hz, as Φ changes from 0_{o} to 30^{o}, then to 60^{o} and 75^{o}, the minimum audible angle ($\Delta\Phi$) increases from about 1^{o} to 1.8^{o}, then to about 3.3^{o} and 7.5^{o}. For signal frequencies between 1.5 kHz and 3 kHz, at $\Phi = 30^{o}$ the value of $\Delta\Phi$ is about 6.5^{o}, becoming indeterminately large at $\Phi = 60^{o}$. Between 4 kHz and 6 kHz, $\Delta\Phi$ is about 5^{o} at $\Phi = 30^{o}$, increases to between 10^{o} and 12^{o} at $\Phi = 60^{o}$, and becomes indeterminately large at $\Phi = 75^{o}$. It is thus apparent that directional discrimination is poor at frequencies much above 1 kHz unless the sound source is directly in front of the listener. If the sound source is located at angles greater than 30^{o} to the side, discrimination of differences in direction is practically impossible at frequencies above 1.5 kHz. If fine differences in angular direction must be resolvable, it is recommended that sources be located at azimuths within $\pm 30^{o}$ and that signal frequencies be no greater than 1 kHz.

3.3. PATTERN DISCRIMINATION: The perceptual patterns present in sound, especially complex sounds, may serve as a means for enhancement of their recognition and differentiation. (For an in-depth review of literature germane to the following see (9) and (10)). Perceptual patterns may be produced either by simultaneously or sequentially sounding acoustic components. The pattern inherent in the frequency relationships of a musical chord, for example, may be sounded either simultaneously or sequentially, or as a temporal progression of octave transpositions. In order for an auditory pattern to be formed, some perceptually invariant feature must be present in the acoustic array. The tonal relations in a chord constitute such an invariant feature--a pitch pattern--which is preserved after

octave transposition even though the absolute pitches of the tonal components are different. The temporal grouping of successively presented sounds may also constitute an invariant feature--rhythm--which may be preserved even after a time transformation that alters absolute durations throughout the sequence, but not the relative temporal relations. Whether the invariant feature that forms an auditory pattern is simultaneously or sequentially present, such patterns contribute to perceptual organization of acoustic information. They may be judiciously used to improve the efficiency of signal recognition, as a means to distinguish the salient elements of a display from its background, or to define the relevant classes of stimuli to be monitored. Foremost, the perceptual organizations inherent in auditory patterns provide a means to reduce display information loads.

3.3.1. STATIONARY PATTERNS: The presence of a perceptually invariant feature within an array of simultaneously sounding acoustic components constitutes a stationary (unvarying in time) auditory pattern. For example the simultaneous sounding of two musical notes, the fundamental frequencies of which stand in the ratio 2.000, 1.498, 1.335, and 1.260, form consonant intervals on the tempered scale that are recognized in music as the octave, fifth, forth, and major third. The perceptual invariants in this case are the intervals. So long as the frequency ratios are kept the same, the intervals heard will remain unchanged even though the absolute frequencies and their pitches may be changed over a wide range. This tendency for the pitches of tones to maintain the same relationship to each other so long as the ratios of their frequencies are equal is known as tonal chroma; i.e., intervals are repeated in successive octaves such that the pitches of tones in one octave stand in the same relation to one another as integral multiples of them do in higher octaves. This cyclical property of the pitches of tones seems to result from the fact that all of the harmonics within the octave coincide with the upper harmonics of the fundamental frequency. Even though stationary musical patterns more complex than two-tone intervals can be readily formed (e.g., triads, sevenths, ninths), musicians usually analyze these chords by determining the basic intervals formed by each tone-pair contained within them. We mention this to emphasize the importance of these intervals in the formation of complex stationary patterns. Because such harmonic tonal complexes are pleasing to hear, distinctive, and easily associated with events, they provide a ready source for the construction of acoustic signals rich in information content. In fact, it is the complexity (in addition to intervals) which seems to enhance the informational value of such sounds. For example, a seventh is more distinctive than a two-tone chord. Likewise, the complex spectra characteristic of individual musical instruments (timbre) renders them readily distinguishable. Essentially the same may be said for the sounds produced by buzzers, engines, and saw-tooth wave generators. Although the spectral components of such sounds are usually not harmonically related, as in the case of musical chords, they are nevertheless distinctive. By comparison, memory

