TECH LIBRARY KAFB, NM
0061195
NASA CR
C.1
LOAN COPY: RETURN TO
AFWL (DOUL)
KIRTLAND AFB, N. M.

NASA CONTRACTOR REPORT

NASA CR-2043

ACCEPTABILITY OF VTOL AIRCRAFT NOISE DETERMINED BY ABSOLUTE SUBJECTIVE TESTING

by Harry Sternfeld, Jr., Ernest G. Hinterkeuser, Roy B. Hackman, and Jerry Davis

Prepared by
THE BOEING COMPANY
Philadelphia, Pa. 19142
for Langley Research Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • JUNE 1972

TECH LIBRARY KAFB, NM

0061195

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CR-2043		

4. Title and Subtitle	5. Report Date
Acceptability of VTOL Aircraft Noise Determined by Absolute Subjective Testing	June 1972
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Harry Sternfeld, Jr. and Ernest G. Hinterkeuser Roy B. Hackman and Jerry Davis	D 210-10392-1
	10. Work Unit No.
9. Performing Organization Name and Address	137-09-02-01
The Boeing Company - Vertol Division Philadelphia, PA 19142	11. Contract or Grant No. NAS1-10,044
	13. Type of Report and Period Covered Contractor Report
12. Sponsoring Agency Name and Address	14. Sponsoring Agency Code
National Aeronautics and Space Administration Washington, D.C. 20546	

15. Supplementary Notes

16. Abstract

A program was conducted during which test subjects evaluated the simulated sounds of a helicopter, a tilt wing and a 15 second, 90 PNdB (indoors) turbojet aircraft used as reference. Over 20,000 evaluations were made while the test subjects were engaged in work and leisure activities. The effects of level, exposure time, distance and aircraft design on subjective acceptability were evaluated.

Some of the important conclusions were:

1. To be judged equal in annoyance to the reference jet sound, the helicopter and tilt wing sounds must be 4 to 5 PNdB lower when lasting 15 seconds in duration.

2. To be judged significantly more acceptable than the reference jet sound, the helicopter sound must be 10 PNdB lower when lasting 15 seconds in duration.

3. To be judged significantly more acceptable than the reference jet sound, the tilt wing sound must be 12 PNdB lower when lasting 15 seconds in duration.

4. The relative effect of changing the duration of a sound upon its subjectively rated annoyance diminishes with increasing duration. It varies from 2 PNdB per doubling of duration for intervals of 15 to 30 seconds, to 0.75 PNdB per doubling of duration for intervals of 120 to 240 seconds.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
aircraft noise VTOL noise subjective response to noise noise	unclassified - unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
unclassified	unclassified	88	$3.00

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

TABLE OF CONTENTS

	Page
SUMMARY	1
INTRODUCTION	2
DESCRIPTION OF PROGRAM	
Concept	3
Test Facility	4
Preparation of Test Sounds	6
Test Subjects	7
Paired Comparison Test	8
Subjective Acceptability Tests	8
DISCUSSION OF RESULTS	
Correlation of Subjective Response with Acoustical Data	9
Effect of Sound Duration	10
Effect of Aircraft Distance	12
Effect of Aircraft Design	13
Effect of Test Subjects' Activities	14
CONCLUSIONS	16
APPENDIX A: Evaluation of Test Methods	19
APPENDIX B: Measurement of Subjective Judgements of Annoyance of VTOL Aircraft Sounds	21
REFERENCES	35

TABLES

Table		Page
I	TEST PROGRAM OUTLINE	37
II	SUBJECTS' EXPOSURE HISTORY	38
III	BASIC ACOUSTIC AND SUBJECTIVE DATA ON HELICOPTER SOUNDS	39
IV	BASIC ACOUSTIC AND SUBJECTIVE DATA ON TILT WING SOUNDS	40
V	CORRELATION OF MEAN SUBJECTIVE RATINGS WITH SEVERAL ACOUSTIC MEASURES	41
VI	FREQUENCY OF "NO RESPONSES"	42
VII	BASIC STATISTICAL DATA ON SUBJECTIVE RESPONSES	43
VIII	ANALYSES OF VARIANCE	48
IX	VARIANCE ACCOUNT	50
X	TOTAL PERCENT OVERLAP	51
XI	RESPONSE STATISTICSFOR JET REFERENCE SOUND	52

ILLUSTRATIONS

Figure		Page
1.	Test Environment - Plan View	53
2.	Test Subjects in Work and Leisure Environment	54
3.	Acoustical System for Sound Generation . . .	55
4.	Subjective Response Recorder	56
5.	Subjective Response Recording System	57
6.	Electronically Synthesized VTOL Aircraft Noise Presented to Test Subjects	58
7.	Tandem Rotor Configuration	59
8.	Tilt Wing Configuration	60
9.	Methods of Tape Preparation	61
10.	Correlation of Subjective Ratings with Helicopter dBC Levels	62
11.	Correlation of Subjective Ratings with Helicopter dBA Lebels	63
12.	Correlation of Subjective Ratings with Helicopter PNdB Levels	64
13.	Correlation of Subjective Ratings with Tilt Wing dBC Levels	65
14.	Correlation of Subjective Ratings with Tilt Wing dBA Levels	66
15.	Correlation of Subjective Ratings with Tilt Wing PNdB Levels.	67
16.	PNdB Level - Duration Trade-off	68
17.	dBA Level - Duration Trade-off	69

Figure		Page
18.	dBC Level - Duration Trade-off	70
19.	Sound Level - Duration Trade-off for Helicopter and Tilt Wing Sounds from Normalized Data of Figure 16. . .	71
20.	Relative Effect of Changing the Duration of a Sound on its Subjectively Rated Annoyance (PNL)	72
21.	Test III Results: Effect of Distance . .	73
22.	PNdB Reduction Trade-off with Distance	74
23.	Helicopter Noise - Design Parameters Trade-off	75
24.	Tilt Wing Noise - Design Parameters Trade-off	76
25.	Sample Problem: Helicopter Noise - Design Parameters Trade-off	77
26.	Paired Comparison Test Results	78
27.	Psychological Scaling of Subjective Responses	79
28.	Psychoacoustical Scaling of Subjective Responses	80
29.	Distribution of Means of Subjective Ratings of VTOL Noise	81
30.	Distribution of Subjective Ratings for VTOL Aircraft Sounds	82
31.	Distribution of Subjective Ratings for Three Levels	84

ACCEPTABILITY OF VTOL AIRCRAFT NOISE DETERMINED BY ABSOLUTE SUBJECTIVE TESTING

By Harry Sternfeld, Jr.
and Ernest G. Hinterkeuser
The Boeing Company, Vertol Division

and Roy B. Hackman and Jerry Davis
Hackman and Associates

SUMMARY

A program was conducted during which test subjects evaluated the simulated sounds of a helicopter, a tilt wing and a 15 second, 90 PNdB (indoors) turbojet aircraft used as reference. Over 20,000 evaluations were made while the test subjects were engaged in work and leisure activities. The effects of level, exposure time, distance and aircraft design on subjective acceptability were evaluated.

Some of the important conclusions were:

1. To be judged equal in annoyance to the reference jet sound, the helicopter and tilt wing sounds must be 4 to 5 PNdB lower when lasting 15 seconds in duration.

2. To be judged significantly more acceptable than the reference jet sound, the helicopter sound must be 10 PNdB lower when lasting 15 seconds in duration.

3. To be judged significantly more acceptable than the reference jet sound, the tilt wing sound must be 12 PNdB lower when lasting 15 seconds in duration.

4. The relative effect of changing the duration of a sound upon its subjectively rated annoyance diminishes with increasing duration. It varies from 2 PNdB per doubling of duration for intervals of 15 to 30 seconds, to 0.75 PNdB per doubling of duration for intervals of 120 to 240 seconds.

5. During leisure activities there were more occasions when the VTOL aircraft sounds were not noticed than during work activities.

6. Approximately equally high correlation was obtained between subjective ratings and acoustical measurements of helicopter and tilt wing VTOL sounds, whether in terms of Peak Perceived Noise Level, dBA or dBC.

7. An extremely high degree of correlation ($r = 0.94+$) was obtained between measured acoustical data and the subjective ratings by the observers, indicating a high level of confidence in the test method used and the results obtained is justified.

INTRODUCTION

The purpose of this program was to determine the relative subjective acceptability of two VTOL aircraft sounds using absolute subjective testing methodology and to investigate the effects of the application of noise criteria to VTOL aircraft. Although the public is the beneficiary of improved transportation systems it has also been found that the advent of advanced systems has generally brought with it increased noise levels which tend to offset some of the total benefits. Experiences of this type, most notably those associated with subsonic jet aircraft, have led to regulatory noise level limits which affect both the aircraft design and operation. Obviously, a proper limitation, for the greatest good, is one which provides adequate noise exposure protection for those on the ground, while still permitting development and operation of the most viable and vigorous transportation system.

In order to serve both of the above interests it is important that the noise level criteria developed should be based on the actual effects on people in their daily activities rather than on abstract opinions or evaluations. It is vital that limits be based on these effects and not be the product of test techniques.

Obviously, the optimum criteria can be established by introduction of the particular service, studying the public response, and then setting the noise limits. This is in fact the noise case history of the jet airplane. It is apparent, however, that such procedures are no longer socially acceptable nor economically feasible, and that evaluations must be made a part of the development cycle.

In a previous study conducted by Boeing-Vertol for NASA (ref. 1) the method of paired comparisons was used to evaluate the relative annoyance of the sounds of several types of V/STOL aircraft. The concept of the presently described program was to investigate a test methodology which attempts to provide subjective evaluations of aircraft noise obtained under more natural conditions than the traditional laboratory tests. This method is based on the interaction between the individual and the effect of noise on his activities and still yields results which are interpretable in terms of criteria.

The concept of obtaining evaluations as affected by personal activity is of particular significance to the VTOL aircraft concept. Because of the ability of VTOL aircraft to hover and fly at extremely low speeds, the potential noise exposure to observers can be many times that resulting from a fly-over of a fixed wing airplane. In such situations it might be expected that the cumulative effects of increased time could be substantially different if the test subject is engaged in activities which absorb his interest, or merely sitting and listening to the sound. In the case of VTOL aircraft the evaluation of acceptable acoustical signatures can dictate the appropriateness of various configurations to serving the commercial market. It is therefore not inconceivable that an artificial conservatism in such an item as time exposure factor could prematurely discourage the development of VTOL aircraft service.

DESCRIPTION OF PROGRAM

Concept

This program was designed to collect the subjective opinions regarding the acceptability of aircraft sounds and to correlate these evaluations with the acoustical environ-

ment. A major goal was to make the procedure as natural as possible so that the evaluation would be based on the manner and degree in which the noise affected the activities of the subject, and the act of performing the rating a secondary task. This is in contrast to another type of evaluation in which the subject offers opinions of how the sound might affect him in various situations and his primary task is listening to noise and filling out an evaluation sheet.

In order to achieve these goals, the test situation was designed to incorporate the following features:

1. The subjects were housed in a "normal" environment and the sounds were generated exterior to the housing structure.

2. The subjects were engaged in activities which had significance to them and the process of indicating their noise evaluations was designed to create a minimum distraction from these activities.

3. The time duration of the sounds and time periods between them were representative of a range which might be expected to occur in commercial VTOL aircraft operations.

Test Facility

The test subjects were housed within a 28 x 10 x 7 ft. (8.5 x 3.0 x 2.1m) office trailer which was located within a 40 x 20 x 20 ft. (12.2 x 6.1 x 6.1m) acoustically isolated chamber. The trailer was of metal frame and skin construction with wood paneling for the interior walls. The floor was completely covered by a carpet. The interior of the trailer was furnished so that one half of it simulated a working environment with office tables and chairs while the other half simulated a leisure environment with a couch, lounge chairs, a television set, tables, etc. No wall was used to separate the areas in order to keep the general atmosphere as spacious as possible. Ventilation was accomplished by tieing into the Acoustical Laboratory air conditioning system thereby providing a pleasant climatic atmosphere without incurring noise levels which would be typical of window air conditioners.

Figure 1 illustrates the general arrangement while Figure 2 shows test subjects in the work and leisure environments. A continuous low level broad-band noise was provided through two speakers located as shown in Figure 1. The spectrum of this noise within the trailer corresponded to the shape of an NC-30 office noise criteria.

The aircraft sounds were played through two speaker systems which were also located as shown in Figure 1. These speakers were mounted above the roof level of the trailer and directed at the trailer wall and roof. Figure 3 illustrates the acoustical systems used to generate the ambient and aircraft noises. The acoustical environment within the trailer was continuously monitored by recording the output of four calibrated microphones on a tape recorder (Figure 1).

In order to collect the subjective evaluations each test subject was equipped with a small box which contained nine pressure sensitive switches imbedded in a thin plastic card as illustrated in Figure 4. These switches were very easy to actuate in that they responded to extremely light finger pressure, had no perceptible motion, and were completely silent. The square format of the push-button arrangement was felt to be advantageous not only because of its compactness but also because it does not have the obvious extreme psychological ends of an in-line check list. It was found that the test subjects rapidly learned to actuate this device with minimal distraction from their other activities.

The actuation of a given response box energized a digital indicator in the control room so that the corresponding numeral was displayed on a read-out indicator. Figure 5 shows a block diagram of the system and the monitor board display. The response data was recorded manually at the end of each test sound and was also recorded automatically every two seconds by a motion picture camera.

A digital clock on the monitor panel was synchronized with two identical clocks inside the trailer so that any observations, made by the psychologist who was present in the trailer with the test subjects, could be correlated with the response data.

Preparation of Test Sounds

Two VTOL configurations were selected for this study: a tandem rotor helicopter and a tilt-wing VTOL. These configurations were selected partly because they embody the types of acoustical signatures which may be most important with respect to several of the configurations under consideration for inter- and intra-urban transportation. More specifically, as shown in Figure 6, the tandem rotor helicopter signature is also quite similar in spectral and temporal characteristics to the main rotor of a single rotor helicopter or to a tilt rotor configuration. The tilt wing, on the other hand, is more similar to the tail rotor of a large single rotor helicopter or to a conventional propeller. Both configurations were powered by turboshaft engines.

Some of the aspects of the acoustical signatures of these VTOL aircraft, which tend to be unique, include:

1. Long exposure times due to low flight speeds and the ability to hover.

2. Spectra which are highly weighted in the low frequency range.

3. Periodic impulsive type sounds.

In addition to these two VTOL configurations, the subjects were exposed to the fly-by noise of a jet airplane.

The basic aircraft designs, noise predictions, and acoustical simulations were taken from those prepared under a previous NASA contract (ref. 1). The aircraft which were essentially sized for 60 passengers and a 500 mile (926 km) range, are illustrated in Figures 7 and 8. Specific details regarding the prediction and simulation of these basic signatures can be found in the reference 1 report. These tapes were reworked using the systems illustrated in Figure 9 to achieve the required variations in time duration, level, and spectrum shape. The levels and spectrum shapes were designed to yield the desired spectra and levels inside the test enclosure.

Tapes were prepared in the above manner to achieve the requirements of the test programs outlined in Table I. The level and spectrum changes for each successive test were

based on a review of the subjective response to the preceding one. In this manner the program could be optimized by concentrating test points in the range of expected responses rather than being forced to cover a broader range with poorer resolution, as would be required to ensure obtaining significant results if the tapes were all made prior to the start of the program.

Test Subjects

The subjects in this program were males between the ages of twenty-one and thirty-four years. They were mainly employed in the fields of business and the professions. All subjects were members of the Swarthmore-Wallingford Chapter of the Pennsylvania Jaycees, a young men's community service organization. Since payment for their services was made directly to the organization and used for their community activities, the subjects were, in an indirect sense, volunteers. Their motivation, as displayed by extremely high attendance (99+%), and highly cooperative attitude were major assets to the program.

Potential test subjects were initially screened by means of audiograms for hearing, and by interview with the psychologists for attitude. Significant deviation from the norm of the group, in either category, was considered justification for disqualification. The remaining candidates were divided into three groups.

1. A control group of five men who attended one night a week for sixteen continuous weeks.

2. Four groups of five men each who attended one night every other week.

3. Alternates for the latter groups. No alternates were permitted for the control group.

Since all tests (Table I) were repeated on at least a two week basis a test group of twenty-five individuals participated in each test.

Prior to the program the subjects were instructed in their functions and the use of the response equipment, and a practice session was held. A verbatim copy of the instructions may be found in Appendix B, Part II.

Paired Comparison Test

Prior to starting the main test program the test subjects participated in a paired comparison test which utilized the jet reference, tandem rotor helicopter, and tilt wing VTOL noise tapes which were the same as those employed during the reference 1 study. The differences was that the test was performed in the trailer which was later to be employed for the subsequent absolute testing and, of course, the new set of test subjects. The primary purpose was to compare the relative evaluations between the stimuli and reference sounds obtained by paired comparisons as opposed to those derived from absolute judgments. The results of this comparison are discussed in Appendix A.

Subjective Acceptability Tests

A typical test session consisted of two 2-hour periods. During the first period the subjects were located at the "work" end of the facility (Figure 2). They performed paperwork and reading activities which had been approved by the psychologists. In general, this work was directly related to the subject's full time occupation and was therefore similar to that he performed during the normal work day. In this manner the tasks being performed were always relevant, and therefore significant to the subject, as opposed to an assigned work effort in which he felt no personal involvement.

The second 2-hour period was generally spent in the more comfortably furnished end of the test room, although it was permissible for subjects to remain in the work area if this was more amenable to their leisure activity. During this time the most frequent activities were watching television, playing cards, conversations, and recreational reading. One of the psychologists was in the trailer with the subjects at all times in order to monitor activities and to act as a general observer.

The subjects had been instructed in the use of the response equipment, as previously described, and quickly learned to respond with a minimum amount of distraction. Periodically, as aircraft sounds were heard they indicated their responses and changed them as often as they desired by pushing a different button. There was no specific period of time at which an evaluation was demanded. The final evaluation for each sound, however, was assumed to be that which was indicated at 10 seconds after the end of the stimulus. At that time the data system automatically disconnected the individuals' response boxes from the indicator board while a final tally was made, and the indicator board was automatically reset to zero in preparation for the next stimulus.

The sounds, which were randomized with respect to test variables, were presented at an average rate of 24 sounds per hour. Since the length of duration (in Tests II, III,IV) was a variable, the time between the end of one exposure and the beginning of the next also varied. This prevented a predictable anticipation of the onset of the next stimulus. The jet airplane sounds were randomly introduced into the mix at a ratio of approximately 1 to 8 with the VTOL sounds.

Table II presents the total exposure history of the entire program and indicates the extensiveness of the data gathered.

DISCUSSION OF RESULTS

Correlation of Subjective Responses with Acoustical Data

The acoustical data, recorded in the test environment was analyzed and the results are presented, along with mean subjective ratings, in Tables III and IV.

Sound pressure levels on the dBA and dBC scales are maximum values attained during each sound. The Composite Perceived Noise Level (PNLC) was calculated using the method and tables of reference 4. In this case the composite value is obtained by using the maximum sound pressure level in each frequency octave band rather than calculations over many time increments. The data reported is that obtained from the microphone located on the work table closest to the

center of the test room. Levels varied somewhat with location of the microphone and with the physical locations of the test subjects. Checks of other microphones and various personnel arrangements showed a non-systematic variation of up to plus or minus 2dB.

Figures 10 through 15 present the basic data from the program. Also shown on each figure is the least square fit line and Pearson's correlation coefficient r (ref. 5). In general the linearity of subjective response with stimulus level is striking with correlation coefficients ranging from .926 to .996 as shown in Table V.

From this table it can be seen that any of the three measures can be used as indicators of the subjects' responses with a high degree of validity. It is, perhaps, noteworthy that while for the tilt wing PNLC performs slightly better than the simpler A and C scale levels, the ratings to helicopter sound correlate equally well with PNLC or with dBC. This is consistent with the trends found by Pearson (ref. 6) and Ollerhead (ref. 7) who indicate that for impulsive noise with low frequency components dBC is a somewhat better predictor of subjective response than is PNLC.

The extremely high correlation coefficients and the apparent insensitivity to the fine points inherent in each different noise indicator, as compared to other studies, are felt to be a substantiation of a testing method which evaluates the effect of these noises on a person as he engages in normal activities. This is in contrast to a test which forces the subject to listen to each sound (or pair of sounds) critically and then make judgments regarding how these sounds might affect him if he were doing something else. In addition, the extremely large number of test exposures helped to stabilize the data and thereby improved the reliability of the results.

Effect of Sound Duration

In order to examine the effect of sound exposure time the information shown in Figures 10 through 15 can be cross-plotted to develop lines of constant subjective response (Figures 16, 17 and 18). This format is quite useful in that it provides an easy method of trading off the sound duration

and noise parameters in a manner which could not be expected to affect subjective response.

Also shown on Figure 16 is the zone bounded by the 75% overlap (Tilton's O, see ref. 3) with the jet reference sound. As discussed in Appendix B on Psychological Scaling, those sounds having an overlap of greater than 75% are essentially judged the same as the jet. The latter represented a 110 PNdB out-of-doors (attenuated to 90 PNdB indoors) sound which, although not an official limit, probably represents a situation which is generally considered as bordering on the unacceptable.

In addition to the 75% overlap, a second limit line corresponding to 45% overlap is also shown. Overlaps of less than 45% are considered as indicative of significant differences, and data which falls below this level would be generally recognized as clearly more acceptable than the reference jet airplane.

In order to assess the specific impact of sound duration on annoyance, the results of subjective testing are usually normalized to one specific duration which, in this and several other investigations, is taken as 15 seconds. Also, since the effects of duration have previously been developed in terms of perceived noise level (ref.'s 8 and 9), only that factor is considered here. By plotting the data of Figure 16 thus normalized, it is apparent that the trends for both the helicopter and tilt wing exhibit roughly the same scatter, as shown in the top and center parts of Figure 19.

These two parts have therefore been combined at the bottom of Figure 19 illustrating the development of an average value of the relative annoyance of both aircraft sounds of several durations based on the means of the scatter of lines of constant subjective rating.

The average trend with duration of this program's subjective ratings is again illustrated as the solid line on Figure 20. Also shown on this figure, as a dotted line, is the ref. 10 data illustrating the relative effect on perceived noise level of changing duration of a noise relative to a duration of 12 seconds. Since the slopes of the lines from the ref. 10 data and from the present study agree over the duration range from 15 to 30 seconds, the data from the present study was shifted downward approximately one-half

PNdB. This was done so that the two sets of data could be properly compared since one set was referenced to 12 seconds and the other to 15 seconds. The important conclusion to be drawn from Figure 20 is that the effects of sound duration continue to be non-linear beyond the durations which had been investigated prior to the present study (64 seconds). Over the range from 15 seconds to 240 seconds, the slopes continue to change from a value of 2 PNdB per doubling of duration at the shorter durations to 0.75 PNdB per doubling of duration for relatively long durations. Although a change in slope of "only" one and one-quarter PNdB per doubling of duration may seem unimportant, it is nevertheless of great importance in the evaluation of the acceptability of proposed VTOL aircraft designs and their inherently larger noise duration characteristics.

Effect of Aircraft Distance

During Test III the sounds presented to the test subjects consisted of re-shaped aircraft spectra according to changes in distance between the aircraft and a hypothetical listener. The changes in spectra consisted of those attributable to spherical spreading of a sound source and the additional reductions in noise due to atmospheric absorptions for an assumed standard day.

