NASA CONTRACTOR REPORT

NASA CR-1598

NASA CR-15
c. 1

STUDY OF THE AUDIBILITY OF IMPULSIVE SOUNDS

by Sanford Fidell and Karl S. Pearsons

Prepared by
BOLT BERANEK AND NEWMAN INC.
Van Nuys, Calif. 91406
for Langley Research Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • MAY 1970

NASA CR-1598

STUDY OF THE AUDIBILITY OF IMPULSIVE SOUNDS

By Sanford Fidell and Karl S. Pearsons

Prepared under Contract No. NAS 2-4963 by
m.e. – BOLT BERANEK AND NEWMAN INC.
Van Nuys, Calif. 91406

for Langley Research Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – CFSTI price $3.00

TABLE OF CONTENTS

LIST OF TABLES

STUDY OF THE AUDIBILITY OF IMPULSIVE SOUNDS

by Sanford Fidell and
Karl S. Pearsons

Bolt Beranek and Newman Inc.

SUMMARY

Six experiments were performed in an anechoic chamber to investigate the effects of various physical parameters on the perceived noisiness of impulsive signals. The parameters investigated included phase, duration, intersignal interval, repetition, and frequency. All data were collected by a computer based adaptive psychophysical procedure called Parameter Estimation by Sequential Testing (PEST).

The major conclusions were as follows:

1) The phase spectrum of an impulsive signal is irrelevant to its perceived noisiness.

2) The ear's sensitivity to noisiness of impulsive signals resembles an energy summation process. No specific time constant of integration was found.

3) The common correction contours (such as AL, NL, and PNL) may undercorrect in the low frequency regions, and thus should be applied with caution to impulsive signals with appreciable low frequency content (for example, sonic booms).

INTRODUCTION

Impulsive noises are assuming increasing significance as irritants in contemporary life. One impulse noise in particular, the sonic boom, threatens to become a major source of noise pollution in a relatively short time. Most efforts at noise abatement have hitherto been directed to the more prevalent steady state noises; i.e., those of durations sufficiently great that continuous measurements provide accurate and repeatable estimates of levels. Standardized procedures for such steady state measurements have been developed over the years. Research programs to reduce steady state noises such as those produced by motor vehicles, industrial machinery, aircraft, etc., have been undertaken on the basis of information provided by these physical measurements.

Since physical measurement is vital for noise reduction programs, it is necessary to develop measurement techniques appropriate to the transient character of impulsive signals before the character and magnitude of impulsive noise pollution problems may be assessed. Fundamental to the development of such techniques is a clear theoretical understanding of human response to transient signals. For example, it is essential to understand the relationships between the physical parameters of transient signals and their effects on subjective judgments of annoyance. Some parameters may have no measurable influence on the perceived noisiness of an impulsive signal, while others may in large part determine the noisiness of the signal.

Accordingly, five experiments were performed to assess the effects of various physical parameters on the perceived noisiness of impulsive sounds. A sixth study was conducted to evaluate the generality of the conclusions reached from the first five studies. Experiments one through five concerned the effects on perceived noisiness of 1) phase, 2) duration, 3) the interval between two identical impulses, 4) repetition of identical impulses, and 5) the interaction between impulse duration and frequency. The final experiment determined the comparative annoyance of both naturally occurring and synthetic impulsive signals. Detailed descriptions of goals and procedures of the experiments may be found in the Purpose and Method Section. The Results and Discussion Section contains the findings of the various experiments and interpretations of them.

LITERATURE REVIEW

One fundamental difficulty encountered in applying measurements derived from steady state noises to transient noises is that the usual unit of measurement is power, or energy per unit time. It is implicitly assumed in measurement of steady state noises that the duration of the signal has no substantial effect upon its power. A primary characteristic of impulse noises however, is that they are of sufficiently short duration that the duration of the measurement procedure has considerable influence on the adequacy of the measurement. Thus, even the prediction of the perceived magnitude of an impulsive signal is not a trivial matter.

One of the easiest and most straightforward ways to account for the apparent magnitude of an impulsive sound is to assume that the ear integrates the acoustic power of an impulse over a finite duration. This period of integration is called the time constant of the ear. To find the apparent magnitude of an impulse, one calculates the acoustic energy of the impulse, finds the effective sound pressure level by dividing this energy by the duration of the time constant, and proceeds as though the sound were steady state.

In effect, the assumption of a finite period of integration allows one to treat both impulsive and steady-state sounds within the same framework. Thus, it is possible to apply all of the information on sone and noy contours to the calculation of the apparent magnitude of impulsive sounds. Obviously the magnitude of the ear's time constant is the crucial quantity in this approach.

One of the first to investigate the time constant of the ear was von Békésy (Ref.1). By varying the duration of an 0.800 kHz sinusoid, he concluded that about 180 msec was needed for it to reach its maximum loudness. Munson (Ref.2) used frequencies of 0.125, 1.000, and 5.650 kHz and estimated the time constant to be about 250 msec. Miller (Ref.3) used broadband noise and estimated the time constant to be about 65 msec at the higher sensation levels. Small, Brandt and Cox (Ref.4) performed a similar experiment but found consistently smaller values. At high levels the time constant estimated by Small et al. appears to be in the range 10 to 20 msec.

Von Port (Ref.5) investigated a variety of brief impulsive sounds. Most were gated noises of different bandwidths, center frequencies, and levels. He estimated that 70 msec was a reasonable value of the time constant for a variety of stimuli. Von Niese (Ref.6) has summarized a number of studies from the Dresden laboratories. He concludes that 25 msec is a good average value and argues that this estimate agrees nicely with the reverberation time estimates. Von Zwicker (Ref.7 and 8) takes issue with the conclusions of the Dresden group and presents more data from the Stuttgart laboratory consistent with the longer estimate, approximately 100 msec. Stevens and Hall (Ref.9) using a magnitude estimation technique, measured the time constant at about 150 msec.

As may be seen, there is remarkably little agreement about the magnitude of the ear's time constant. There are at least two reasons for the lack of agreement. First, the judgments are extremely difficult for an observer to make. Practically every investigator of impulsive sounds has commented on this aspect of the problem. Further, because the task is difficult, great variation among judgments of different subjects is found. Garner (Ref.10) for example, published a complete data for six subjects. Three of his observers show complete energy summation up to 300 msec. The other three give loudness judgments that are nearly independent of the signal duration. For the impulses of 10 msec duration used by Garner, the difference in the judgment of these two groups is about 10 dB. Von Reichardt, von Niese, and Müller (Ref.11) show a distribution of judgments for forty subjects. The distribution of judgments is essentially normal but ranges over more than 20 dB. Regrettably, this same variability among observers was found in some of the current experiments. In one experiment (the Repetition experiment), two observers differed in their judgments by about 25 dB.

Secondly, there is fairly strong evidence that judgments of "annoyance" and/or "noisiness" are somewhat different from judgments about "loudness" of impulsive sounds (von Zwicker, Ref.8). This difference appears to be particularly important when the judgments are made of repetitive impulsive sounds. Despite the potential importance of this variable it has not been systematically studied. Two recent reports, one of von Zwicker's study and one by Shepard and Sutherland (Ref.12) are among the first published studies dealing specifically with the problem of instructions. Since our instructions used "annoyance" and "noisiness" rather than "loudness", this may account for the absence of a time constant difference which had been previously found in other experiments which used "loudness" instructions.

PURPOSE AND METHOD

General Procedure

Data collection was governed by an adaptive psychophysical procedure called Parameter Estimation by Sequential Testing (PEST), described in detail in Appendix A. All observers' hearing was screened to within 15 dB of the proposed ISO standard threshold (Ref.13). The observers were selected from a group consisting primarily of college students ranging in age from 17 to 32 years, with a median age of 20 years. One half hour's familiarity with the PEST system was required of observers before their participation in experimentation. Observers were instructed to depress one of two lighted response switches corresponding to the more noisy of a pair of signals. The test instructions are reproduced in Appendix B. Each pair of signals contained in random order an invariant standard signal and a variable comparison signal. Table 1 summarizes the signal sets employed in each experiment. The intent and design of each of the experiments are discussed in order below. The equipment employed in generating the stimuli and collecting and analyzing the data is described in Appendix C.

Phase Experiment

The purpose of the first experiment was to determine whether variations in phase spectra of impulsive signals influence their perceived noisiness. Although human observers can under some circumstances make discriminations on the basis of phase information, it was not known whether phase contributed to the annoyance of impulsive signals. If phase relationships could be shown to exert no measurable influence

TABLE 1: SUMMARY OF EXPERIMENTAL CONDITIONS

Experiment	Standard Signal(s)	Comparison Signals
PHASE	Stds. 1&2: two synthetic waveforms	25 synthetic signals in five waveform families
DURATION	1000 msec sample of an octave band of noise between 0.600 and 1.200 kHz	5 white noise spectra of six temporal durations
INTERVAL	1000 msec sample of an octave band of noise between 0.600 and 1.200 kHz	2 synthetic waveforms presented twice at six temporal intervals
REPETITION	1000 msec sample of an octave band of noise between 0.600 and 1.200 kHz	2 synthetic waveforms presented 1,2, 4, 6 and 8 times at 333 msec intervals
INTERACTION	Std.1: 10 msec sample of an octave band of white noise from 2.400 to 4.800 kHz	10 msec sample of an octave band of identical white noise between 0.075 and 0.150 kHz; also, 1000 msec sample of an octave band of identical white noise between 0.075 and 0.150 kHz; also, 1000 msec sample of an octave band of white noise from 2.400 to 4.800 kHz
	Std.2: 1000 msec sample of an octave band of white noise from 2.400 to 4.800 kHz	10 msec sample of an octave band of white noise from 2.400 to 4.800 kHz; also, 1000 msec sample of an octave band of white noise from 0.075 to 0.150 kHz
EVALUATION	1000 msec sample of an octave band of noise between 0.600 and 1.200 kHz	12 naturally occurring impulsive sounds
	1000 msec sample of an octave band of noise between 0.600 and 1.200 kHz	8 synthetic waveforms

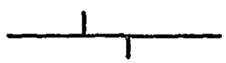
on annoyance judgments, there would be little reason to consider phase differences among signals in subsequent experimentation.

In order to assess the effects of phase on annoyance judgments a set of synthetic signals was constructed, such that families of waveforms of nearly identical power spectra displayed widely divergent phase spectra. Ten observers compared each of the set of twenty-five synthetic impulses with two members of the set which were selected as standard signals.

Five basic transient waveforms were employed in construction of the synthetic impulsive sounds. The basic waveforms were 1) an ideal N-wave, 2) an N-wave with leading and trailing segments of approximately one msec duration, 3) a triangular waveform, 4) a square waveform, and 5) a doublet waveform. Each basic transient wave was analyzed by digital Fast Fourier Transform techniques. A new transient waveform was then generated by scrambling the phase relationships of the original transient. The initial transient waveform was designated the prototype and was employed in the production of four phase-distorted replicas. Thus, the prototypical triangular waveform is designated T0, and its four phase-scrambled replicas are designated T1, T2, T3, and T4.

Although any two members of a waveform family (such as the triangular waveforms T0, T1, T2, T3 and T4) have distinctly different shapes, they share nearly identical power spectra. The durations of any prototype and its four replicas are identical (approximately 26 msec; except D0 which is 13 msec). Further details on the procedures employed in generating and measuring the synthetic impulse sounds are discussed in Appendix D. Table 2 lists all twenty-five synthetic impulses employed in the first experiment and describes their general spectral characteristics.

TABLE 2: DESCRIPTION OF SYNTHETIC IMPULSE SOUNDS EMPLOYED IN PHASE EXPERIMENT

Impulses		Prototype Waveform	Description	Spectral Characteristics
Ideal N-Wave	*IN0 IN1 IN2 IN3 IN4		Prototype and four phase distorted replicas of an N-wave with infinitesimally short leading and trailing times	Spectrum peaks at 40 Hz and falls off at 6 dB/oct for higher and lower frequencies
N-Wave	N0 N1 N2 N3 N4		Prototype and four phase distorted replicas of an N-wave with less abrupt leading and trailing times	Spectrum same as for "IN" series except slope at frequencies above 250 Hz falls at 12 dB/oct
Triangular Wave	T0 T1 T2 T3 T4		Prototype and four phase distorted replicas of a triangular wave	Spectrum peaks at 40 Hz then falls at 12 dB/oct for higher frequencies
Square Wave	S0 S1 S2 S3 S4		Prototype and four phase distorted replicas of a square wave	Same spectrum as "IN" series
Doublet Wave	*D0 D1 D2 D3 D4		Prototype and four phase distorted replicas of a doublet waveform	Spectrum is flat above 40 Hz

*Standard Signals

Note: All synthetic signals are of 26.0 msec duration except D0 which is 13.0 msec.