for the pitches of single tones is very poor in most people. Even the assignment of singly-presented tones to the broadly defined pitch categories "high" and "low" may not be reliably accomplished (except with pitches at the extremes) unless the tones are presented in close succession (see section 3.3.2.). Spectrally complex acoustic arrays thus are more desirable for the composition of stationary signals than simple tones, especially if they contain some perceptually invariant features such as consonant intervals, timbre, etc. Such sounds can be used to signify events, actions, places, etc., with minimal risk of confusion. They are particularly useful under conditions where signal duration must be brief, the number of signals to be individually recognized is large, and the responses to such acoustic signals must be rapid and accurate.

3.3.2. SEQUENTIAL PATTERNS: The presence of a perceptually invariant feature within an array of successively sounding acoustic components constitutes a sequential pattern, i.e., the pattern develops as a function of time. For example, a succession of tones forming a melody is a sequential pattern bound by certain "contours" such as direction of pitch change, interval size, and pitch range. All of the characteristics of stationary patterns may be incorporated into sequential patterns as either temporal contrasts (e.g., one timbre followed by a different timbre) or as progressions (e.g., the notes of a chord may be sounded individually in succession). The perceptual coherence of sequential patterns depends not only on the temporal order of presentation of component sounds, but also on other factors including melodic contours (mentioned above), frequency disparity between components, timbre disparity, rate of component presentation, rhythm, etc. Through appropriate manipulation of these factors, perceptually coherent configurations may be formed, i.e., certain components in acoustic arrays are phenomenally "grouped" together while other components are excluded. Interaction of the various factors that control grouping of sequential acoustic events into coherent patterns may be illustrated with simple tone series. For example, a series of temporally contiguous tones presented to a listener at a rate of about 10/sec will be heard as a unitary "stream" of connected sounds provided that the tones do not differ in frequency by more than about 15 percent. Tones in the series that do differ in frequency by much more than 15 percent will be perceptually isolated and heard as unrelated tone segments. If alternately presented tones are derived from two sets of tones, where the sets differ in frequency by much more than 15 percent, the listener will hear two simultaneous streams that appear to overlap in time and seem to originate from different places in auditory space. Pitch and rhythmic patterns can be heard only within streams. The frequency disparity between sets of tones forming different streams can be reduced if the rate of tone presentation increases. Likewise, frequency differences within sets must be reduced at high rates of presentation to achieve coherence. Streaming at slow rates of presentation is possible if the number of related tones is increased. Time gaps that break the rhythm of successive tone presentations tend to destroy