Two levels of each aircraft sound of five different durations each were subjectively rated. Results from all five durations have been combined in Figures 21a and 21b. In order to achieve an improvement in subjective acceptability it is, of course, necessary to reduce the Perceived Noise Level. By taking the baseline aircraft design and the distance at which it is judged equivalent to an outdoor 110 PNdB jet aircraft sound, a relationship between lower perceived noise level and the amount of distance between the aircraft and a listener can be established. Such a relationship is illustrated in Figure 22 for the Tilt Wing (part a) and the Helicopter (part b). The slopes of the lines of PNdB reduction versus distance for the two aircraft are somewhat different: They vary from 4 to 8 PNdB per doubling of distance for the Tilt Wing and from 4 to 11 PNdB per doubling of distance for the Helicopter. It is important to keep in mind that the two slopes illustrated hold only for the baseline aircraft design and that the slopes for other, perhaps more noise conscious designs, would be different.

The information presented in these figures may also be used interchangeably in conjunction with that contained in the figures presented in the following section on aircraft design, since both distance and component noise control information should be used to complement each other in an effort to provide acoustically acceptable commercial VTOL aircraft service.

Effect of Aircraft Design

In Test IV the sounds presented to the test subjects consisted of re-shaped spectra of the Helicopter and Tilt Wing aircraft according to a re-design of each aircraft in several stages. The levels of each aircraft sound component which make up a complete aircraft spectrum were adjusted, somewhat dependent on each other, from the basic design of the aircraft to yield a more acceptable sound signature. For both aircraft, far field spectra contain components as follows: rotor or propeller discrete rotational frequencies, rotor or propeller random broad-band "vortex" noise, and random noise components emitted by the engine from its inlet, usually in the four-thousand and eight-thousand Hertz octave band. No engine pure tone frequencies needed to be taken into account since they were placed beyond the upper limit of the last octave band (8 KHz) in the original design.

Two levels of each aircraft at each of the five sound durations used in this program were subjectively rated. The results of the duration effect on subjective response from this test series has been included in a description of the overall effect of duration from the entire four part test series as has been described elsewhere in this report. This section concentrates on the effects of Test IV results on the practical implications for the aircraft designer.

The format of Figures 23 and 24 and Sample Problem, Figure 25 is useful in that it describes the necessary noise control required for each aircraft noise component, and the resulting effect on aircraft noise reduction is illustrated by the top and left most axis labeled "Required PNdB Reduc - tion". As indicated by the arrow to the right of the scale, a combination of all three major noise components is required, to varying degrees, to achieve an overall aircraft noise reduction. Each of the three sub-components in turn, as illustrated by the down-pointed arrows, is responsive to one or

more changes in an aircraft noise sensitive design parameter such as tip speed, power, etc.

Figure 25, which is a reproduction of Figure 23, will be used to illustrate how the detailed noise reduction requirements may be determined for the Helicopter. If it is determined, from subjective tests for example, that an 11 PNdB reduction is required, the top three figures yield the values of required component noise reduction: 13 dB of rotational noise, 15 dB of vortex noise, and 6 dB of engine noise. Each of these component noise reductions in turn yield the required change in design component values to achieve the required component noise reductions. For the rotational noise to be reduced 13 decibels, a blade tip speed reduction of 200 feet per second (61 m/s) would be necessary.

In addition to the 10 dB reduction achieved by reducing the tip speed, an extra blade could be added to the rotor system for an additional 3 dB. For the vortex noise to be reduced by 15 dB, all three major design factors (tip speed, blade performance, and blade area) should be optimized first, or noise reduction is to be achieved by means of tip shape and plan form of the blade. Similarly, to achieve 6 dB of engine noise reduction, compressor tip speed could be reduced by approximately 125 feet per second (38 m/s).

The symbols "+/OR" on Figures 23 through 25 indicate that, for the design parameter trade-off, various combinations of parameters may be used to achieve a required amount of component noise reduction. Also, the dashed extrapolation of the lines shown indicate that beyond a certain limit in a parameter an impractical or impossible design change would have to be made. The concept shown can be extended to cover additional design parameter-component noise tradeoffs possible, but for clarity, only the major noise-influential parameters are shown.

Effect of Test Subjects' Activities

Responses to aircraft sounds during work and leisure environmental conditions.- In each experimental session, the test subjects spent the first two hours in a "work" environment and the second two hours in a "leisure" environment.

During both conditions, the subjects rated their annoyances to the aircraft sounds (see Appendix B). They chose a variety of work activities during the first two hours such as writing technical reports or research papers, professional study, or preparing for tests. For the second two hours, the subjects also selected a diversity of activities such as playing cards, watching television, casual conversation, or reading. They were not required to engage in the same "work" or "leisure" activities during the sixteen weeks of experimentation. The activities engaged in by the subjects were recorded by the psychologist present at each session.

Table VI shows the results for the frequency of "no response" (zero) for the two aircraft under conditions of work and leisure. There are striking differences in the proportion of "no responses" for both aircraft for each experiment. Overall, there are approximately twice as many no responses under leisure conditions as there are under work conditions. The differences shown for each test are highly significant statistically ($X^2 < .001$). It should be noted that detailed analysis of the data for the individual tests shows higher frequencies of no response for the shorter durations - particularly for 15-second and 30-second durations and for lower volume levels. This holds true for both aircraft.

Although there are detailed variations in consistency, in general the data indicates that when the "no responses" are included in the medians, the aircraft noises are usually rated more annoying during the leisure period than during the work period. However, if the data is examined without including "no responses" as a numerical zero value, the trend tends to reverse, showing either (a) little difference between work and leisure responses, or (b) the aircraft rated are less annoying during leisure. In evaluating this apparent contradiction, it could be concluded that during leisure, the observers tend to be less aware of the presence of the aircraft sound, especially for short-duration, lower level sounds.

Examination of observers' responses according to leisure activities. - Results were obtained for observers engaged in different leisure activities, comparing mean annoyance ratings under work and leisure conditions for each aircraft separately. Data were used from tests III and IV. Activities consisted of listening to and watching television, playing cards, or talking. Combinations of these three activities were then classi-

fied into five groups: 1. television; 2. talking; 3. playing cards; 4. television and talking; and 5. television and playing cards.

Based on the analysis of the results of the subjective response distributions for each of the foregoing activity categories, it is not obvious to what degree the aircraft noise intruded. Of the patterns observed, the two most consistent were: a) both levels of helicopter sounds were judged less desirable when the television was on compared to when there was talking or playing of cards; and b) the lower level of tilt wing noise was judged more acceptable during conversation than during card play or watching television. An explanation of these results in terms of speech interference level or some other such communications efficiency measure is not readily apparent.

CONCLUSIONS

As a result of the psycho-acoustical tests described in this report, the following conclusions are drawn:

1. Helicopter and Tilt Wing sounds of 15 seconds duration are judged equal in annoyance to a jet reference sound when they are 4 to 5 PNdB lower than the jet. This conclusion is based on the amount of overlap in the distribution of the subjective responses exceeding 75 per cent, which implies that there are probably no real differences between those sounds subjectively.

2. A Helicopter sound of 15 second duration is judged significantly more acceptable than a jet when its perceived noise level is 10 PNdB below that of the jet. Similarly, a 15 second Tilt Wing sound is so judged when 12 PNdB lower than the jet. These conclusions are based on a 45 per cent overlap of subjective response distributions.

3. The relative effect of changing the sound duration on its subjectively rated annoyance diminishes with increasing duration. This study verified that for low-frequency VTOL aircraft sounds the effect amounts to a 2 PNdB per doubling of duration between

15 and 30 seconds. For larger durations, however, this study concluded that this slope is no longer valid and diminishes to 1.5 PNdB per doubling of duration between 30 and 60 seconds; 1 PNdB per doubling of duration between 60 and 120 seconds; and 0.75 PNdB per doubling of duration between 120 and 240 seconds. This asymptotic approach to a zero slope implies that there exists a level of aircraft sound sufficiently low for the duration effects to be completely negligible. For the background noise (NC-30) used in this study, this level of aircraft noise is from 5 to 7 PNdB below a 15 second PNdB level taken as reference.

4. When the test subjects were engaged in leisure activities, low level-short duration VTOL aircraft sounds were noticed less often than when the subjects were engaged in work activities. Apparently, the leisure activities detracted the test subjects to the extent that they were less aware of the presence of an aircraft noise during leisure than during the work sessions.

5. Equally good correlation was obtained between subjective ratings and several acoustical measures used to predict subjective reaction: Peak Perceived Noise Level (PNLC), dBA, and dBC. These results in essence agree with those of previous investigators (ref.'s 6 and 7) who indicate that for low-frequency, impulsive noises, such as those of the Helicopter used in this study, dBC is a somewhat better predictor of subjective response than PNLC.

6. An extremely high correlation coefficient ($r = 0.94+$) was obtained in this study between the measured acoustical data and the subjective responses given by the test subjects. The equally high correlation of the various noise indicators implies an insensitivity to the finer points inherent in the more complex ones, such as PNLC, for the low-frequency impulsive types of noises.

7. The high correlation coefficients obtained are indicative of a substantiation of the test method used in this program to evaluate the effect of the aircraft noises on a person as he engages in normal work and leisure activities.

The Boeing Company
Vertol Division
Philadelphia, Pa., January 10, 1972

APPENDIX A

EVALUATION OF TEST METHODS

As discussed in the section on Test Methods, the test subjects participated in a paired comparison test, identical with that given to other subjects as part of the reference 1 study. The only difference was that this second program was conducted in the trailer which was to be used for the absolute test program. The results, however, were substantially different than the one which had been conducted earlier. Since it was suspected that there might be discrepancies due to differences in sound reproduction or room acoustics, a re-test was conducted using fewer subjects. This second test reaffirmed the previous results and are shown in Figure 26. No firm explanation can be offered which can point to variation being largely due to room acoustics, subject attitudes, or other factors.

Also shown on Figure 26 are points which were derived from the results of the absolute testing but with the data treated in the manner of paired comparisons. That is, if an individual rated a sound with a higher response than the jet it was considered as being judged worse than the jet reference. This is in effect another type of comparative testing although not involving specific paired comparison judgments.

Another procedure for judging the equivalence of the jet reference and various VTOL sound was suggested in the section on Duration Effects. This was based on the 75% overlap of the Tilton's O criteria.

APPENDIX B

MEASUREMENT OF SUBJECTIVE JUDGMENTS OF ANNOYANCE OF VTOL AIRCRAFT SOUNDS

INTRODUCTION

Psycho-acoustic studies of the subjective responses of people to aircraft sounds in terms of their acceptability have been conducted in field studies and in the laboratory. Various experimental designs, using psychophysical and psychological scales have been employed, depending on the purpose of the study. Each method has its own advantages and disadvantages, depending on the type of data being studied. Subjective ratings have been correlated with the physical characteristics of sounds in order to study the relationships of such factors as volume and duration to perceived annoyance.

It is extremely important to describe and evaluate the methods used in any experiment which depend on the responses of observers to some form of sensory stimulation. The purpose of this Appendix is to list and briefly describe available methods, describe the psychological procedures used in the present experiment, and report findings from studies of this method.

PART I

PSYCHOPHYSICAL AND PSYCHOLOGICAL SCALING METHODS*

The following methods and their variations have been used to study subjective responses to aircraft noise in laboratory investigations:

---

* Excellent discussions of psychophysical and psychological scaling methods can be found in references 11 through 13.

APPENDIX B

Methods Using Comparative Judgments

Methods of constant stimulus differences. - This variation of the Paired Comparison Method is often referred to as the Constant Method. The procedure is as follows: (1) a set of stimuli is prepared in which the magnitude or size of the stimuli proceeds in regular steps. Each step is well below the expected threshold; (2) a stimulus near the middle of the series is selected to serve as a standard; (3) the standard stimulus is presented to the observer along with each other stimulus, either successively or simultaneously; (4) the observer reports which stimulus of the pair has more or less of the specific attribute (e.g., more annoying); (5) scale values are obtained for the stimuli.

Method of paired comparisons. - This method consists essentially of the following: (1) Stimuli to be evaluated on a psychological scale are usually presented to the observer in all possible pairs in counter-balanced order; (2) The observer reports which one of the pair is of greater magnitude (e.g., more annoying) than the other, according to his specific instructions; (3) the object is to give each stimulus a single value on a scale. It is hoped that the scale will have the properties of a linear scale.

Ratio Methods

Ratio scales have the advantage of having an absolute zero. The zero stands for none of the properties represented by the scale. However, these scales are probably more difficult to use because they require the observer to carry out the complex task of judging ratios. This method has been used to some extent in the study of the acceptability of aircraft sounds.

Examples of Ratio Methods are as follows: (1) Method of Fractionation - stimuli are selected and the observer is asked to find a stimulus which is some ratio to another (e.g., one-half, one-third, etc.); (2) Method of Multiple Stimuli - the observer is instructed to find a stimulus which

is two or three (etc.) times as great as another; (3) Method of Constant Sums - the observer is presented with two stimuli and is instructed to report how many times as great one is over the other.

Method of Successive Categories

This psychological scaling method is also called the Method of Single Stimuli. It includes all rating methods in which categorical judgments are made. The judgments are absolute, rather than comparative. This approach will be discussed in some detail, since it was used in the present investigation.

Assumptions.-

1. That the observer is capable of rating or sorting stimuli on an equal-interval scale.

2. That the categories are in correct order and that their boundary lines are stable.

3. That numerically equal distances on the scale stand for equal distances on some aspect of the stimuli (e.g., annoyance vs. PNdB).

Method.-

1. A set of stimuli are prepared (e.g., aircraft sounds).

2. The observer is supplied with a rating scale, consisting of discrete steps or categories which are assigned numbers. The steps on the scale are generally defined in some way.

3. Each stimulus is presented to each subject one or more times in a random order.

4. Using the scale the observer assigns one and only one number to each stimulus presented to him, using some form of recording device.

5. Scale values are generally obtained by computing the means and standard deviations of the subjective judgments (ratings) for each individual stimulus (aircraft sound); either for a single observer or for a group of observers.

Replication.- Because of variability in responses, it is necessary to obtain repeated measurements over trials with a single judge or over judges in order to arrive at more stable estimates of scale values and to measure the reliability of the scale.

Instructions to the observers.- The outcome of an experiment of this type depends to no small extent on what instructions are given to the observers. Without specific instructions, an observer is free to handle the rating scale in any way he chooses. This introduces an unknown variance factor which is extremely difficult to take into account because of individual differences between observers. Through the instructions given to the observers, the experimenter can, to some extent at least, minimize this problem and stabilize the scale units. (The instructions utilized in this experiment appear in Part II of this Appendix.)

Reliability.- Measures of subjective responses to known stimuli are of questionable value without some measures of the reliability or consistency of individual ratings. Generally speaking, measures of the variance of ratings, internal consistency, test-retest reliability, and inter-judge agreement are used to judge the reliability of subjective ratings.

Establishment of a reference scale.-

1. A reference scale is established by use of a response function. A mean and variance of scale values is estimated from the data.

2. The sample mean plus and minus two standard errors is regarded as an interval which includes the population parameter with a probability of at least .95. This is the coefficient of confidence for a bound.

APPENDIX B

General evaluation of the method of successive categories.-

1. Bock and Jones (ref. 11, pg. 213) state that results from the Method of Successive Categories have been found repeatedly to correlate high with those obtained from the Method of Paired Comparisons.

2. Torgerson (ref. 12, pg. 53) states that we would expect substantially the same rank order for a given set of stimuli for a specified attribute, whether we use Paired Comparison, Ranking, Single Stimulus Rating, or sorting into successive intervals.

3. Guilford (ref. 13, pg. 297) in a general evaluation of rating methods as compared to paired comparisons, points out that: (a) ratings require less time; (b) the procedure is far more interesting to observers; (c) the procedure can be used with psychologically naive observers; (d) best judgments are made when stimuli are presented singly; and (e) the experimental operations for obtaining judgments using the method of successive categories are simple and economical. From this point of view, the method has everything in its favor.

4. In an experiment where it is desired to study the absolute annoyance value of aircraft sounds, it is desirable to use the method of successive categories. The basic problem from a practical point of view is how annoying the sounds are to an individual person (who in a real sense has his own scale) and not how sounds or persons compare to each other. The real concern of course is, given a level of annoyance to certain aircraft sounds, will the person complain about it or take some kind of social action. The present study did not investigate this problem.

5. All in all, the method of successive categories, using the best instructions and rating scale that we could devise, appeared to be the most suitable method for the present experiment. A succeeding section of this Appendix discusses some of the findings concerning its characteristics and use.

APPENDIX B

PART II

THE OBSERVERS AND THEIR INSTRUCTIONS

Subjects

The subjects in this experiment were all males between the ages of twenty-one and thirty-four. They were engaged in teaching, business and the professions. All were volunteers and members of the Jaycees. All compensation for the services of the subjects was paid directly to the Jaycees chapters involved.

Instructions to Subjects

"(1) During the next four hours, you will be involved in a study regarding the subjective acceptability of aircraft noise.

(a) The first two hours is defined as a "work" period. Bring your own work with you. Examples: writing technical reports or papers, research, professional study, preparing for tests, etc.

(b) The second two hours is defined as a "leisure" period. Such activities as watching television, playing cards, conversation, or recreational reading are suitable.

(2) You will be hearing a variety of aircraft sounds during the evening. Your primary task is to record your reactions to each sound you hear on the instrument in front of you.

(3) Note the color of the instrument. You are expected to use the same instrument throughout the four-hour period. Also, note the various colored electrical outlets in the room. It is essential that you match the color of your instrument with the color of the outlet.

(4) You will also notice that your recording instrument has nine numbers, arranged in three rows. Record your reaction by pushing a number from one (#1) to nine (#9).

(a) For example, push number one, two or three (bottom row) when the sound your hear is in an acceptable or non-annoying range.

#1 is least annoying on the scale.

(b) Or, push number seven, eight or nine (top row) when the sound is in an annoying or irritating range.

#9 is most annoying on the scale.

(c) Further, push number four, five or six (middle row) when the sound you hear is in a moderately annoying range.

(d) Each time, push the button you choose firmly and hold it for a few seconds.

(5) If you decide to change your response during the sound or immediately afterward, you may do so by pressing another number. The last number you push is the one recorded by the monitor.

Any Questions?"

Note: The psychologist who was acting as the test administrator at the time read the instructions to the observers at the beginning of the experiment and whenever necessary afterwards. He demonstrated the use of the apparatus and checked whether the responses were being recorded on the monitor. Also at the beginning of each session, each observer pressed each button on his instrument in order to check whether or not the response was being shown on the monitor. The test administrator in the trailer was in direct telephonic communication with the technician in the separate electronics laboratory, so that immediate steps could be taken to correct any errors in the recording system at any time.

APPENDIX B

PART III

SPECIAL STUDIES

Reliability

There are three general approaches to the problem of measuring the reliability of psychological measuring instruments: (1) test-retest reliability; (2) measures of internal consistency; and (3) inter-judge agreement. Of these the first method was used in the present study. The reliability of sample statistics such as the mean and correlation coefficients was established by means of the standard errors of the statistics in question (See Tables VII (a) to (d)). However, the latter is based on the variance of the data as a whole, rather than the performance of individual observers.

The test-retest coefficients were computed by correlating the sum of the first two exposures to a stimulus with the sum of the last two exposures - person by person. It should be remembered that each stimulus was exposed in a random position in the total series at least four times during the particular two-hour work or leisure session (four tapes, each one-half hour long). On the basis of the following results, we may conclude that the procedure does, in fact, have very high reliability at the beginning (Test I) and the end of the whole investigation (Test IV).

The results were as follows: (1) Test I - Work, $r = .90$; Leisure, $r = .93$; (2) Test IV - Work, $r = .91$; Leisure, $r = .90$.

The standard errors of the means are shown in the basic tables of results for each experiment (Tables VII (a) to (d)). As pointed out above, the confidence bands for a mean at the .95 level of confidence is plus and minus two standard errors. This is an important consideration in interpreting the data, particularly in considering the average differences between stimuli.

APPENDIX B

Characteristics of the Scale*

Standard score equivalents are commonly used in psychological testing of all sorts - they are considered to be equidistant from a psychological point of view. They are obtained by dividing the distance of a score from the mean by the standard deviation of the distribution. Using the total distribution of ratings for Test I (ratings for all observers, all conditions, and all stimuli), it was found that there is an apparent linear relationship between original scale values and the corresponding standard score values. The interval between scale values is approximately .40 of a standard deviation. It is interesting to note that the overall mean is 4.56 with a standard deviation of 2.41. The middle of the original scale is, of course, five (5). Figure 27 shows the relationship of original scale scores to standard scores.

Using the mean subjective ratings for each condition in Test I, it was found that there is also an apparent linear relationship between annoyance ratings and PNdB levels of stimuli. If desired, curves of this type can be used to estimate the scale values of any new stimulus located within the range of the original set. Figure 28 shows the relationship between mean ratings and PNdB values.

Analysis of Variance Results

Separate analyses of variance were performed for all observers under all experimental conditions in Test I as follows: Helicopter - Work; Helicopter - Leisure; Tilt Wing - Work; Tilt Wing - Leisure. There were twenty-five (25) observers and five (5) volume (PNdB) levels. Time duration for the stimuli was constant at one minute.

The purposes of the analyses were as follows:

1. to determine whether differences between means of observers were statistically significant, using repeated observations of the same stimuli.

---

* Test I was used in these studies as standard.

2. to determine whether differences between means of aircraft sounds differing in volume were statistically significant;

3. to determine the relative proportion of the total variation in a specific experimental condition (e.g., Helicopter - Work) that could be attributal to individual differences between observers, differences between stimuli, and error (unaccounted for variance). The separate analyses of variance are shown in Table VIII.

Results. - The differences between the means of observers were highly significant statistically ($p < .01$), indicating that there are individual differences between observers in their subjective ratings of the annoyance of aircraft sounds. Table VIII shows the results for the analysis of variance and Table IX shows the proportion of the variance accounted for. Since Individual Differences between Observers accounts for only about 15% of the total variance, it is seen that the subjective scale is not primarily measuring (testing) general annoyance to aircraft sounds -- in other words it is not just another attitude scale without any basis upon which to interpret it. Figure 29 shows the distribution of the means of the twenty-five subjects for the four experimental situations. Ordinarily, distributions of psychological measurements in terms of individual differences distribute themselves normally, as they do here. The distribution shows no systematic bias in the subjective ratings for these subjects. It does bring out very clearly that some subjects are very sensitive to aircraft noise and are annoyed by it, even at low levels of volume. Others are apparently much less annoyed, regardless of the level (volume). However, the bulk of the mean ratings are in the middle range (similar to findings in some other studies). The mean of 4.85 and standard deviation of 1.03 are further evidence of the validity of the scale.

The differences between the means of annoyance ratings were also highly significantly statistically ($p < .01$). Thus, there appear to be real differences between the annoyance ratings for the aircraft sounds. Figures 30 (a) and (b) show graphically the distributions of ratings for different levels (volumes) of aircraft sounds for the Helicopter and Tilt Wing. The reference jet sound is also shown. Differences between the means of stimuli account for about 66% of

the variance, indicating that the main source of variation is attributable to differences between ratings of sounds (as compared to individual differences). The validity of the subjective rating scale in this type of psycho-acoustical study appears to be high.

The error component accounts for about 16% of the variance under the conditions of Test I.

