NASA CONTRACTOR REPORT

NASA CR-1779

STUDY OF THE EFFECTS OF THE DOPPLER SHIFT ON PERCEIVED NOISINESS

by Karl S. Pearsons, Ricarda Bennett, and Sanford Fidell

Prepared by
BOLT BERANEK AND NEWMAN, INC.
Canoga Park, Calif. 91406
for Langley Research Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • JULY 1971

1. Report No. NASA CR-1779	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle STUDY OF THE EFFECTS OF THE DOPPLER SHIFT ON PERCEIVED NOISINESS		5. Report Date July 1971
		6. Performing Organization Code
7. Author(s) Karl S. Pearsons, Ricarda Bennett, Sanford Fidell		8. Performing Organization Report No.
9. Performing Organization Name and Address Bolt Beranek and Newman, Inc. Canoga Park, California		10. Work Unit No.
		11. Contract or Grant No. NAS1-9427
		13. Type of Report and Period Covered Contractor Report
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

Subjective judgments of the effects of Doppler shifts on perceived noisiness were made by 20 college students in an anechoic chamber. The stimuli heard in the tests included both recorded and simulated aircraft flyovers. Computer controlled generation of the simulated flyovers permitted independent variation of the source frequency, apparent altitude, amplitude, and Doppler patterns. Data collection was governed by a computer based adaptive technique known as Parameter Estimation by Sequential Testing (PEST).

The major finding was that Effective Perceived Noise Level (EPNL) is a fairly accurate predictor of noisiness of flyovers containing Doppler shifts; except perhaps at altitudes of less than 500 feet, for which flyovers it underestimates apparent noisiness. Compensation for the rise time of the stimuli (an onset correction) did not improve the accuracy of prediction provided by Effective Perceived Noise Level. The frequency of pure tone components in the stimuli did not influence the subjects' judgments.

17. Key Words (Suggested by Author(s)) Aircraft Noise; Subjective Response to Noise; Perceived Noise Level; Doppler Shift	18. Distribution Statement Unclassified - Unlimited

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 54	22. Price* $3.00

*For sale by the National Technical Information Service, Springfield, Virginia 22151

ACKNOWLEDGMENTS

The authors express their thanks to Dr. David M. Green for his assistance in planning the experimental design, and to Mr. Brian Curtis for his assistance in stimulus preparation and data collection.

TABLE OF CONTENTS

STUDY OF THE EFFECTS OF THE DOPPLER SHIFT ON PERCEIVED NOISINESS

By Karl S. Pearsons, Ricarda Bennett, and Sanford Fidell

Bolt Beranek and Newman Inc.

SUMMARY

Subjective judgments of the effects of Doppler shifts on perceived noisiness were made by 20 college students in an anechoic chamber. The stimuli heard in the tests included both recorded and simulated aircraft flyovers. Computer controlled generation of the simulated flyovers permitted independent variation of the source frequency, apparent altitude, amplitude, and Doppler patterns. Data collection was governed by a computer based adaptive technique known as Parameter Estimation by Sequential Testing (PEST).

The major finding was that Effective Perceived Noise Level (EPNL) is a fairly accurate predictor of noisiness of flyovers containing Doppler shifts; except perhaps at altitudes of less than 500 feet, for which flyovers it underestimates apparent noisiness. Compensation for the rise time of the stimuli (an onset correction) did not improve the accuracy of prediction provided by Effective Perceived Noise Level. The frequency of pure tone components in the stimuli did not influence the subjects' judgments.

INTRODUCTION

Previous studies of the noisiness of aircraft flyovers have investigated the effects of spectral shape, duration, modulation and multiple tones. In these prior studies the test stimuli were systematically varied along a single dimension so that the effects of these parameters could be studied in an orderly fashion. Recent work by Ollerhead at the Wyle Laboratories (Ref.1) and Nixon, von Gierke, and Rosinger (Ref.2) at Wright-Patterson Air Force Base have suggested that Doppler shift and/or the related temporal pattern of an aircraft flyover may influence the noisiness of these flyovers. However, the conclusions of these two groups are somewhat contradictory. Ollerhead's data implies that the presence of a Doppler shift in an aircraft flyover increases its perceived noisiness more at low altitudes than at high altitudes. Conversely, one interpretation of Nixon's data reveals that those aircraft exhibiting a long onset duration have a higher altitude and hence produce an increase in the judged noisiness of the flyover.

The main purpose of the present study was to gather more information on the judged noisiness of signals containing the Doppler shift. Among the test stimuli are some with time profiles similar to those used by Nixon and some actual recordings of flyovers similar to those used by Ollerhead. A number of flyovers were simulated using a special computer program that allowed separate control of the important stimulus parameters. Thus in the first series of tests the altitude, and the Doppler and time patterns were independently varied. The second test employed realistic time and Doppler patterns but the source frequency and altitude were varied. The results of these tests are compared with the conclusions of previous data and are analyzed in terms of a variety of procedures designed to predict the apparent noisiness of the flyovers.

PREVIOUS RESEARCH

Nixon, in his study at Wright-Patterson AFB, investigated the psychological response judgments of annoyance to "approaching" and "receding" sounds. His data supported the hypothesis that a continuous increase in noise frequency and intensity as the aircraft appears to approach the observer creates anxiety thus magnifying the annoyance or objectionableness of the noise. The stimuli employed were mainly pure tones at 125, 1000, and 4000 Hz slowly varying monotonically in frequency by plus or minus one-half an octave in 15 seconds. Also, the intensity of the tones was varied over a range of 30 dB in approximately 15 seconds. The temporal patterns of the two principal stimuli Nixon used are illustrated in Figure 1. It might be noted that the pattern of frequency variation is similar to that of amplitude variation. Those sounds which increase slowly to a peak and then turn off relatively quickly Nixon defines as "approaching", and those with sharp onset and gradual decline are termed "receding" flyovers. Different combinations of increasing and/or decreasing frequency and intensity, along with additional temporal patterns were used for his study. The results suggested that "approaching" sounds were judged more annoying than "receding" sounds regardless of whether the tone frequency was increasing or decreasing during the stimulus presentation. The approximate magnitude of the difference was 5 dB. That is, the maximum level of the "approaching" signal had to be set 5 dB less than the level of the "receding" signal in order to be judged equally annoying.

Kryter, citing Nixon's data (Ref.3), suggests that an onset correction be applied to the EPNL of sounds in accordance with the function shown below:

$$\text{Onset correction} = \begin{cases} 0 & \text{for } T \leq 3.5 \text{ sec.} \\ 10 \log \left(\frac{T_{15}}{3.5}\right) & \text{for } 3.5 < T < 35 \text{ sec.} \\ 10 & \text{for } T \geq 35 \text{ sec.} \end{cases}$$

The onset correction is based on the amount of time the sound takes to reach a peak from a level 15 dB below the maximum level. Kryter mentions other definitions for onset time in his original report, however, the one stated above will be used in the analysis of the data for this study.

The tests conducted by Ollerhead at Wyle Laboratories demonstrated that judgments of noisiness based on actual recordings of aircraft flyovers appear to require a Doppler correction. He suggested the following formula:

$$\text{Doppler correction} = 5 \log \left(\frac{\text{alt. in feet}}{2200}\right)$$

This correction uses the altitude of an aircraft and hence relates indirectly to the duration of the aircraft noise. In essence the Doppler shift correction is an attenuated duration correction (1.5 dB/doubling rather than 3 dB/doubling).

In Ollerhead's test of recordings of actual flyovers, the EPNL measure which incorporates a 3 dB/doubling duration correction, did not predict the results as well as the

Doppler correction. For the simulated flyovers, which contained no Doppler shift, the data appeared to require a full 3 dB/doubling duration correction. Thus Ollerhead's results might be interpreted as indicating the necessity for a different duration correction for flyovers containing a Doppler shift.

The following sections present the results of two experiments designed to explore various parameters of simulated Doppler flyovers. Altitude, duration, temporal and Doppler patterns are independently varied. The results are compared with the onset and Doppler corrections as well as the standard duration corrections.

TEST DESCRIPTION

Stimuli Preparation

Considerable care was taken to try to realistically simulate the noise produced by aircraft flyovers. Independent computer control of the Doppler pattern, amplitude of the noise and the amplitude of the pure tone components allowed precise manipulation of the important acoustic parameters and insured an accurate simulation of various real flyover patterns. In addition the control allowed the simulation of various flyovers in which one or another of the elements was unrealistic.

The equipment utilized for preparing the test samples in the judgment experiments is indicated by the block diagram in Figure 2. The stimuli were generated in three phases as suggested by parts [a], [b], and [c]. Part [a] shows that apparatus used to generate the Doppler shift. First, two tones were recorded on channel one and a shaped band of noise was recorded on channel two. This tape was then played back on tape recorder #1 whose playback speed was controlled by the digital computer. The computer was programmed so that the frequency change corresponded to a Doppler shift produced by an aircraft at a specified speed and altitude, such as shown in Figure 3. Playing the original tone and noise tape on tape recorder #1 and recording it at the computer controlled speed on tape recorder #2 allowed a Doppler shifted noise track and a Doppler shifted tone track. This tape was then utilized in the amplitude shaping apparatus shown in the [b] portion of Figure 2. The digital computer was programmed to dictate the levels of the voltage controlled amplifier in the implementation of the different types of time history patterns. Here the amplitude shaping was done on each channel independently to allow for different amplitude shaping for the tone and noise portions of the stimulus. In addition, the noise stimulus which was predominantly low frequency, peaked at a different time relative to the time of the maximum tone amplitude. This simulated the directional characteristics associated with the normal turbofan engine.

The final portion of the stimuli preparation process is indicated by part [c] of Figure 2. Here the tone and noise channels were mixed to provide appropriate tone to noise ratios. This combination was then low frequency modulated by the voltage controlled amplifier. The low frequency control for this operation was derived from a noise generator, a third octave band filter set at 63 Hz, and a linear to log converter for smoothing purposes. This modulation was added to simulate the normal fluctuations present in aircraft flyover noise due to atmospheric and turbulence effects. The final recording produced after this process was then employed in the tape cartridge machine described under Stimulus Presentation Equipment in Appendix A. Also included in Appendix A are details of the equipment employed in the sound analysis.

Test Stimuli

The test stimuli employed in the two tests are given in Tables I and II respectively. For Test I various combinations of Doppler shifts and time histories were employed. The stimuli were grouped according to the various Doppler patterns (i.e., "realistic", "approach-hover", etc.) and ordered according to altitude from the lowest to the highest. All samples employed tone frequencies at about 3000 Hz. In contrast, the stimuli 1 through 16

TABLE I

STIMULI USED IN TEST I

	Doppler	Doppler Pattern	Time Pattern	Altitude (in feet)	Stimuli Number[1]
SIMULATED[2]	Realistic		∧	250	1
		↓	↓	400	4
				800	5 (STD)
				2000	10
	Approach-Hover		/‾	250	2
		↓	↓	800	6
				2000	11
	Hover-Recede		‾\	250	3
		↓	↓	800	7
				2000	12
	Unrealistic		∧	800	8
			/‾	800	9
	None		∧	250	13
		↓	↓	800	14
				2000	15
ACTUAL RECORDINGS[3]	Real-life		∧	500	16
		↓	↓	600	17
				1850	18

1. As designated in Table B-I of Appendix B.
2. Simulated velocity of stimuli 1-12 is 200 mph.
3. Recorded flyovers at estimated altitudes.

Note: Stimuli 1-15 contain 3000 and 3300 Hz tones.

TABLE II

STIMULI USED IN TEST II

Stimuli	Altitude (in feet)	Tone Frequency f_1	f_2	Doppler Pattern[1]	Time Pattern
1	250	NONE			
2	250	100	107		
3	250	300	330		
4	250	1000	1070		
5	250	3000	3300		
6	800	NONE			
7	800	100	107		
8	800	300	330		
9	800	1000	1070		
10 (STD)	800	3000	3300		
11	2000	NONE			
12	2000	100	107		
13	2000	300	330		
14	2000	1000	1070		
15	2000	3000	3300		
16[2]	500	----			
17	---	Octave Band (1000)			
18	---	1000			
19	---	Octave Band (1000)			
20	---	1000			

1. Simulated velocity of stimuli 1-15 is 200 mph.
2. Recorded flyover at estimated altitude.

for Test II have the "realistic" Doppler shift pattern along with varying altitudes and tone frequencies. For the stimuli 17 through 20 which were simulated to resemble the ones employed by Nixon, the tone and amplitude are uncorrelated. A one-third octave band analysis of the stimuli is presented in Appendix B.