Duration Experiment

The principal intent of the second experiment was to study the effects of signal duration on perceived noisiness. In particular, signal durations from 10 to 1000 msec were employed in an attempt to establish the existence and value of the ear's time constant (as discussed in the Literature Review). A secondary goal of the duration study was to investigate the effects of spectral composition of impulsive signals on the ear's time constant; i.e., to determine whether the value of the time constant was a function of frequency.

Ten observers compared five filtered white noise spectra of six temporal durations with a standard 1000 msec burst of an octave band of white noise between 0.600 and 1.200 kHz. The band limits and durations of the signals employed in this study may be found in Table 3. The order of administration of each cell of the factorial design illustrated in Table 3 was randomized independently for each observer.

Interval Experiment

Whether repeated impulsive signals (such as the leading and trailing portions of a sonic boom) are perceived as single or multiple events obviously depends upon both the temporal interval between repetitions and the time constant of the ear. More specifically, impulsive sounds separated by an interval greater than the ear's time constant would be perceived as multiple events, whereas impulsive sounds separated by an interval less than the time constant would be perceived as single events. If the observer's

TABLE 3: TABULAR FORM OF FACTORIAL DESIGN EMPLOYED IN DURATION EXPERIMENT

Rows: Signal Spectra
Columns: Signal Durations

Comparison Signals

Band Limits in kHz	Duration in Msec					
	10	33	67	100	333	1000
0.150-0.300						
0.300-0.600						
0.600-1.200						X
2.400-4.800						
0.020-5.000						

Standard Signal

Octave band of filtered white noise from 0.600-1.200 kHz of 1000 msec duration (denoted by "X" in table) noisiness judgments are directly related to the energy of the signals presented, then one would expect a 3 dB increase in noisiness of two signals separated by an interval greater than the ear's time constant.

In order to explore this hypothesis, four observers were asked to compare two synthetic waveforms (INO and DO) presented twice at six intersignal intervals (33, 67, 100, 200, 333, and 1000 msec) with a standard 1000 msec sample of an octave band of noise between 0.600 and 1.200 kHz.

Repetition Experiment

If multiple repeated impulses are perceived as temporally discrete events then the annoyance of a set of multiple impulses should be functionally related to the annoyance of a single impulse by the number of impulses in the repeated set. If the combined annoyance of multiple impulses follows an energy-like summation, one would anticipate a 3 dB increase in noisiness to be associated with a doubling of the number of elements in the set of repeated impulses.

To test whether a general rule of energy-like summation of annoyance describes sensitivity to multiple impulses, eight observers compared two sets of 1, 2, 4, 6 and 8 impulses separated by an interval of 333 msec with a standard 1000 msec burst of filtered white noise between 0.600 and 1.200 kHz. The synthetic signals INO and DO served as the multiple repeated signals.

Interaction Experiment

The purpose of the Interaction experiment was to determine whether the common frequency weighting curves (which were developed for steady state measurement)provide an adequate basis for impulsive measurements. In particular, the experiment was intended to reveal whether the product of signal intensity and time was constant at two different frequency ranges separated by about 20 dB on the weighting curves. Two octave bands were compared with one another at two different durations. The two-by-two table below (Table 4) summarizes the conditions of this experiment.

TABLE 4: SUMMARY OF EXPERIMENTAL CONDITIONS OF INTERACTION STUDY

Duration	Frequency: Low	Frequency: High
Short		Standard 1
Long		Standard 2

Twelve observers participated in the interaction study. The standard and comparison signals were filtered white noise of two durations in two octave bands. The durations were 10 msec (short) and 1000 msec (long); the octave bands were 0.075 to 0.150 kHz (low) and 2.400 to 4.800 kHz (high). The first standard signal was of high frequency and short duration. The signals with which it was compared were of high frequency, long duration; low frequency, long duration; and low frequency, short duration. The second standard signal was of high frequency and long duration. The signals with which the second standard was compared were of high frequency, short duration; and of low frequency, long duration.

The comparison signals were computer generated filtered white noise of a special character: each noise sample of the same duration and band limits was an identical waveform. This measure was taken to insure that fluctuations in the low frequency, short duration comparison signals, due to their small number of effective degrees of freedom (Ref.14), would not per se introduce variance.

Evaluation Experiment

The last experiment was conducted to evaluate the generality and applicability of the results of the first five studies (which employed synthetic sounds) to naturally occurring impulsive sounds. Ten observers compared each of 12 naturally occurring impulses (described in Table 5) with a standard signal consisting of a 1000 msec sample of an octave band of filtered white noise between 0.600 and 1.200 kHz.

The same ten observers compared the identical standard signal with eight synthetically generated impulsive signals. The synthetic signals employed were selected from the set of synthetic signals employed in the Phase experiment. The eight signals were IN0, IN1, D0, D1, D4, N0, S0 and T0.

TABLE 5: DESCRIPTION OF NATURALLY OCCURRING IMPULSIVE SOUNDS EMPLOYED AS COMPARISON SIGNALS IN EVALUATION EXPERIMENT

Impulse	Duration (msec)	Identification	Approximate Spectral Characteristics
1	300	Automobile Door Slam	Peaks at 0.500 kHz
2	150	Paper Tearing	Near flat spectrum to 10.0 kHz
3	425	Hand Clap	Rises and falls about 0.800 kHz
4	450	Two Bottles Clinking Together	Highly leptokurtic at 4.0 kHz
5	580	Chain Collapsing on Itself	Near flat spectrum to 1.0 kHz, falls slowly at higher frequencies
6	480	Nocturnal Animal Noise	Complex spectrum peaked at 0.125 and 2.5 kHz
7	180	Squeaky Release of Air through a Valve	Peaks at 0.800, 1.600, and 5.000 kHz
8	400	Balloon Bursting	Predominantly low frequency, falls slowly above 0.200 kHz
9	600	Balloon Bursting	Peaks at 0.200 kHz
10	180	Automobile Horn	Discrete frequency peaks concentrated between 0.300 and 1.0 kHz
11	1200	Simulated Sonic Boom	Predominantly low frequency, falling steeply from 0.125 kHz
12	900	Basketball Bounce in highly Reverberant Environment	Energy concentrated between 0.200 and 1.600 kHz
Standard	1000	White Noise Burst	Octave band from 0.6 - 1.2 kHz

RESULTS AND DISCUSSION

The results of the experiments described in the preceding section are presented and interpreted in order. The basic datum of the statistical and graphical analyses in this section is the level of the comparison signal at which the observer was indifferent between the perceived noisiness of the comparison and standard signals. Data are plotted in the form comparison level minus standard level. Comparison signals judged more noisy than the standard are therefore represented by points lying below the zero dB reference, while comparison signals judged less noisy than the standard correspond to points lying above the reference line. Extreme differences in judged levels of some standard and comparison signals (on the order of 40 dB), as well as individual differences among observers, sometimes necessitated adjustment of absolute levels of standard or comparison signals within experiments. Justification for such shifts in signal levels within experiments is available both from informal studies (in which level adjustments were found to cause negligible effects) and from von Zwicker (Ref.7).

Phase Experiment

The results of the Phase experiment are presented in Figures 1, 2 and 3. The figures show the levels at which each of the synthetic signals were judged equal in noisiness to the standard signal. The level of the standard signal serves as the zero dB reference line in all of the figures. In Figures 1 and 2, the signals of the same waveform family are connected by line segments.

If all of the signals of the same waveform family were judged equally noisy to each other, the line segments joining data points would form a straight line of zero slope. As can be seen from Figures 1 and 2, containing judgments made against the two standard signals DO and INO respectively, there is no systematic trend toward a slope other than zero. Further, the total range of variability within waveform families is typically on the order of ±1.5 dB. From these data one may conclude that phase differences among impulsive signals, even if they are discernable, do not contribute significantly to perceived noisiness.

Analyses of variance were performed on the data of each of the waveform families to determine whether any differences among signals were of significant proportions. When the prototype and the four replicas of each waveform family were compared, only the DO family showed a significant difference ($F[4,45]$ 4.6; $P < .05$). The significant difference in levels of the members of the D waveform family seems due simply to the duration difference between DO (13 msec) and its replicas (26 msec). Figure 3 clearly illustrates this effect; even with an N-weighting curve applied, DO is much farther above the reference line than any of its replicas. In order to compensate for DO's short duration, observers apparently increased its magnitude.

Figure 3 presents combined judgments for all signals to both standard signals, INO and DO. All data derived by comparison to the DO standard were normalized to the INO standard, which is represented by the zero reference line. The impulsive signals paired with DO are represented by open symbols. Those paired with INO are identified by closed symbols. The line segments drawn between phase scramblings illustrate the average level for each signal compared alternately with the two standards - INO and DO. The N-weighting curve applied to the data in Figure 3 both compressed the differences in levels between waveform families and also rendered them closer in level to the standard signals. Special attention should be directed to the relative efficiency of the N transformation on the various waveform families. The triangular waveforms, which contained the greatest concentration of low frequency energy of the five waveform families, were less affected by the N transformation than any of the other waveforms. The importance of this observation will be discussed at length later.

Duration Experiment

The data of the Duration Experiment are displayed in Figures 4, 5, 6 and 7. They are plotted in terms of dB of difference in noisiness between the comparison signals and the standard signals. Data points for judgments of the same octave band of noise are joined by lines.

In general, it may be seen that short duration, low frequency, narrow band noises are greater in level when judged equally noisy to the standard signal, than are long duration, high frequency, wide band noises. The 3 dB per duration doubling increase in noisiness, predicted by a simple energy detection model of sensitivity to impulse noise, is represented in Figure 4 by the slope of the shaded area. The fit of the 3 dB per duration doubling line to the data appears quite reasonable. Further, there is no systematic divergence from the 3 dB per doubling line as a function of frequency range. The generality of the 3 dB per duration doubling increase in noisiness may be seen most clearly in Figure 5, in which the results of the current Duration experiment are plotted on the same scale as those of another study recently completed by BBN for the FAA (Contract FA68WA-1978). The data from the FAA study, including signal durations greater than 1000 msec, were collected under experimental conditions similar to those of the current studies. The 3 dB per duration doubling rule provides an excellent fit across four decades of time.

Figure 6 contains the same data as Figure 4 transformed to the A-scale. Figure 7 shows these same data on an N-scale. Once again, the compression of variability afforded by the A and N transforms is evident. The N transform offers slightly greater reduction in variance than the A transform. The mean ratio of A-weighted to N-weighted standard deviations for the data at each duration is 1.4.

Interval Experiment

The results of the Interval experiment are plotted in Fig. 8. Due to considerable individual differences among observers' responses (as discussed in the Literature Review section) it was necessary to subtract each observers' mean from his own data in preparing Figure 8. As an example of the range, responses of two observers to the INO signal at the 33 msec interval ranged from 3 dB to 25 dB; at the 1000 msec interval the same two observers' responses were 3 dB to 24 dB.

There does not appear to be any systematic trend in the data displayed in Fig. 8. This suggests strongly that the noisiness of two impulses is independent of the duration of the interval between them. However, a combined repeated measures analysis of variance (one which specifically did not consider variation due to individual differences among observers) revealed the existence of a significant difference as a function of interval ($F[5,24] = 6.0$; $P < .01$).

Nonetheless, the non-monotonicity of the data prevents the inference of effects due to intersignal interval. The absence of a unique inflection point in the data prevents the inference of a specific value for the time constant of the ear over the range of intersignal intervals tested.

Repetition Experiment

The findings of the Repetition experiment are illustrated in Fig. 9. Again, due to gross differences (of about 20 dB) among observers' responses, the mean of each observer's responses was subtracted from his own data. With the observers' variability accounted for, a graphical inspection of the data revealed that the noisiness of multiple impulses was directly proportional to the number of impulses in the comparison signal. The close fit of the data points of Fig. 9 afforded by the line of 3 dB per impulse doubling slope suggests that noisiness judgments are determined by the total energy present in the compound signal.

Separate repeated measures analyses of variance (with 4 and 20 degrees freedom) for comparison DO and INO yielded F ratios of 82.0 and 19.0, respectively ($P < .01$), confirming the significance of the trend in the data. The repeated measures design was necessitated by the broad range of individual differences among the observers.

Interaction Experiment

All observers' data from the Interaction experiment are presented in Table 6. The cell entries represent a combination of information derived from the short and long duration standard signals. To render the responses made to the different standards comparable, it was necessary to normalize the data generated in comparisons with the long duration standard to data generated in comparisons with the short duration standard. This transformation was achieved by finding the factor which represented the difference in the responses to the two standards at the level of subjective equality. This factor was added to the data generated by the long duration standard. The normalized data and the data associated with the short duration standard were then averaged. The results of the averaging are presented in Table 6.

TABLE 6: NORMALIZED N-LEVEL DATA FOR INTERACTION EXPERIMENT

Duration (msec)	Octave Band (kHz) Low 0.075-0.150	Octave Band (kHz) High 2.400-4.800
(Short) 10	81.5dB	STD: 89.5dB 91.5dB
(Long) 1000	69.5dB	79.5dB

As may be seen from Table 6, there is a consistent 10 dB difference in levels between low and high frequency signals. The N correction contour should have reduced the difference between low and high frequencies to zero. It appears, therefore, that the contour undercorrects at low frequencies. The same inference may be drawn from the under-correction of the low frequency triangular waveforms of the Phase experiment.