Sensitivity of the scale. - It has been shown from the analyses of variance that differences between means of sounds are statistically significant. Using the data shown in Figures 30 (a) and (b), and applying Chi Square, it was found that the distributions of ratings were significantly different from each other ($p < .01$) and that the correlation (coefficient of contingency) between the magnitude of the sounds and that of the ratings was high: $C = .77$. These findings are further evidence of the validity of the scale, since under fairly standard conditions (Test I), the scale not only appears to be linear, but the relationship between scale values and physical values is high and positive.

Overlapping. - A very useful way of looking at the data is in terms of the extent to which the distributions overlap. Dunnette (ref. 2, p. 142-143) states that calculation of means does not really yield much information for accurately identifying (or differentiating between) different groups. He holds that the percentage of scores obtained by one group that may be matched by scores obtained in the other group is an index of the test's accuracy (validity) for separating any two groups. He recommends the use of Tilton's O statistic (ref. 3) and suggests criteria for interpretation of the amount of total overlap between two distributions.

The question here is, to what extent the distributions of annoyance ratings for the different stimuli overlap each other. If they overlap completely, there is obviously no difference between them - they are equally annoying. Conversely, if there is no overlap, the two distributions are completely different. Also, the question may be raised as to how many scale separations are necessary in order to be confident that the two sounds really differ in subjective annoyance.

APPENDIX B

Tables X (a) and X (b) show for Test I, the total percent of overlap (Tilton's O) of the jet reference sound with each of the other sounds and also that for each level with the other levels. For both aircraft, it is readily apparent that the overlap decreases markedly as the stimuli become farther apart in volume.

According to Dunnette (ref. 2, p. 149), very high overlap ($> 75\%$) indicates that there is probably no real difference in the ratings - that the annoyance value is the same for practical purposes. For both aircraft, the jet and Levels 1 and 2 sounds show very high overlap, the highest being with Level 1. Also Levels 4 versus 5 show high overlap for both aircraft. Moderate overlap ($> 45\%$ and $< 75\%$) indicates that there are real, but moderate differences in the ratings of annoyance for the levels concerned. For both aircraft, Level 1 versus Level 2 falls into this category. Finally, low overlap ($< 45\%$) indicates large differences in the ratings. For both aircraft, for example, Level 1 versus Level 4 and Level 5 falls into this category.

In terms of the findings, Level 1, 3 and 5 appear to be clearly different in subjective ratings for annoyance for both aircraft. For example, for the helicopter Figure 31 shows the distributions of subjective responses for these three levels. Overall, the correlation between the annoyance scale and the physical scale (PNdB) is .83 when these levels are used. When all five levels are taken into account, this correlation, in terms of the contingency coefficient C, is only slightly lower.

Results for jet reference sound.- Every work session and leisure session included a number of jet reference sounds (PNdB = 90) which were inserted at random in the sequence of stimuli. Table XI shows the means, medians, and standard deviations for work during the course of the four tests for the same jet reference sound. It is apparent that there were no large changes in the average annoyance ratings for the jet sound over the 16-week period.

APPENDIX B
CONCLUSIONS

The nine-step scale for the subjective rating of annoyance to aircraft sounds is highly reliable and is clearly related to the physical characteristics of the aircraft sounds heard by the observers under the conditions of the present tests. As far as can be judged, the scale intervals appear to be equal and linear. The scale appears to be suitable for investigations of this type.

REFERENCES

1. Hinterkeuser, E.G. and Sternfeld, H. Jr., "Subjective Response to Synthesized Flight Noise Signatures of Several Types of V/STOL Aircraft", The Boeing Company, NASA Contractor Report CR-1118, August 1968.
2. Dunnette, M.D., "Personnel Selection and Placement", Belmont, California, Wadsworth Pub. Co., 1966.
3. Tilton, J.W., "The Measurement of Overlapping", Journal of Education Psychology, 28, pp.656-662, 1937.
4. Anonymous, "Definitions and Procedures for Computing the Perceived Noise Level of Aircraft Noise", Society of Automotive Engineers, Inc., Report 865A, August 15, 1969.
5. Freund, J.E., "Mathematical Statistics", Prentice-Hall Pub. Co., 1962.
6. Fidell, S. and Pearsons, K.S., "Study of the Audibility of Impulsive Sounds", Bolt Beranek and Newman Inc., NASA Contractor Report CR-1598, May 1970.
7. Ollerhead, J.B., "An Evaluation of Methods for Scaling Aircraft Noise Perception", Wyle Laboratories, NASA Contractor Report CR-1883, May 1971.
8. Kryter, K.D. and Pearsons, K.S., "Some Effects of Spectral Content and Duration on Perceived Noise Level", Journal of the Acoustical Society of America, 35, pp.866-883, 1963.
9. Pearsons, K.S., "The Effects of Duration and Background Noise Level on Perceived Noisiness", Report FAA-ADS-78, Bolt Beranek and Newman Inc., Van Nuys, Cal., 1966.
10. Kryter, K.D., "The Effects of Noise on Man", Academic Press, pg. 299, 1970.
11. Bock, R.D. and Jones, L.V., "The Measurement and Prediction of Judgement and Choice", San Francisco, Holden-Day Pub. Co., 1968.
12. Torgerson, W.S., "Theory and Methods of Scaling", New York, Wiley and Sons, 1967.
13. Guilford, J.P., "Psychometric Methods", New York, McGraw-Hill Pub. Co., 1954.

TABLE I

TEST PROGRAM OUTLINE

Test	Purpose	Durations	Levels
1	Comparison of subjective test techniques	60 Seconds	5 Volume levels
2	Trade-off of sound level with sound duration	15,30,120 and 240 Sec.	3 Volume Levels
3	Trade-off of sound level with aircraft distance	15,30,60,120 and 240 Sec.	2 Volume levels spectrum shaped for distance
4	Trade-off of sound level and aircraft design	15,30,60,120 and 240 Sec,	2 Volume levels spectrum shaped for component noise control

TABLE II
SUBJECTS' EXPOSURE HISTORY

Test Series	1	2	3	4	Total
Test Duration - Weeks	2	6	4	4	16
Number of times each stimulus occurs/2 hr session	5	2	2	2	
Total exposures/ stimulus	10	12	8	8	
Total observations/ stimulus (3 sessions/ wk. X 5 subjects/ session)	150	180	120	120	
Number of stimuli (levels and duration)	5	12	10	10	
Number of exposures/ aircraft/ environment	750	2160	1200	1200	5310
Number of exposures-2 aircraft 2 environments	3000	8640	4800	4800	21240

TABLE III

BASIC ACOUSTIC AND SUBJECTIVE DATA ON HELICOPTER SOUNDS

SOUND NO.	TEST SERIES No.	SOUND LEVEL No.	SOUND DURATION (Sec.)	dB-C LEVEL (dB)	dB-A LEVEL (dB)	PNLC (PNdB)	SUBJECTIVE RATING MEAN (WORK)
1	1	1	60	84	71	86	7.57
2		2		81	67	82	5.72
3		3		77	62	78	4.88
4		4		71	56	71	3.03
5		5		67	49	66	2.25
6	2	1	15	89	77	91	7.27
7			30	89	76	91	7.61
8			120	87	74	89	7.25
9			240	82	69	84	6.61
10		2	15	80	68	81	3.55
11			30	80	66	81	5.33
12			120	74	61	76	3.91
13			240	68	54	69	3.19
14		3	15	66	52	68	1.78
15			30	67	52	67	2.15
16			120	62	45	63	1.69
17			240	62	47	64	1.63
18	3	1	15	84	72	86	6.30
19			30	87	77	90	6.64
20			60	86	76	89	6.08
21			120	86	74	87	6.22
22			240	82	71	84	5.57
23		2	15	72	56	71	1.74
24			30	72	58	73	2.20
25			60	67	52	67	2.22
26			120	68	49	63	2.36
27			240	69	56	71	3.29
28	4	1	15	87	73	88	7.50
29			30	86	72	87	7.53
30			60	83	70	85	7.39
31			120	83	70	84	7.48
32			240	78	63	79	6.18
33		2	15	68	55	70	2.37
34			30	68	54	70	2.24
35			60	68	50	67	2.19
36			120	65	51	67	2.24
37			240	70	58	71	3.80

TABLE IV

BASIC ACOUSTIC AND SUBJECTIVE DATA ON TILT WING SOUNDS

SOUND NO.	TEST SERIES No.	SOUND LEVEL No.	SOUND DURATION (Sec)	dB-C LEVEL (dB)	dB-A LEVEL (dB)	PNLC (PNdB)	SUBJECTIVE RATING MEAN (WORK)
38	1	1	60	86	68	84	7.52
39		2		81	61	78	5.86
40		3		78	58	75	4.40
41		4		71	51	68	2.75
42		5		66	46	65	1.88
43	2	1	15	88	72	86	7.08
44			30	90	72	88	7.73
45			120	84	66	81	7.20
46			240	82	64	80	6.80
47		2	15	77	61	76	3.56
48			30	78	60	76	4.80
49			120	74	55	71	3.90
50			240	69	50	67	3.26
51		3	15	64	46	63	1.69
52			30	65	56	64	2.41
53			120	60	45	59	1.95
54			240	62	43	61	1.50
55	3	1	15	84	65	82	5.51
56			30	83	64	81	5.93
57			60	82	64	81	6.11
58			120	81	62	77	5.82
59			240	76	58	71	4.99
60		2	15	72	52	69	2.23
61			30	69	51	67	2.69
62			60	68	48	66	2.43
63			120	69	51	65	2.50
64			240	67	47	65	2.33
65	4	1	15	78	64	79	4.71
66			30	77	64	80	5.16
67			60	76	63	78	5.22
68			120	74	61	77	4.93
69			240	69	58	73	4.85
70		2	15	71	56	71	2.83
71			30	65	50	66	2.35
72			60	60	45	62	2.01
73			120	61	46	63	1.89
74			240	64	44	62	1.81

TABLE V

CORRELATION OF MEAN SUBJECTIVE RATINGS WITH SEVERAL ACOUSTIC MEASURES

			Duration t - Seconds				
Aircraft	Acoustic Measure	All Durations	15 Sec.	30 Sec.	60 Sec.	120 Sec.	240 Sec
Helicopter	PNL C	.940	.956	.964	.951	.961	.958
"	dB A	.926	.942	.956	.937	.962	.982
"	dB C	.943	.944	.978	.960	.968	.967
Tiltwing	PNL C	.950	.974	.990	.988	.968	.983
"	dB A	.942	.967	.961	.982	.980	.996
"	dB C	.939	.965	.992	.952	.964	.943

TABLE VI

FREQUENCY OF "NO RESPONSES"

HELICOPTER

	WORK			LEISURE		
Test	N	Zero	%	N	Zero	%
I	750	51	7	750	99	13
II	2040	288	14	2040	473	23
III	1378b	129	9	1385	288	21
IV	1471	93	6	1485	262	18
Total	5639	561	10	5660	1122	20

TILTWING

	WORK			LEISURE		
Test	N	Zero	%	N	Zero	%
I	750	63	8	750	150	20
II	2040	204	10	2040	489	24
III	1307	121	9	1440	461	32
IV	1434	172	12	1433	380	27
Total	5531	560	10	5663	1480	26

TABLE VII

BASIC STATISTICAL DATA ON SUBJECTIVE RESPONSES

Part a: SUMMARY TABLE-TEST I

One Minute Duration

	Helicopter						Tilt Wing					
	Work			Leisure			Work			Leisure		
Level	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.
1	7.57	1.46	0.12	7.54	1.47	0.12	7.52	1.24	0.11	7.75	1.34	0.12
2	5.72	1.55	0.20	6.20	1.79	0.15	5.86	1.51	0.14	5.73	1.62	0.16
3	4.88	1.63	0.14	5.13	1.73	0.15	4.40	1.56	0.13	4.50	1.64	0.15
4	3.03	1.30	0,11	3.34	1.58	0.15	2.75	1.39	0.12	3.20	1.40	0.13
5	2.25	1.13	0.10	2.38	1.26	o.12	1.88	1.10	0.09	2.15	1.23	0.11
Jet	6.91	2.00		7.55	1.49							

$\bar{X}$ = Mean
S.D.= Standard Deviation
S.E.= Standard Error

TABLE VII - BASIC STATISTICAL DATA ON SUBJECTIVE RESPONSES - Continued

Part b: SUMMARY TABLE TEST II

Level 1

	Helicopter						Tilt Wing					
	Work			Leisure			Work			Leisure		
Duration	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.
15 Sec.	7.27	2.24	0.16	8.15	1.41	0.12	7.08	1.85	0.14	7.17	1.78	0.14
30 Sec.	7.61	1.55	0.12	7.81	1.63	0.13	7.73	1.82	0.14	7.91	1.35	0.12
120 Sec.	7.25	1.68	0.13	7.56	1.52	0.12	7.20	1.27	0.10	7.02	1.66	0.12
240 Sec.	6.61	1.96	0.15	7.11	1.64	0.14	6.80	1.63	0.12	6.48	1.69	0.13

Level 2

	Helicopter						Tilt Wing					
	Work			Leisure			Work			Leisure		
DURATION	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.
15 Sec.	3.55	1.73	0.14	4.17	1.84	0.18	3.56	1.71	0.14	4.03	1.76	0.17
30 Sec.	5.33	1.35	0.11	5.72	1.98	0.13	4.80	1.81	0.15	5.58	1.75	0.17
120 Sec.	3.91	1.66	0.13	3.93	1.76	0.15	3.90	1.58	0.12	3.90	1.55	0.13
240 Sec.	3.19	1.55	0.12	3.04	1.64	0.13	3.26	1.56	0.12	3.17	1.47	0.12

$\bar{X}$ = Mean
S.D.= Standard Deviation
S.E.= Standard Error

TABLE VII - BASIC STATISTICAL DATA ON SUBJECTIVE RESPONSES - Continued

Part b: SUMMARY TABLE TEST II - Concluded

Level 3

	Helicopter						Tilt Wing					
	Work			Leisure			Work			Leisure		
Duration	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.
15 Sec.	1.78	1.02	0.09	1.89	1.13	0.12	1.69	1.14	0.10	1.91	1.36	0.15
30 Sec.	2.15	1.16	0.12	2.63	1.53	0.14	2.41	1.35	0.10	2.19	1.20	0.15
120 Sec.	1.69	0.97	0.08	1.79	1.09	0.10	1.95	1.22	0.10	1.80	1.51	0.13
240 Sec.	1.63	0.99	0.08	1.85	1.44	0.12	1.50	0.84	0.07	1.69	1.26	0.10

TABLE VII - BASIC STATISTICAL DATA ON SUBJECTIVE RESPONSES - Continued

Part c: SUMMARY TABLE-TEST III

Level 1

	Helicopter						Tilt Wing					
	Work			Leisure			Work			Leisure		
Duration	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.
15 Sec.	6.30	1.71	0.16	7.13	1.50	0.16	5.51	1.63	0.16	5.85	1.47	0.16
30 Sec.	6.64	1.50	0.12	7.05	1.47	0.13	5.93	1.40	0.11	5.57	1.34	0.11
60 Sec.	6.08	1.50	0.14	6.60	1.67	0.17	6.11	1.33	0.13	5.85	1.46	0.11
120 Sec.	6.22	1.58	0.13	6.58	1.50	0.12	5.82	1.46	0.11	5.32	1.46	0.12
240 Sec.	5.57	1.87	0.18	5.69	1.59	0.15	4.99	1.77	0.17	4.50	1.51	0.15

Level 2

	Helicopter						Tilt Wing					
	Work			Leisure			Work			Leisure		
Duration	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.
15 Sec.	1.74	1.04	0.10	2.45	1.45	0,17	2.23	1.12	0.11	2.84	1.51	o.18
30 Sec.	2.20	1.34	0.11	2.89	1.59	0.15	2.69	1.44	0.12	2.83	1.25	0.12
60 Sec.	2.22	1.69	0.18	2.69	1.60	0.18	2.43	1.31	0.13	2.27	1.34	0.15
120 Sec.	2.36	1.33	0.11	2.70	1.61	0.13	2.50	1.22	0.10	2.36	1.46	0.12
240 Sec.	3.29	1.75	0.17	3.43	1.49	0.15	2.33	1.38	0.12	2.26	1.33	0.13

TABLE VII - BASIC STATISTICAL DATA ON SUBJECTIVE RESPONSES - Concluded

Part d: SUMMARY TABLE-TEST IV

Level 1

	Helicopter						Tilt Wing					
	Work			Leisure			Work			Leisure		
Duration	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.
15 Sec.	7.50	1.72	0.16	7.65	1.73	0.17	4.71	1.27	o.12	5.37	1.56	0.16
30 Sec.	7.53	1.44	0.11	7.55	1.41	0.11	5.16	1.45	0.11	5.47	1.27	0.12
60 Sec.	7.39	1.42	0.13	7.06	1.71	0.17	5.22	1.26	0.12	5.80	1.35	0.13
120 Sec.	7.48	1.33	0.10	7.18	1.47	0.11	4.93	1.33	0.10	5.01	1.52	0.12
240 Sec.	6.18	1.51	0.14	6.10	1.72	0.16	4.85	1.58	0.15	4.93	1.74	o.17

Level 2

	Helicopter						Tilt Wing					
	Work			Leisure			Work			Leisure		
Duration	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.	$\bar{X}$	S.D.	S.E.
15 Sec.	2.37	1.12	0.11	2.63	1.27	0.14	2.83	1.18	0.12	3.27	1.15	0.18
30 Sec.	2.24	1.12	0.11	2.95	1.35	0.12	2.35	1.35	0.11	2.66	1.60	0.16
60 Sec.	2.19	1.17	0.11	2.42	1.28	0.13	2.01	1.53	0.16	2.08	1.39	0.18
120 Sec.	2.24	1.20	0.09	2.40	1.26	0.10	1.89	1.20	0.10	2.05	1.23	0.11
204 Sec.	3.80	1.70	0.16	3.65	1.70	0.16	1.81	1.22	0.12	2.05	1.28	0.13

TABLE VIII

ANALYSES OF VARIANCE

TEST I

Part a: Helicopter-Work

Source of Variation	d.f.	Sum of Squares	Mean Square	F
Total	124	633		
Between Observers	24	111	4.62	4.625
Between Levels	4	426	106.50	106.50
Error	96	96	1.00	

Part b: Tilt Wing-Work

Source of Variation	d.f.	Sum of Squares	Mean Square	F
Total	124	657		
Between Observers	24	91	3.79	5.19
Between Levels	4	496	124.00	169.86
Error	96	70	.73	

TABLE VIII (Concluded)

Part c: Helicopter-Leisure

Source of Variation	d.f.	Sum of Squares	Mean Square	F
Total	124			
Between Observers	24	152	6.33	3.66
Between Levels	4	327	81.70	47.23
Error	96	167	1.73	

Part d: Tilt Wing-Leisure

Source of Variation	d.f.	Sum of Squares	Mean Square	F
Total	124	667		
Between Observers	24	138	5.75	7.27
Between Levels	4	453	113.25	143.35
Error	96	76	.79	

TABLE IX

VARIANCE ACCOUNT

PROPORTION OF VARIANCE ACCOUNTED FOR

AIRCRAFT-ACTIVITY	BETWEEN OBSERV-ERS	BETWEEN SOUNDS	ERROR
Helicopter - Work	18%	65%	15%
Tilt Wing - Work	13%	75%	12%
Helicopter-Leisure	24%	51%	25%
Tilt Wing-Leisure	21%	68%	11%
Combined	15%	66%	16%

TABLE X

TOTAL PERCENT OVERLAP

TEST I

Part a: Helicopter-Work

Level	1	2	3	4	5
Stimulus	86PNdB	82PNdB	78PNdB	71PNdB	66PNdB
Jet 90 PNdB	86%	74%	57%	25%	14%
Level 1 86 PNdB		54%	39%	10%	2%
Level 2 82 PNdB			79%	35%	20%
Level 3 78 PNdB				53%	34%
Level 4 71 PNdB					75%

Part b: Tilt Wing-Work

Level	1	2	3	4	5
Stimulus	84PNdB	78PNdB	75PNdB	68PNdB	65PNdB
Jet 90 PNdB	85%	76%	37%	22%	10%
Level 1 84 PNdB		55%	26%	7%	2%
Level 2 78 PNdB			64%	28%	14%
Level 3 75 PNdB				58%	35%
Level 4 68 PNdB					73%

TABLE XI

RESPONSE STATISTICS FOR JET REFERENCE SOUND

Test	$\bar{X}$	S.D.	Md.
Test 1 Helicopter and Tilt Wing	6.91	2.06	7.82
Test 2 Helicopter and Tilt Wing	6.88	1.86	7.62
Test 3 Helicopter	7.05	1.57	7.73
Test 3 Tilt Wing	7.24	1.56	8.83
Test 4 Helicopter	7.45	1.43	8.08
Test 4 Tilt Wing	7.27	1.36	7.88