Subjects

Twenty-two college students were used as subjects for Test I and twenty subjects were employed in Test II. All subjects were screened to within 15 dB of the proposed ISO standard threshold (Ref.4). The total group consisted of approximately an equal number of males and females ranging in age from 17 to 27 years, with a median age of 21 years. Fifteen subjects participated in both test series.

Procedure

The judgment tests were conducted in an anechoic chamber 8' x 10' x 7 1/2' high. The testing method employed for this study was a modified form of the paired comparison testing procedure. Recent developments in methodology have produced several improved adaptive testing procedures, one of which is called Parameter Estimation by Sequential Testing (PEST) (Ref.5). This method, adapted for use at Bolt Beranek and Newman Inc. is described in detail in Appendix C. The method utilized an on-line computer as shown in Figure 4 to select "standard" and "comparison" stimuli iteratively. The subject decides which of the two is noisier (complete subject instructions are presented in Appendix D). The computer records the subject's response and adjusts the level of the comparison stimulus in a direction contingent upon the response, in order to make the sounds more equally noisy on the next trial. Order effects are automatically averaged by randomization of the order of presentation of the standard and comparison signals. This technique is repeated until the subject's answers converge on a prescribed level of performance.

Analysis

Upon the completion of a series of judgments for each subject the computer prints out the relative levels of three measures for the comparison signal. These levels are those at which the subject judged the comparison and standard stimuli to be equally noisy. Averaging the levels across all subjects for each stimulus provides that level for which the group judged the comparison to be equal to the standard. This is the dependent variable used in reporting the results. The average test results for each stimulus are given in terms of various measures in Appendix E.

RESULTS

Test I

The results of Test I are plotted as a function of altitude as shown in Figure 5. Here the standard level is represented by a heavy horizontal line. The measure employed in the top portion of the figure is tone corrected PNL (PNLT). The tone correction employed for this measure is that suggested by the FAA (Ref.6). In general as the altitude increases the level of the sounds which were judged equal to the standard decreases. This is to be expected since the duration of the sound also increases with the altitude. Results of the data which incorporates a duration correction such as Effective Perceived Noise Level (EPNL) (Ref.6) are plotted in the lower portion of Figure 5. Notice that the measurements appear to be in better agreement with the standard (they lie closer to the standard level). However, there does appear to be more scatter at the 250 foot altitude.

In reference to the stimuli that contain Doppler patterns related to Nixon's "approach" and "receding" stimuli, it appears that the "approach-hover" flyovers were judged noisier than the "hover-receding" flyovers even if both have the same EPNL value. This is true for the 250 and 800 feet altitudes. At the 2000 feet altitude they are practically identical. Thus, the results at least at the low altitudes are consistent in direction but not magnitude with Nixon's findings. It should be noted that even though the "approaching-hover" flyovers were judged to be noisier than the "hover-recede" the EPNL measure does a reasonable job of predicting their noisiness relative to the standard stimulus.

Test II

The results of Test II are plotted in Figure 6. On the top part of Figure 6 the results using PNLT as a measure are plotted. The results at the various altitudes seem to cluster quite well indicating that there is no unusual effect due to the different frequencies of the pure tones employed in the simulated flyovers. Also, on the graph is a portion of the results of Test I which are indicated by the closed symbols. The agreement between the two tests was quite good (within 1.5 dB). In general, PNLT appears to predict the data at 250 and 800 feet but underestimates the noisiness at 500 and 2000 feet. If a duration correction is added as shown by EPNL in the lower part of Figure 6, it is seen that this duration correction tends to over compensate for the results at 250 feet. Thus, at this altitude EPNL underestimates the noisiness of the stimuli. The recorded aircraft flyover at an altitude of 500 feet is also underestimated.

The last four stimuli of Table II closely resemble those used in Nixon's study. The results for these stimuli are presented in Figure 7. The heavy horizontal line represents the level of the standard stimulus. Thus, a point plotted at -5 dB indicates that that stimulus was 5 dB less than the standard at judged equal noisiness according to the measure indicated on the abscissa. As stated earlier, Nixon found that sounds with a gradual onset were judged to be noisier than those with a fast onset and a slow decay. The results plotted for PNLT, EPNL, and $EPNL_{O.C.}$ (with onset correction) for Test II do not show appreciable differences between the slow rise and slow decay stimuli. However, all measures underestimated the noisiness of these samples by about 8 dB. It is not evident as to why these measures should fail to predict the noisiness of these stimuli. Several alternatives are plausible. First, these stimuli unlike all others employed in the test, and particularly the standard, are relatively narrow band. Even the pure tone correction was designed to work with spectra containing lines in the presence of a broadband background. Another explanation could be attributed to the context effect since almost all of the other stimuli employed in this test had "realistic" time and Doppler patterns. And essentially there was no difference (±1 dB) between the "realistic" and other simulated Doppler shifted stimuli in Test I. This aspect of the results cannot be directly compared with Nixon's since all his measurements involved comparisons within this set of stimuli.

Further Analyses of Data

In the remaining discussion of the results a comparison of certain parts of the data will be made with some of the conclusions and correction formulas suggested by previous research.

In Test I there were three stimuli that had no Doppler shift. Let us compare these three stimuli to those similar in all other respects and containing a Doppler shift. There are actually two such sets - one from Test I and the other from Test II. The top part of Figure 8 shows this comparison in terms of PNLT. The stimuli without a Doppler shift show a more pronounced effect of altitude than for those containing a Doppler shift. Since altitude is correlated with duration, we have in effect supported Ollerhead's conclusion that flyovers without a Doppler shift require more duration correction than those containing a Doppler shift. A replot of the same data in the lower portion of Figure 8 using EPNL as a measure shows the same trend. The duration correction incorporated in EPNL, however, has done a reasonable job of correcting both sets of stimuli.

Let us now apply Ollerhead's Doppler correction to all the "realistic" Doppler pattern stimuli both simulated and real life for the two tests; and use the simple duration correction on the results of the stimuli with no Doppler shift. The upper portion of Figure 9 shows that this correction ($PNLT_R$) does indeed appear to bring the tone corrected perceived noise level more into line with the judgment results. This effect appears to be true regardless of whether a tone is present or not in the Doppler shifted stimuli. This is noted for the data represented by the open triangle (△) which does not appear to differ greatly from the other points on this figure.

If the duration of an aircraft flyover is proportional to its altitude, then the Doppler correction amounts to a correction of duration at the rate of 1.5 dB/doubling. The results of applying this modified duration correction ($PNLT_{.5/D}$) are shown in the lower part of Figure 9. As can be seen either measurement scheme provides fairly good predictions of the data.

A less successful correction is the onset correction suggested by Kryter. Figure 10 shows the results using an onset correction ($EPNL_{O.C.}$) applied to all "realistic" Doppler pattern stimuli both simulated and real life. It is clear that the onset correction does not predict the data very well. Naturally, since the standard is at 800 feet, the effective onset correction for this altitude is nearly zero.

Finally let us consider the "realistic" flyovers of Figure 10 in terms of two popular measures of noisiness. Figure 11 represents an analysis of the data in terms of EPNL. This measure appears to predict the data fairly well; the major discrepency being those flyovers with an apparent altitude of 250 feet. Here EPNL somewhat underestimates the noisiness of the flyovers.

Figure 12 shows a similar analysis for A-level. This measure also is fairly successful in predicting the data. The noticeable difference again being at the 250 feet altitude where AL overestimates the noisiness of the flyovers.

CONCLUSIONS

An analyses of the results of the tests conducted in this study indicated that:

1) Effective Perceived Noise Level provides a fairly accurate predictor of the noisiness of airplane flyovers containing Doppler shift except perhaps for flyovers at 500 feet or less (nominal velocity 200 mph) where it somewhat underestimates the apparent noisiness.

2) An onset correction which suggests that sounds with a long onset duration are noisier than those with a short onset duration does not appear to be applicable for the stimuli employed in this study. In particular, the onset correction does not improve the results provided by Effective Perceived Noise Level.

3) The frequency of the source tones does not appear to influence the judgment results for the Doppler shift patterns tested.

APPENDIX A

PLAYBACK AND ANALYSIS EQUIPMENT

PLAYBACK AND ANALYSIS EQUIPMENT

In order to perform the types of judgment tests referred to in this report, different equipment systems were employed in the various phases of preparation and analysis. Two divisions of the equipment systems are I) Stimulus Playback Equipment and II) Sound Analyses Equipment. Equipment used in the preparation of the stimuli was discussed in the body of this report.

Stimulus Playback Equipment

The stimulus playback equipment employed for this test is indicated by a block diagram in Figure A-1. The multiple cartridge tape recorder supplied the sound stimuli for the test. Each cartridge has two channels, one with the signal on it and the other with cue-tones which are used to control an electronic switch to prevent objectionable tape hiss between sound samples. In addition, the cue-tones are utilized to stop the cartridges and to indicate to the computer when to select another cartridge or await the subject's response. The rise-decay time of the electronic switch is 250 milliseconds to prevent any undesirable click in the signal. The subject response box in the anechoic chamber allows the subject to choose which of the two sounds he thought was the more annoying. This response is stored in the computer for use in determining the level of the next comparison signal. The computer then randomly selects another pair of sounds to present to the subject. The loudspeaker in the chamber is placed in front of the subject while the test stimuli are presented.

Sound Analysis Equipment

As indicated in the block diagram in Figure A-2, a stimulus analysis was performed in two parts; first, the signal was recorded while being played back in the anechoic chamber; and second, this tape was analyzed using the sound spectrum analyzer in conjunction with a digital computer. The microphone used for measurement in the chamber was a one-half inch condenser microphone. This microphone was placed approximately where the subject's ear normally would be located in the chamber. However, all measurements were performed without any subjects in the chamber. Sweep frequency tones were placed on the tape to measure the frequency responses of the recording system. To account for the frequency response, one-third octave band corrections were applied by the computer during the final analysis of the data. The signals on the tape were then played back through the spectrum analyzer and the band sound pressure levels read by the digital computer which in turn generated a permanent paper tape record of each signal. The paper tape which contained a one-third octave band analyses made every one-half second of the stimulus duration was read back into the computer. From this information the computer determined perceived noise level measures for each of these one-half second intervals and the effective perceived noise level of the total stimulus. These were the primary measures employed in the analyses of the subjective data.