streaming, as do frequency glides between tones. The perceptual organization inherent in tonal streams may be utilized in various interesting applications. For example, distracting tones can be eliminated from interfering in patterns by adding tones that group with them and cause a separate stream to be formed, thus stripping the distractive tones away from the tones that form the pattern of interest. Likewise, if the pitch categories "high" and "low" are assigned special significance such that the occurrence of target signals in these categories must be detected and the correct category recognized, accuracy of performance may be greatly enhanced by inserting the target tones into an ongoing stream the pitch of which is intermediate between the high and low categories. Target tones the frequencies of which are more than 15 percent greater or less than the tones forming the central stream will be heard as clearly belonging to the high or low categories. In this case, the stream of intermediate-frequency tones provides more than just a central point of reference relative to which the pitches of targets are judged. Rather, the central stream organization excludes the targets in the correct directions thereby rendering them at once distinguishable. It should be noted that optimal target recognition occurs at relatively slow rates of tone presentation, e.g., rates less than about 10/second. Still another application of sequential organization involves the emergence of pitch patterns within streams. A sequence of tones containing two patterns, the individual components of which alternate, may be heard as having no discernable pattern if the rate of presentation is too slow (or fast) to permit the formation of two separate streams. However, as rate increases to the point that two coherent streams are formed, the pitch pattern contained in each stream emerges and both appear to be simultaneiously present. The optimal rate of presentation for stream formation seems to depend on tone durations, inter-tone intervals, and frequency differences between tones within and across streams. The preciseness with which the onsets and offsets of successive tones are synchronized also may influence stream formation. The perceptual organization responsible for the grouping of successive tonal components into streams appears also to account for the grouping of successive sounds on the basis of timbre. For example, sounds produced by the same kind of musical instrument are heard together even though they may play different notes, while sounds produced by instruments of very different timbres are heard as separate Furthermore, the order in a sequence of sounds may not be heard if the components in the sequence differ in timbre. The differences in spectral distributions of acoustic energy (overtone structure) responsible for the recognizable timbre differences between musical instruments thus provide the structural basis not only for the formation of stationary patterns (see section 3.3.1.), but also sequential patterns. The successive sounds of different instruments appear to originate from different spatial locations, those of the same timbre being grouped together, the temporal patterns in each group being heard separately. Given the capability of modern technology to generate electronically sounds with definitive timbres and onset-offset characteristics, the

opportunity now exists for unique applications of this technology in acoustic display systems designed to transmit non-musical information. Two other factors that contribute to the perceptual grouping of successive sounds necessary for the formation of sequential patterns are direction of pitch change and rhythm. Continuation of a unidirectional patterned change in a sequence of tones (e.g., where each successive tone either increases, or decreases, in pitch) results in a perceptual grouping of the tones such that the order of successive pitch changes is more readily identified than if the pitch changes are bidirectional. Furthermore, the coherence of tonal sequences can be achieved at faster rates of presentation if the pitch changes are unidirectional. The temporal structure of successive components, i.e., rhythm, also may contribute to the organization of sound sequences into perceptual groups. For example, a succession of sounds will be grouped into rhythmic units if each unit contains an accented component followed by several unaccented components. The optimal rate of presentation for this kind of organization is about 3/second. Accents appear to be effective in marking off rhythmic units because they differ from other components in the sequence along some discernable dimension (pitch, loudness, duration, and/or timbre), i.e., accents are distinctive. The temporal separation of successive components also contributes to the perception of rhythmic patterns. Lastly, it should be noted that highly distinctive sequential patterns may be composed by combining various of the organizational factors discussed above into the same pattern, e.g., unidirectional pitch changes of a given timbre and rhythm.

REFERENCES

01. ANSI S3.4 1968, R1972 American National Standard Procedure for the Computation of Loudness of Noise.

02. ANSI S3.2 1960, R1976 American National Standard for Measurement of Monosyllabic Word Intelligibility.

03. ANSI S3.24 1977 American National Standard for Rating Noise with Respect to Speech Interference.

04. ANSI S3.5 1969, R1978 American National Standard Methods for the Calculation of the Articulation Index.

05. Fletcher, H. F. and Munson, W. A. Loudness, its definition, measurement and calculation. Journal of the Acoustical Society of America, 1983, 5, 82-108.

06. International Standard ISO 389-1975 Acoustics, "Standard reference zero for the calibration of pure tone audiometers."

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

08. Mulligan, B. E. and Elrod, M. Monaural detection and filtering. Journal of the Acoustical Society of America, 1970, 1947, 1548-1556.

09. Mulligan, B. E., Goodman, L. S., Mitchell, T. M., Crosby, T. N., McBride, D. K., Gleisner, D. P., Stewart, K. D. and Hitchcock, L. Human Engineering Guide to Auditory Displays; Vol. I: Elements of Perception and Memory Affecting Auditory Displays. Government Printing Office, in press.