$\bar{X}$ = Mean
S.D. = Standard Deviation
Md. = Median

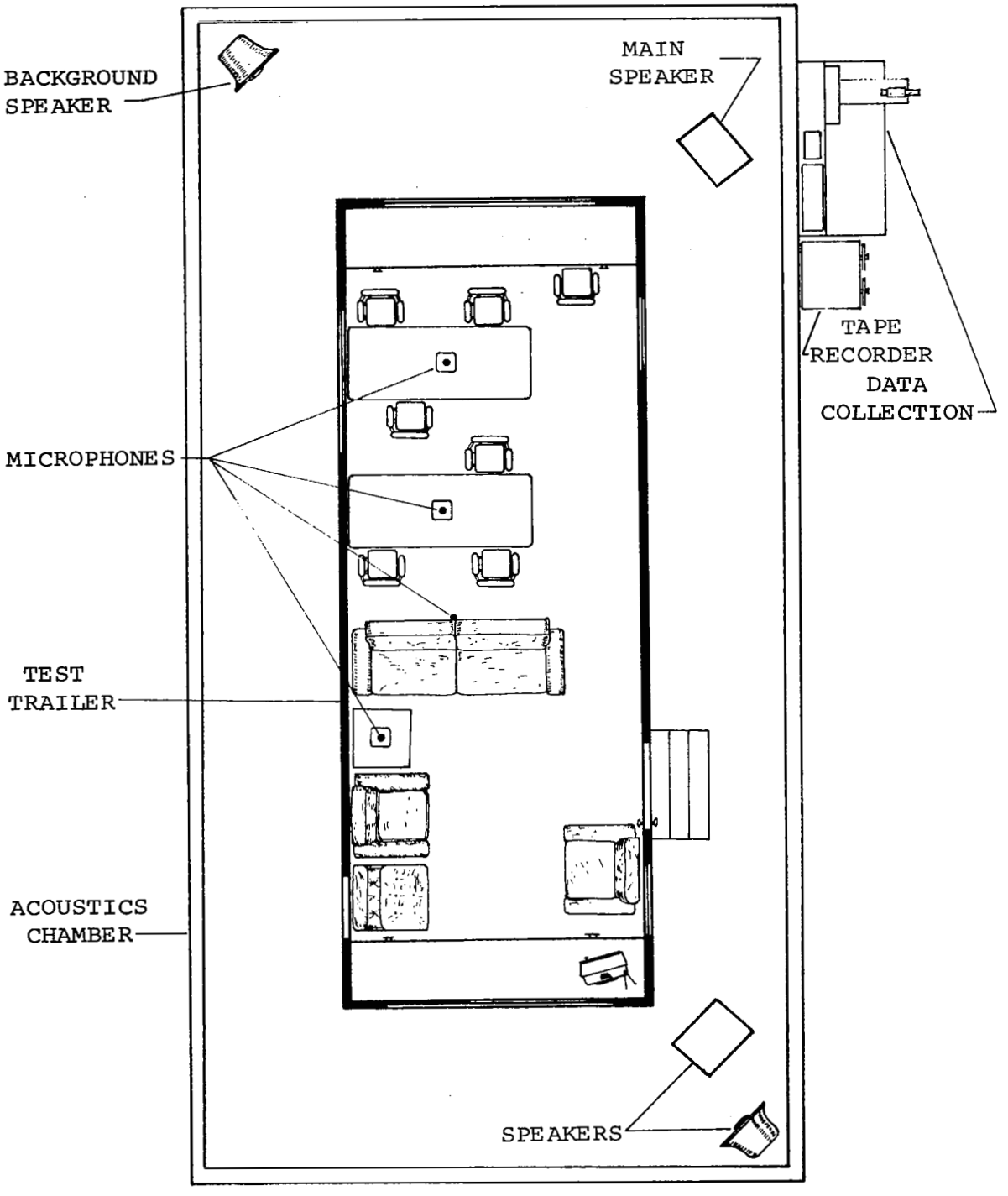


Figure 1. Test Environment - Plan View

WORK ENVIRONMENT



LEISURE ENVIRONMENT

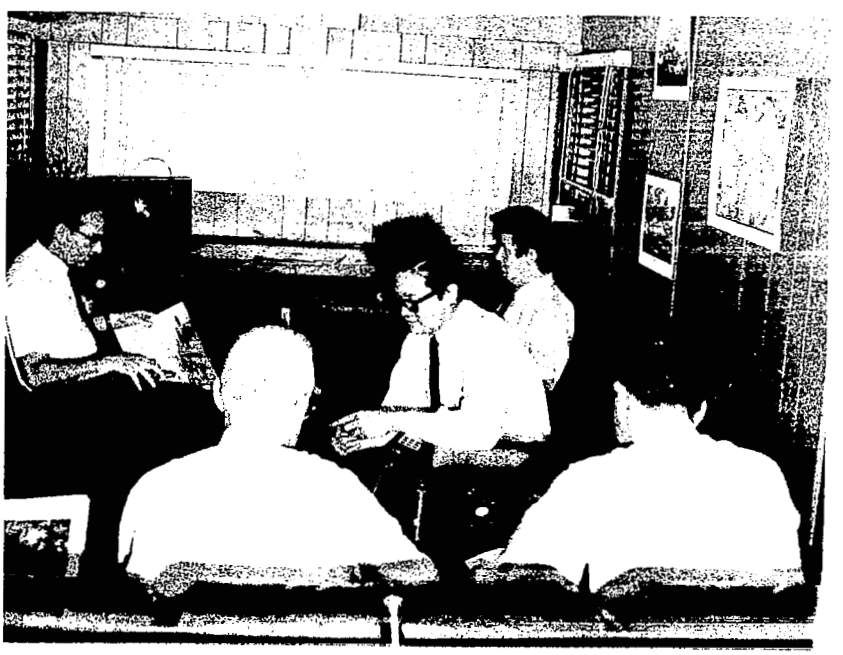
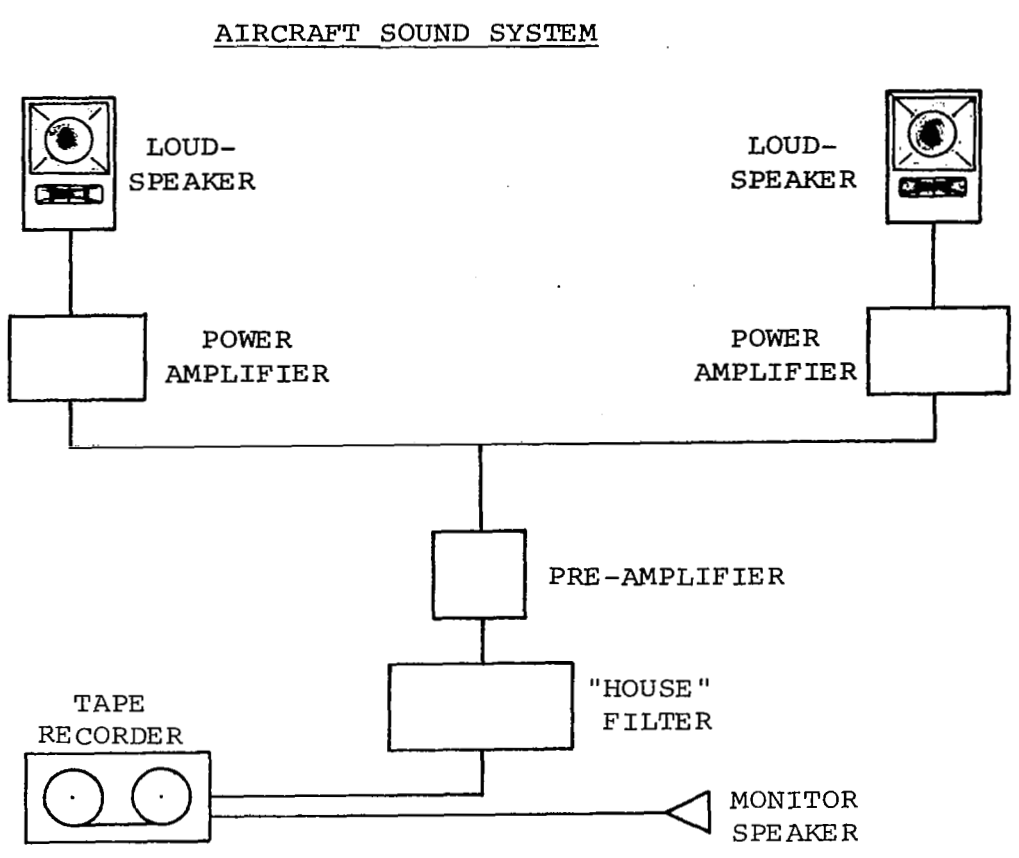


Figure 2. Test Subjects in Work and Leisure Environment.



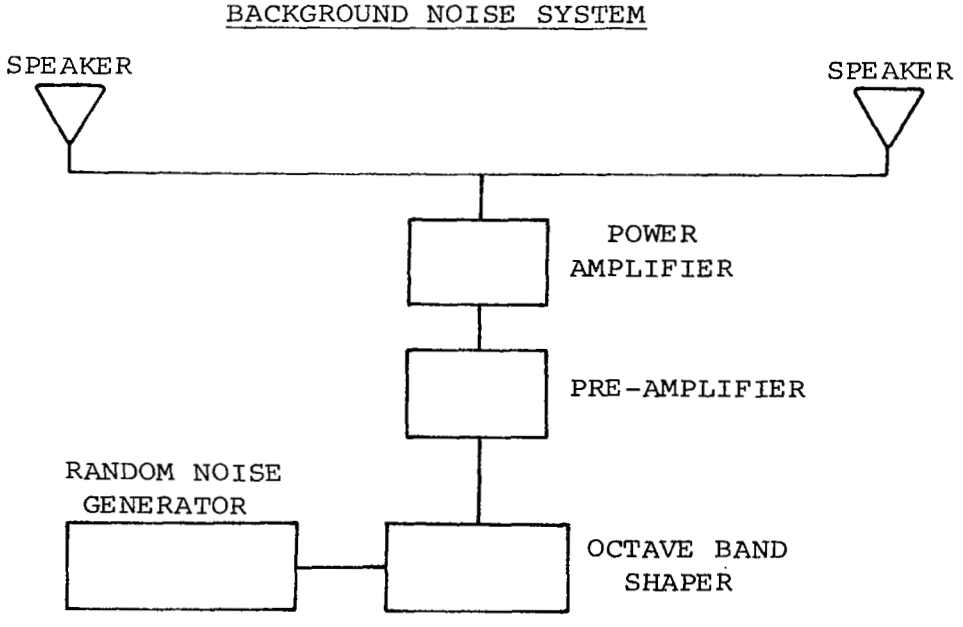
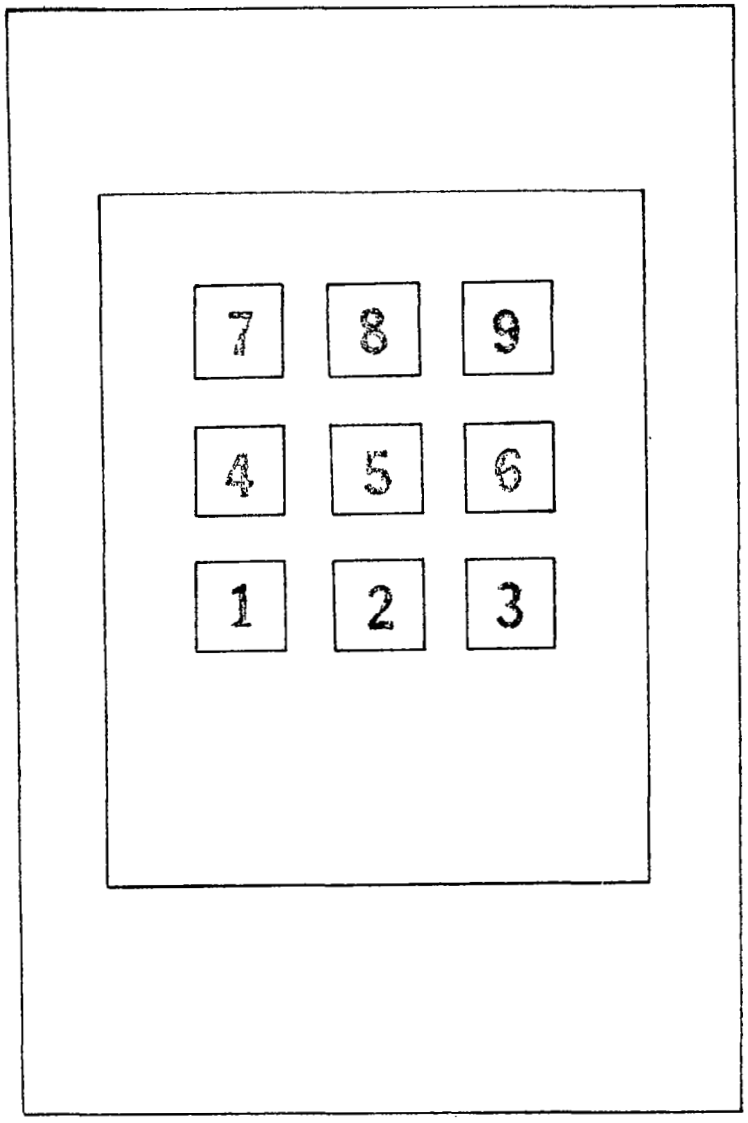


Figure 3. Acoustical System for Sound Generation

(FULL SIZE)



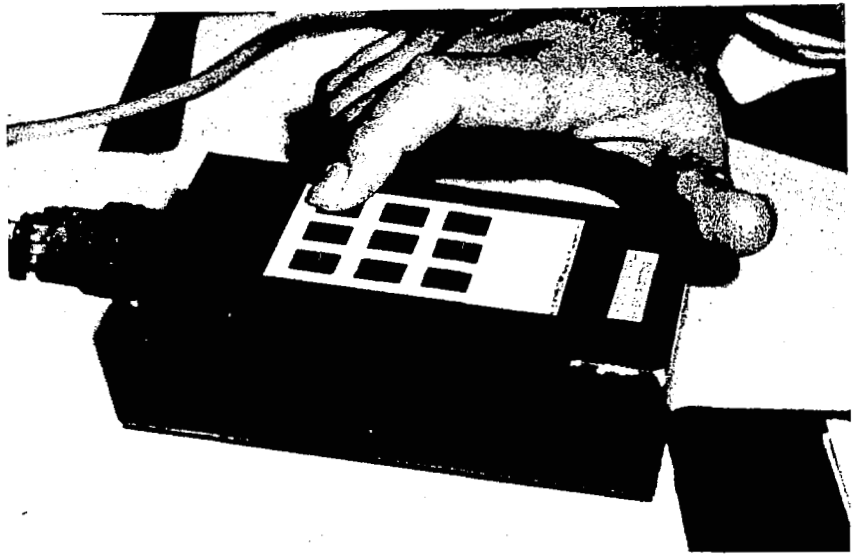
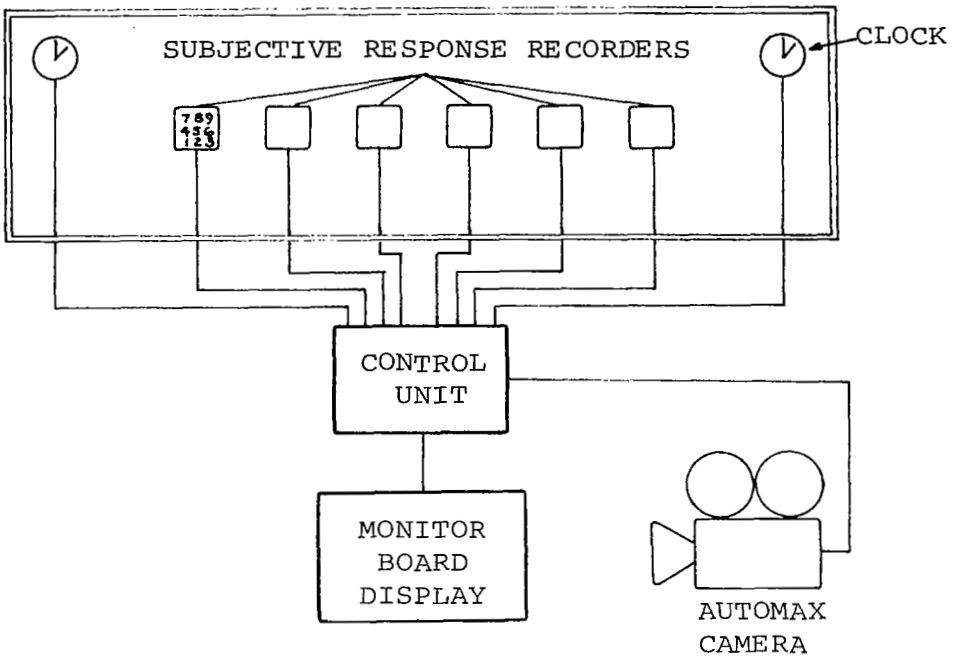


Figure 4. Subjective Response Recorder

TEST ENVIRONMENT



SUBJECTIVE RESPONSE MONITOR BOARD DISPLAY

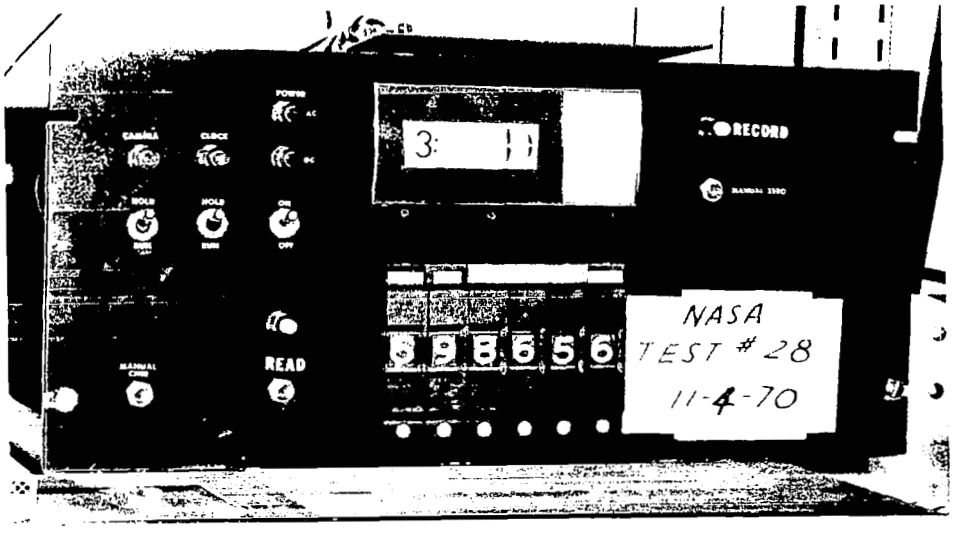


Figure 5. Subjective Response Recording System

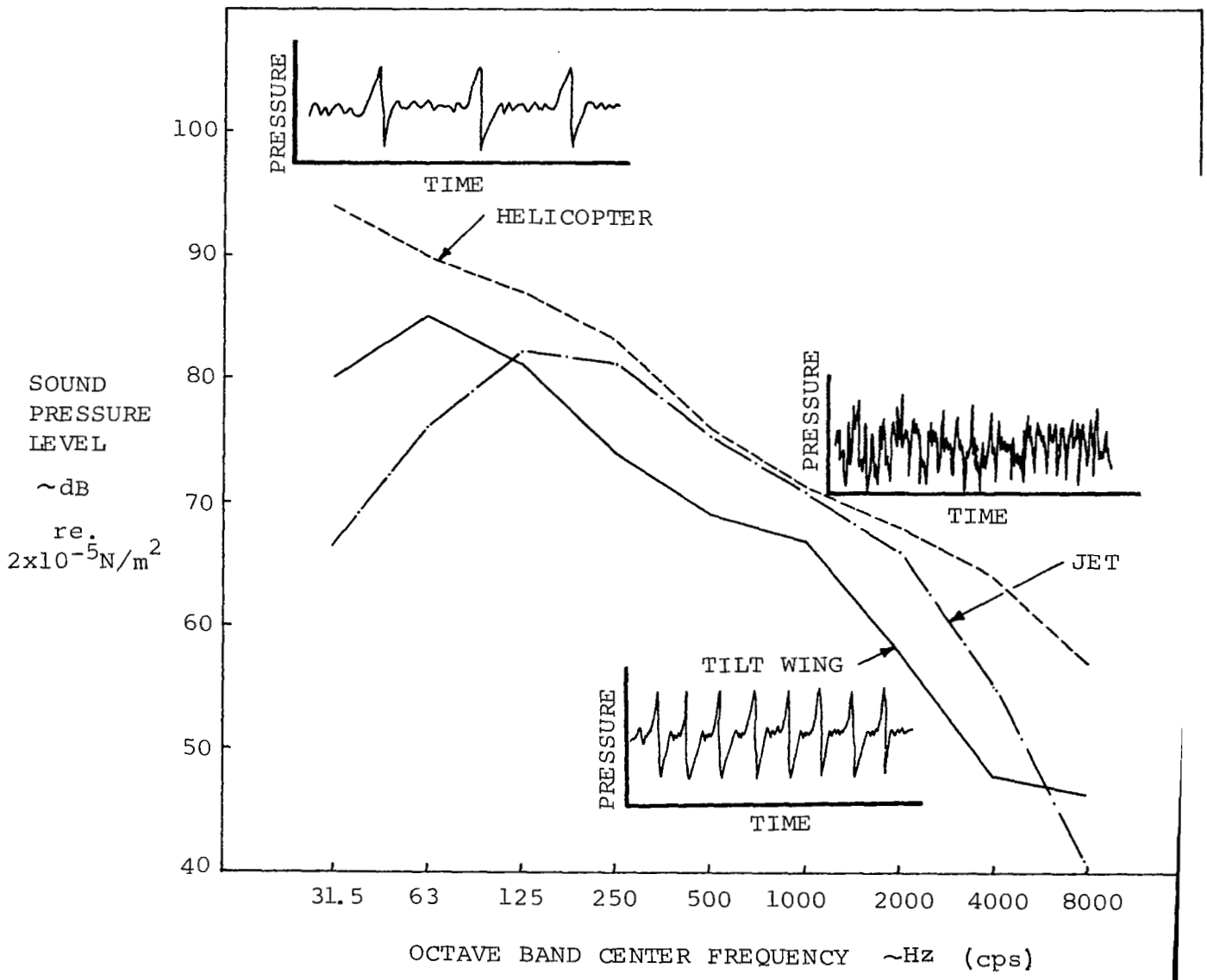
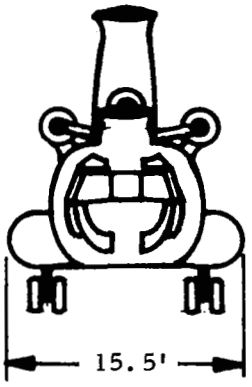
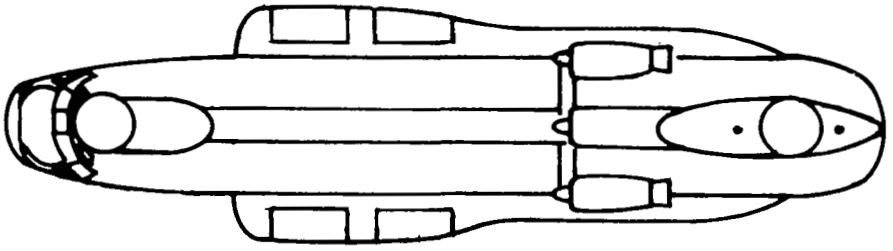


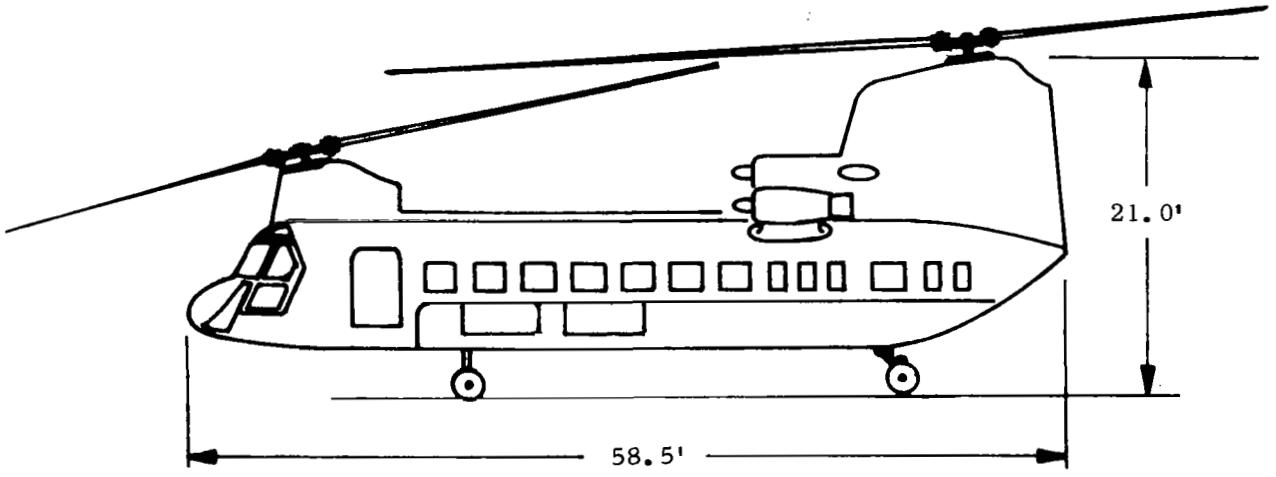
Figure 6 . Electronically Synthesized VTOL Aircraft Noise Presented to Test Subjects