Additional confirmation of this hypothesis is available from the Duration experiment. Figure 10 shows the data of the Duration experiment extrapolated to one lower octave band

and re-plotted to remove the effects of duration (by averaging across durations). A 10 dB difference in N-levels between the low and high frequency regions is observed. Similar findings were obtained for PNL corrections. The results of the Evaluation experiment (see below) provide yet another basis for believing that the nature of the error in the N-level correction contour is undercorrection at the low frequencies.

The effects of signal duration in the Interaction experiment are not as clear as in the previous studies. On the basis of a simple energy detection scheme, one would expect a 20 dB difference between the short and long duration signals, independently of any frequency effects. The data show a difference in the proper direction, but of only 12-14 dB, depending on the correction contours applied to them. It is possible that context effects (such as the recency of judgments of a different nature, or the levels of the standard and comparison signals) may have influenced the observers' decisions.

Evaluation Experiment

The results of the Evaluation experiment, represented as overall (OASPL), A-level (AL), N-level (NL), and perceived noise level (PNL) are plotted in Figure 11. The points plotted as circles represent values of noisiness judgments of naturally occurring impulses averaged over all observers. The figures plotted as squares represent similar values for synthetic signals. The noise measures for the impulses were integrated to account for the disparity between standard and comparison signal durations. For example, a 25 msec sample of a 1000 msec signal would be corrected by 16 dB (10 log 25/1000).

These data permit clear comparisons of the effectiveness of the various correction contours on several bases. First, it should be noted that for all of the correction procedures, the synthetic and naturally occurring impulses are intermingled. The intermingling demonstrates that correction procedures applied to artificially generated signals in laboratory research can yield similar results when applied to real-life impulsive sounds.

Secondly, one may compare the means and variances of the distributions of noisiness judgments under the various correction procedures. In general, AL appears to compress the range of variability among signals of different spectral content more than NL or PNL. At the same time however, NL and PNL results offer better mean predictions of noisiness, as evidenced by the closer proximity of the NL and PNL distributions to the zero dB reference line. Since there is little difference in the variability of the distributions of the three correction procedures (the standard deviations range from about 3.5 to 5.0 dB), there may be some advantage in employing NL or PNL for the better accuracy of prediction which they provide.

Third, the relative effects of the transformations on signals composed primarily of low frequency energy must be noted. The open square in each of the distributions of Figure 11 corresponds to the mean noisiness of the synthetic signal T0. T0 has the spectrum most heavily weighted in the low frequency regions of any of the impulses tested. Its position in the AL, NL, and PNL distributions is consistently farther from the zero dB reference line than any of the other signals' position. The position of T0 in the uncorrected (OASPL) distribution however, is intermediate.

One might therefore conclude that weightings associated with AL, NL, and PNL are inadequate in the low frequency regions.

Corliss and Winzer (Ref.15) have reported findings which corroborate the apparent undercorrection of low frequencies by the AL and NL contours. They found better agreement between overall sound pressure measurement and subjectively judged impact noise than between subjective judgments and the Stevens or von Zwicker phon scales.

CONCLUSIONS

The three major conclusions which may be drawn from the present series of experiments are:

1) Variations in the phase spectra of impulsive signals of similar amplitude spectra do not affect subjective judgments of perceived noisiness. Two signals of identical amplitude spectra but different phase spectra may sound dissimilar, but they do not vary in judged noisiness.

2) The human ear appears to function as an energy detector in evaluating the noisiness of impulsive signals. Support for this conclusion was found in the Duration experiment (3 dB per duration doubling) and in the Repetition experiment (3 dB per doubling of signal number). There was insufficient evidence to substantiate the existence of a specific value for the ear's time constant.

3) A-weighted and N-weighted correction curves should be applied with caution to impulsive signals if the power spectra of the signals contain appreciable low frequency energy (as for example, sonic booms). Both the A-level and N-level contours appear to have too great a slope in the low frequency regions. Additional research is necessary to determine whether the A and/or N scales are directly applicable to impulse measurement.

APPENDIX A

PEST PROCEDURES EMPLOYED

FOR THE JUDGMENT TESTS

PEST PROCEDURE EMPLOYED FOR SUBJECTIVE JUDGMENT TESTS

Parameter Estimation by Sequential Testing (PEST) is a computer based adaptive psychophysical procedure which administers an iterative form of the standard paired comparison task to human observers. PEST is called an adaptive procedure because the sequence of signals heard by an observer is not fixed in advance, but rather is determined by his ongoing responses. PEST thus preserves many of the advantages of the paired comparison method while gaining the speed and convenience of an adjustment method.

BBN's implementation of the PEST is based on an interactive teletype conversation between the experimenter and the computer-based system. The system acquires information needed for conduct of an experiment by inquiring of the experimenter the values of a series of parameters which determine the course of the PEST procedure. Initially, the computer requests identification of the observer, the signals employed, and the experimental session. The next questions posed by the computer concern the relative levels at which signals are presented to the observer on subsequent trials.

The experimenter may then specify a standard operating procedure consisting of predetermined values of a dozen parameters such as the intersignal interval, intertrial interval, initial step size, maximum step size, degree of confidence in the observer's responses, anticipated direction of first step, and region of interest of the psychometric function.

A final question serves to delay onset of a trial series until the experimenter and observer are ready to procede. Upon receiving an affirmative response to the question "READY?", the computer types a data heading and awaits final confirmation in the form of "START" switch depression by an observer in an adjacent anechoic chamber.

The trial procedure is a two interval forced choice, in which one signal (the standard) is invariant over trials, while the other signal (the comparison) may change in level. Approximately one second after START switch closure, the computer presents a pair of signals and waits for the observer to decide on his preference for the signal of the first or second interval. Upon receipt of the observer's response, the computer calculates the level at which the comparison signal will be presented on the next trial. After another pause of approximately one second, the computer initiates the next trial by presenting a modified signal pair.

PEST determines the increment in comparison signal level as follows (Ref.16)

1. On every reversal of step direction, halve the step size.

2. The second step in a given direction, if called for, is the same size as the first.

3. Whether a third successive step in a given direction is the same as or double the second depends on the sequence of steps leading to the most recent reversal. If the step immediately preceding that reversal resulted from a doubling, then the third step is not doubled; while if the step leading to the most recent reversal was not the result of a doubling, then this third step is double the second.

4. The fourth and subsequent steps in a given direction are each double their predecessor (except that large steps may be disturbing to a human observer and an upper limit on permissible step size of 16 dB is maintained).

The system provides information about the progress of each run in the form of "UP" and "DOWN" lights (signifying the direction of change of comparison signal level on the current trial), and also in two digital counters which cumulate numbers of trials and of decision reversals.

A run, composed of a variable number of trials, is terminated when the system determines that sufficient information has been collected. The general stopping criterion for a run is satisfied when the anticipated step size is 1 dB. When a run terminates, the computer prints the number of the run, the level of the comparison signal on the last trial of the run, the number of trials in the run and the mean response latency. Throughout all testing, the program was set to determine the point of subjective equality, or the level at which observers judged the standard and comparison signals equally noisy.

Five consecutive PEST runs were administered for each pair of signals. Experimental sessions lasted 30 minutes, and were followed by five minute enforced rest periods in which observers left the anechoic chamber. The rate of pay for Study II was $2.50 per hour. In the remainder of the studies, the pay rate was contingent upon the observers' performance, as may be seen from the "Revised Pay Procedure" instructions in Appendix B.

APPENDIX B

INSTRUCTIONS USED FOR JUDGMENT TESTS

INSTRUCTIONS

The purpose of this test is to gather information about the relative noisiness of various sounds. The test is part of a program of research designed to obtain information that will be of aid in the planning of airports, airplanes, and for noise control purposes in general.

The computer will present a series of pairs of sounds. After each pair of sounds is presented, your task is to decide which of the two sounds, the first or the second, is the more noisy. Regardless of how you have previously defined noisy, by noisy, we mean that sound which is the more annoying, unacceptable, objectionable and disturbing if heard in your home during the day and night. Pick that sound which you would less like to have in your home, even though you might not want either of them.

The computer varies the characteristics of the two sounds in each pair on each trial. If you think the first sound of a pair is the more noisy, push button 1 on the metal response box. If you think the second sound is the more noisy, press the button labeled 2. It is important that you judge each pair of sounds on its own merits regardless of any similarities or differences you may hear among successive pairs of sounds. There are no right or wrong answers. We are interested only in how noisy or unnacceptable the sounds seem to you.

The response buttons will light up when the computer has been informed of your decision. The computer will wait for you to reach a decision about each pair of sounds before it will present the next pair of sounds. Therefore, you control the pace of the experiment directly. The more quickly you decide which sound was more noisy the more quickly the experiment will end. Most people find that they can make good decisions within a second or two after hearing the second sound of a pair.

The START button commands the computer to present the first pair of sounds. I will tell you when to push START. If you push the STOP button the computer will interrupt the test series. There should be no ordinary reason for pushing the STOP button during a series of trials. If you do have a reason for pushing STOP, please tell me before pushing START again. I will tell you when a series of trials has ended.

In summary, select the sound (the first or the second) which you feel is the more noisy, unacceptable, or disturbing. Remember to listen carefully to each pair of sounds, and to base your decision solely upon the current pair. If you have any questions, please feel free to discuss them with me at the end of a test series.

ABOUT THE EXPERIMENT

The experiments in which you are participating were designed to be administered by a computer. A PDP-8 digital computer in a room down the hall will be presenting a series of sounds which you will listen to in the anechoic chamber. The machine will control the presentation of the sounds; it will also record your responses and monitor your behavior in the experiment. For this reason, it is important that you follow the detailed instructions carefully, since the machine will interrupt the progress of the experiment if your responses seem inconsistent.

You will be able to communicate with the machine by pushing one of the buttons on a response box in the anechoic chamber. The machine will store the information you provide and then process the information in a manner specified by a program which we have written for it. Among other things, the machine will measure the amount of time it takes for you to reach a decision, your choices on each trial, the number of times you push each of the response buttons and so forth.

Of course, the experimenter will always be at hand to answer any questions that you might have and to take care of any problems that may arise.

REVISED PAY SCHEDULE

A new pay schedule has been established. The rate of pay, previously a flat $2.50 per hour, has been revised to encourage and reward alert performance. The hourly pay rate will henceforth be related to the number of runs which the observer completes in each hour, according to the schedule below. A run is completed when the computer lights the stop button.

NUMBER OF COMPLETED RUNS	PAY RATE PER HOUR
18 or fewer	$1.75
20	1.90
23	2.00
27	2.25
32	2.50
40	3.00
50 or more	3.50

The calculations for the new rates are based on two five minute breaks per hour for the observer and ten minutes of preparation time for the computer operator. The computer will keep track of the number of completed runs and print the total at the end of each experimental session.

The number of runs which an observer completes in an hour depends upon several factors, notably:

1. The consistency of his decisions
2. The amount of time required to reach each decision
3. The amount of "dead time" the observer introduces by pushing the stop button, responding too quickly, etc.

In general, an observer can make his most consistent judgments about the relative annoyance of two sounds by consciously concentrating his attention on the signals and preventing his mind from wandering. Haphazard decisions prolong the duration of each run, since the computer is forced to acquire more information about the inconsistencies in the observer's decisions.

Paying careful attention to the pairs of sounds will permit the observer to make the greatest amount of money in the least amount of time. The decision times and trial lengths will take care of themselves. The computer is programmed to administer experimental conditions as quickly and efficiently as the observer will permit.

APPENDIX C

EQUIPMENT FOR JUDGMENT TESTS

Error

An error occurred while processing this page. See the system log for more details.

Figure C-2. The signals were presented in the anechoic chamber and recorded on magnetic tape for further analyses. The microphone during the recording was located in the position of the observer's ear during a standard test run. The recorded signals were analyzed using "N" and "A" weighting networks in conjunction with a true RMS volt meter. Because of their impulsive nature, some of the sounds were repeated in concatenated form to measure steady state rather than impulsive signals. Detailed analyses, including narrow band and 1/3 octave band measures, were undertaken by a Fast Fourier transform program in our Cambridge office. Perceived noise levels were calculated from the 1/3 octave band values.

APPENDIX D

MEASUREMENT PROCEDURES

SIGNAL GENERATION AND MEASUREMENT PROCEDURES

As discussed in the Purpose and Method Section, the Phase Experiment required generation of replicas of several prototype signals, each having the same magnitude spectrum but widely different phase spectra. Several digital methods of achieving the phase distortion of the prototype signal without alteration of its spectral magnitude were attempted.