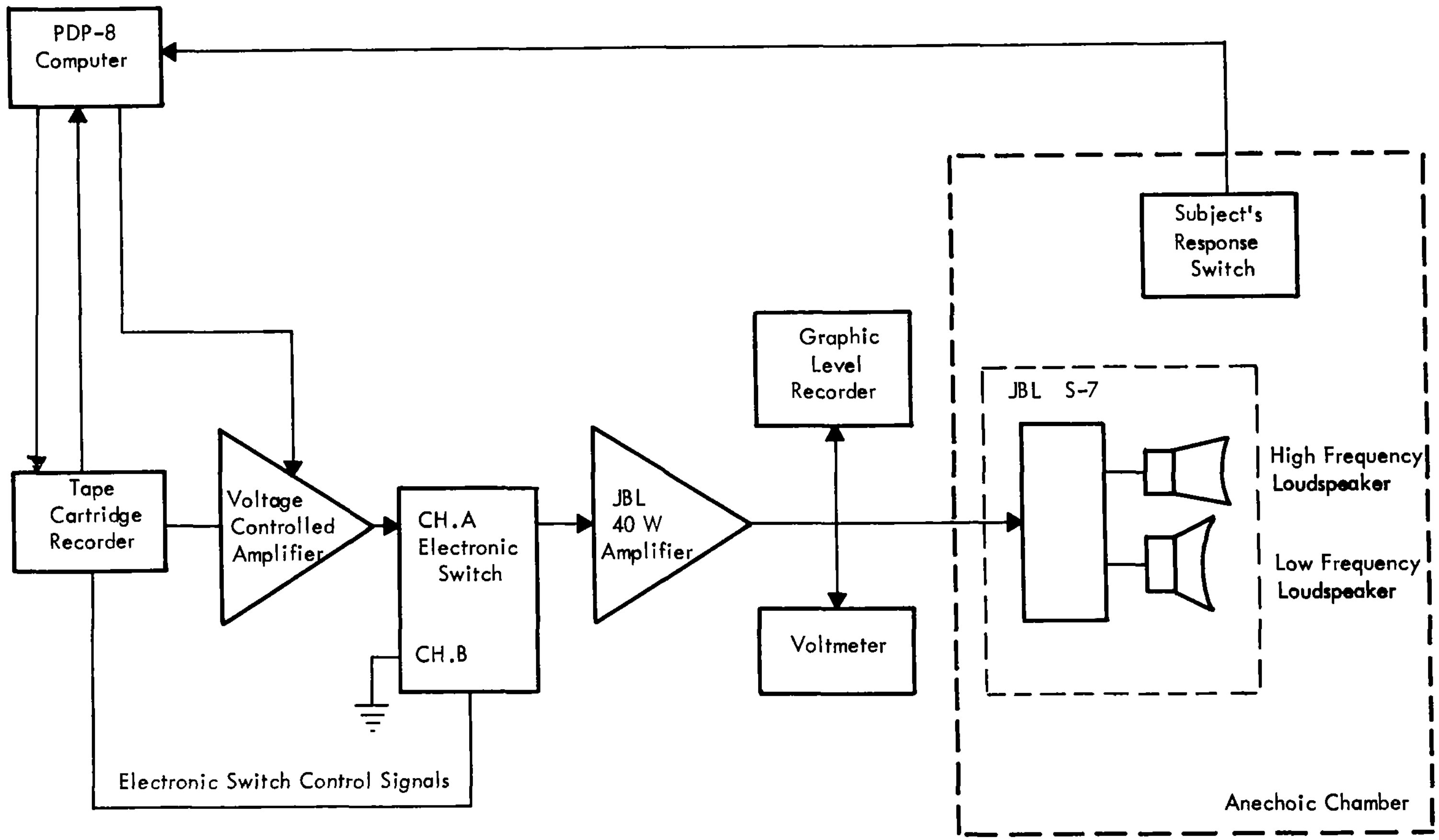


FIGURE A-1. BLOCK DIAGRAM OF PLAYBACK SYSTEM FOR JUDGMENT TESTS

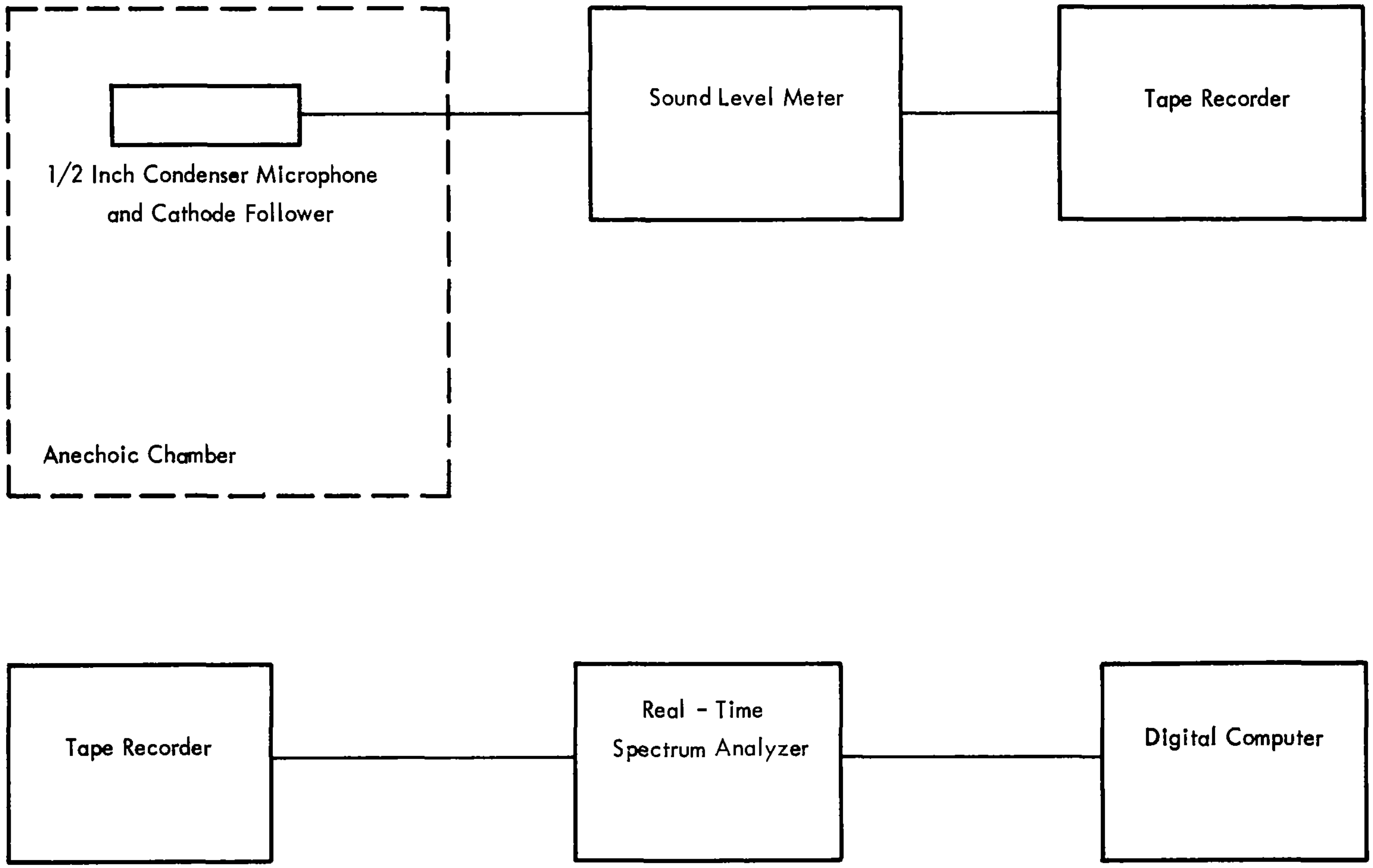


FIGURE A-2. BLOCK DIAGRAM OF STIMULUS ANALYSIS EQUIPMENT FOR JUDGMENT TESTS.

APPENDIX B

ONE-THIRD OCTAVE BAND LEVELS OF STIMULI USED IN JUDGMENT TESTS

TABLE B-I

STIMULI FOR TEST I

MEASURED AT 10 dB BELOW MAXIMUM LEVELS

		Sound Pressure Level in dB re 0.0002 dyn/sq. cm. One-Third Octave Band Center Frequency, Hz.																								
Comp.	Altitude (in feet)	OASPL	50	63	80	100	125	160	200	250	315	400	500	630	800	1000	1250	1600	2000	2500	3150	4000	5000	6300	8000	10000
1	250	93.0	56.5	68.0	73.0	80.5	86.0	83.5	84.0	84.0	81.0	79.0	77.0	74.5	71.5	71.5	69.5	67.5	65.5	72.0	76.5	72.0	54.5	52.0	49.5	48.0
2	250	95.0	62.5	65.0	75.5	83.5	89.5	87.0	87.5	87.0	85.5	82.5	82.0	78.5	75.5	75.5	74.5	72.0	70.0	74.0	82.5	77.5	61.0	60.0	54.5	52.5
3	250	95.0	62.5	69.5	75.5	84.0	90.0	87.5	87.0	87.5	85.5	83.0	82.0	78.5	75.5	75.5	75.0	72.5	70.5	75.0	83.0	64.5	61.5	61.0	55.5	53.0
4	400	92.0	61.0	69.0	72.0	80.5	85.5	83.0	82.0	82.0	81.5	79.0	77.5	74.0	71.0	70.0	68.5	67.0	65.0	72.5	75.5	68.5	52.0	49.5	46.5	42.5
5	800	94.0	61.5	72.5	76.5	82.5	88.5	85.5	84.5	84.5	83.0	80.5	79.5	76.5	73.0	72.5	70.5	69.5	66.5	71.5	77.5	69.5	55.0	52.5	48.5	46.5
6	800	95.0	61.5	65.0	74.5	83.0	89.0	87.0	86.5	87.5	86.0	82.5	81.5	78.5	75.5	75.5	74.5	72.5	70.0	72.0	80.5	71.0	61.5	59.5	55.5	52.0
7	800	95.0	63.0	72.0	76.5	85.0	91.0	87.5	87.0	87.5	86.5	83.5	82.0	72.5	77.0	77.0	75.5	73.0	70.5	74.5	81.0	64.0	61.5	59.5	55.0	51.0
8	800	95.0	62.0	63.0	73.5	82.0	87.5	86.5	88.0	87.5	86.0	82.5	82.5	79.0	76.0	75.5	75.0	74.0	69.5	67.5	77.5	68.5	61.0	59.5	53.5	47.5
9	800	95.0	55.0	66.0	75.5	83.5	89.0	87.5	87.5	87.0	86.0	82.5	82.0	79.5	76.5	76.0	75.5	73.0	71.0	76.0	81.5	64.0	61.5	58.0	55.5	52.0
10	2000	94.0	60.5	72.0	74.5	83.0	89.5	86.0	84.5	84.5	83.0	80.5	79.5	76.0	72.5	72.5	71.0	69.0	66.0	68.5	72.5	60.0	54.5	51.5	47.5	43.0
11	2000	95.0	56.5	65.5	76.5	83.0	89.5	86.5	87.0	87.0	86.0	83.5	82.0	79.0	76.0	76.5	75.0	73.0	70.5	68.5	75.0	63.5	61.0	57.5	54.0	47.5
12	2000	94.0	59.0	71.0	75.0	83.0	87.0	86.0	85.5	85.5	84.5	81.0	81.0	77.0	74.0	74.0	73.0	71.0	69.0	68.0	73.0	61.5	60.0	56.5	53.5	47.5
13	250	94.0	57.5	62.5	72.5	80.5	86.5	83.0	84.0	83.5	82.0	79.0	78.5	75.0	72.5	72.0	71.5	69.5	67.0	70.5	79.5	61.0	58.5	55.5	52.0	49.5
14	800	94.0	54.5	64.0	74.0	81.5	87.5	85.5	85.0	85.0	84.0	81.5	80.5	77.0	73.5	73.5	73.0	70.5	68.0	68.5	76.5	60.5	58.5	55.0	51.5	46.0
15	2000	95.0	54.5	64.5	73.5	81.0	87.5	85.0	86.0	85.5	85.0	81.5	80.5	77.0	74.0	74.0	74.0	72.0	69.0	64.5	71.5	62.0	60.5	56.5	53.0	43.5
16	500	93.0	66.0	71.0	71.5	79.5	85.0	83.0	84.0	83.5	83.0	81.5	81.0	77.0	75.0	75.5	77.0	76.5	76.0	78.0	80.0	78.0	74.0	72.5	68.5	58.5
17	600	91.0	55.0	55.5	57.5	69.5	74.5	73.0	72.0	73.0	70.5	70.0	70.5	67.0	65.5	72.5	76.0	74.0	83.0	87.0	79.0	76.5	76.0	70.5	66.0	55.0
18	1850	93.0	65.0	66.5	77.0	85.0	89.5	83.5	84.0	85.5	84.0	81.0	80.5	78.0	75.5	74.5	76.0	76.5	74.5	76.5	83.5	67.0	60.0	60.0	50.0	48.5
STD*	800	92.0	59.5	70.5	74.5	80.5	86.5	83.5	82.5	82.5	81.0	78.5	77.5	74.5	71.0	70.5	68.5	67.5	64.5	69.5	75.5	67.5	53.0	50.5	46.5	44.5