FRONT VIEW



TOP VIEW



SIDE VIEW

Figure 7. Tandem Rotor Configuration

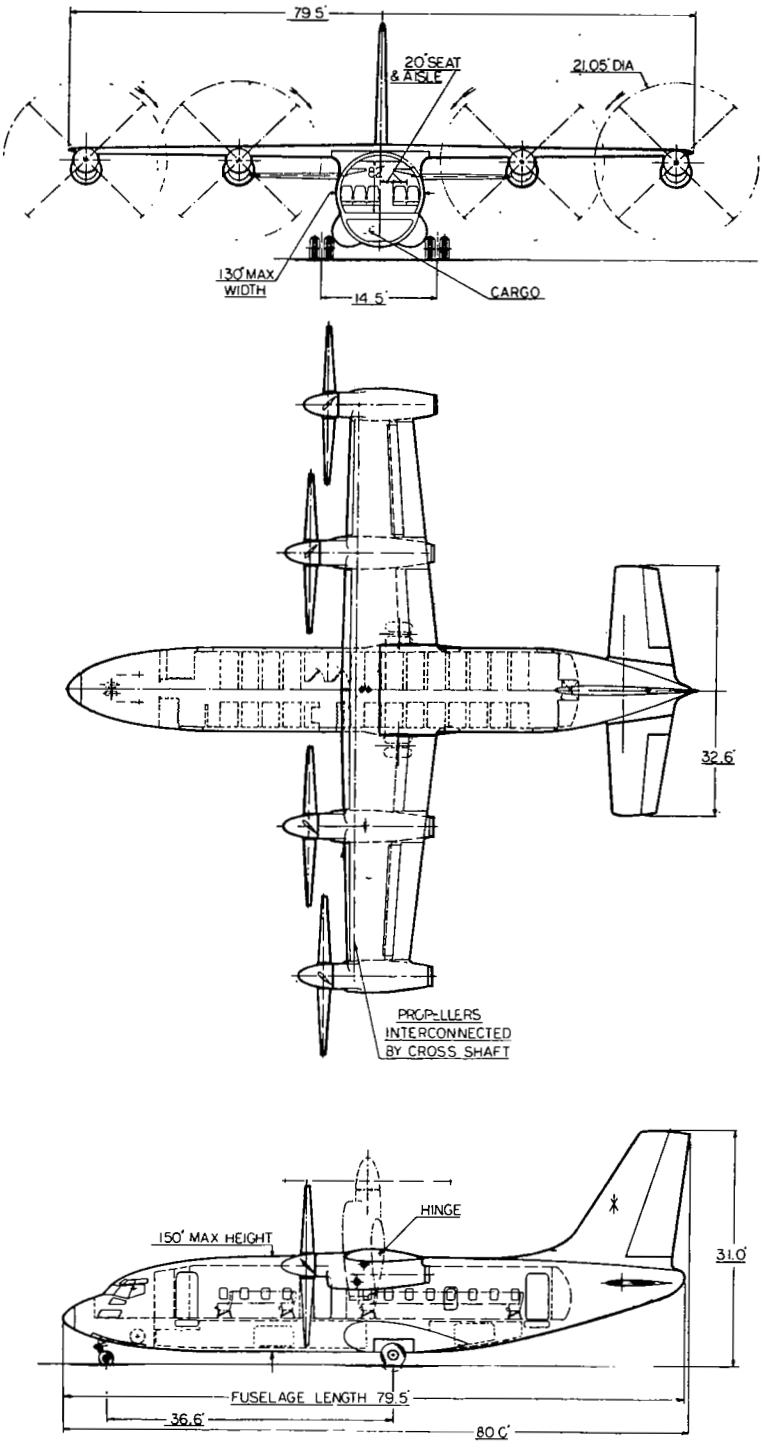


Figure 8. Tilt Wing Configuration

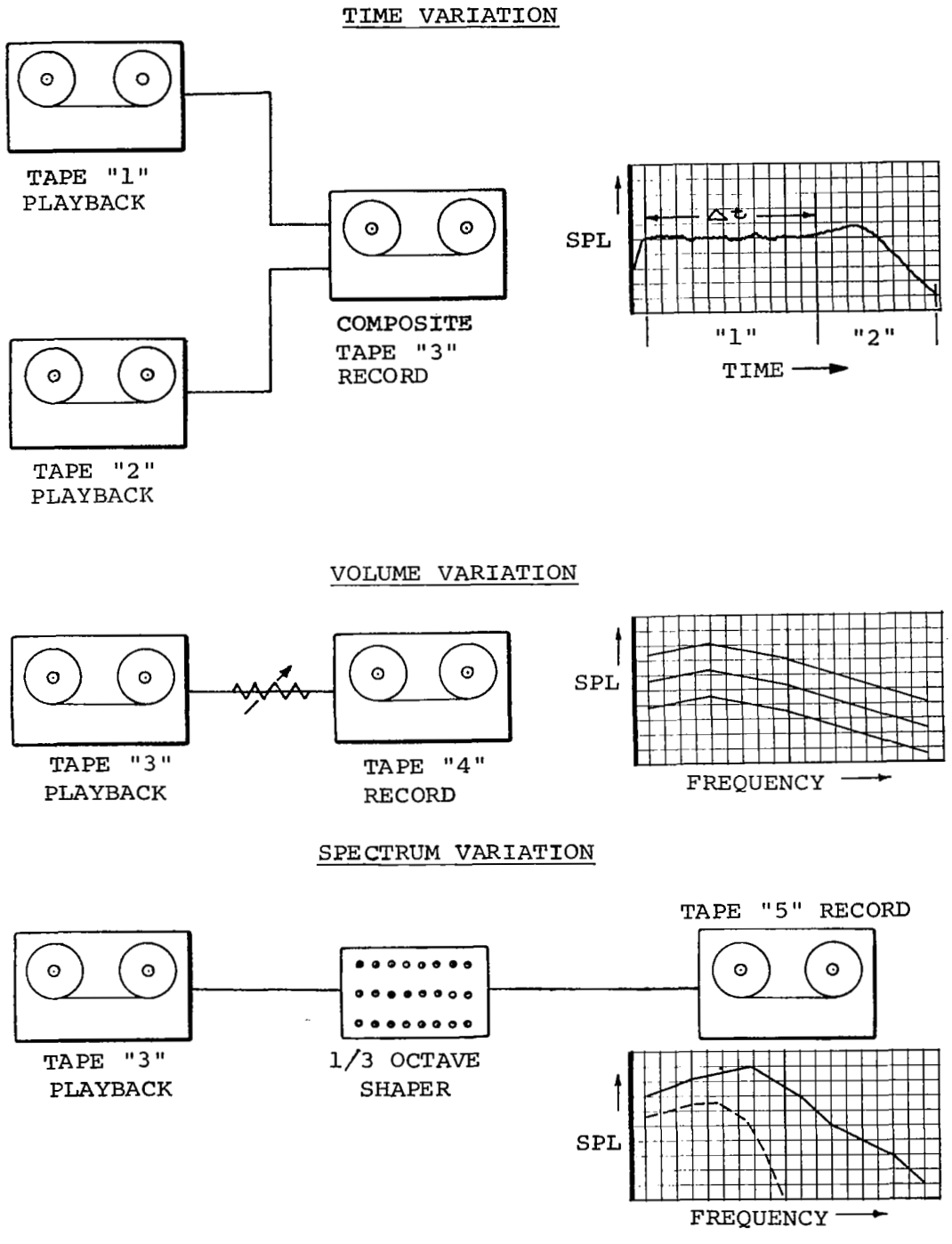


Figure 9. Methods of Tape Preparation

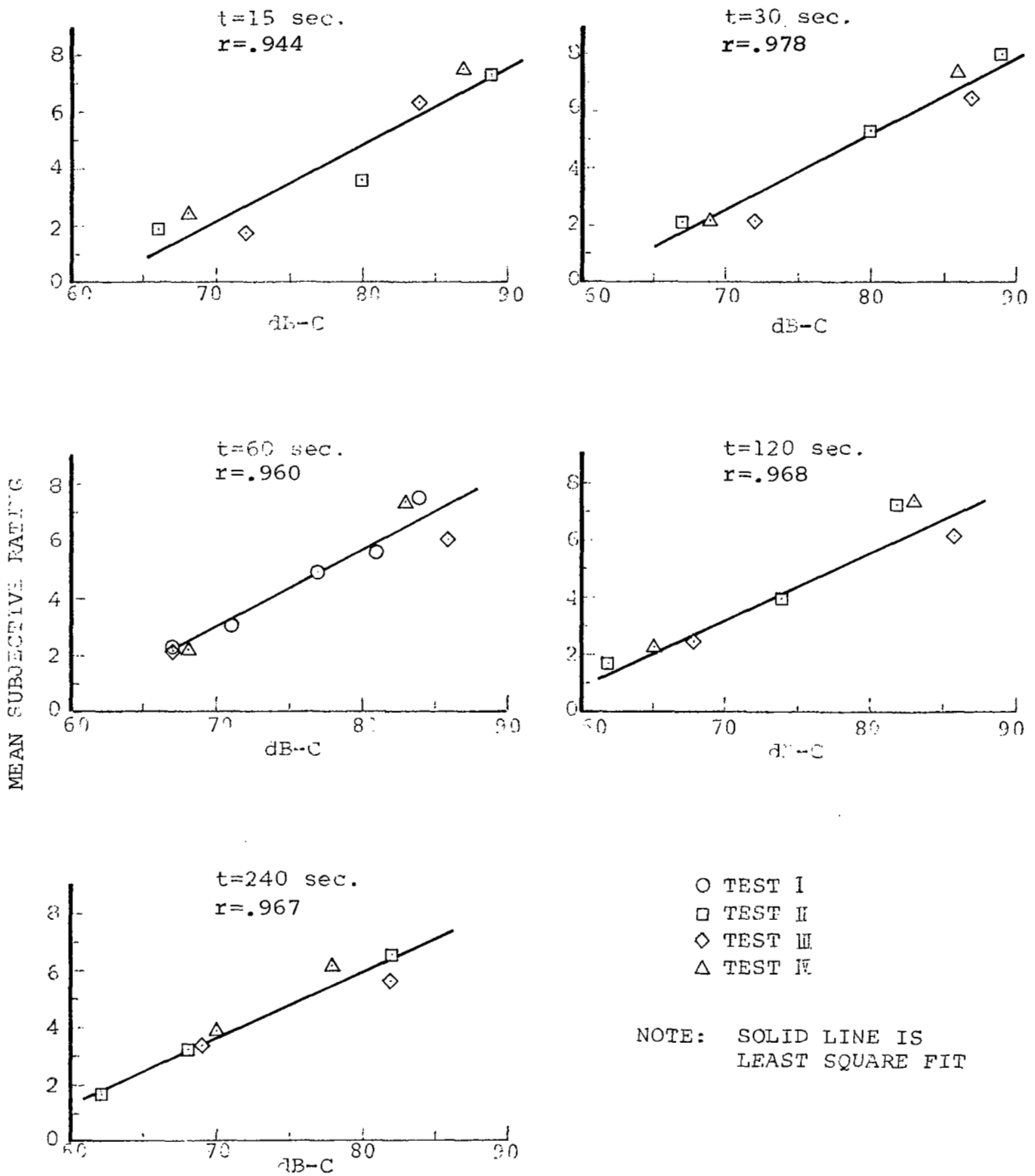
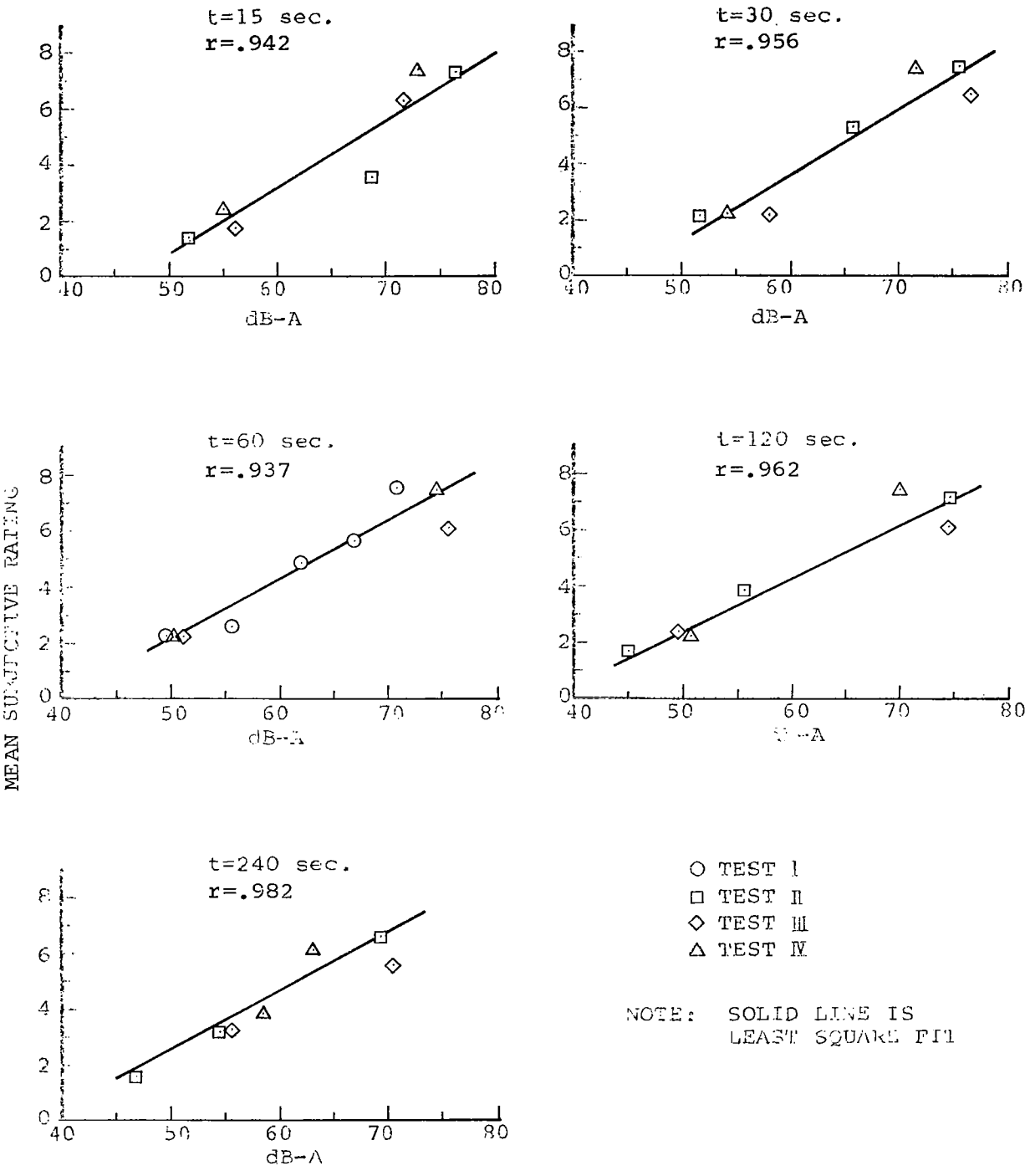