10. Mulligan, B. E., Goodman, L. S., Mitchell, T. M., Crosby, T. N., McBride, D. K., Gleisner, D. P., Stewart, K. D. and Hitchcock, L. Human Engineering Guide to Auditory Displays; Vol. II: Elements of Signal Reception and Resolution Affecting Auditory Displays. Government Printing Office, in press.

11. Robinson, D. W. and Dadson, R. S. A re-determination of the equal-loudness relations for pure tones. British Journal of Applied Physics, 1956, 7, 166-181.

12. Scharf, B. Loudness. In E. C. Carterette and M. P. Friedman (Eds.), Handbook of Perception IV. New York:Academic Press, 1978, 196.

13. Shower, E. G. and Biddulph, R. Differnetial pitch sensitivity of the ear. Journal of the Acoustical Society of America, 1931, 3, 275-287.

14. Stevens, S. S. Psychophysics. New York: Wiley and Sons, 1975, 310-311.

15. Zwicker, E., Flottorp, G. and Stevens, S. S. Critical bandwidth in loudness summation. Journal of the Acoustical Society of America, 1957, 29, 548-557.

ALGORITHM I

Procedure to enhance detectability of signals in noise

APPENDIX A

BASIC COMPUTER LISTING FOR ALGORITHM I

```
1    HOME
10   REM ALGORITHM I: PROCEDURE TO ENHANCE DETECTABILITY OF SIGNALS
     IN NOISE
20   REM VERSION 06 FEB 84
50   PRINT "PROCEDURE TO ENHANCE DETECTABILITY"
60   PRINT "OF SIGNALS IN NOISE"
70   PRINT
80   PRINT
100  PRINT "DETERMINE","DETERMINE"
110  PRINT "SIGNAL","NOISE"
120  PRINT "SPECTRUM","SPECTRUM"
130  PRINT
140  PRINT TAB( 4)"COMPARE SPECTRA"
150  PRINT
1000 PRINT "CAN SPECTRUM LEVEL OF NOISE"
1010 PRINT "BE REDUCED IN REGION OF SIGNAL"
1020 GOSUB 5000
1030 ON N GOTO 1600, 1100
1100 PRINT "IS NOISE SPECTRUM FLAT"
1110 GOSUB 5000
1120 ON N GOTO 1300,1200
1200 PRINT "CAN SIGNAL SPECTRUM BE"
1210 PRINT "SHIFTED TO DIFFERENT FREQUENCY REGION"
1220 GOSUB 5000
1230 ON N GOTO 2100,1300
1300 PRINT "CAN SIGNAL LEVEL BE INCREASED SAFELY"
1310 GOSUB 5000
1320 ON N GOTO 1700,1400
1400 PRINT "CAN SIGNAL BE PHASE-SHIFTED INTERAURALLY"
1410 GOSUB 5000
1420 ON N GOTO 2200,1500
1500 PRINT "SUBSTITUTE NON-ACOUSTIC SIGNAL"
1510 END
1600 PRINT "REDUCE NOISE LEVEL MAXIMALLY"
1610 PRINT
1700 PRINT "ADJUST SIGNAL LEVEL TO ACHIEVE"
1710 PRINT "OPTIMAL S/N RATIO"
1720 PRINT
1810 PRINT "IS SIGNAL DETECTABILTY ADEQUATE"
1820 GOSUB 5000
1830 ON N GOTO 1900,2000
1900 PRINT "SIGNAL IS ACCEPTABLE"
1910 END
2000 PRINT "SELECT NEW  SIGNAL AND REPEAT PROCEDURE"
2010 END
2100 PRINT "MOVE SIGNAL SPECTRUM TO LEAST"
2110 PRINT "INTENSE REGION OF NOISE SPECTRUM
2120 PRINT
2130 GOTO 1700
2200 PRINT "PHASE SHIFT INTERAURALLY BY 180 DEGREES"
```

```
2210 PRINT
2300 PRINT "WAS PHASE SHIFT EFFECTIVE"
2310 GOSUB 5000
2320 ON N GOTO 1700, 1500
5000 PRINT
5010 PRINT "(Y=YES, N=NO) ";: INPUT A$
5020 IF A$ ="Y" THEN N = 1: GOTO 5100
5030 IF A$ ="YES" THEN N = 1: GOTO 5100
5040 IF A$ = "N" THEN N = 2: GOTO 5100
5050 IF A$ = "NO" THEN N = 2: GOTO 5100
5060 GOTO 5010
5100 HOME : RETURN
```