First, we synthesized all-pass networks (with symmetrically located pairs of poles and zeroes) and filtered our prototype signals through them. This produced replicas of identical magnitude spectra. Unfortunately, the degree of phase coherence in the networks was too high, so that the phase distortion accomplished was insufficient. In other words, the replicas were too much like the prototypes.

Next, we tried incorporating more phase incoherence by increasing the number of poles and zeroes, using a method first proposed by Huffman. In spite of improvement in the results, the degree of phase distortion obtained was still considered insufficient and the method was abandoned.

The third method which we adopted made use of the inverse transform capabilities of the available Fast Fourier Transform program. This method consisted of computing the Digital Fourier Transform of a prototype signal, changing the phase of each spectral component and generating a replica by computing the Inverse Digital Fourier Transform. More specifically, given a prototype digital signal of N samples (data points), the method consisted of:

1) computation of the N/2 complex values (magnitude and phase) corresponding to the Digital Fourier Transform of the prototype;

2) generation and addition of a pseudo-random phase angle to each of the N/2 phase angles of the prototype; and

3) application of the inverse Digital Fourier Transform algorithm to obtain a new waveform having the same spectral magnitude but a radically different and totally incoherent phase spectrum.

The results obtained with this method were considered satisfactory.

The number of samples, N, was determined by the experimental conditions that, on the one hand required signals of about 20 milliseconds duration, and on the other hand involved sound reproduction equipment that limited the high frequency response to about 10 kHz. Consequently, with a Nyquist sampling frequency of 20,000 samples per second, N had to be about 400. Given the requirements of the Fast Fourier Transform program, the number finally adopted was 512.

The basic plan was followed throughout the study, but several modifications were necessitated. The most important one arose from a difference between the actual playback conditions of the signals and the one assumed by the method of calculating the spectra. The Digital Fourier Transform algorithm actually computes the spectrum corresponding to a periodic waveform obtained by concatenating (adjoining end to end and ad infinitum), the basic N-sample waveform. The actual playback conditions, however, imply widely spaced occurrences of the waveform in time. The consequences of this temporal disparity are important for our purposes for the reasons discussed below.

Figure D-1 (a) represents one of our real impulsive waveforms (T0) and its spectrum in actual playback conditions, that is, isolated in time. Calculating the spectrum by means of Digital Fourier Transform techniques implies extraction of samples from the waveform at a rate equal to or greater than twice the maximum frequency present in the signal. If we extract samples from the signal only during the epoch for which it differs significantly from zero, (interval 0, T in Fig. D-1 [a]) we have the case depicted in Fig. D-1 (b). Here the calculated spectrum $F_A(\omega)$ consists of N/2 sampled values, where N is the number of samples extracted from the waveform.

The separation between spectrum samples is $\frac{2\omega_{MAX}}{N}$, where ω_{MAX} is the maximum angular frequency contained in the signal. Since no new information is added to the waveform by adding to it a "guard zone" of 0-valued samples (or by adding to an acoustic signal a "guard zone" of silence), we did not expect to see any changes in the spectrum other than the ones due to the increased resolution. If the proper precautions are adopted, namely, if the addition of the "guard zone" is effected without introducing spurious high frequency components, this is indeed what is observed. But inadvertant prolongation with 0's of a waveform that does not start (and end) on a 0-value may indeed introduce this high frequency distortion. All waveforms having zero d.c. value can be made to start on zero value by simply introducing a time delay equal to the distance to their first zero-crossing. In other words, the discontinuity can be eliminated at the cost of an additional linear change of the phase spectrum. In Fig. D-1 (c) we have represented what happens after one such careful prolongation; namely, increased spectral resolution without spectral distortion.

It must be noticed that measures of noise-levels generally operate on average spectral power within bands defined in a logarithmic frequency scale, for example, 1/3 octave bands. In such cases, increased resolution may introduce changes in the results that appear surprising at first sight. If the frequency resolution of $2\omega MAX/N$ is about the same as the width of the lower 1/3 octave bands, energy may or may not be passed through such a filter, as in Fig. D-1 (b) and (c). In other words, $2\omega MAX/N$ may be sufficiently great at low frequencies that errors in energy measurements may occur as a function of the bandwidth of the measurement procedure. Thus, differences in low frequency measurements among signals of very similar spectra may be observed simply because their spectra are not as closely defined in the low frequency region as elsewhere. In the case of the signals generated for the current study, individual spectral components were spaced approximately 40 cycles apart. Since 1/3 octave bands at low frequencies may be as narrow as 10 or 20 cycles, calculations (such as PNdB) performed on the basis of 1/3 octave analyses could be appreciably biased by the presence or absence of a single spectral component.

Measurement Procedures

After playback through the experimental sound reproducing system, signals were recorded on magnetic tape and made available for digital analysis and measurement.

In order to interpret the numbers obtained by digital signal processing, signal calibration and the basic relation between analog voltages and digital values must be explained. When the tapes were played back into the computer, the playback gain was adjusted so that a 124 dB calibration tone (with an attenuator set at 100 dB) resulted in a sinusoidal signal of amplitude 10 volts. This voltage corresponds to the maximum amplitude that the A/D converter can digest, and results in a converted value of $2047 = 2^{11}-1$. The level of 100 dB was arbitrarily assigned to this value, so that if P is some number representing the power of a signal, its relative dB level will be given by

$$dB = 10 \log P + C$$

where C is such that

$$100 = 20 \log 2047 + C. \quad C = 33.8$$

The signals were digitized at a sampling rate of 20,000 samples per second, equivalent to a sampling interval of 50 microseconds. Since the signals have a nominal duration of 20 msecs and 4096 samples were taken, there is roughly a 10:1 ratio sampling aperture and signal duration.

The measures obtained were:

a) Signal power level

b) A-scale value

c) N-scale value

The signal power level (SPL) is the dB level of the average of the sum of the squares of the signal values, that is

$$SPL = 10 \log \left(\frac{1}{4096} \sum_{n=1}^{4096} f^2 (nT) \right) + C$$

where the f(nT) are the sampled values.

$$\text{The A-scale value} = 10 \log \left(\sum_{n=1}^{2048} F(\tfrac{n}{4096T}) S_A (\tfrac{n}{4096T})^2 \right) + C$$

where $F(\frac{n}{4096T})$ is the magnitude of the nth value of the discrete Fourier transform of f(nT) and $S_A(\frac{n}{4096T})$ is the value of the "A" frequency response interpolated from the published 3rd-octave values.

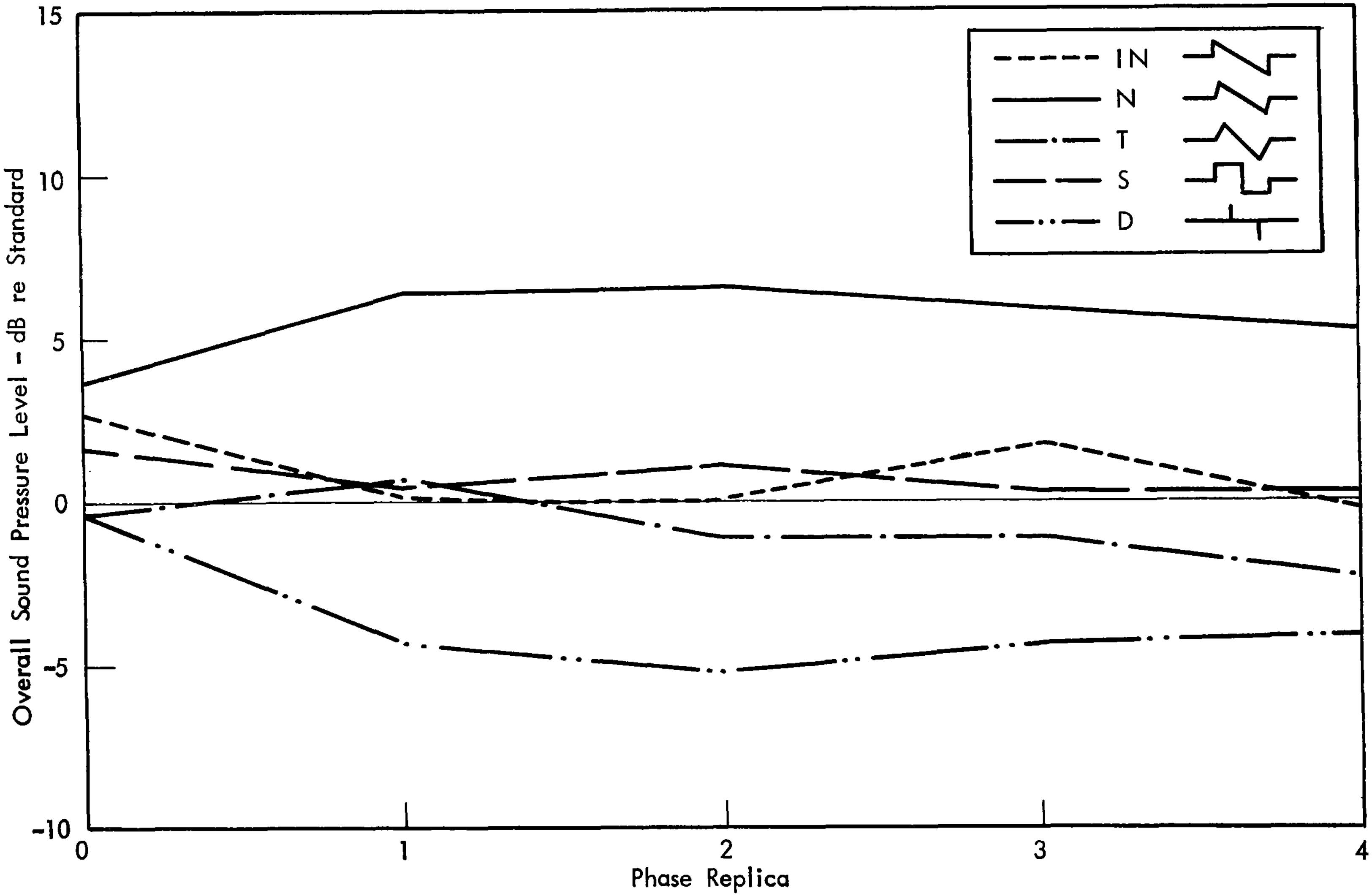


FIGURE 1. RESULTS OF PHASE EXPERIMENT IN TERMS OF OVERALL SOUND PRESSURE LEVEL
Impulses of Various Phase Spectra Judged Equally Noisy to Doublet (D0) Standard

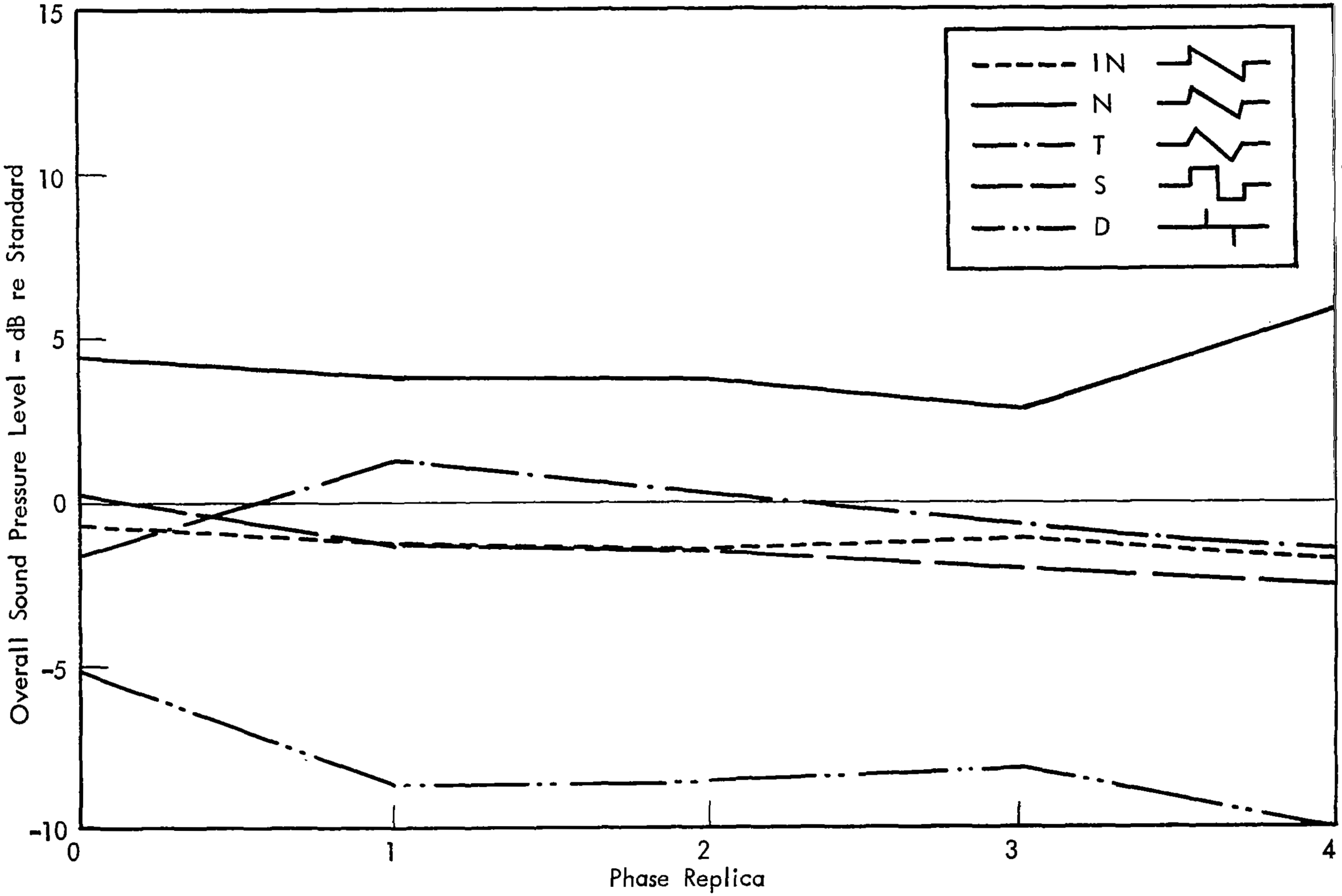


FIGURE 2. RESULTS OF PHASE EXPERIMENT IN TERMS OF OVERALL SOUND PRESSURE LEVEL
Impulses of Various Phase Spectra Judged Equally Noisy to Ideal N Wave (IN0) Standard