Figure10.Correlation of Subjective Ratings With Helicopter dB-C Levels



Figurell.Correlation of Subjective Ratings With Helicopters dB-A Levels

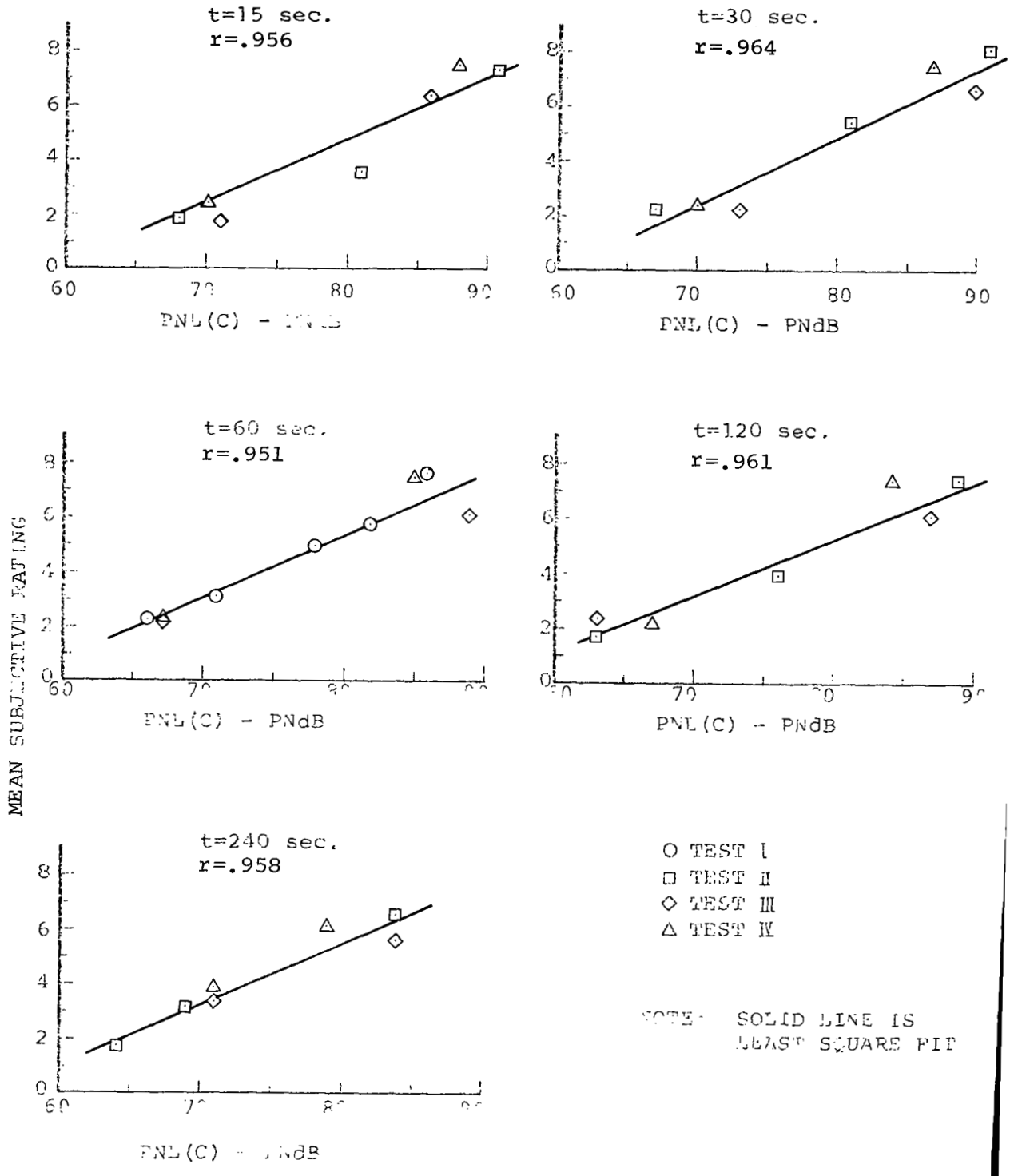


Figure 12. Correlation of Subjective Ratings With Helicopter PNdB Levels

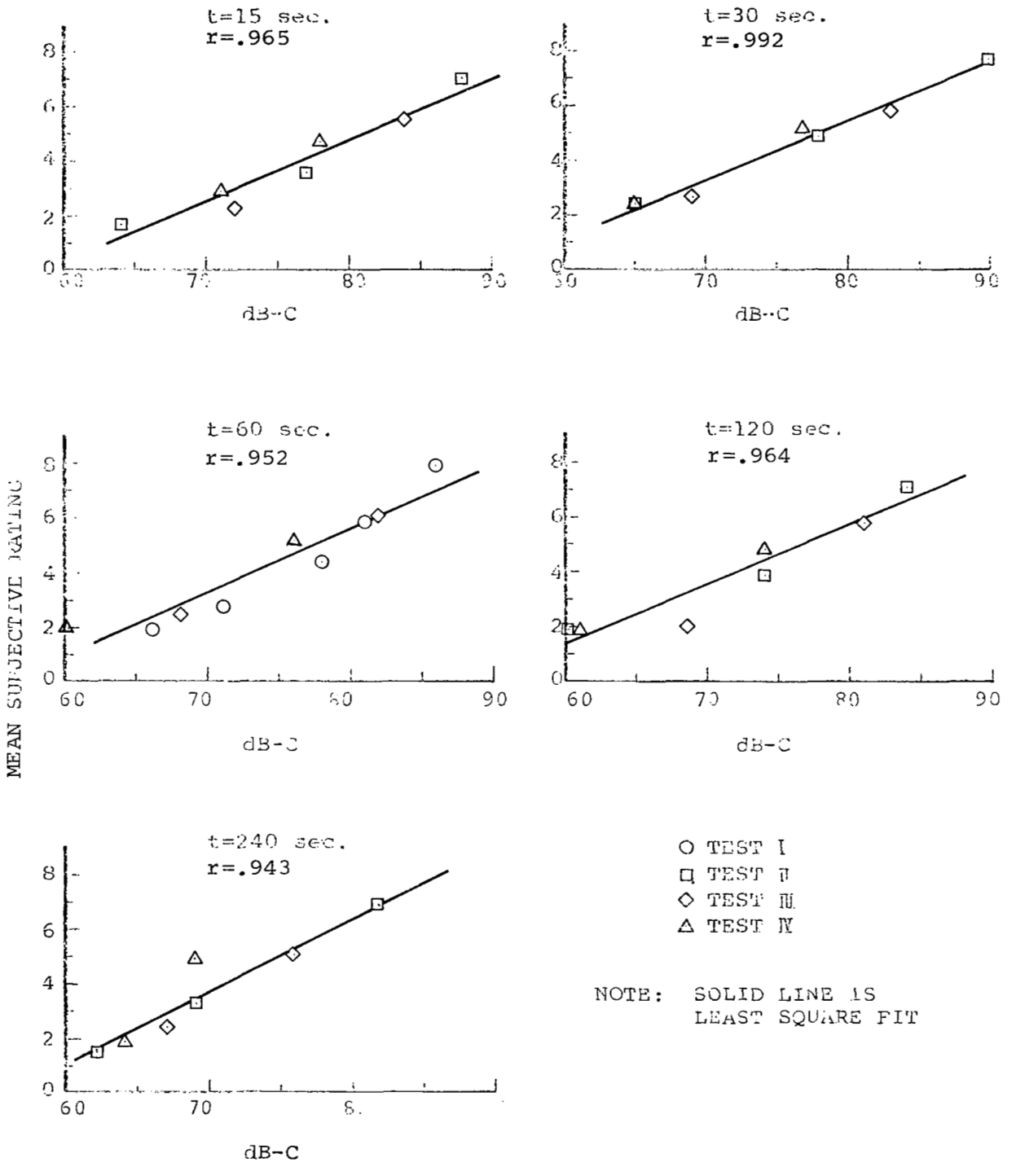


Figure 13 Correlation of Subjective Ratings With Tilt Wing dB-C Levels

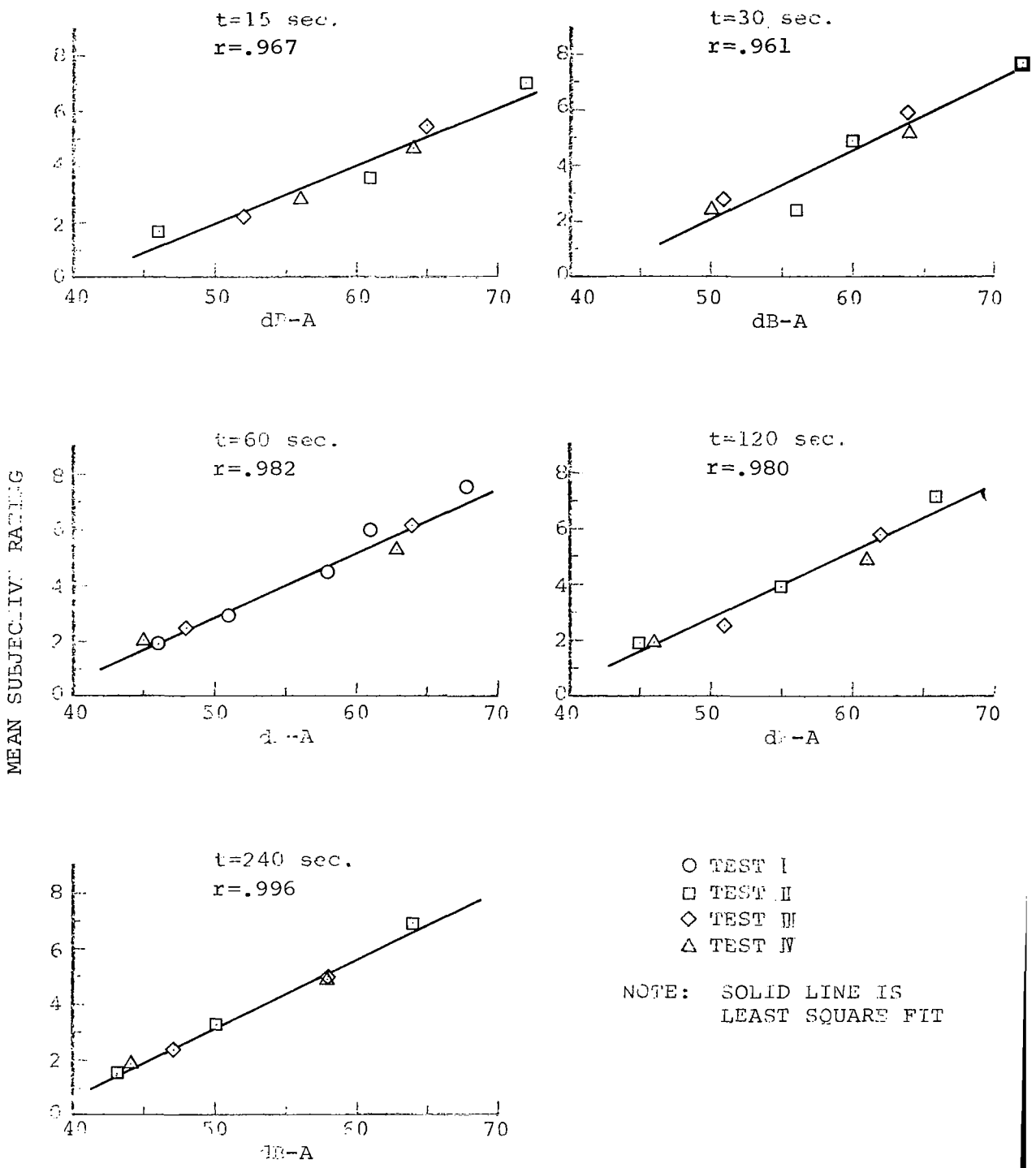


Figure14.Correlation of Subjective Ratings With Tilt Wing dB-A Levels

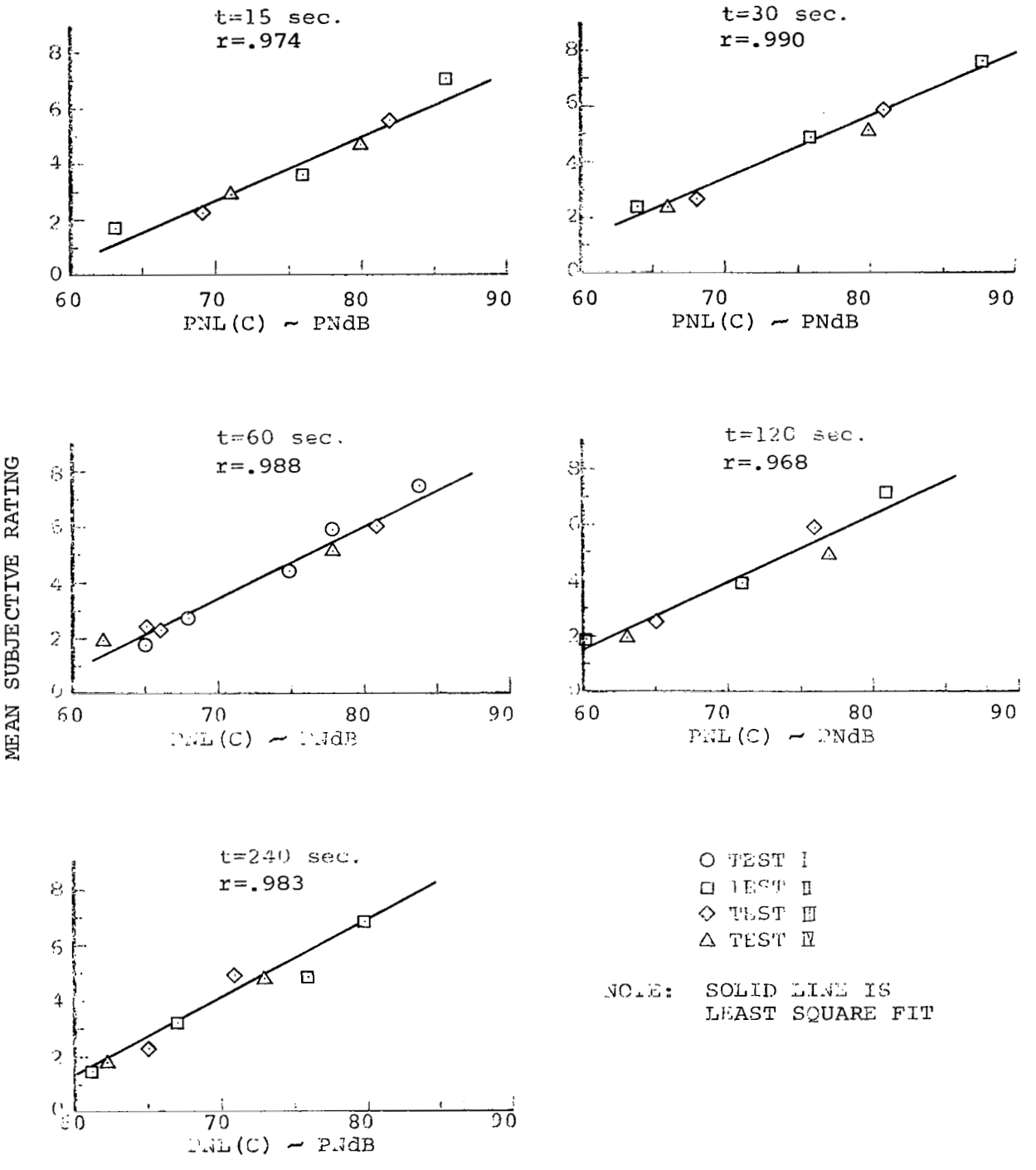


Figure15. Correlation of Subjective Ratings With Tilt Wing PNdB Levels

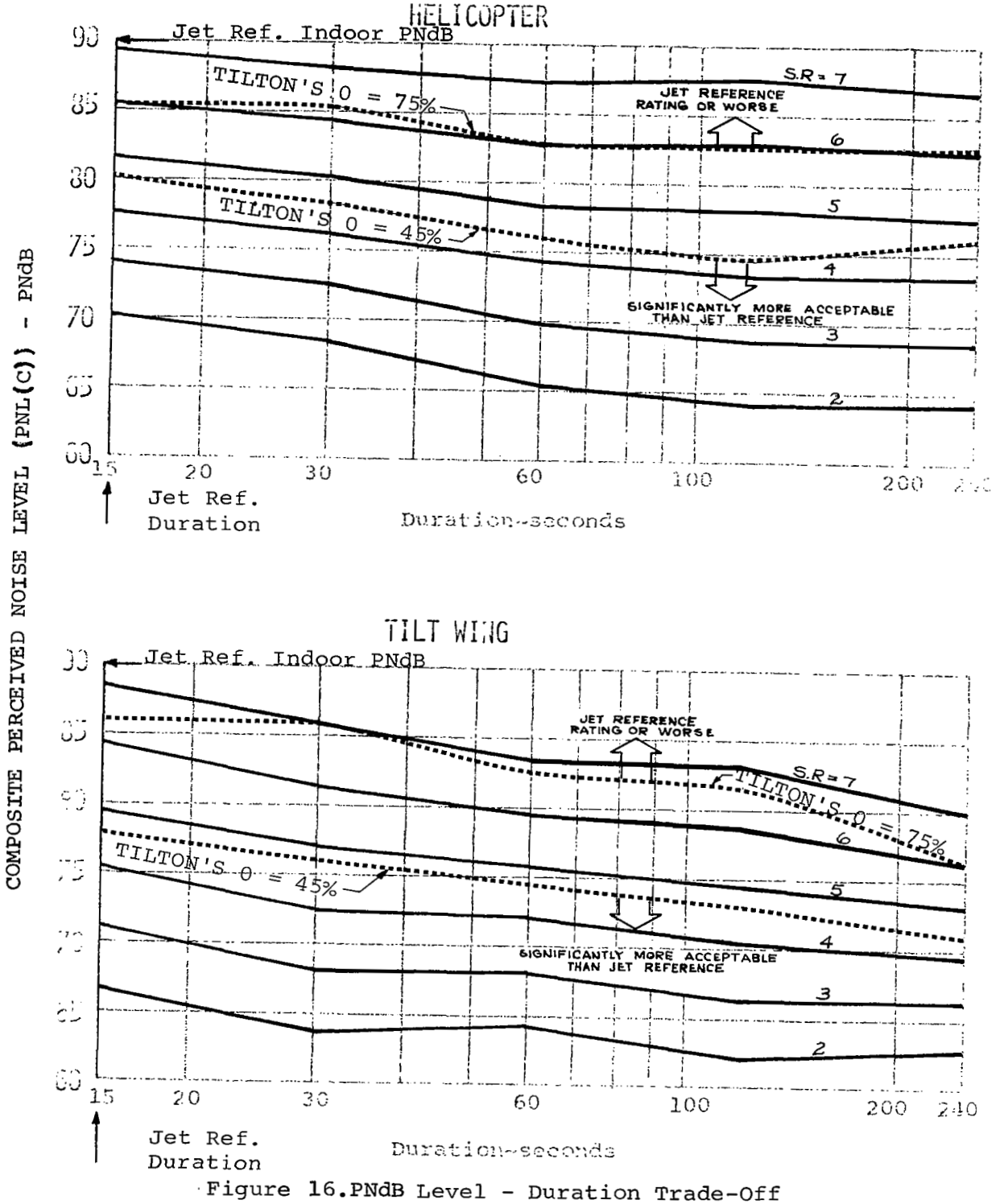


Figure 16.PNdB Level - Duration Trade-Off

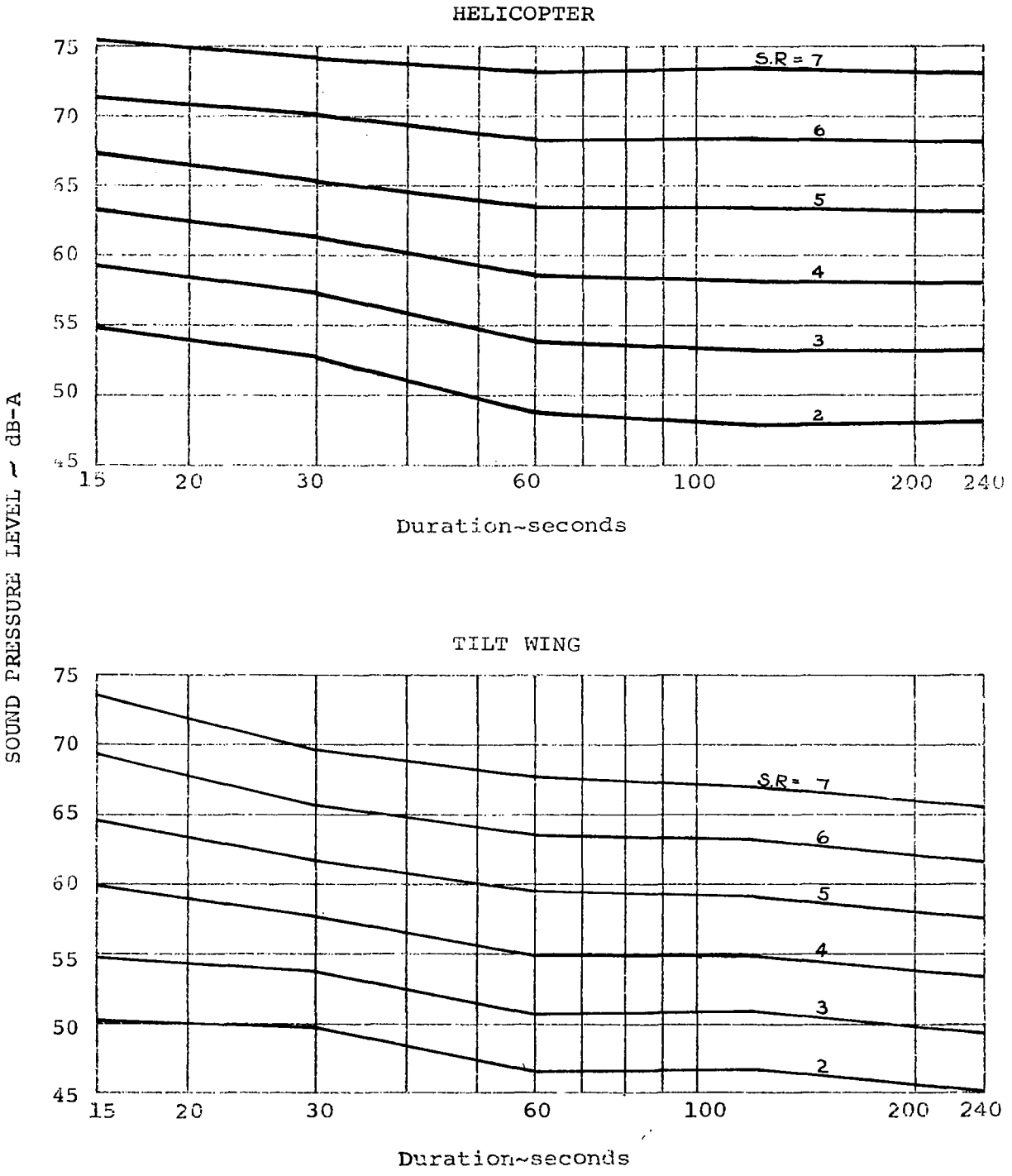


Figure 17. dBA Level - Duration Trade-Off

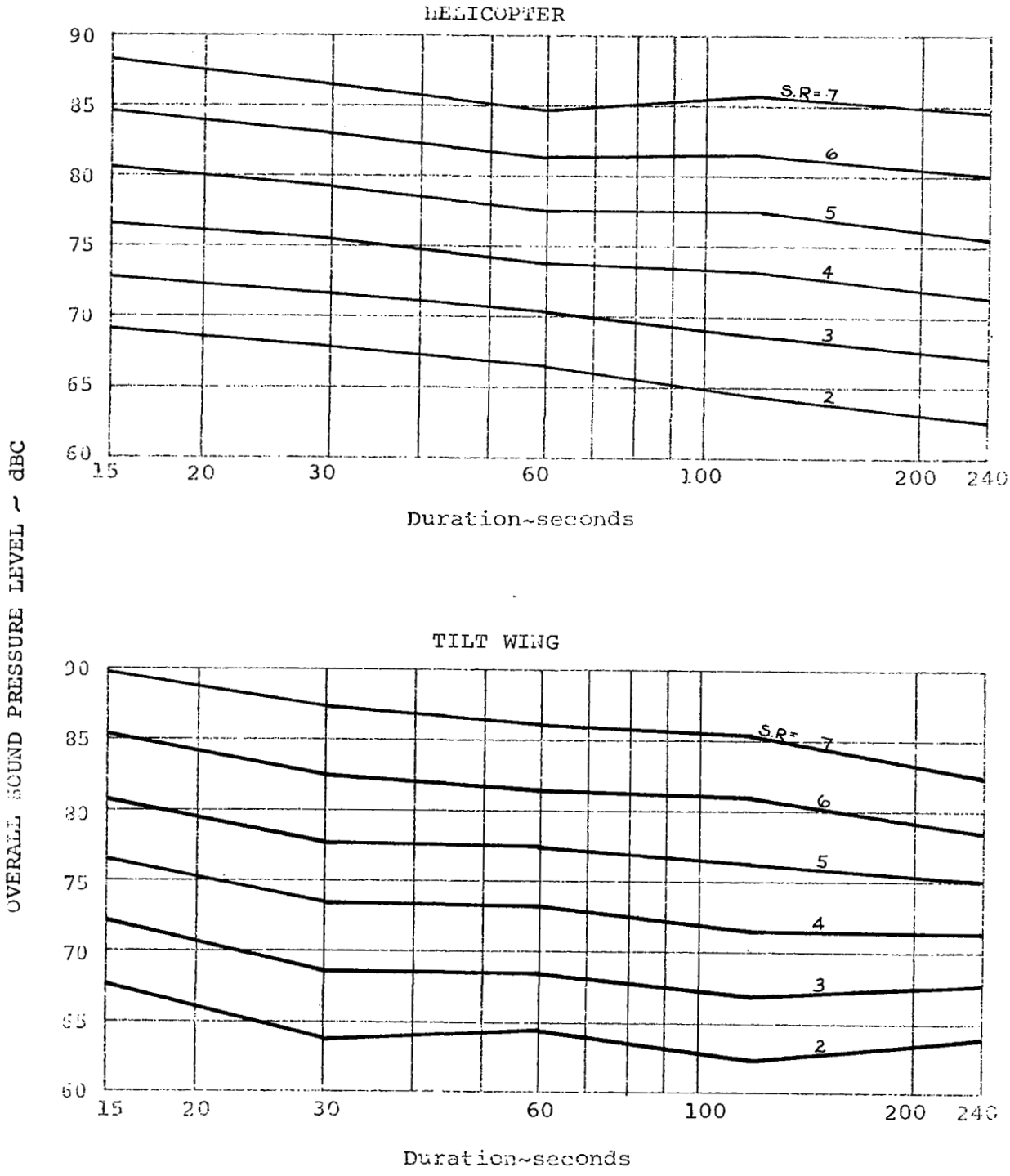
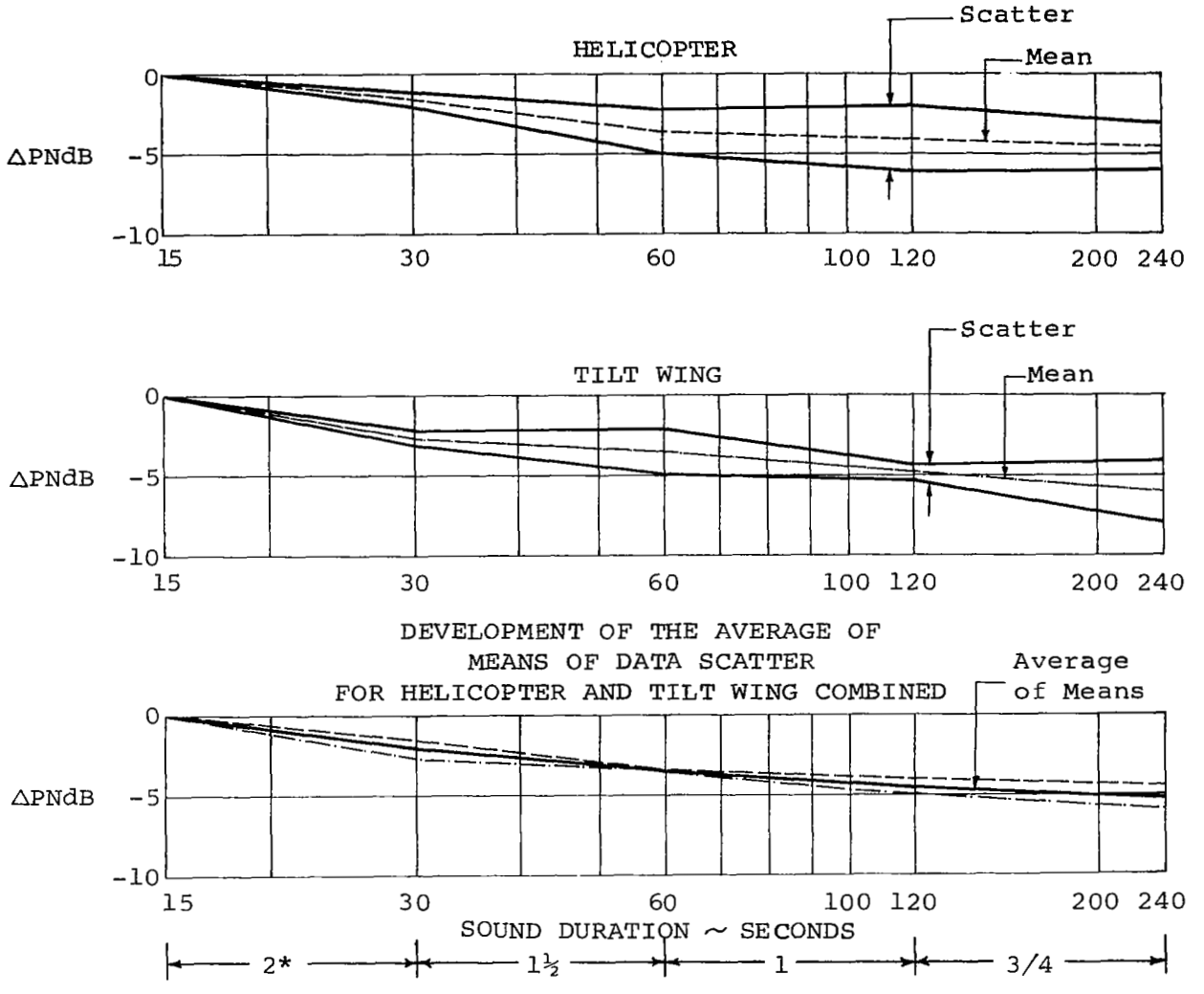


Figure 18. dBC Level - Duration Trade-Off



* THESE NUMBERS DENOTE THE CONSTANT Δ PNdB PER DOUBLING OF DURATION INTERVAL

Figure 19. Sound Level-Duration Trade-Off for Helicopter And Tilt Wing Sounds From Normalized Data of Figure 16.