*Actual Presentation Level

TABLE B-TI

STIMULI FOR TEST II

MEASURED AT 10 dB BELOW MAXIMUM LEVELS

Stimuli	Altitude (in feet)	Sound Pressure Level in dB re 0.0002 dyn/sq. cm. One-Third Octave Band Center Frequency, Hz. OASPL	50	63	80	100	125	160	200	250	315	400	500	630	800	1000	1250	1600	2000	2500	3150	4000	5000	6300	8000	10000
1	250	92.0	60.0	68.5	74.0	79.5	85.5	85.0	83.0	84.5	82.0	79.5	77.5	74.5	71.5	71.0	69.5	68.5	66.5	61.5	60.0	59.5	58.0	58.5	54.5	47.5
2	250	93.0	52.0	60.5	77.5	87.0	91.0	75.5	74.5	74.5	72.5	70.0	68.5	65.0	62.0	61.0	60.5	59.5	57.0	52.5	50.5	50.5	49.5	50.0	47.5	42.0
3	250	95.0	52.0	59.5	65.0	73.0	77.5	75.5	77.0	89.0	91.5	85.0	70.0	66.5	65.0	65.5	61.0	60.5	58.0	53.0	51.5	51.5	50.0	50.5	47.0	42.0
4	250	92.0	55.5	67.5	72.0	78.0	84.5	81.0	81.0	81.0	80.0	77.5	74.5	72.0	83.5	87.0	87.0	72.0	64.0	60.0	60.0	58.0	54.5	54.5	50.5	44.0
5	250	93.0	57.0	68.5	73.0	80.5	86.0	84.0	84.0	84.0	82.0	80.0	77.5	75.5	72.5	72.5	70.5	69.0	67.5	72.5	78.5	75.0	57.0	56.5	52.5	49.5
6	800	93.0	61.0	70.0	74.5	82.5	88.5	86.0	85.0	84.0	83.0	80.5	78.5	75.5	72.5	72.0	70.5	69.5	68.0	62.5	60.5	60.5	58.5	58.0	53.5	46.5
7	800	94.0	56.5	66.5	77.5	92.0	92.5	83.0	82.0	82.0	80.5	77.5	75.5	72.0	70.5	69.0	67.5	67.0	64.5	60.0	57.5	57.5	55.0	54.5	51.5	44.0
8	800	97.0	58.0	66.0	73.0	81.0	85.5	83.5	82.5	92.0	96.5	86.0	76.0	72.5	70.5	70.5	67.0	66.0	64.0	59.0	56.5	55.5	53.0	52.0	46.5	40.5
9	800	93.0	61.0	68.5	73.5	82.5	88.0	85.0	84.0	85.0	83.5	80.5	78.5	75.5	81.0	88.0	86.0	69.5	68.0	62.5	62.0	60.0	58.5	58.0	53.5	46.5
10	800	93.0	57.5	71.5	75.0	82.0	88.0	85.5	84.0	83.5	83.0	80.0	79.0	76.0	73.5	72.5	70.5	69.5	67.0	71.5	77.5	69.5	55.5	54.5	51.0	49.0
11	2000	93.0	59.5	68.5	75.0	82.5	88.0	86.5	83.5	84.5	83.0	81.0	78.0	74.5	72.5	71.5	69.5	69.0	67.5	61.5	59.5	59.0	56.0	53.5	49.0	41.5
12	2000	94.0	62.0	70.5	77.5	92.0	90.5	86.0	85.0	85.5	84.0	82.0	78.5	76.0	73.0	73.0	71.5	70.5	68.5	64.0	61.5	61.0	58.0	56.0	52.5	44.5
13	2000	95.0	61.0	70.5	75.0	81.5	88.0	85.5	84.0	89.5	94.0	80.0	78.0	75.0	72.0	71.5	70.5	69.5	67.5	63.0	61.0	60.5	58.0	56.0	52.0	45.0
14	2000	93.0	61.0	70.5	75.5	82.5	88.0	86.0	84.5	84.5	83.0	80.5	78.5	75.0	75.0	83.5	79.5	69.5	67.5	62.5	60.0	59.5	57.0	55.0	51.5	43.5
15	2000	94.0	60.5	72.0	74.5	83.0	89.5	86.0	84.5	84.5	83.0	80.5	79.5	76.0	72.5	72.5	71.0	69.0	66.0	68.5	72.5	60.0	54.5	51.5	47.5	43.0
16	500	93.0	65.0	70.5	71.5	80.0	85.5	84.0	85.0	84.0	83.5	81.5	82.0	78.0	76.5	76.5	78.0	78.5	78.5	79.5	82.0	81.0	78.5	78.5	75.0	65.0
17	---	91.0	48.5	50.0	43.0	42.5	49.0	50.0	51.5	54.0	55.0	55.5	65.0	76.0	85.5	86.0	86.5	81.0	67.0	56.5	54.5	49.0	45.0	44.5	42.5	39.5
18	---	95.0	46.0	50.0	41.0	39.5	43.5	35.0	36.0	36.5	37.0	35.5	38.0	48.0	73.5	81.0	95.5	93.5	72.5	57.0	53.5	62.5	53.0	50.5	49.0	43.5
19	---	88.0	67.0	68.5	63.5	63.5	63.5	47.0	47.5	50.0	51.5	53.0	63.0	73.5	83.0	84.5	85.0	79.5	65.5	54.0	52.0	47.5	43.5	43.0	42.0	39.0
20	---	90.0	42.5	49.0	34.0	39.5	46.0	33.5	35.5	38.0	37.0	35.0	41.0	52.0	71.5	91.5	73.5	53.0	57.5	45.0	55.5	40.5	48.0	44.0	45.0	41.0
STD*	800	92.0	55.0	69.5	73.5	80.0	85.5	83.5	83.0	82.0	80.5	78.0	77.0	74.5	71.5	71.0	68.5	67.5	65.0	69.5	76.0	68.5	54.0	53.0	49.5	47.5

*Actual Presentation Level

APPENDIX C

PEST PROCEDURES EMPLOYED FOR THE JUDGMENT TESTS

PEST PROCEDURES EMPLOYED FOR THE JUDGMENT TESTS

The method used for the subjective judgment tests, Parameter Estimation by Sequential Testing (PEST) (Ref.5), is an iterative procedure that adjusts the level of a comparison stimulus until the subject's responses indicate that it is subjectively equal to some standard stimulus. This computer program used a paired-comparison paradigm in presenting the standard stimulus at some fixed level and the comparison stimulus at a variable level. The presentation order (i.e., whether the standard is presented first or second in the pair) is randomized to counterbalance for various order errors. The subject indicates which of the two sounds is the noisier and the computer adjusts the level of the comparison stimulus for its next presentation depending on the subject's response. Thus, this procedure preserves the advantages of the paired comparison technique while utilizing the speed and convenience of the adjustment procedure.

As PEST was used in these experiments, five PEST runs were intermingled so that the signals were presented randomly against the standard. This technique was employed so that signals of different duration would occur within the same test session and subjects would be encouraged not to respond simply on the basis of peak level. The maximum presentation levels for all stimuli were determined prior to the actual experimentation. At the beginning of each set of five comparison stimuli and standard stimulus, the experimenter entered into the computer the corresponding maximum levels. The test commenced with the presentation of the first pair of stimuli. On the basis of the subject's response as to which was the more annoying, the computer adjusted the level of the comparison signal by varying step sizes. The exact way in which the computer selects the step size for adjustment of the comparison signal is as follows:

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 12 dB is maintained.

After a degree of consistency in a subject's responses is attained in agreement with a preset criteria, the computer terminates the run and records the new absolute levels of the comparison signal and the relative levels of the comparison to the standard for three measurement schemes such as A-level, PNL and EPNL.

The complete history of the order of presentation of the signals is printed after all the runs have terminated. In addition the computer records the details of each trial including:

1) Whether the standard signal was presented first or second;

2) the subject's response; and,

3) whether the level of the comparison signal increases or decreases when next presented.

The experimenter at all times had the option of determining the number of stimuli to be presented to the subject during a test session, ranging from one to five different signals. The computer was programmed to allow him to specify the maximum levels of the stimuli in terms of three measures such as A-level, PNL, and EPNL as well as the initial levels of presentation of the standard and comparison signals. The experimenter also had the choice of changing other parameters such as the maximum number of trials to be allocated for each stimulus run, the step size of the final increment in the comparison signal level, the region of interest of the psychometric function, and the degree of confidence in the subject's responses.

APPENDIX D

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 more 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 unacceptable 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.

APPENDIX E

RESULTS OF THE JUDGMENT TESTS

TABLE E-I

RESULTS OF TEST I

Levels of Comparison Judged Equally Noisy to Standard

Comp.	Altitude (in feet)	Duration[1]	Rel. Comp. Level[2]	OASPL	AL	PNL	PNLT	EPNL	$PNLT_R$	$PNLT_{.5/D}$
1	250	5.5	- 3.0	90.0	82.0	95.5	98.0	92.5	93.5	96.0
2	250	8.0	-10.5	84.5	78.5	93.5	99.0	96.5	94.5	97.5
3	250	8.5	- 7.5	87.5	81.5	97.0	102.5	99.5	98.0	101.0
4	400	7.0	- 0.5	91.5	83.5	96.0	100.0	94.0	96.0	98.0
5	800	10.5	- 3.0	91.0	83.0	95.0	99.5	95.0	97.5	98.5
6	800	9.5	-11.0	84.0	77.0	92.0	96.0	94.5	94.0	95.0
7	800	12.5	- 9.5	85.5	79.5	94.0	99.0	96.5	96.5	98.0
8	800	10.5	- 6.0	89.0	82.0	94.0	97.0	93.5	95.0	96.0
9	800	8.5	- 9.5	85.5	79.5	94.0	99.0	97.0	97.0	97.5
10	2000	19.5	- 2.5	91.5	82.5	95.0	97.5	96.5	97.5	98.0
11	2000	16.5	- 7.5	87.5	79.5	93.0	96.0	96.5	95.5	96.0
12	2000	19.0	- 5.0	89.0	82.0	93.5	96.5	95.5	96.0	96.5
13	250	5.5	- 3.5	90.5	82.5	96.5	102.0	95.0	97.0	99.0
14	800	11.0	- 4.5	89.5	82.5	93.5	98.0	94.0	96.0	97.0
15	2000	17.5	- 5.5	89.5	81.5	93.0	93.5	93.5	93.5	94.5
16	500	7.5	- 7.5	85.5	81.5	96.0	97.5	92.0	94.5	96.0
17	600	8.0	- 7.5	83.5	83.5	96.5	102.0	95.5	99.0	100.5
18	1850	14.0	- 6.0	87.0	83.0	97.5	102.5	97.5	102.5	101.0
STD	800	10.5	0.0	92.0	84.0	96.0	100.5	96.0	98.5	99.5

1. Duration is the amount of time in seconds the signal is within 10 dB of maximum level.
2. Relative Comparison Level - comparison gain judged equality in dB re standard gain.