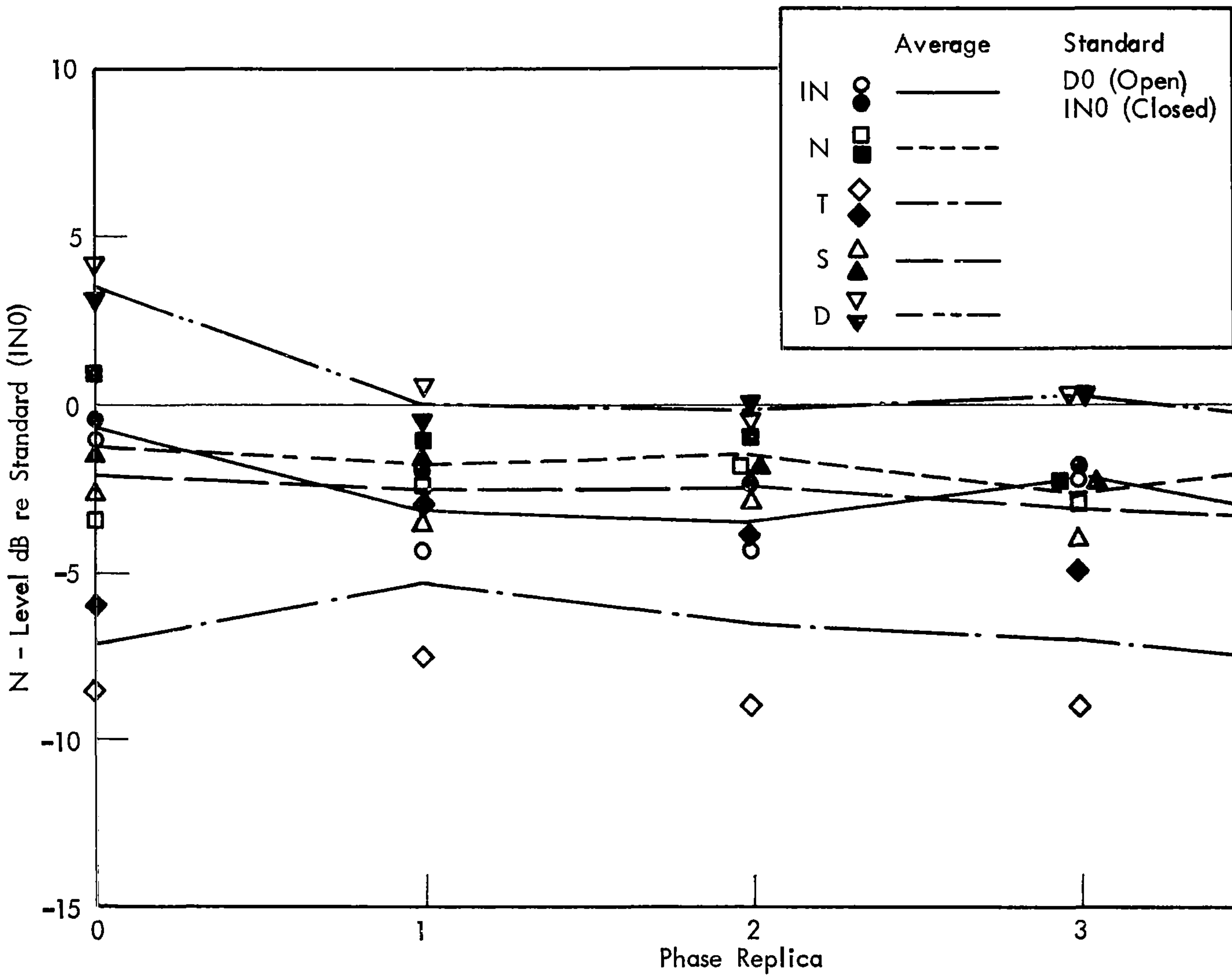


FIGURE 3. RESULTS OF PHASE EXPERIMENT IN TERMS OF N-LEVEL RELATIVE TO IN0 STANDARD
Impulses of Various Phase Spectra Judged Equally Noisy to Ideal N Wave (IN0) or Doublet (D0)

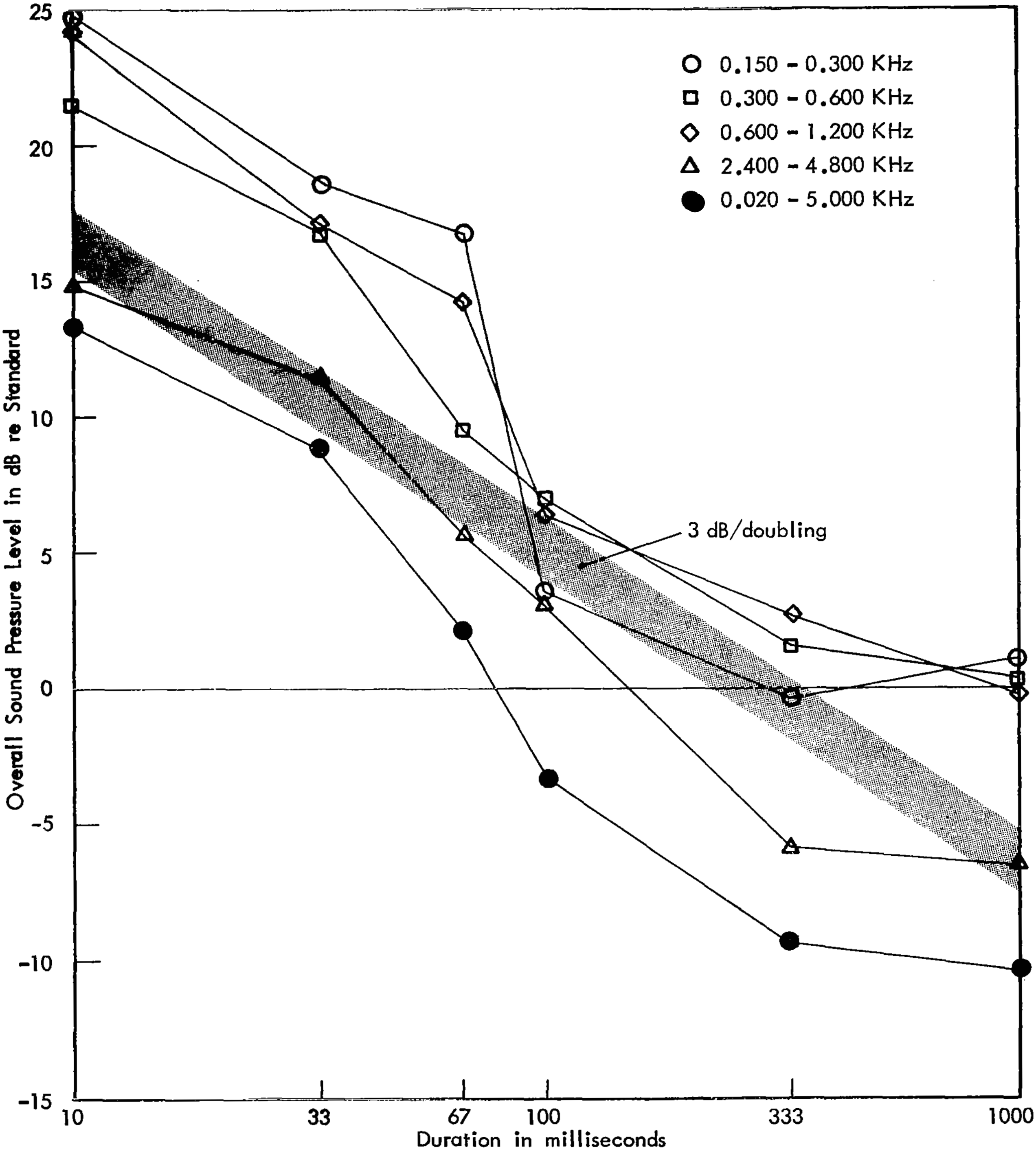


FIGURE 4. RESULTS OF DURATION TEST IN TERMS OF OVERALL SOUND PRESSURE LEVEL
Bands of Noise Judged Equally Noisy to Octave Noise Band Standard Between 0.600 and 1.200 KHz

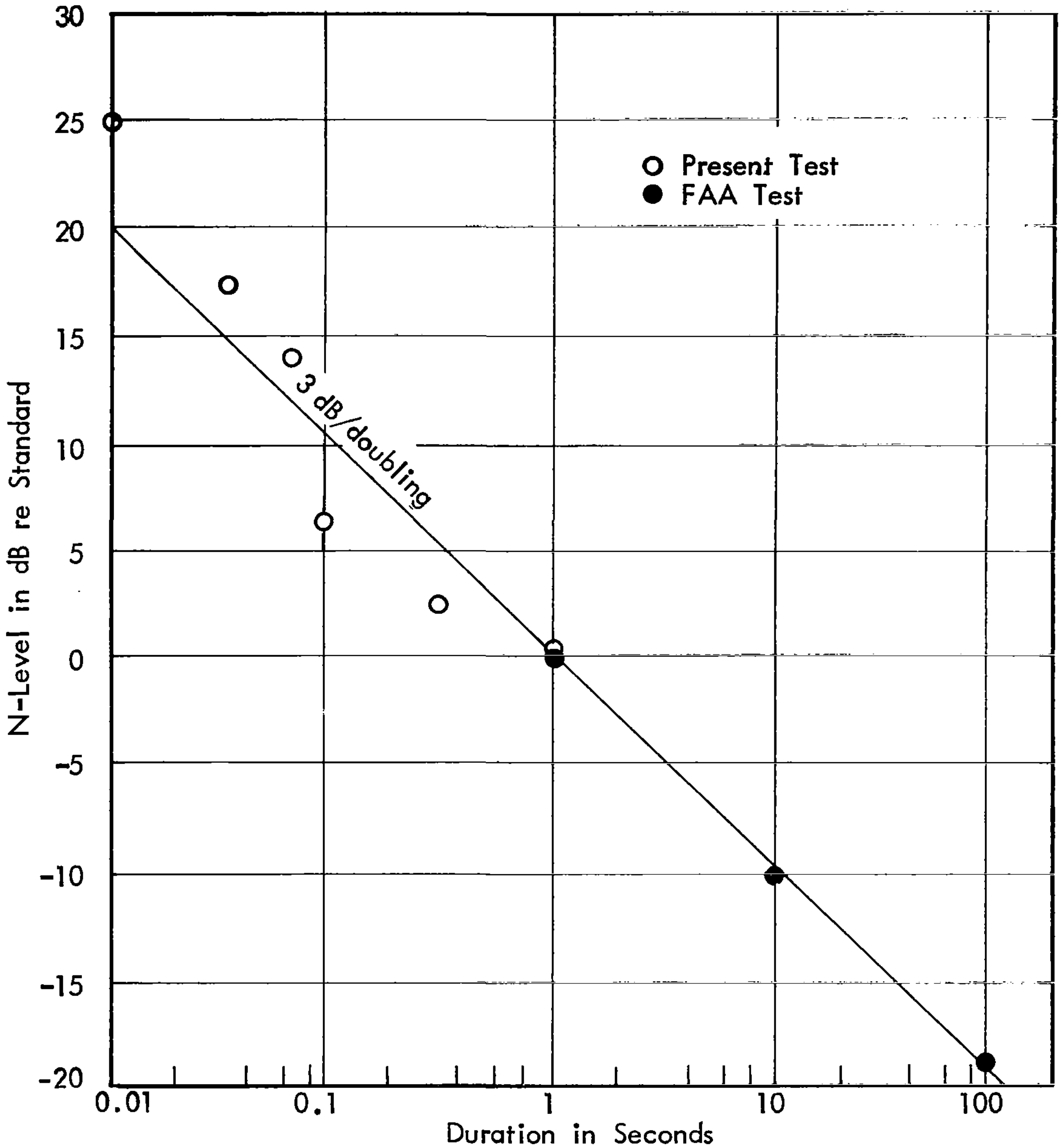


FIGURE 5 . RESULTS OF PRESENT DURATION TEST COMBINED WITH LONGER DURATION STIMULI FROM FAA TEST
Octave Bands of Noise of Various Durations Judged Equally Noisy to 1 Second Octave Band of Noise (Standard)