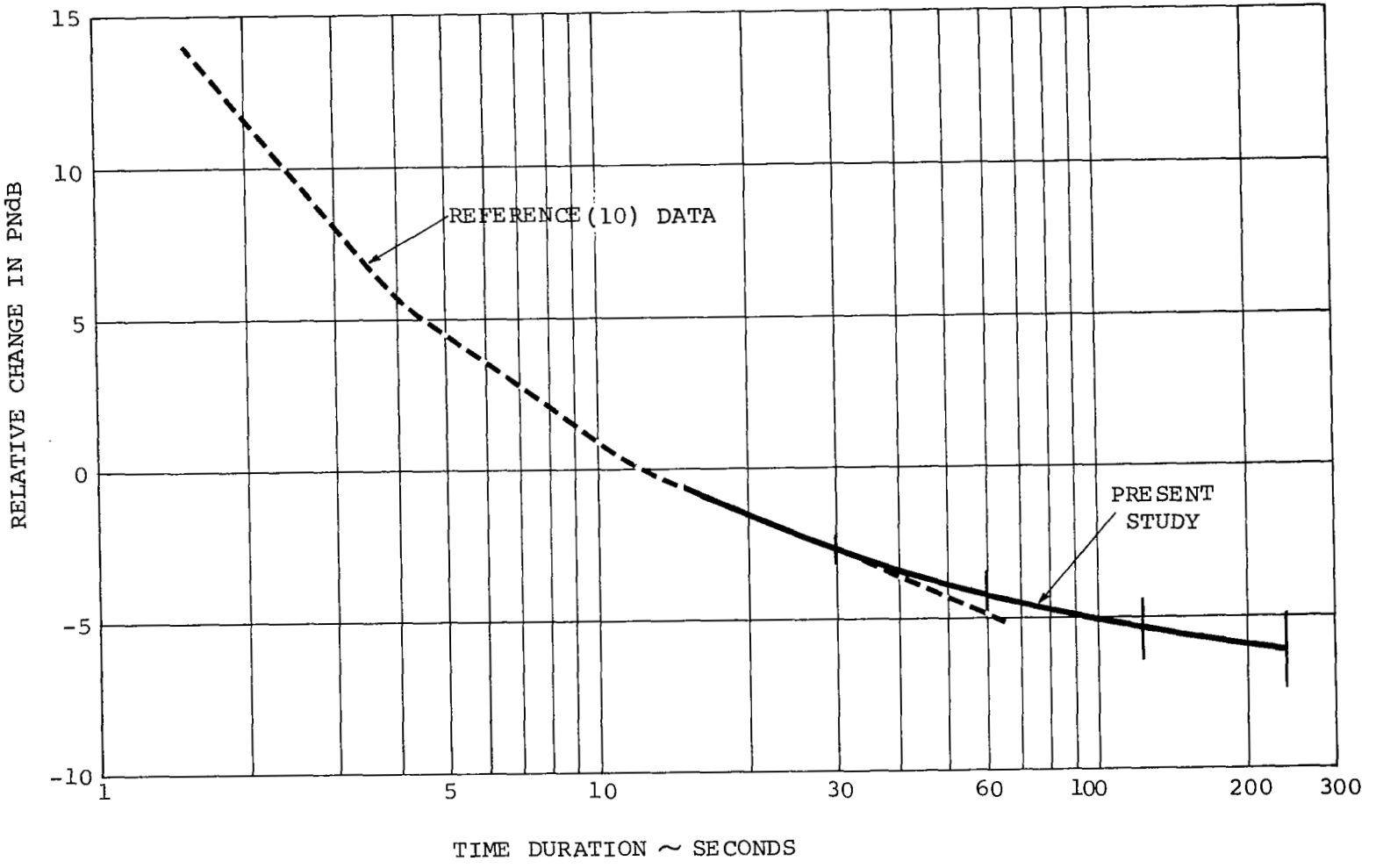
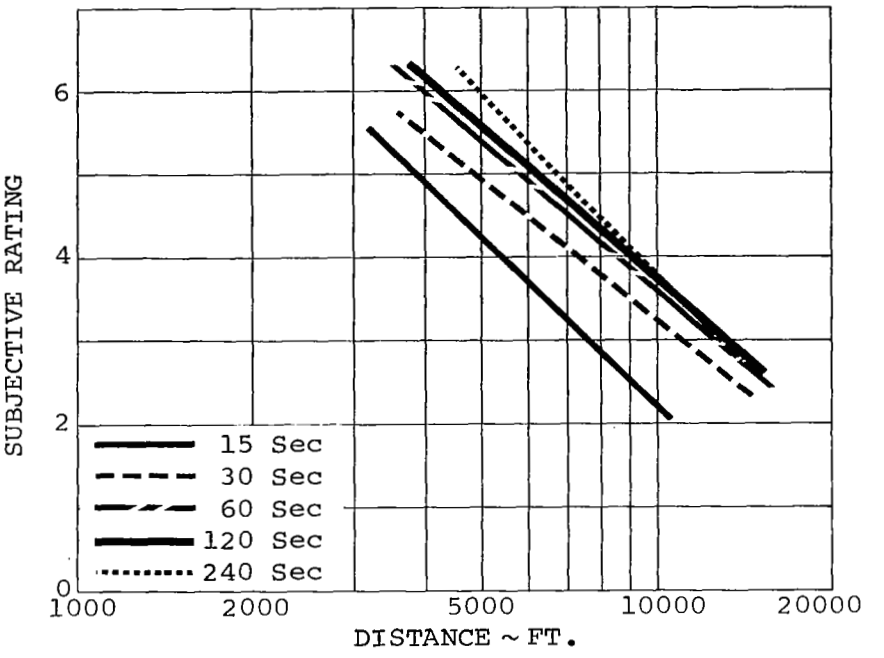


Figure 20. Relative Effect of Changing the Duration of a Sound on its Subjectively Rated Annoyance (PNL).

Part a: TILT WING



Part b: HELICOPTER

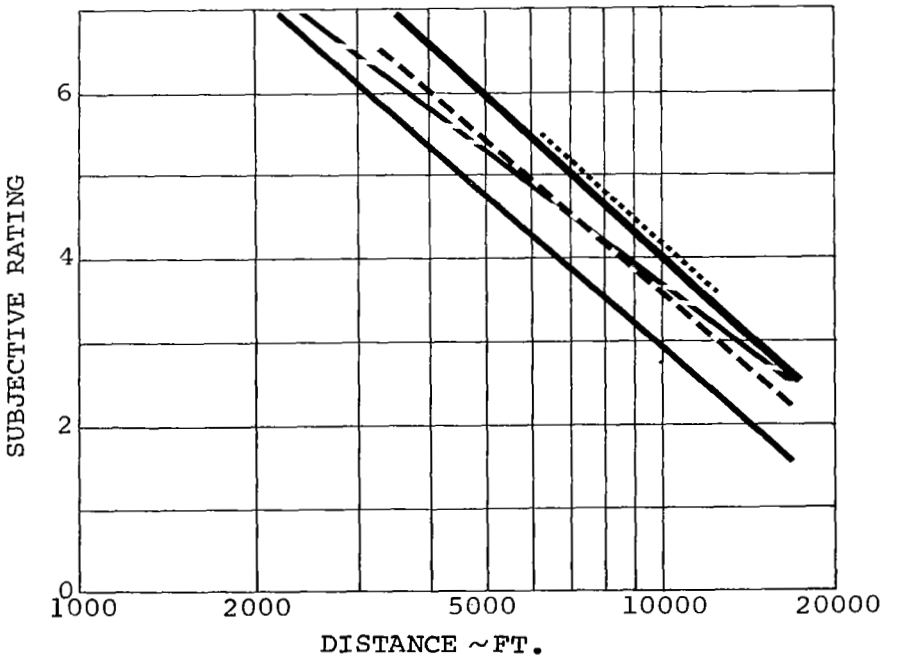
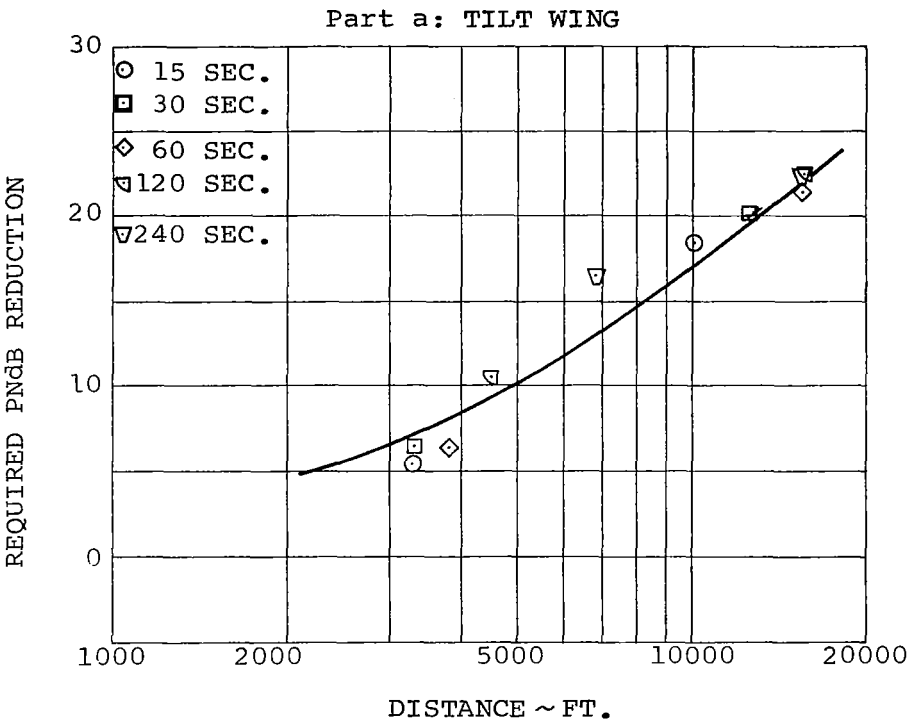