TABLE E-II

RESULTS OF TEST II

Levels of Comparison Judged Equally Noisy to Standard

Comp.	Altitude (in feet)	Duration[1]	Rel. Comp. Level[2]	OASPL	AL	PNL	PNLT	EPNL	$PNLT_R$	$PNLT_{.5/D}$
1	250	6.5	1.5	93.5	85.5	98.0	99.0	93.0	94.0	97.0
2	250	6.0	6.0	99.0	83.0	98.5	102.0	94.5	97.0	99.5
3	250	5.0	0.0	95.0	88.0	97.5	99.5	92.0	94.5	97.0
4	250	5.0	- 4.5	87.5	86.5	93.5	99.5	91.5	94.5	96.5
5	250	5.5	- 3.0	90.0	83.0	96.5	99.5	93.5	94.5	97.0
6	800	13.0	0.0	93.0	85.0	97.5	98.0	95.5	96.0	97.5
7	800	13.0	- 1.0	93.0	81.0	95.0	98.0	94.0	96.0	97.5
8	800	7.5	- 6.0	91.0	85.0	95.5	98.5	93.5	96.5	97.0
9	800	9.5	- 4.0	89.0	85.0	95.0	100.0	96.0	98.0	99.0
10	800	11.5	- 4.0	89.0	81.0	94.5	98.5	94.0	96.0	97.5
11	2000	21.5	- 2.0	91.0	82.0	95.0	95.5	94.5	95.0	96.0
12	2000	21.0	- 4.0	90.0	81.0	94.0	95.0	95.5	94.5	95.5
13	2000	18.0	- 5.0	90.0	83.0	95.0	98.0	96.0	98.0	98.5
14	2000	21.0	- 4.5	88.5	80.5	92.5	94.5	95.0	94.5	95.5
15	2000	19.5	- 3.0	91.0	82.0	94.5	97.0	95.5	96.5	97.5
16	500	7.5	-10.5	82.5	81.5	95.0	96.0	91.0	92.5	94.0
17	---	7.0	- 4.5	86.5	86.5	92.0	93.0	88.0	---[3]	91.5
18	---	5.5	-14.5	80.5	81.5	88.5	93.0	86.5	---	91.0
19	---	8.5	- 3.5	84.5	84.5	91.5	93.0	88.5	---	91.5
20	---	8.0	-11.5	78.5	78.5	83.0	89.5	85.0	---	88.0
STD	800	11.0	0.0	91.0	82.0	96.5	100.5	96.5	98.5	99.5

1. Duration is the amount of time in seconds the signal is within 10 dB of maximum level.
2. Relative Comparison Level - comparison gain judged equality in dB re standard gain.
3. Doppler Correction (R) not applicable to stimuli 17-20.

REFERENCES

1. Ollerhead, J.B., "Subjective Evaluation of General Aircraft Noise," Wyle Laboratories, Contract No. FA67WA-1731, Federal Aviation Administration Technical Report No. 68-35, 1967.

2. Nixon, C.W., von Gierke, H.E., and Rosinger, G., "Comparative Annoyances of 'Approaching' versus 'Receding' Sound Sources," J. Acoust. Soc. Am., 45, 330, 1969.

3. Kryter, K.D., "Possible Modifications to the Calculation of Perceived Noisiness," National Aeronautics and Space Administration, NASA CR-1636, 1970.

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

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

6. Sperry, William C., "Aircraft Noise Evaluation," Federal Aviation Administration Report No. FAA-NO-68-34, 1968.

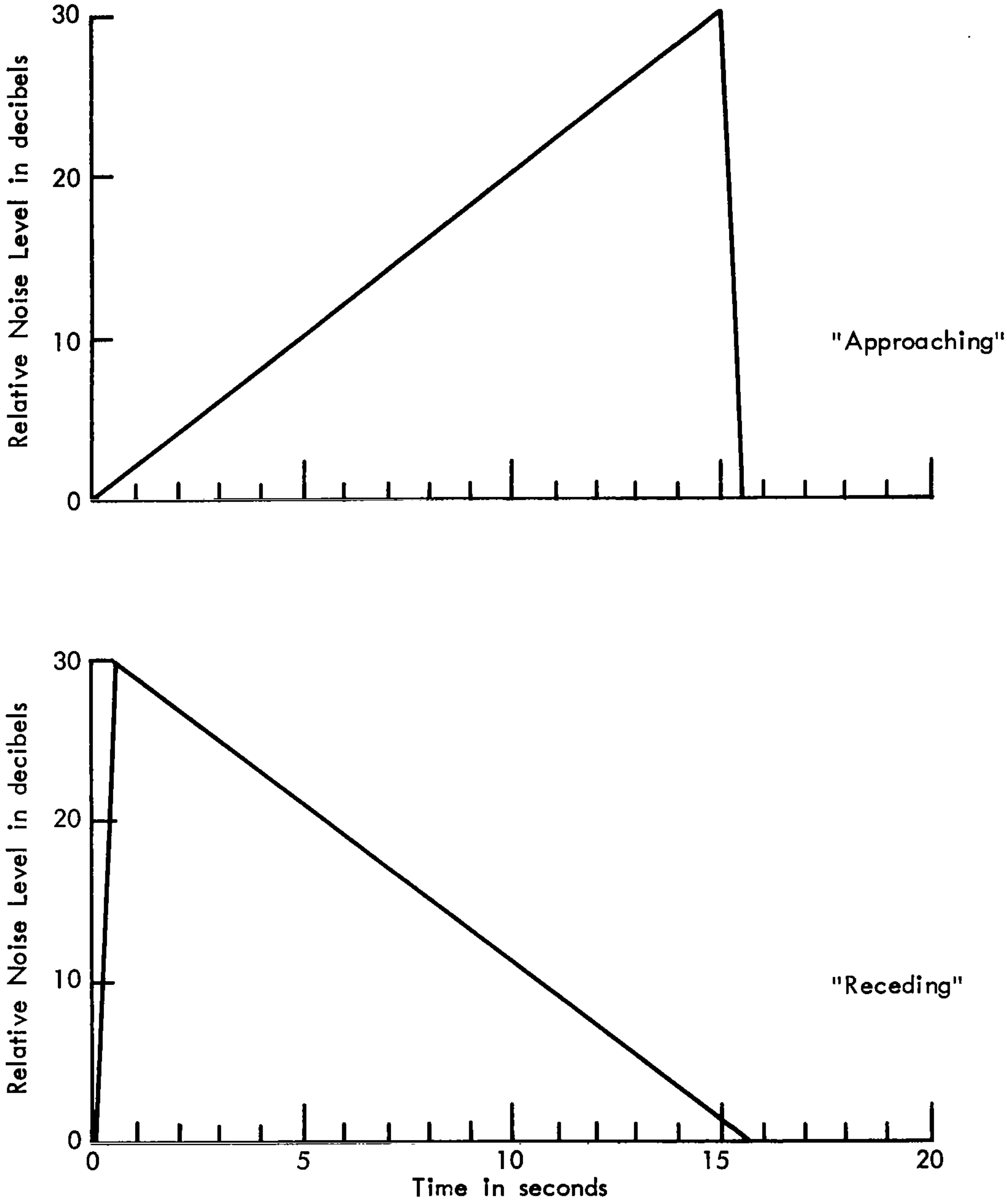
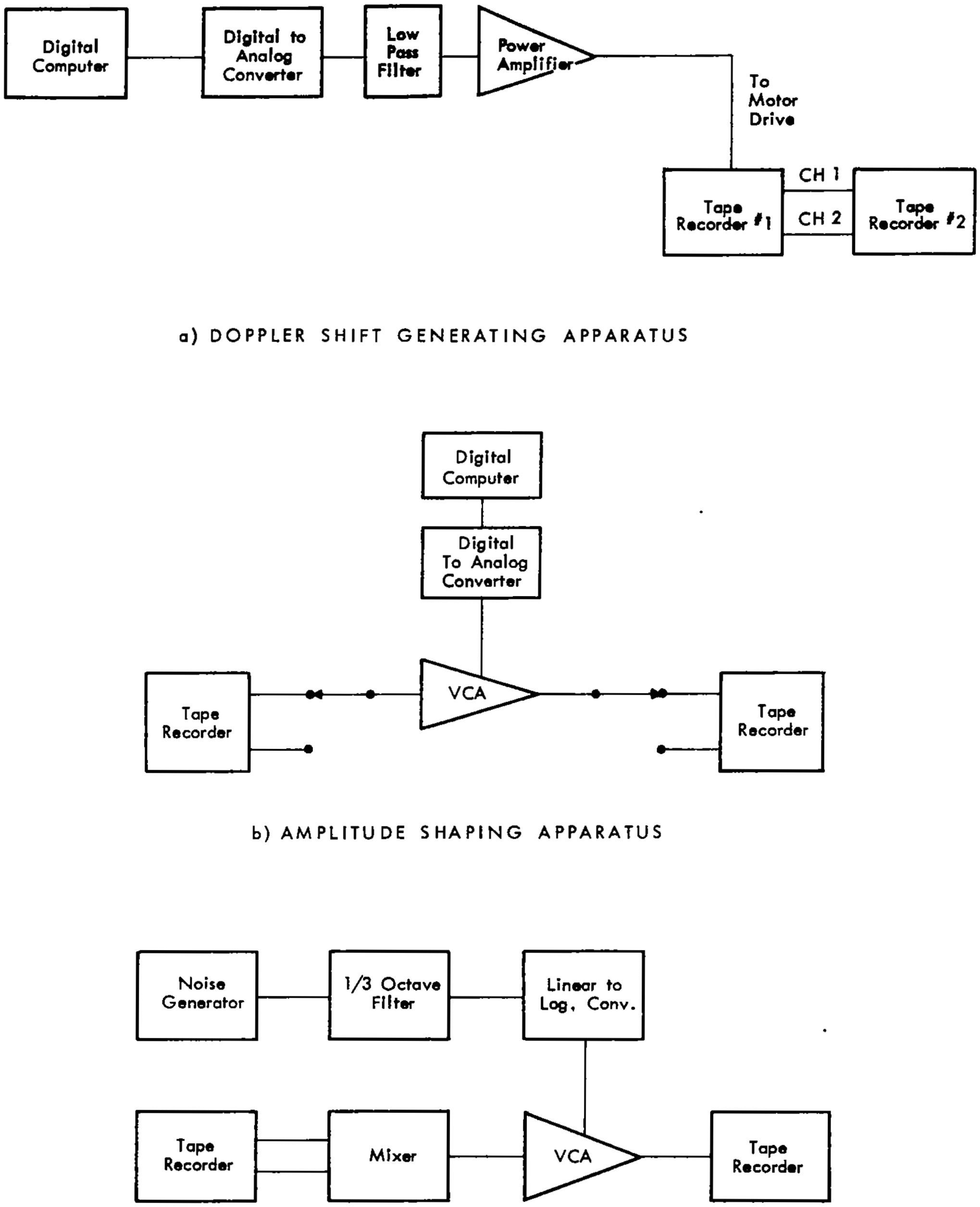


FIGURE 1. TIME HISTORIES OF STIMULI EMPLOYED IN INVESTIGATION OF APPROACHING AND RECEDING SOUND (Ref 2)