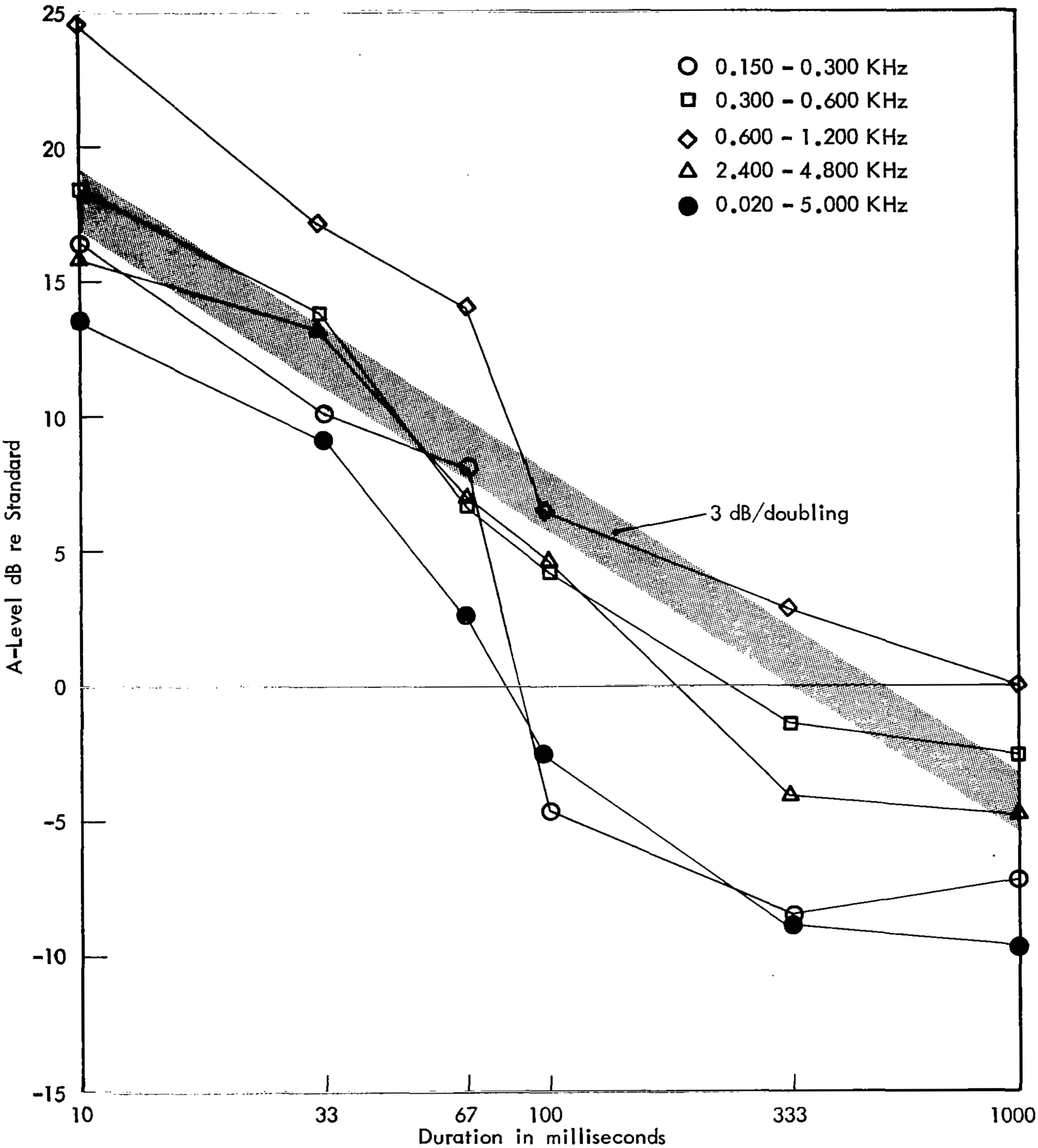


FIGURE 6. RESULTS OF DURATION TEST IN TERMS OF A-LEVEL
Bands of Noise Judged Equally Noisy to Octave Noise Band Standard Between 0.600 and 1.200 KHz

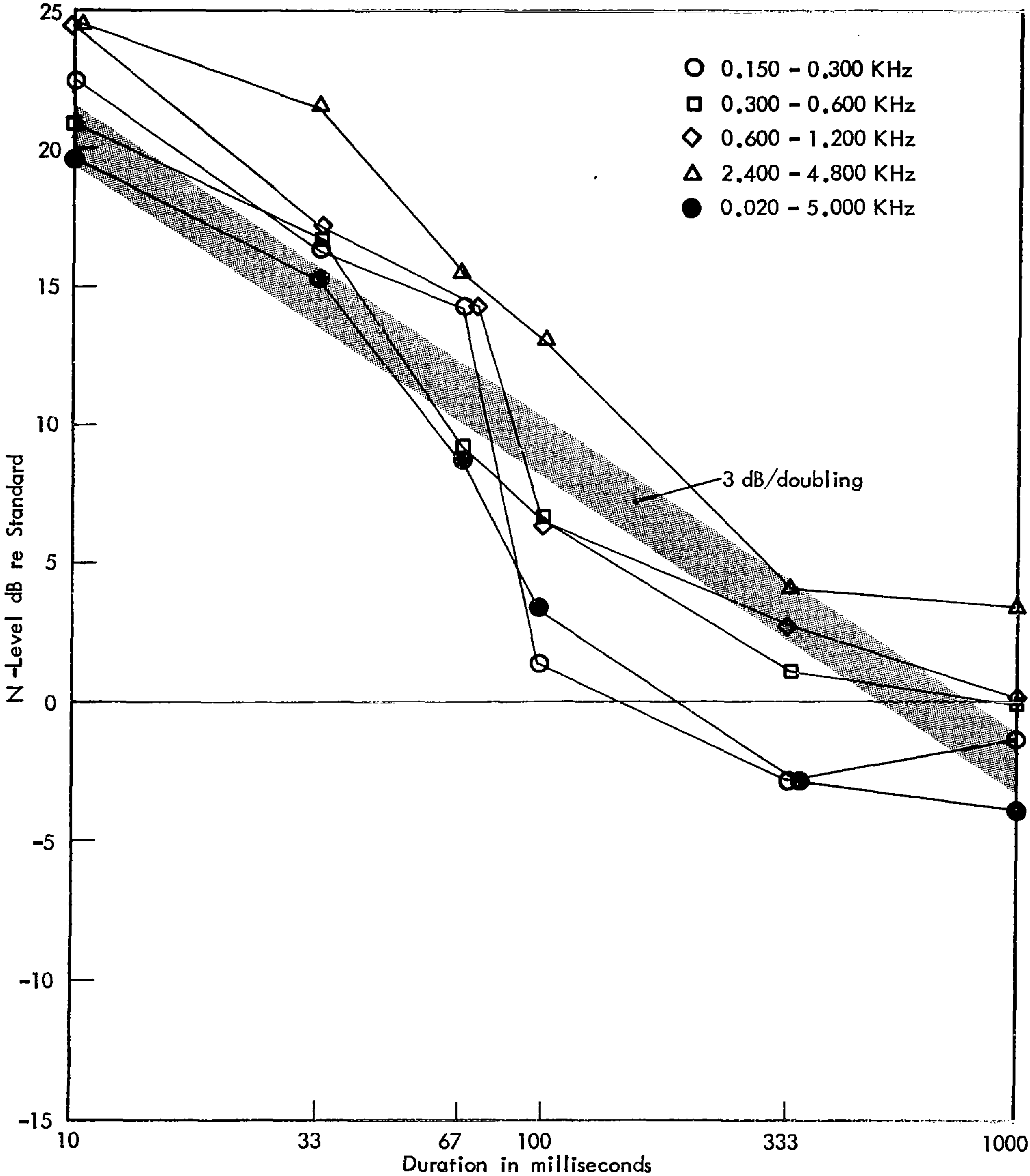


FIGURE 7. RESULTS OF DURATION TEST IN TERMS OF N-LEVEL
Bands of Noise Judged Equally Noisy to Octave Noise Band
Standard Between 0.600 and 1.200 KHz

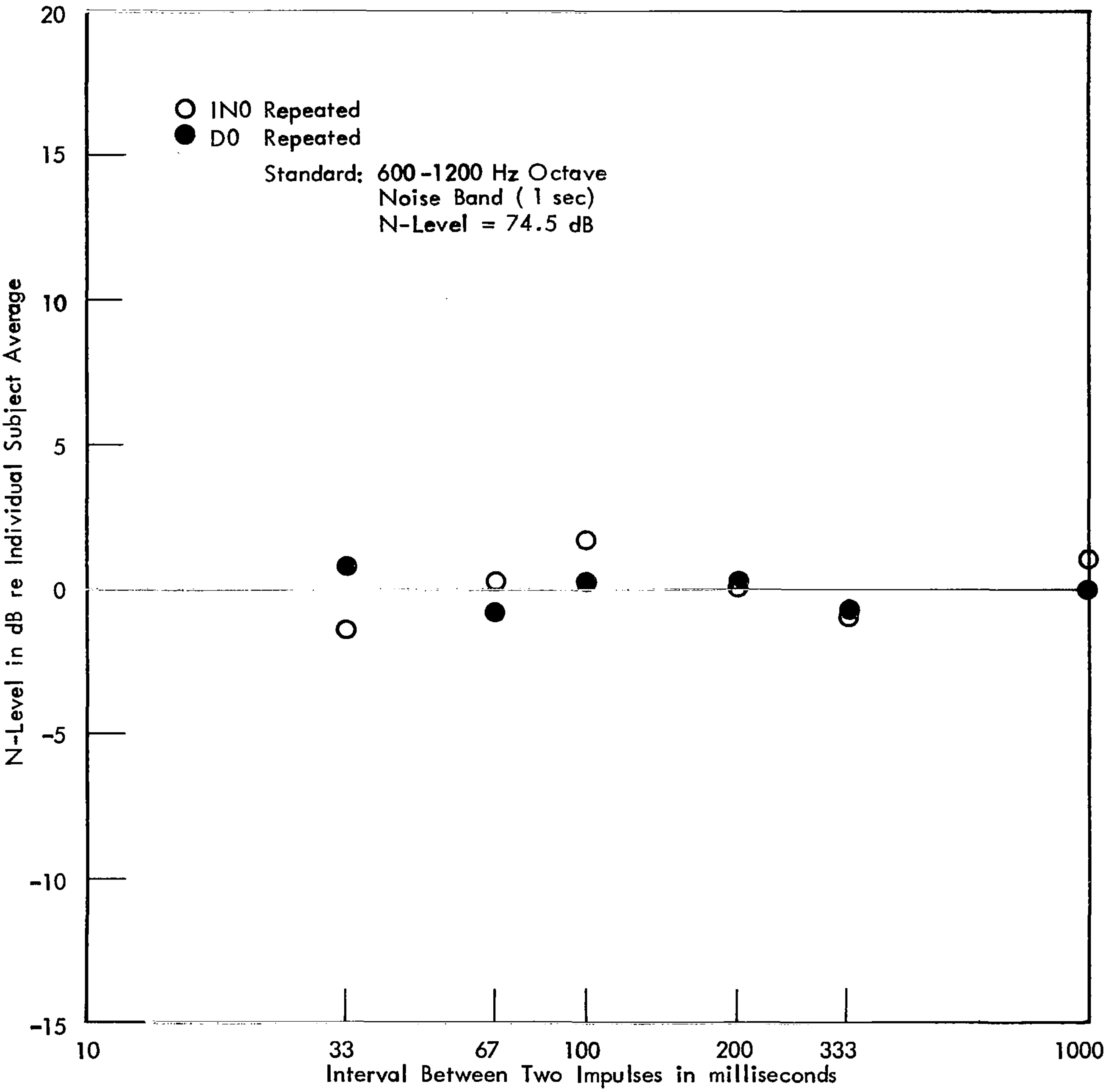


FIGURE 8. RESULTS OF INTERVAL EXPERIMENT
Two Impulses with Various Interval Spacing Judged Equally Noisy to Octave Band of Noise