ALGORITHM II

Procedure to increase loudness without increasing signal level

APPENDIX B

BASIC COMPUTER LISTING FOR ALGORITHM II

```
Ø1   HOME
1Ø   REM ALGORITHM II: PROCEDURE TO INCREASE LOUDNESS WITHOUT
     INCREASING  LEVEL
1ØØ  PRINT "ALGOTITHM II: PROCEDURE TO INCREASE"
11Ø  PRINT "LOUDNESS WITHOUT INCREASING LEVEL"
12Ø  PRINT  : PRINT
5ØØ  PRINT "WHEN SIGNAL LEVEL EXCEEDS 8ØDB SL"
6ØØ  PRINT : PRINT
1ØØØ PRINT "IS SIGNAL BINAURAL"
1Ø1Ø GOSUB 5ØØØ
1Ø2Ø ON N GOTO 14ØØ,11ØØ
11ØØ PRINT "CAN SIGNAL BE PRESENTED BINAURALLY"
111Ø GOSUB 5ØØØ
112Ø ON N GOTO 12ØØ,14ØØ
12ØØ PRINT "PRESENT SIGNAL BINAURALLY"
121Ø PRINT
13ØØ PRINT "IS LOUDNESS OF SIGNAL ADEQUATE"
131Ø GOSUB 5ØØØ
132Ø ON N GOTO 4ØØØ,14ØØ
14ØØ PRINT "DOES SIGNAL DURATION EXCEED 2 SECONDS"
141Ø GOSUB 5ØØØ
142Ø ON N GOTO 15ØØ,16ØØ
15ØØ PRINT "CAN INTERVAL BETWEEN SUCCESSIVE"
151Ø PRINT "SIGNALS BE INCREASED"
152Ø GOSUB 5ØØØ
153Ø ON N GOTO 17ØØ,21ØØ
16ØØ PRINT "IS SIGNAL DURATION LESS THAN Ø.5 SECONDS"
161Ø GOSUB 5ØØØ
162Ø ON N GOTO 19ØØ,18ØØ
17ØØ PRINT "SET INTERVAL AT TWICE"
171Ø PRINT "THE SIGNAL DURATION"
172Ø PRINT
18ØØ PRINT "IS SIGNAL LOUDNESS ADEQUATE"
181Ø GOSUB 5ØØØ
182Ø ON N GOTO 4ØØØ,21ØØ
19ØØ PRINT "CAN SIGNAL DURATION BE INCREASED"
191Ø GOSUB 5ØØØ
192Ø ON N GOTO 2ØØØ,18ØØ
2ØØØ PRINT "SET SIGNAL DURATION BETWEEN"
2Ø1Ø PRINT "Ø.5 AND 2.Ø SECONDS"
2Ø2Ø PRINT : GOTO 18ØØ
21ØØ PRINT "IS SIGNAL A SINGLE TONE"
211Ø GOSUB 5ØØØ
212Ø ON N GOTO 24ØØ,22ØØ
22ØØ PRINT "IS SIGNAL A MULTI-TONE COMPLEX"
221Ø GOSUB 5ØØØ
222Ø ON N GOTO 25ØØ,23ØØ
23ØØ PRINT "IS SIGNAL SPECTRUM CONTINUOUS"
231Ø GOSUB 5ØØØ
```

```
2320 ON N GOTO 26ØØ,31ØØ
24ØØ PRINT "CAN OTHER TONES BE ADDED TO SIGNAL"
241Ø GOSUB 5ØØØ
242Ø ON N  GOTO 27ØØ,3ØØØ
25ØØ PRINT "CAN TONE FREQUENCIES BE ALTERED"
251Ø GOSUB 5ØØØ
252Ø ON N GOTO 27ØØ,3ØØØ
26ØØ PRINT "CAN WIDTH OF SPECTRUM BE INCREASED"
261Ø GOSUB 5ØØØ
262Ø ON N GOTO 28ØØ,3ØØØ
27ØØ PRINT "SEPARATE TONAL COMPONENTS 1 TO 2"
271Ø PRINT "OCTAVES KEEPING OVER ALL LEVEL"
272Ø PRINT "AT CONSTANT SL"
273Ø PRINT : GOTO 29ØØ
28ØØ PRINT "INCREASE SIGNAL SPECTRUM TO 2 OR 3"
281Ø PRINT "TIMES THE WIDTH OF THE CRITICAL"
282Ø PRINT "BAND KEEPING OVERALL LEVEL CONSTANT"
283Ø PRINT
29ØØ PRINT "IS SIGNAL LOUDNESS ADEQUATE"
291Ø GOSUB 5ØØØ
292Ø ON N GOTØ 4ØØØ,3ØØØ
293Ø PRINT "SUBSTITUTE NON-ACOUSTIC SIGNAL"
3Ø1Ø END
31ØØ PRINT "TERMINATION OF ALGORITHM"
311Ø PRINT "REFER TO MANUAL"
312Ø END
4ØØØ PRINT "SIGNAL IS ACCEPTABLE"
4Ø1Ø END
5ØØØ PRINT
5Ø1Ø PRINT "(Y=YES, N=NO)";: INPUT A$
5Ø2Ø IF A$ = "Y" THEN N = 1: GOTO 51ØØ
5Ø3Ø IF A$ = "YES" THEN N = 1: GOTO 51ØØ
5Ø4Ø IF A$ = "N" THEN N = 2: GOTO 51ØØ
5Ø5Ø IF A$ = "NO" THEN N = 2: GOTO 51ØØ
5Ø6Ø GOTO 5Ø1Ø
51ØØ HOME : RETURN
```