a) DOPPLER SHIFT GENERATING APPARATUS

b) AMPLITUDE SHAPING APPARATUS

c) FINAL MIXING AND AMPLITUDE MODULATION APPARATUS

FIGURE 2. BLOCK DIAGRAM OF STIMULUS PREPARATION EQUIPMENT FOR JUDGMENT TESTS

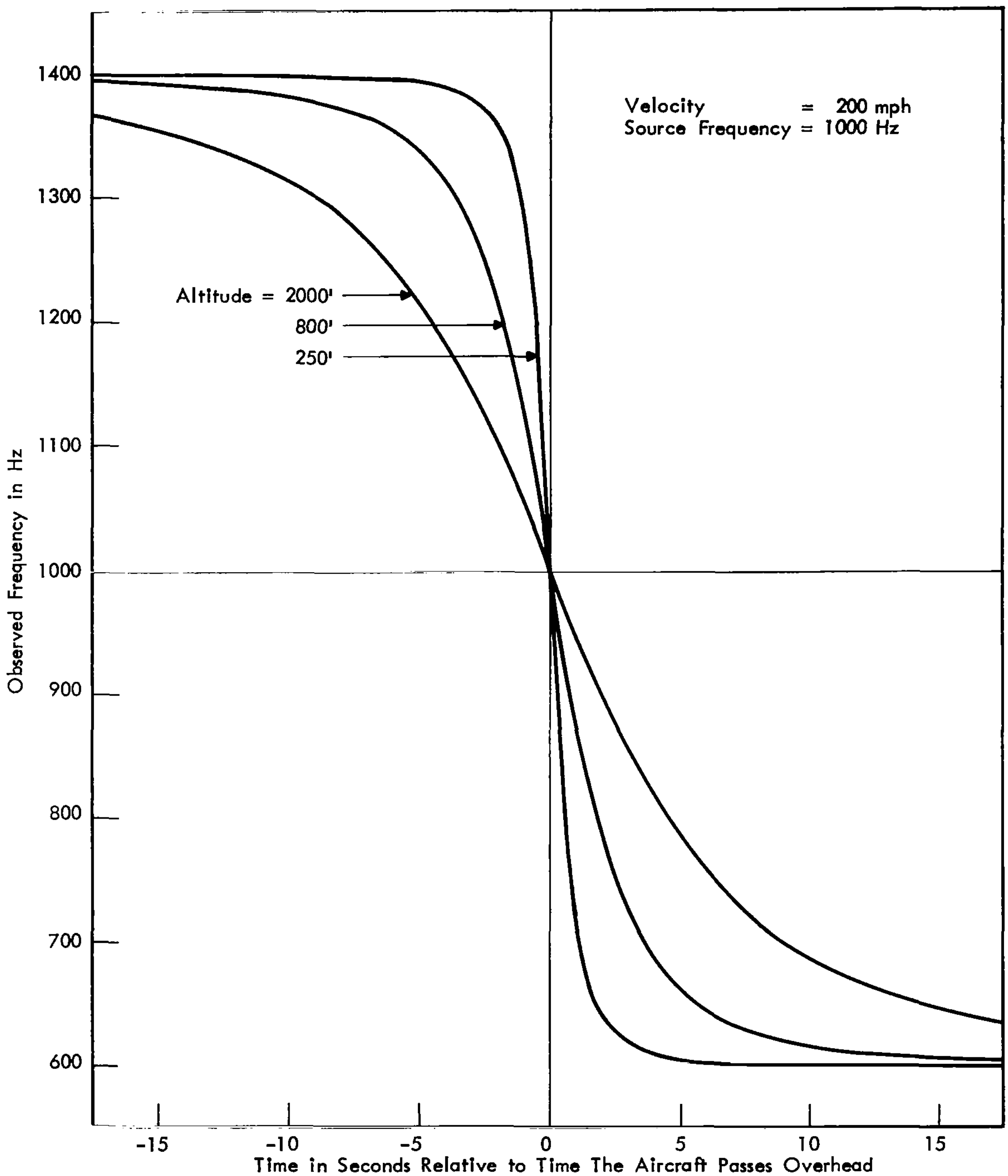
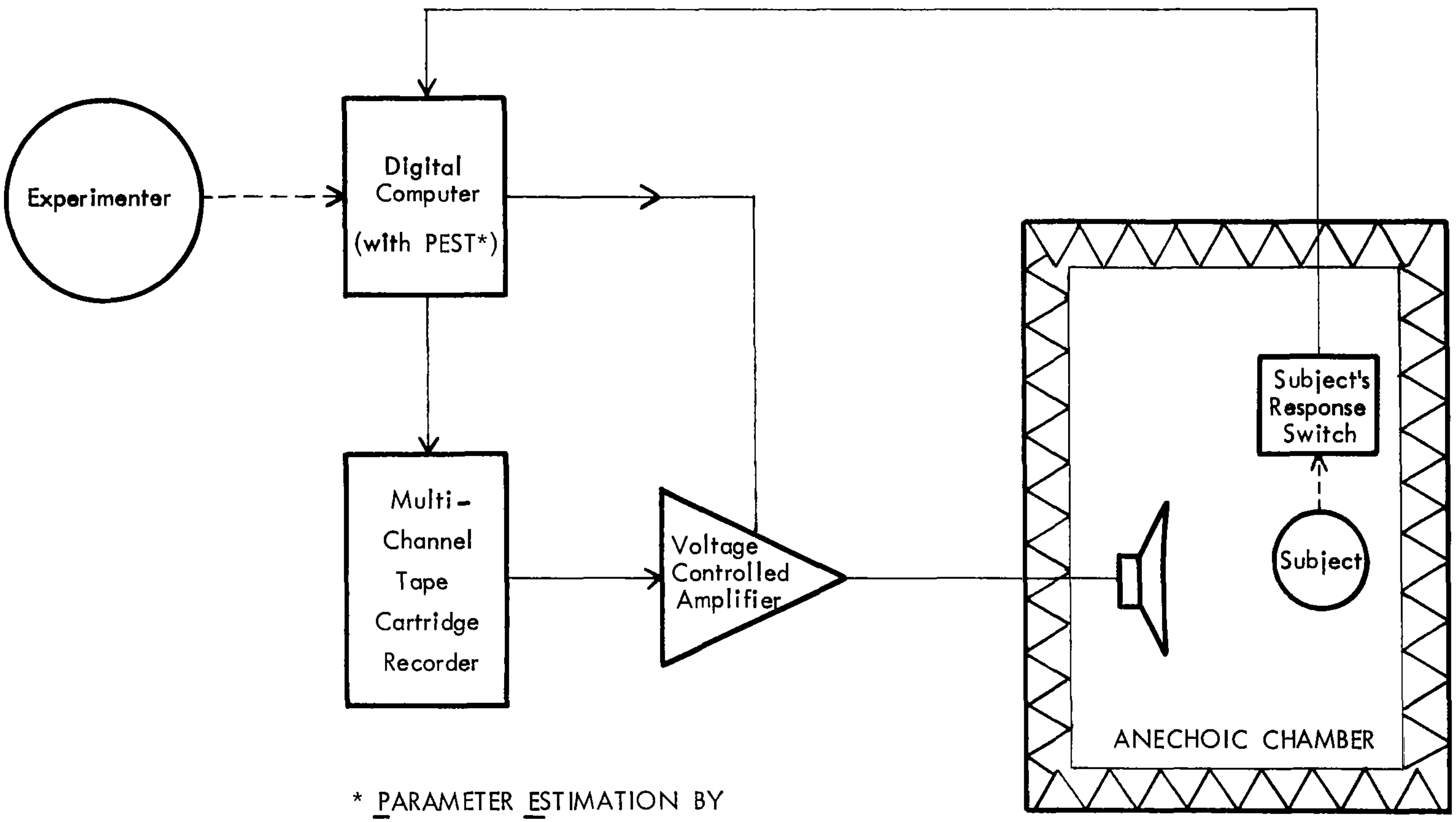


FIGURE 3. DOPPLER PATTERN FOR VARIOUS ALTITUDES



* PARAMETER ESTIMATION BY
SEQUENTAL TESTING

FIGURE 4. SYSTEM FOR JUDGMENT TESTS

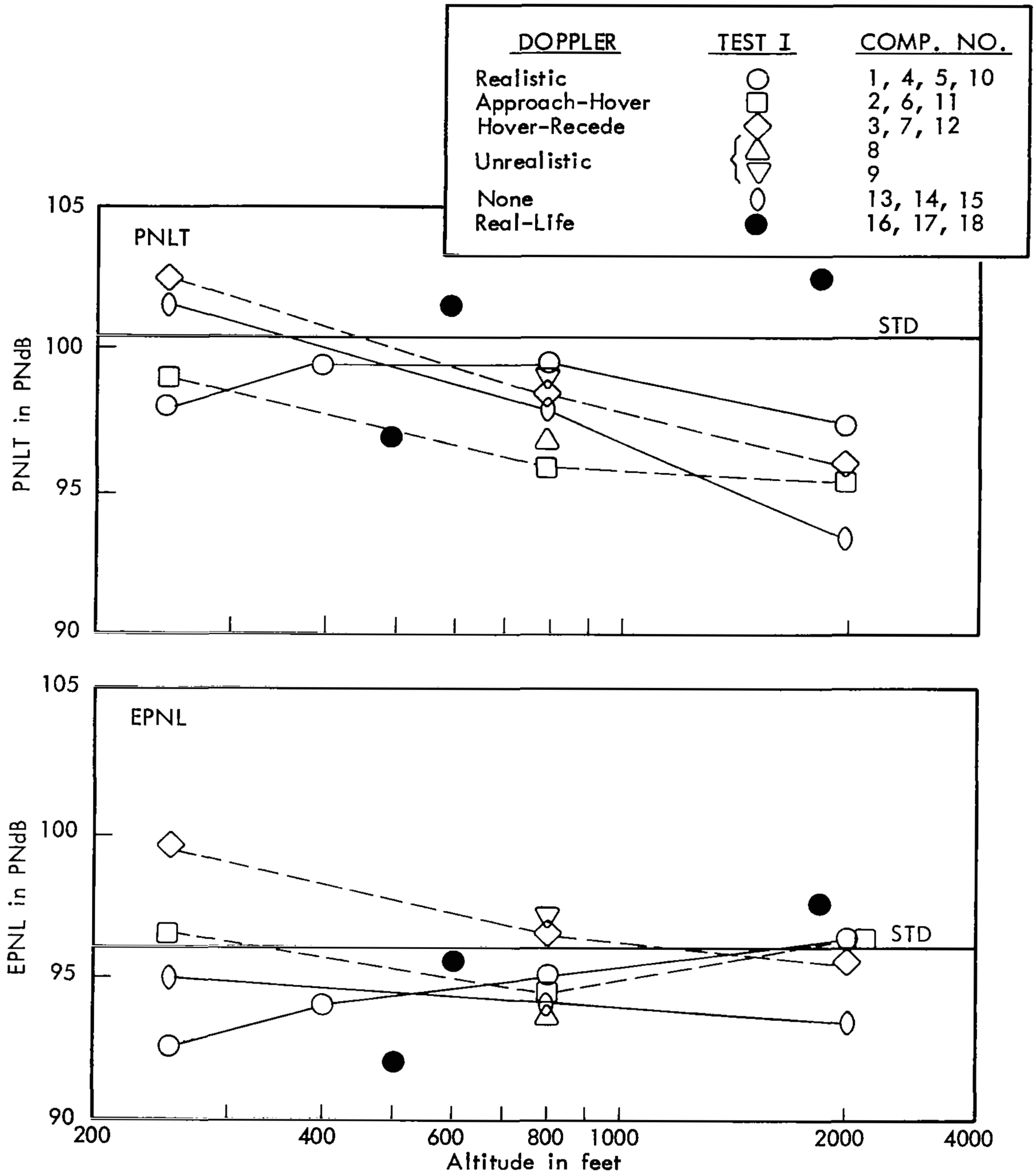


FIGURE 5. RESULTS OF TEST I IN TERMS OF PNLT AND EPNL
Stimuli Judged Equally Noisy to Simulated Flyover
at 800 ft. Altitude