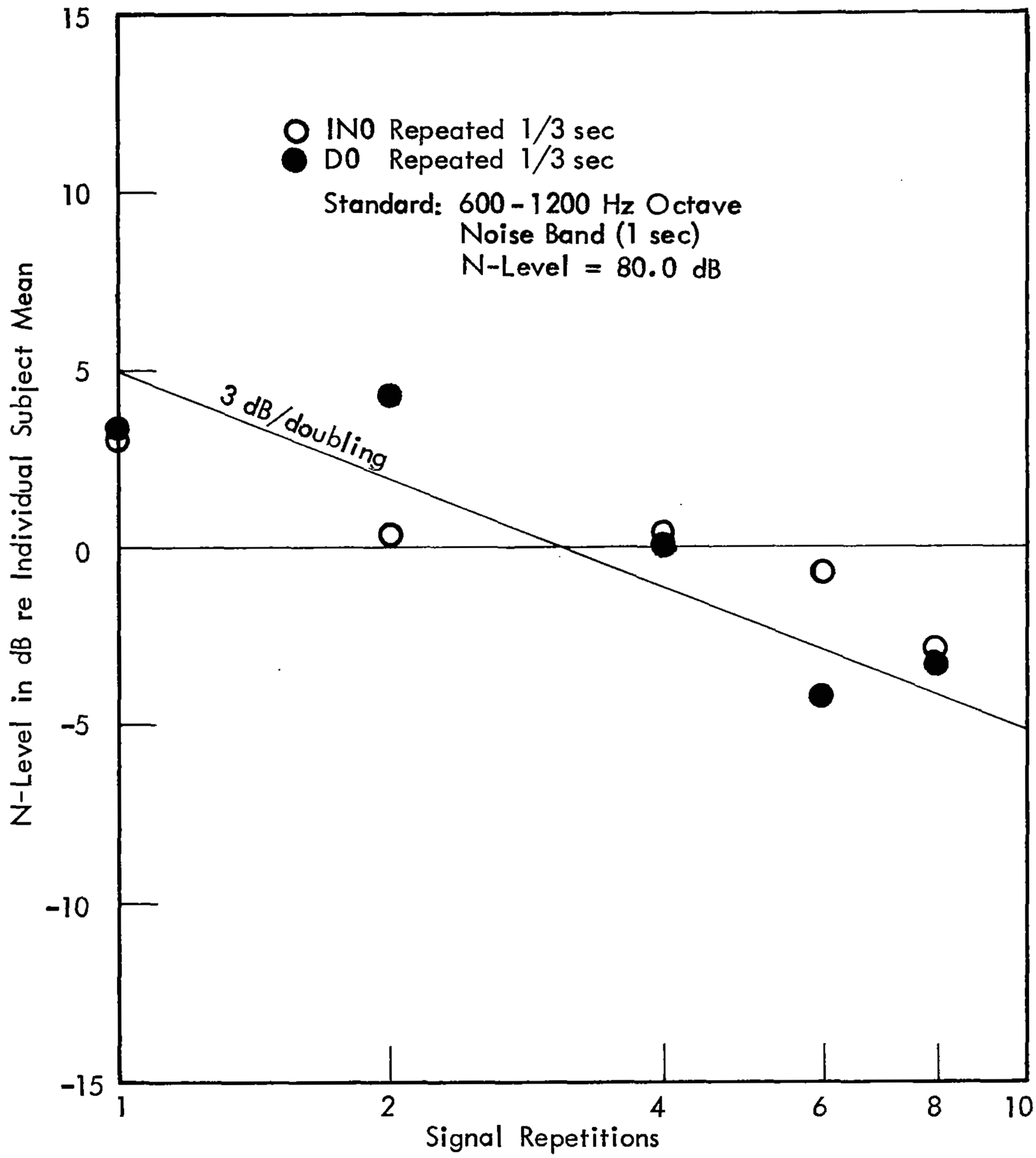


FIGURE 9. RESULTS OF REPETITION EXPERIMENT
Repeated Impulses Judged Equally Noisy to Octave Band of Noise

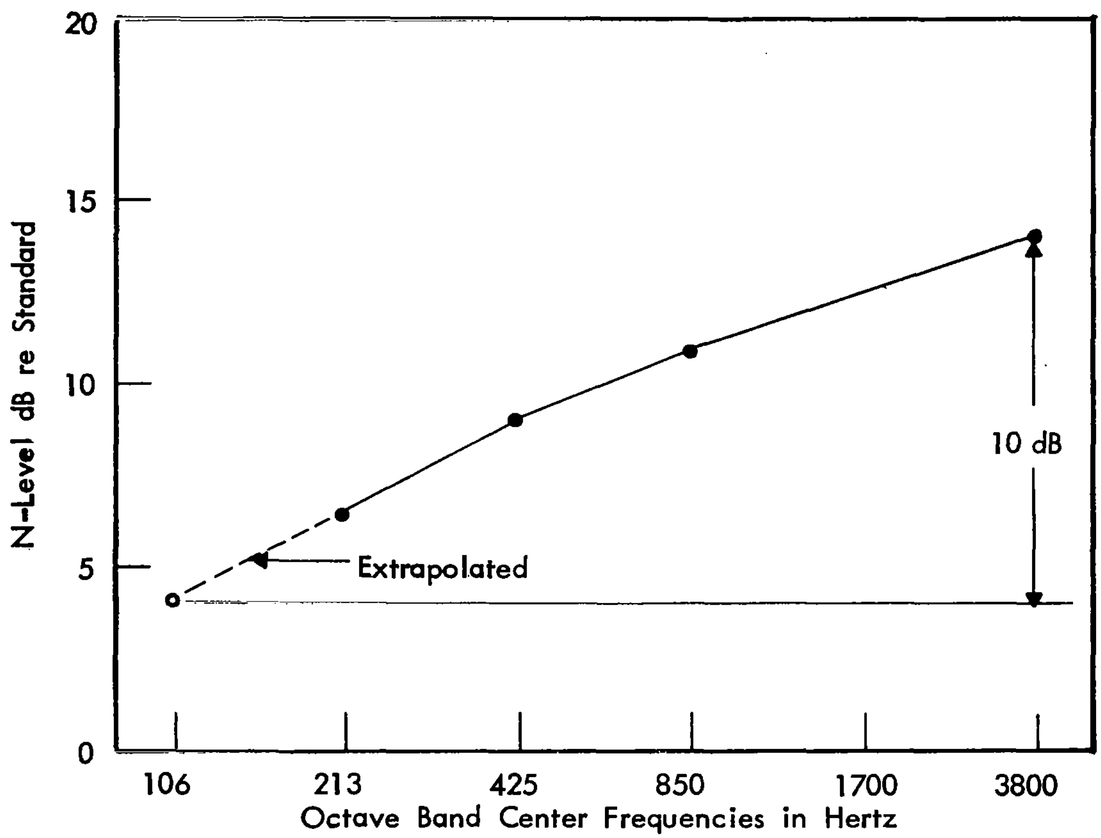


FIGURE 10. RESULTS OF DURATION TEST AVERAGED FOR EACH OCTAVE NOISE BAND

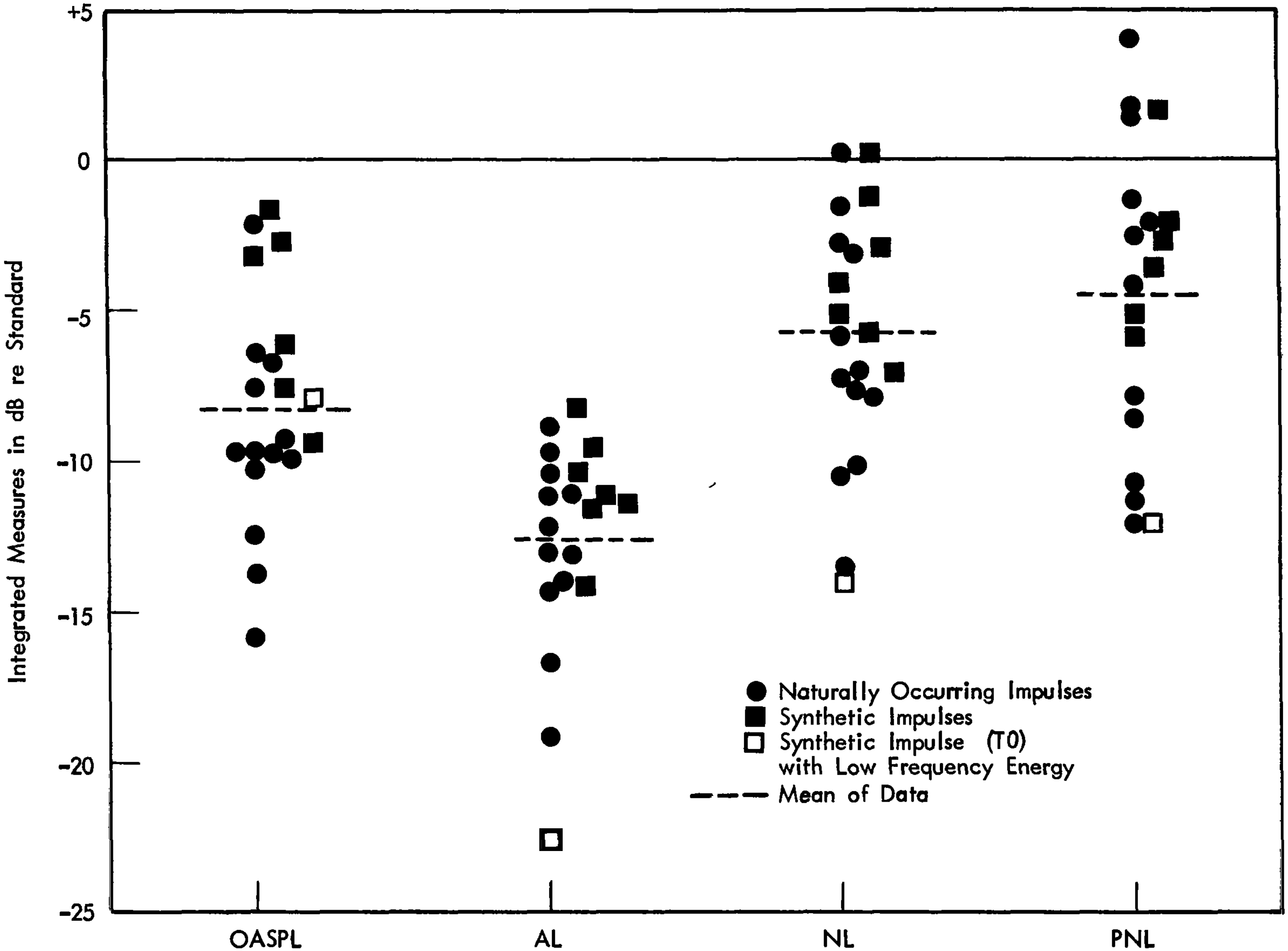


FIGURE 11. RESULTS OF EVALUATION EXPERIMENT IN TERMS OF VARIOUS MEASURES
Impulse Noises Judged Equally Noisy to 1 Second Octave Band of Noise (Standard)