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Special Report 84-1	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A design guide for nonspeech auditory displays		5. TYPE OF REPORT & PERIOD COVERED Interim 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) B. E. Mulligan,* D. K. McBride and L. S. Goodman		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Aerospace Medical Research Laboratory Naval Air Station, Pensacola, FL 32508-5700		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Medical Resarch and Development Command Naval Medical Command, National Capitol Region Bethesda, Maryland 20814		12. REPORT DATE September 1984 13. NUMBER OF PAGES 45
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Dr. B. E. Mulligan's current address is Department of Psychology, University of Georgia, Athens, Georgia 30602

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Attention,
Auditory Auditory Signals,
Loudness
Detection
Nonspeech acoustic displays.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Vast attention has been devoted to the investigation of various sensory and perceptual characteristics of the human auditory system. It is not often obvious, however, how the aggregate findings provided by these efforts might effectively be utilized to design auditory displays of information. This report -

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

S N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract (continued)

condenses and synthesizes critical research findings on the (1) detection , (2) loudness, and (3) distinctiveness of non-speech auditory displays. The format of this report provides a unique guide for the design of nonspeech auditory displays.

Eight tables and two algorithms (in flow-chart form) were developed and are provided to assist the auditory display engineer in (1) increasing the detectability of signals presented in noise and (2) increasing the loudness of signals without increasing signal level. The algorithms are coded in the BASIC computer language and are enclosed as appendices.

The scope of this report and the algorithms provided are limited to three important areas of auditory display engineering. Similar attention should be devoted to other critical aspects of audition, such as, reaction time, stimulus-response compatibility, attention, recognition, and memory.

supplied keywords included, → front