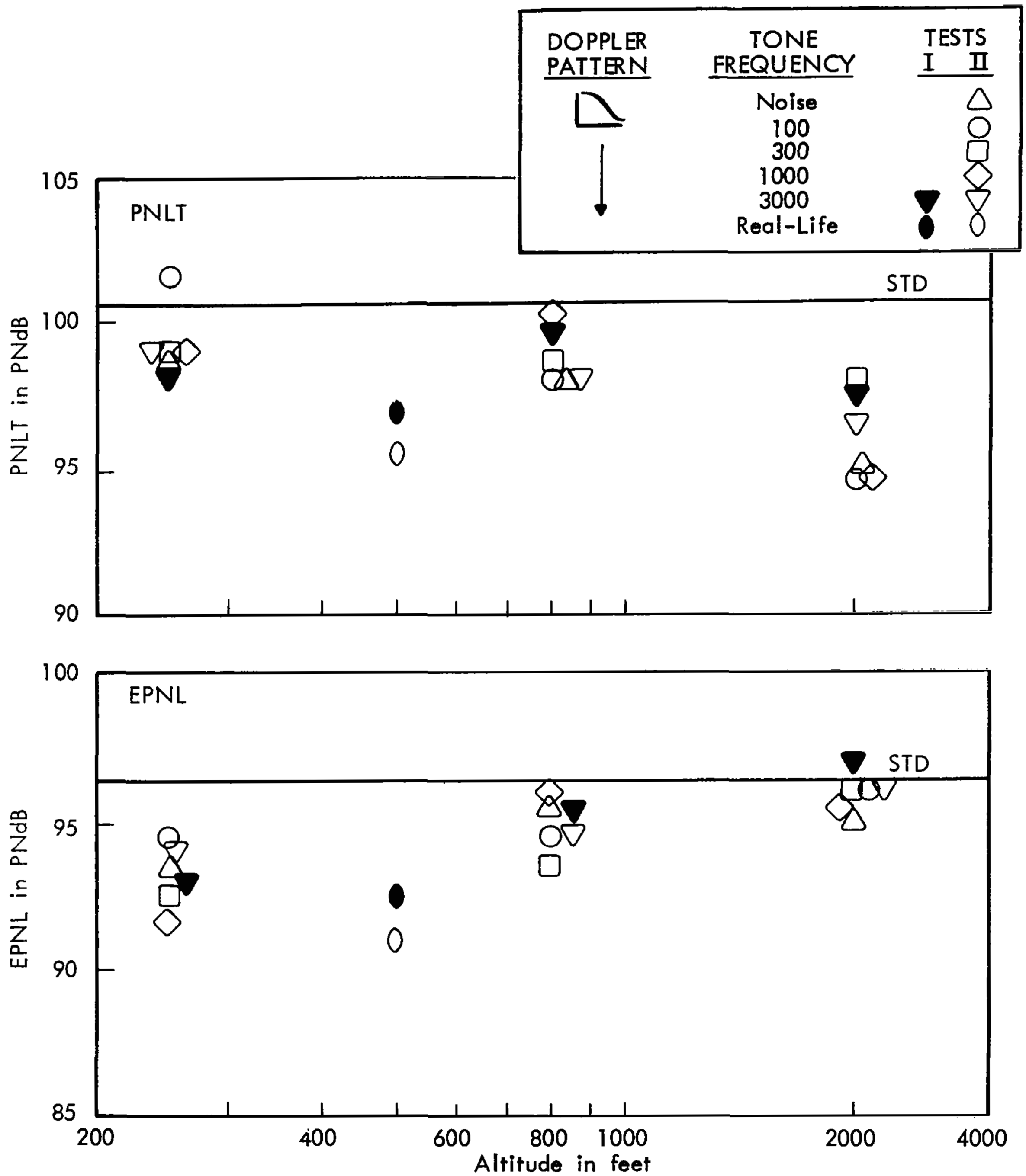


FIGURE 6. RESULTS OF TEST II IN TERMS OF PNLT AND EPNL
Stimuli Judged Equally Noisy to Simulated Flyover at 800 ft. Altitude

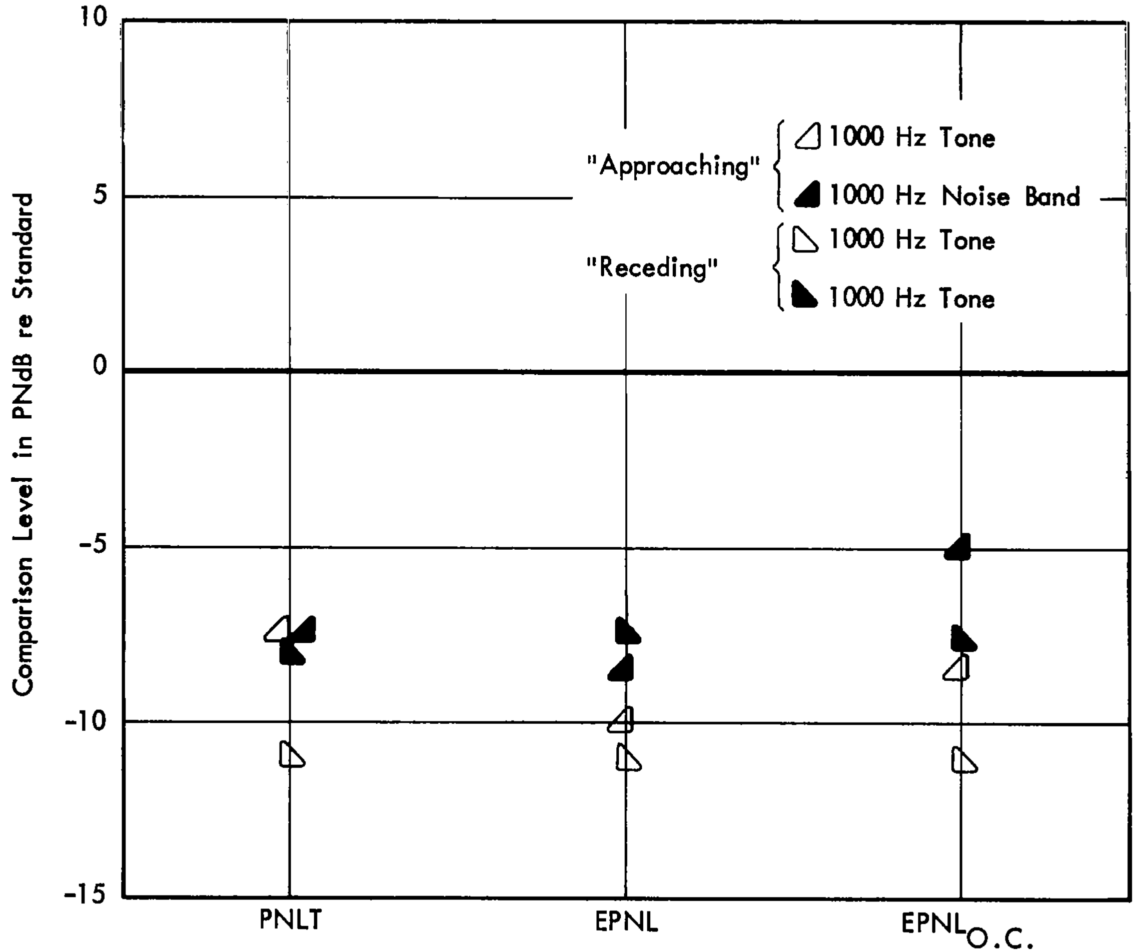


FIGURE 7. RESULTS USING STIMULI SIMILAR TO THOSE OF NIXON ET. AL. (REF. 2)
Stimuli Judged Equally Noisy to Simulated Flyover at 800 ft. Altitude

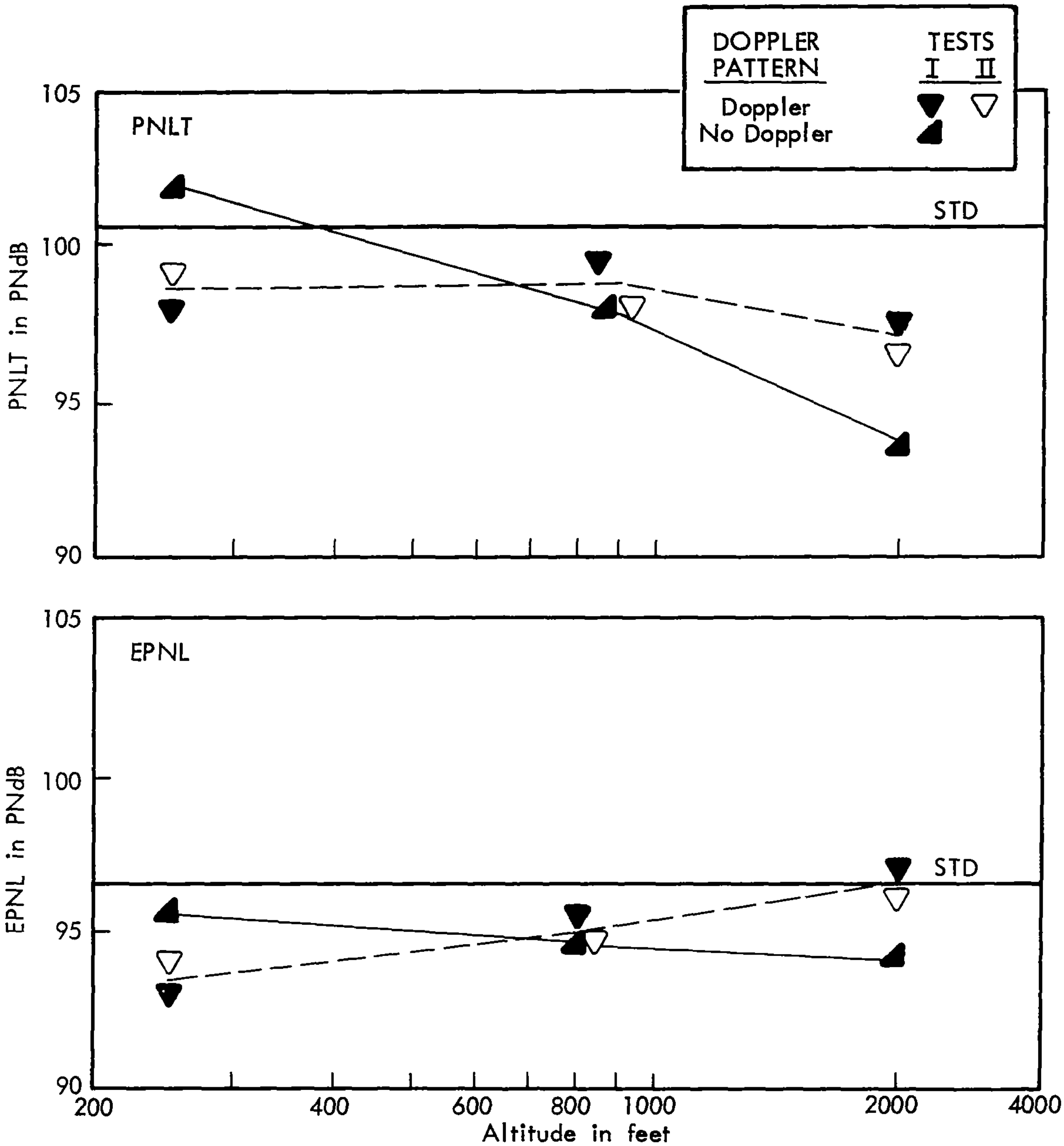


FIGURE 8. RESULTS SHOWING EFFECT OF DOPPLER SHIFT IN TERMS OF PNLT AND EPNL
Stimuli Judged Equally Noisy to Simulated Flyover at 800 ft. Altitude