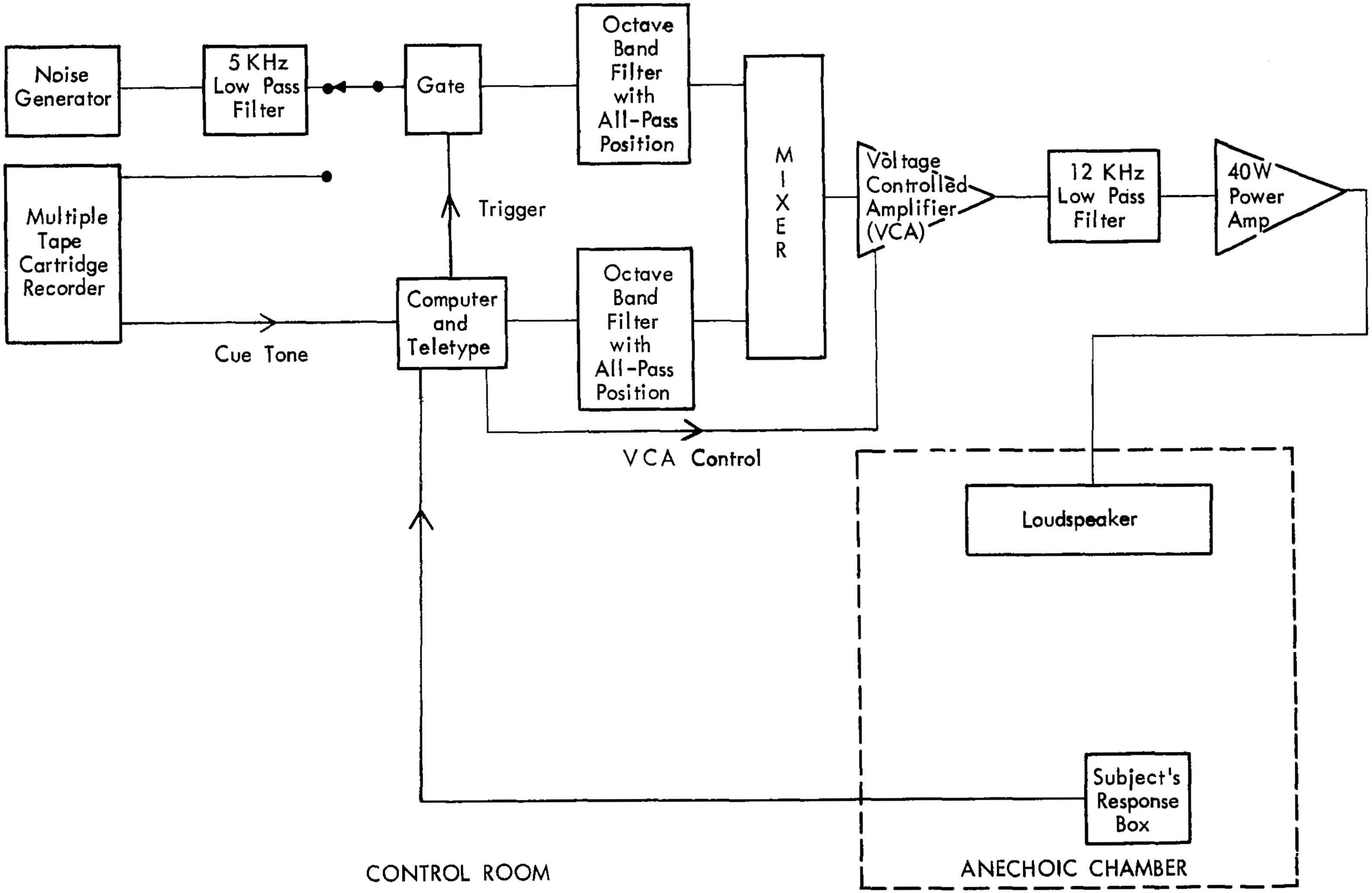


FIGURE C-1. BLOCK DIAGRAM OF STIMULUS PLAYBACK SYSTEM FOR JUDGMENT TESTS

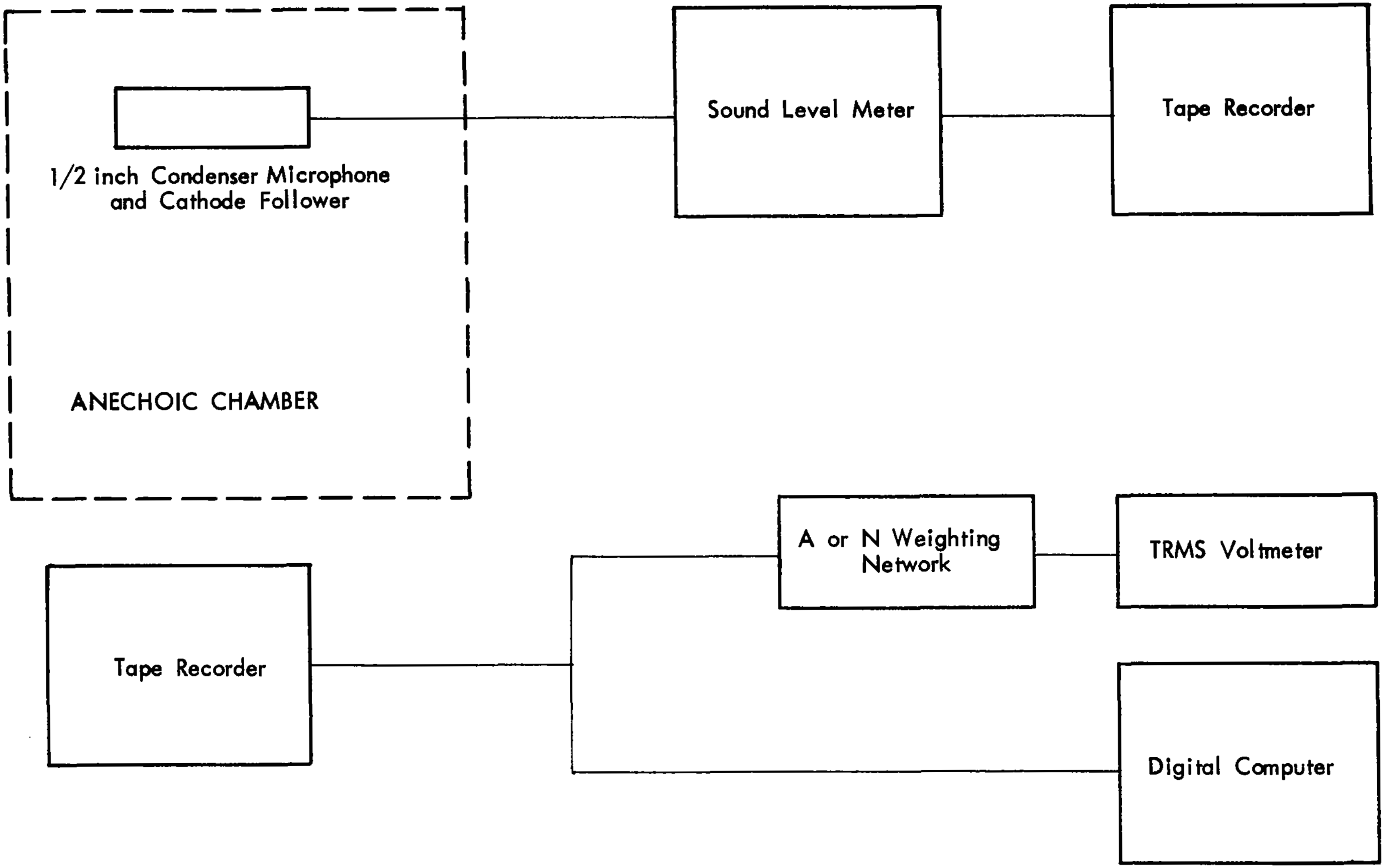


FIGURE C-2. BLOCK DIAGRAM OF STIMULUS ANALYSIS EQUIPMENT FOR JUDGMENT TESTS

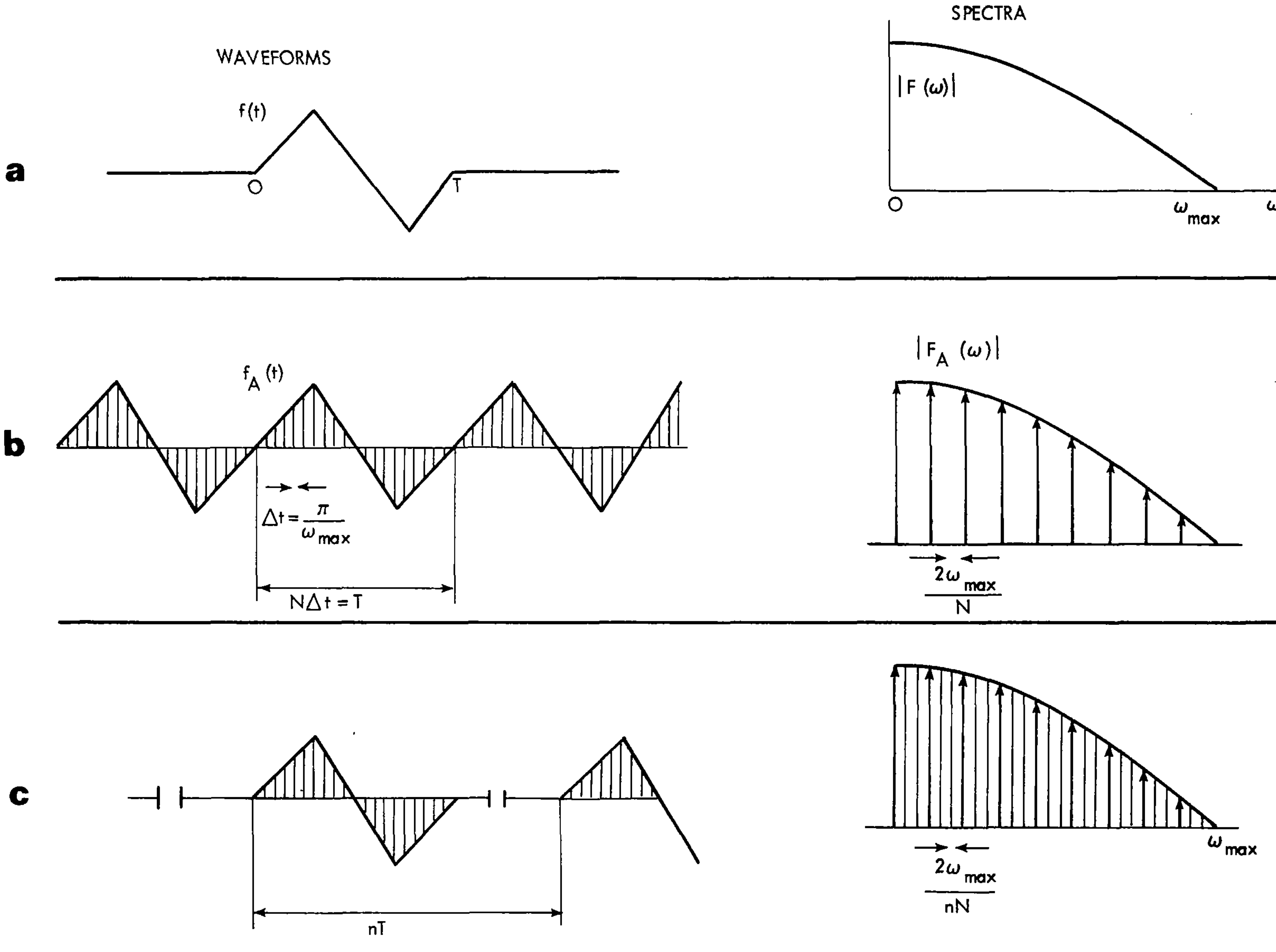


FIGURE D-1 ILLUSTRATIONS OF WAVEFORMS AND SPECTRA FOR VARIOUS SAMPLING TECHNIQUES

REFERENCES

1. von Bekesy, G., "Zur Theorie des Horens. Ueber die Bestimmung des einem reinen Tonempfinden entsprechenden Erregungsebietes der Basilarmembran vermittels Ermudungserscheinungen" Phys. Z. 30, 115-125, 1929.

2. Munson, W.A., "The Growth of Auditory Sensation," J. Acoust. Soc. Am., 19, 584-591, 1947.

3. Miller, G.A., "Perception of Short Bursts of Noise," J. Acoust. Soc. Am., 20, 160-170, 1948.

4. Small, A.M., Jr., J.F. Brandt, and P.G Cox, "Loudness as a Function of Signal Duration," J. Acoust. Soc. Am., 34, 513, 1962.

5. von Port, E., "Uber die Lautstarke Einselner Kurzer Schallimpulse," Acustica, 13, 211-223, 1963.

6. von Niese, H., "Methode zur Bestimmung der Lautstarke Beliebiger Gerausche," Acustica 15, 117-126, 1965.

7. von Zwicker, E., "Ein Beitrag zur Lautstarkemessung impulshaltiger Schalle," Acustica, 17, 11-22, 1966.

8. von Zwicker, E., "Ein Beitrag zur Lautstarkemessung impulshaltiger Schalle," Acustica, 17, 22-25, 1966.

9. Stevens, J.C. and J.W. Hall, "Brightness and Loudness as a Function of Stimulus Duration," Perception and Psychophysics, 1, 319-327, 1966.

10. Garner, W.R., "Loudness of Short Tones," J. Acoust. Soc. Am., 21, 398-403, 1949.

11. von Reichardt, W., H. von Niese, and M. Müller, "Die Rauhigkeilskomponente der Lautstarke von Pulsen und ihre Nachbildung im Impulsschallpegelmesser," Hochfrequenstechnik u. Electrakustik, 75, 31-38, 1966.

12. Sheperd, L.J. and W.W. Sutherland, "Relative Annoyance and Loudness Judgments of Various Simulated Sonic Boom Waveforms," NASA Report CR-1192, 1968.

13. International Standards Organization, "Standard Reference Zero for the Calibration of Pure-Tone Audiometers," ISO Recommendation R-389.

REFERENCES (cont'd)

14. Blackman, R.B., and J.W. Tukey, "The Measurement of Power Spectra," Dover Publication Inc., New York, N.Y., 1959.

15. Corliss, E. and G.E. Winzer, "Study of Methods for Estimating Loudness," J. Acoust. Soc. Am., 38, 424,428, 1965.

16. Taylor, M.M. and C.D. Creelman, "PEST: Efficient Estimates on Probability Functions," J. Acoust. Soc. Am., 41, 782-787, 1967.