Figure 21. Test III Results: Effect of Distance



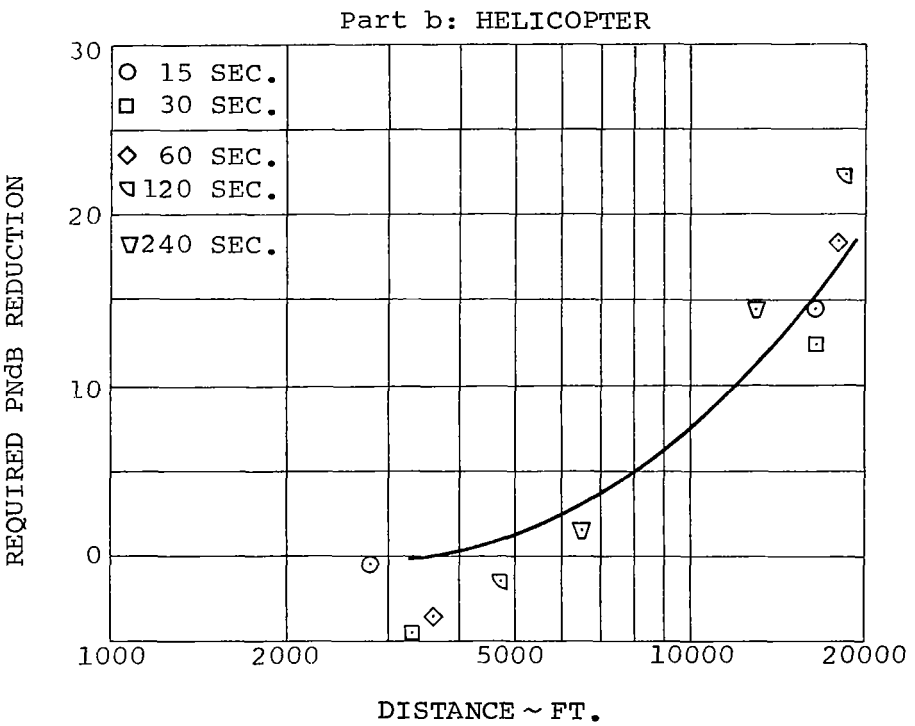


Figure 22. PNdB Reduction Trade-Off with Distance

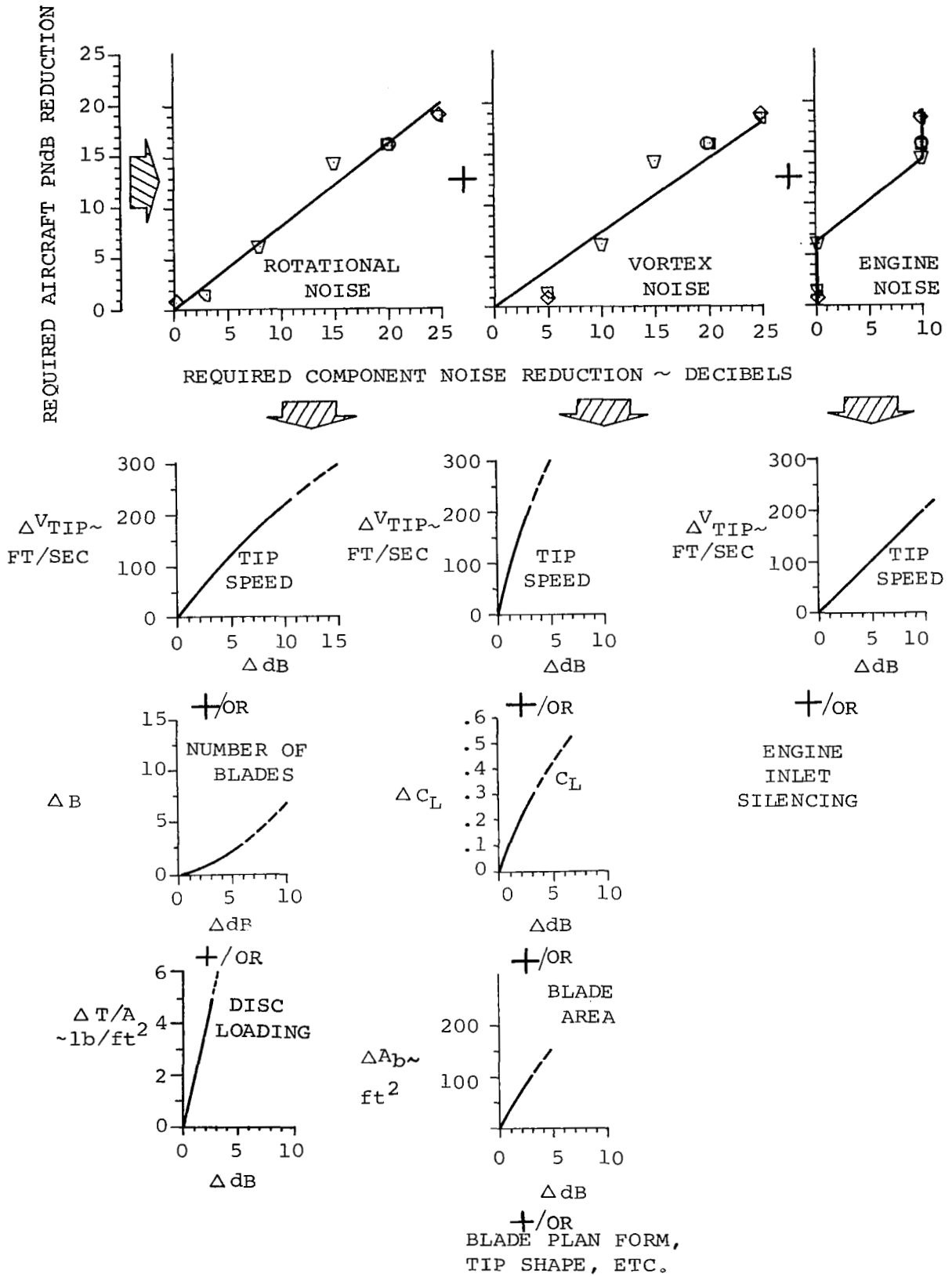


Figure 23. Helicopter Noise - Design Parameters Trade-Off

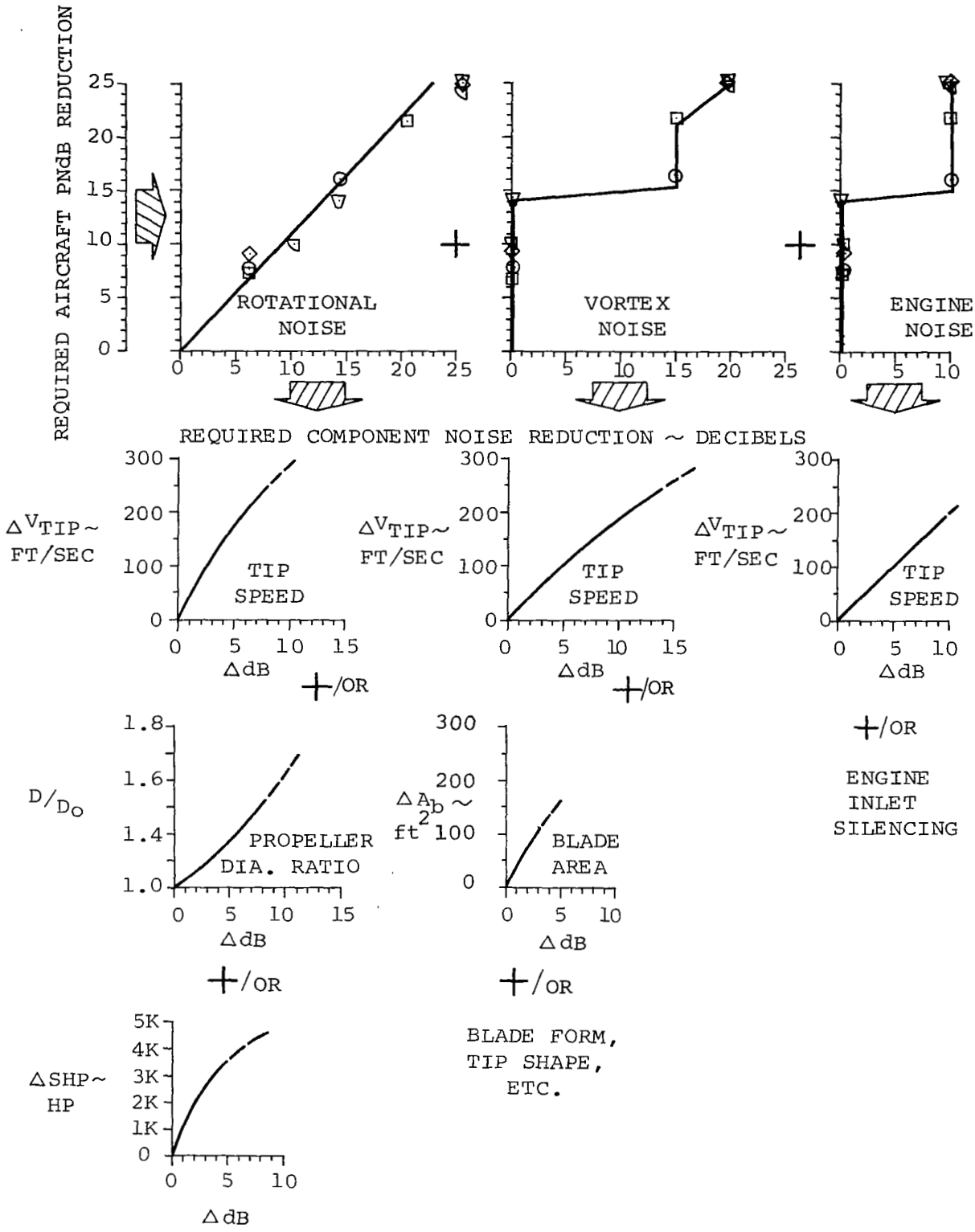


Figure 24. Tilt Wing Noise - Design Parameters Trade-Off

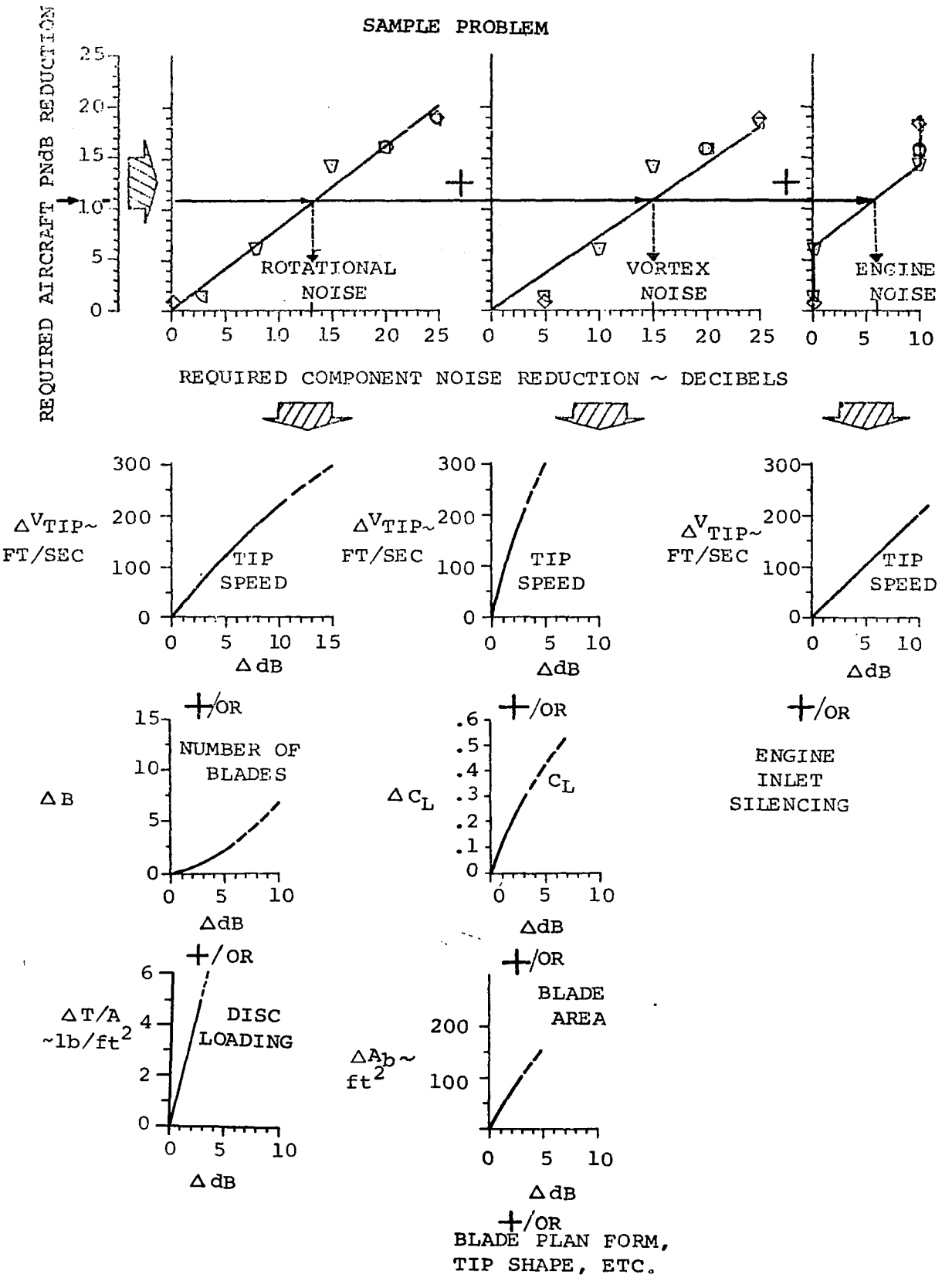


Figure 25. Sample Problem: Helicopter Noise - Design Parameters Trade-Off

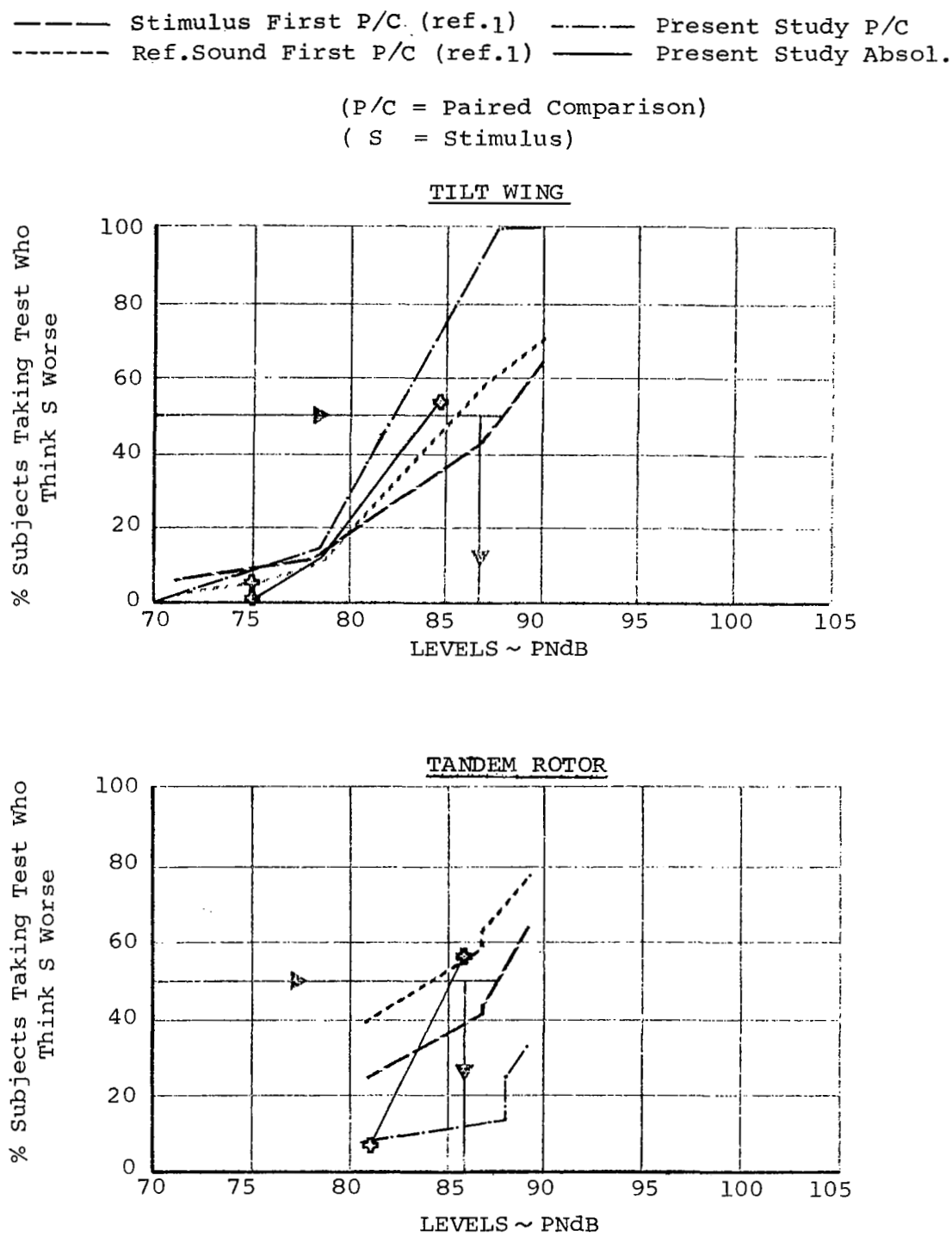


Figure 26. Paired Comparison Test Results

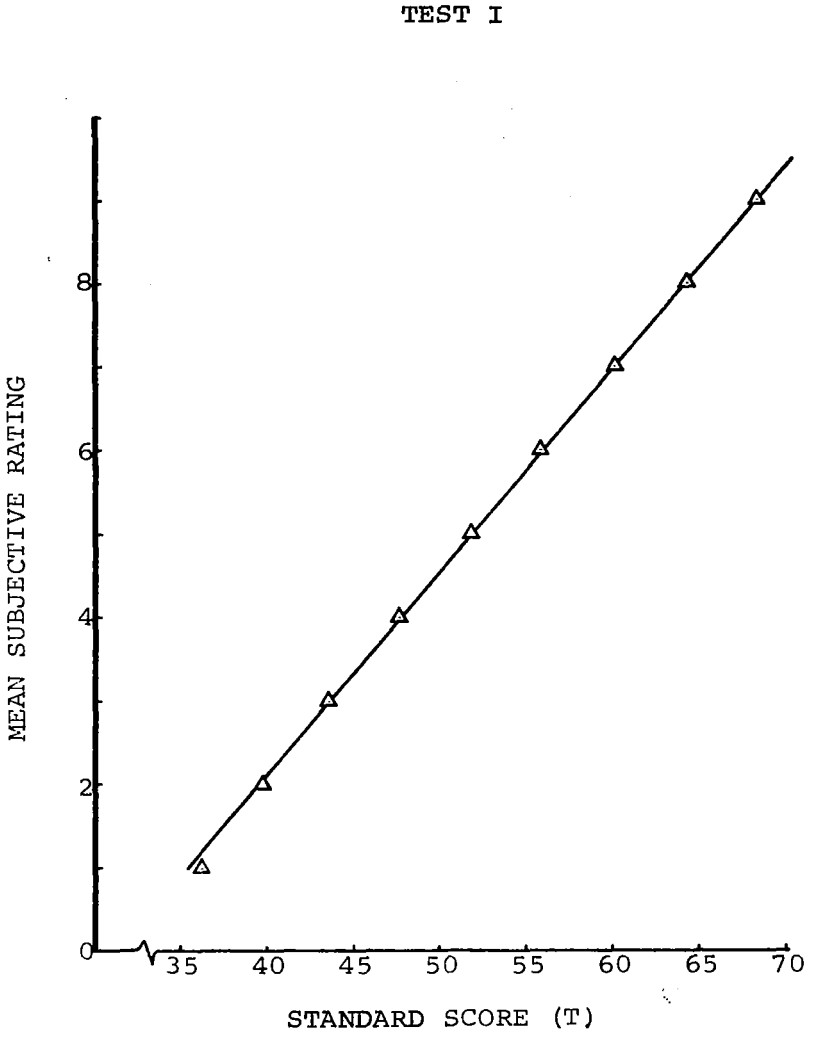


Figure 27. Psychological Scaling of Subjective Responses

TEST I

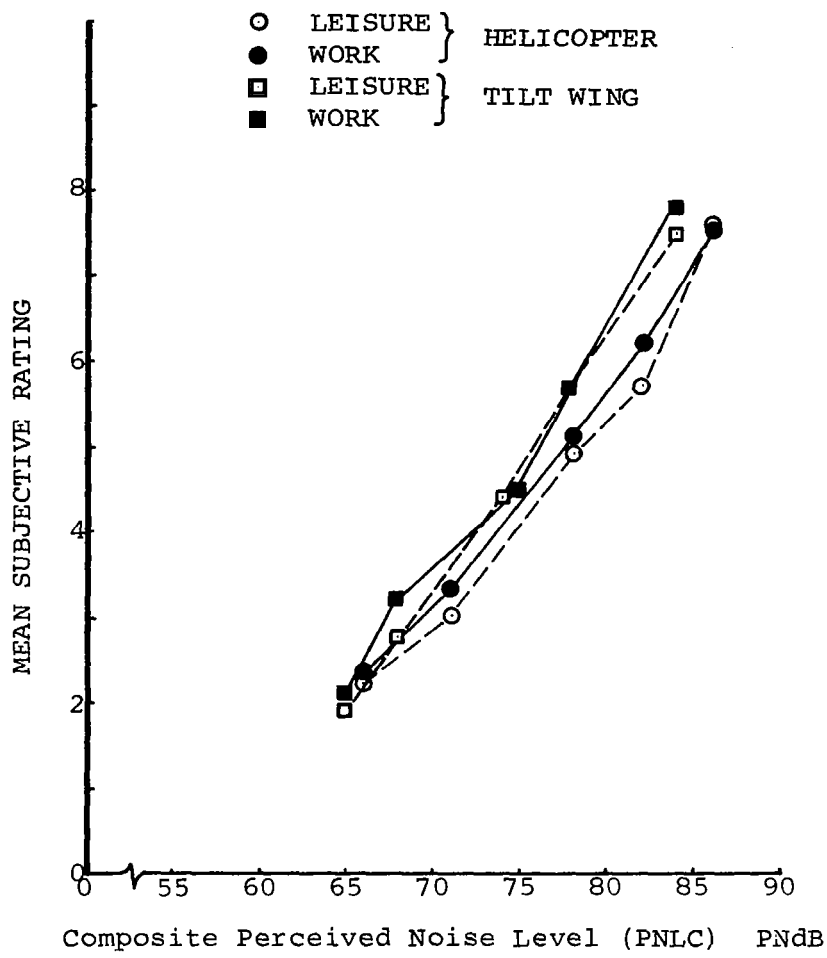


Figure28.Psychoacoustical Scaling of Subjective Responses

INDIVIDUAL DIFFERENCES

TEST I

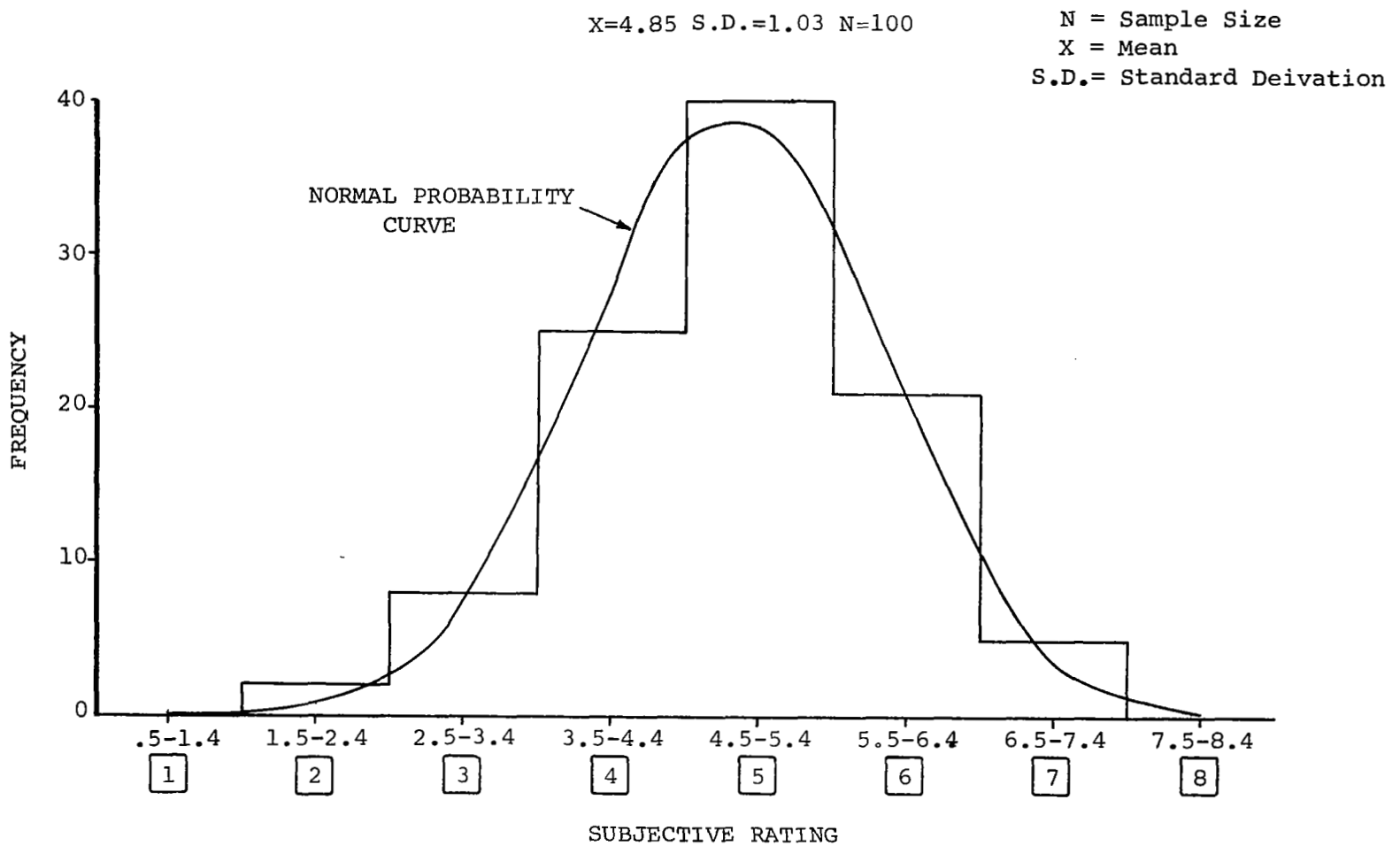


Figure 29. Distribution of Means of Subjective Responses of VTOL Noise

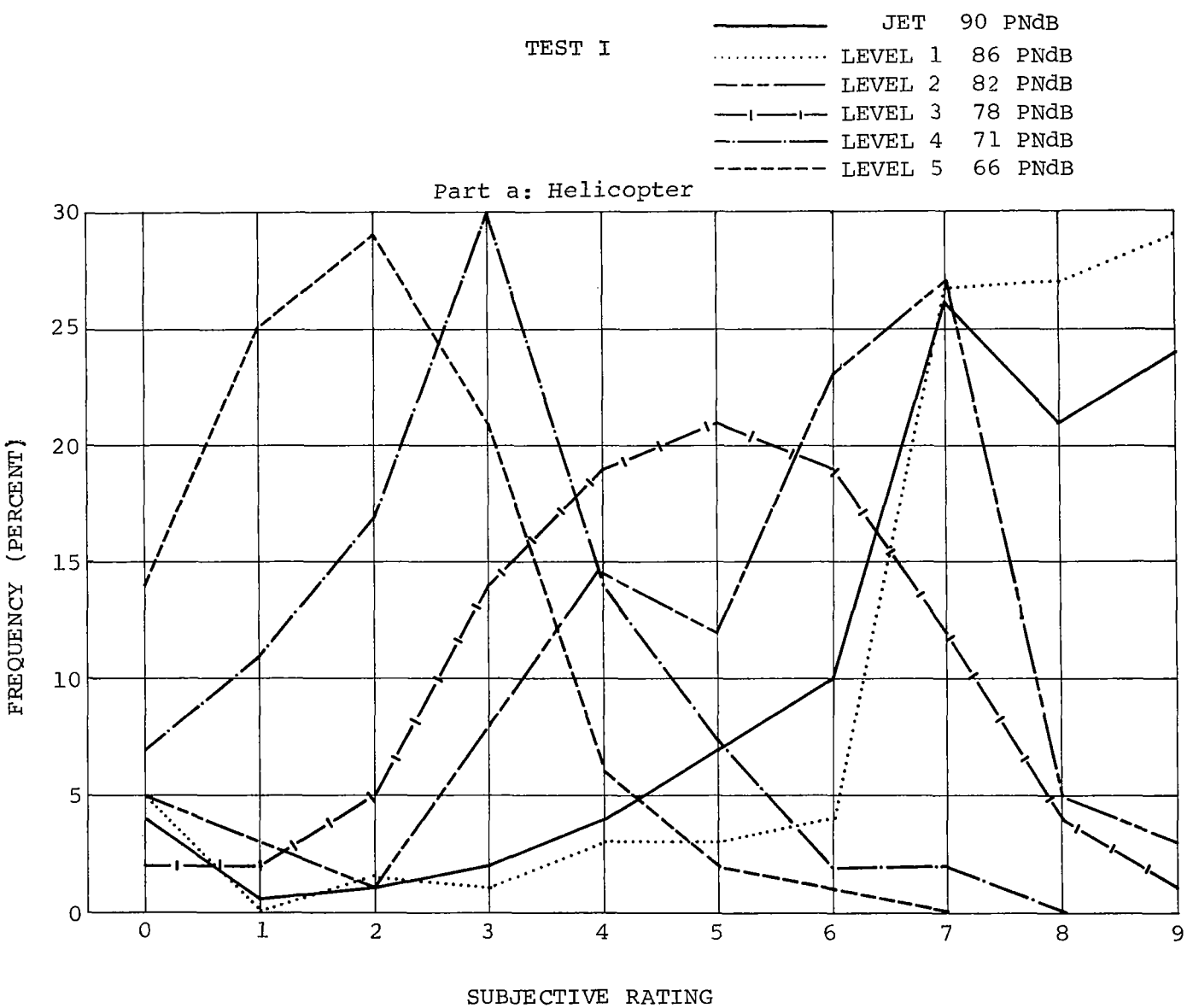


Figure 30. Distribution of Subjective Ratings for VTOL Aircraft Sounds

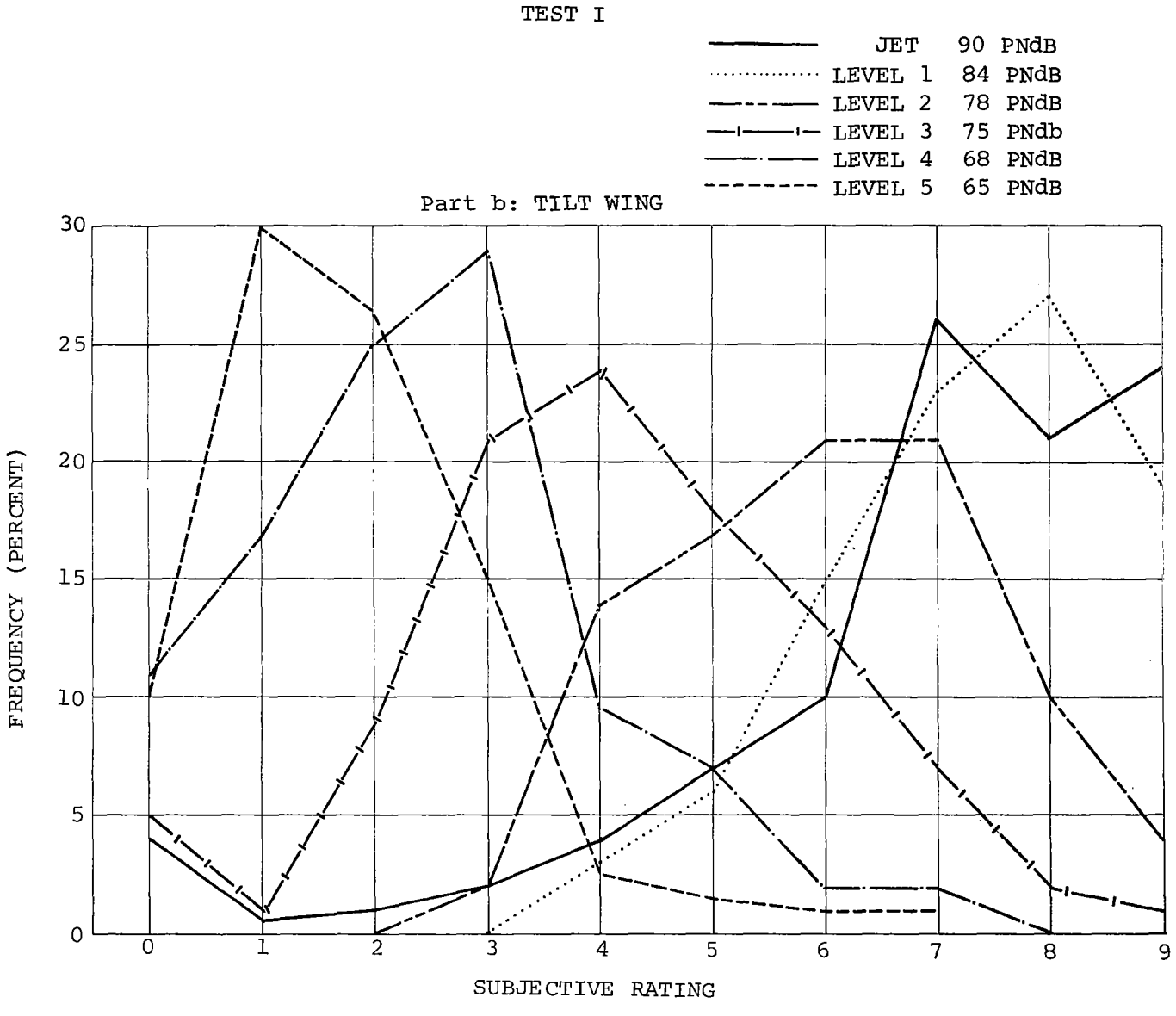


Figure 30. - Concluded

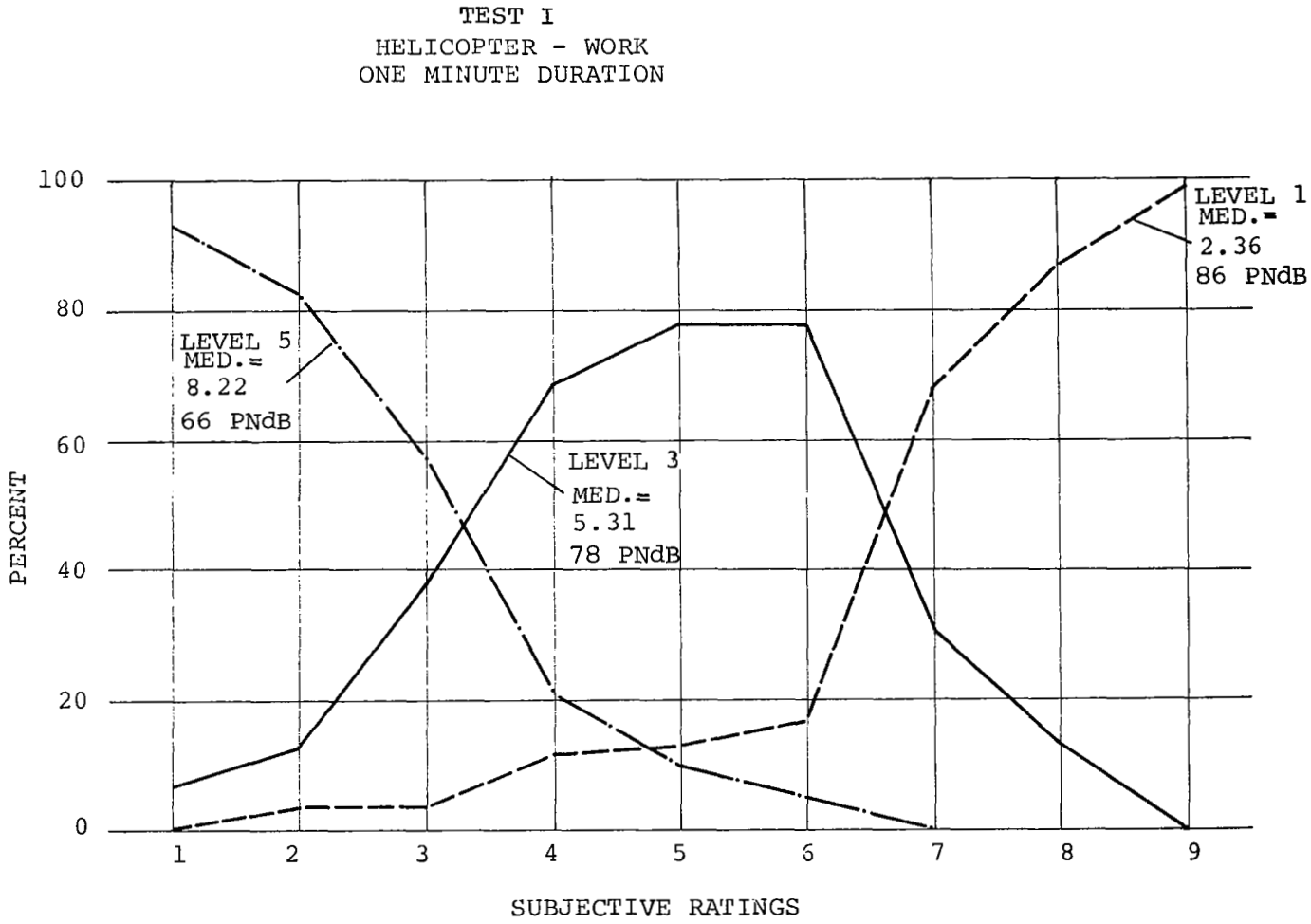


Figure31.Distribution of Subjective Ratings for Three Levels