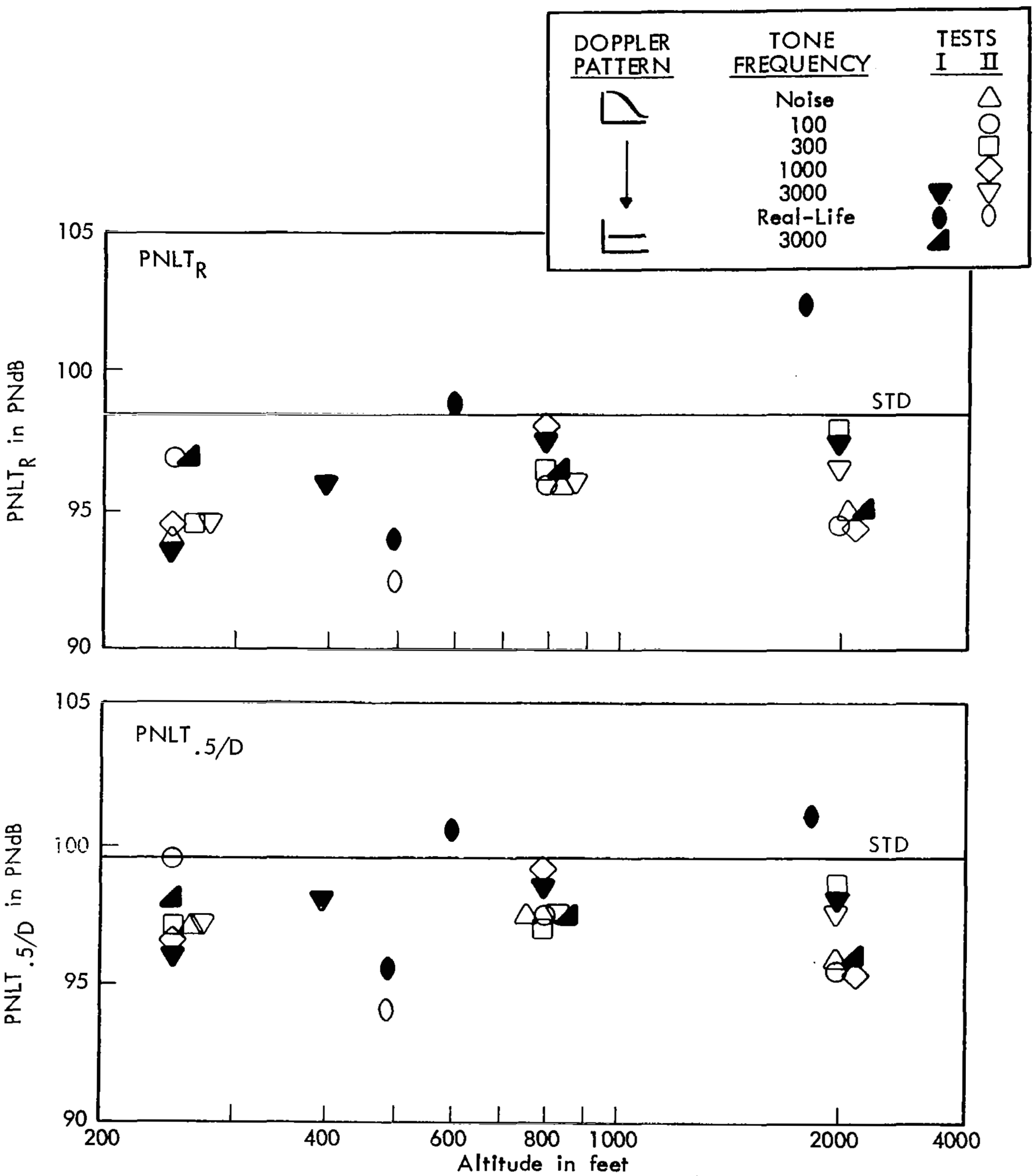


FIGURE 9. RESULTS SHOWING EFFECT OF DOPPLER CORRECTION AND MODIFIED DURATION CORRECTION
Stimuli Judged Equally Noisy to Simulated Flyover at 800 ft. Altitude

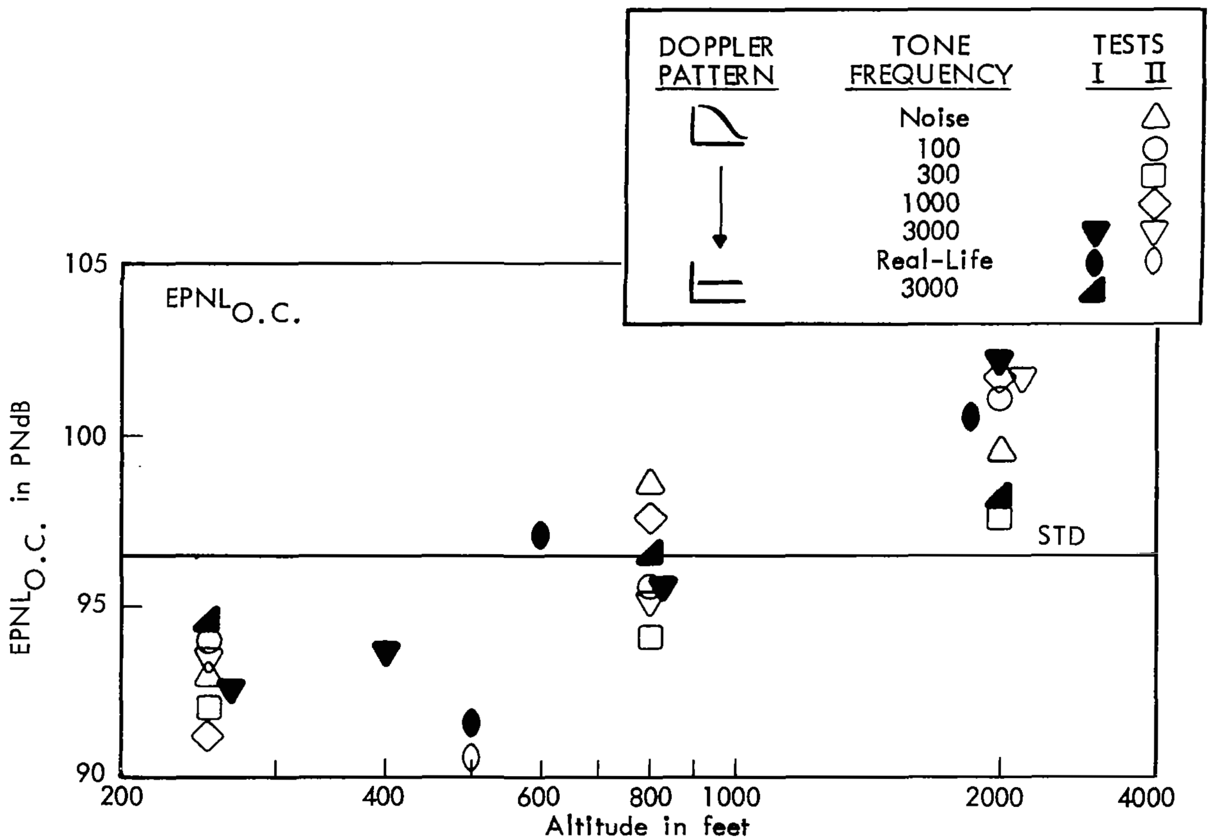


FIGURE 10. RESULTS SHOWING EFFECT OF ONSET CORRECTION
Stimuli Judged Equally Noisy to Simulated Flyover at 800 ft. Altitude

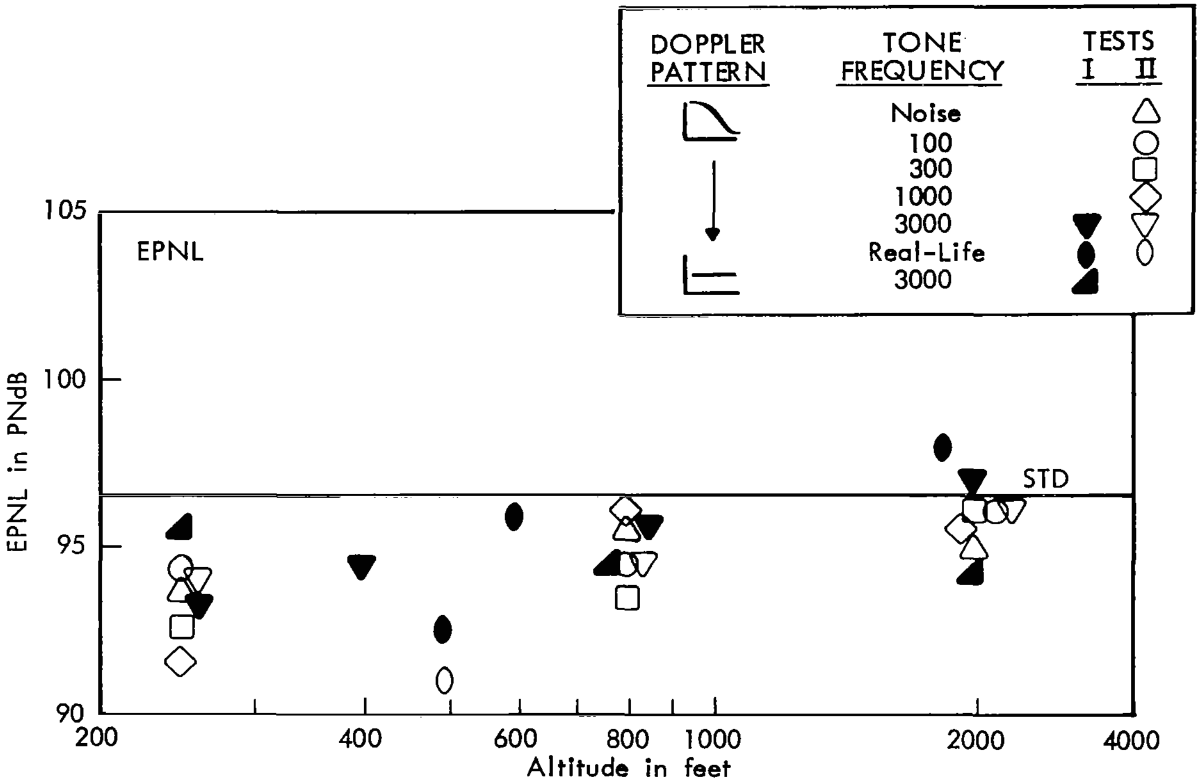


FIGURE 11. RESULTS IN TERMS OF EPNL
Stimuli Judged Equally Noisy to Simulated Flyover at 800 ft. Altitude

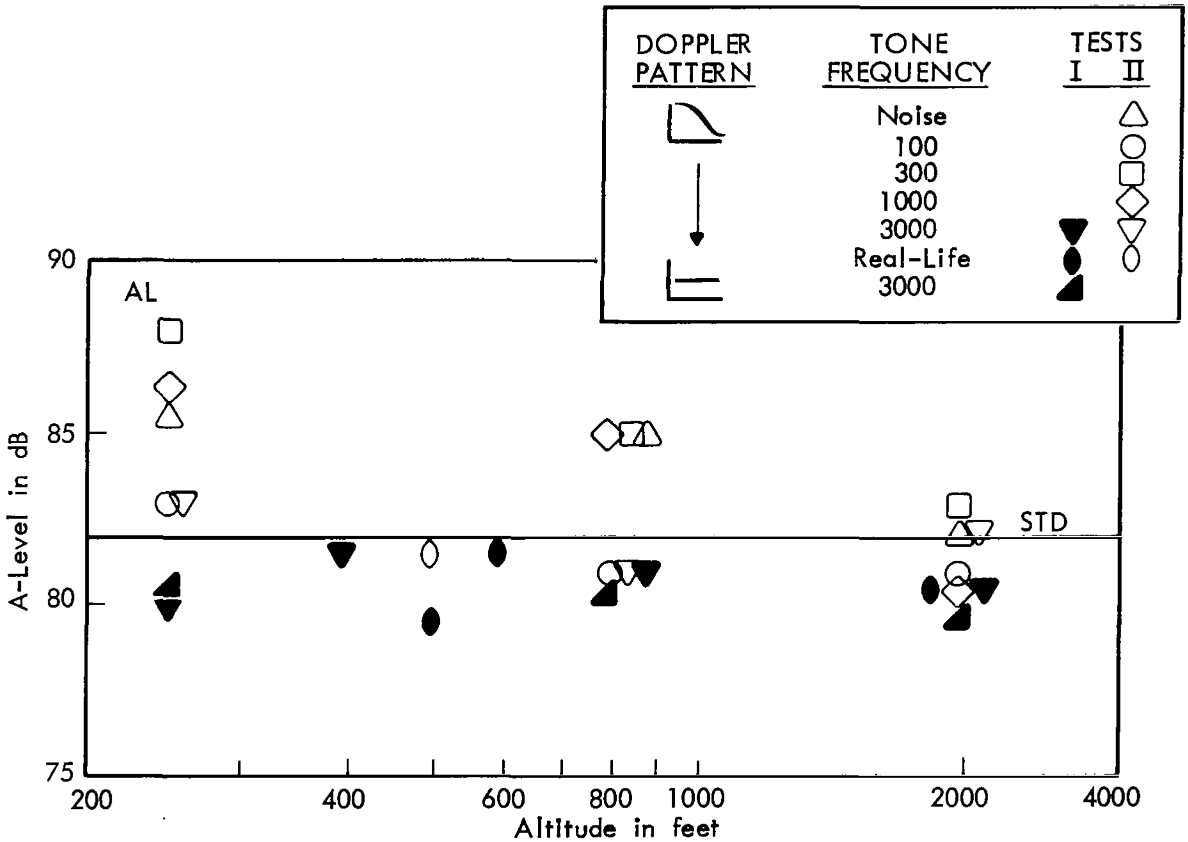


FIGURE 12. RESULTS IN TERMS OF A-LEVEL
Stimuli Judged Equally Noisy to
Simulated Flyover at 800 ft. Altitude

NASA-Langley, 1971 —— 4 CR-1779