FAA NO-68-35

# SUBJECTIVE EVALUATION OF GENERAL AVIATION AIRCRAFT NOISE

TECHNICAL REPORT



APRIL 1968

By
WYLE LABORATORIES RESEARCH STAFF
Huntsville, Alabama 35807

Prepared For
DEPARTMENT OF TRANSPORTATION
FEDERAL AVIATION ADMINISTRATION
OFFICE OF NOISE ABATEMENT
Washington, D. C. 20590
Under Contract FA67WA-1731





107

# SUBJECTIVE EVALUATION OF GENERAL AVIATION AIRCRAFT NOISE

TECHNICAL REPORT
FAA NO-68-35

April 1968

Prepared by
J. B. Ollerhead

WYLE LABORATORIES RESEARCH STAFF
7800 Governors Drive, W.
Huntsville, Alabama 35807

for

DEPARTMENT OF TRANSPORTATION
FEDERAL AVIATION ADMINISTRATION
OFFICE OF NOISE ABATEMENT

Under Contract FA67WA-1731

## ABSTRACT

An experiment was conducted, in a progressive wave field, in which a group of subjects evaluated the noisiness of recorded general aviation aircraft sounds in comparison to jet transport flyover noise. The aircraft studied were piston engined, turboprop and turbojet aircraft in the weight range 2000 to 13,000 lbs. Twenty-eight noise rating scales were evaluated and it was found that of currently used scales, Perceived Noise Level, corrected for pure tone content, gave the best correlation with the subjective results. For the signals studied, duration appeared to have little influence on the subjective noisiness of flyover sounds. By comparing the results for a number of simulated flyover sounds it was concluded that an explanation lies in the influence of the Doppler frequency shift, which tends to cancel the effects of duration. A Doppler correction for current noise rating methods is presented which generally improves their correlation with observed results.

## TABLE OF CONTENTS

## LIST OF FIGURES

LIST OF FIGURES (Continued)

BLANK PAGE

## LIST OF TABLES

LIST OF TABLES (Continued)

## GLOSSARY

Some terms appearing in this report are listed and defined here for convenience. These and other symbols which are used only locally are defined where they first occur in most cases.

| Symbol/term | Definition |
|---|---|
| $a_o$ | Atmospheric speed of sound (ft/sec) |
| b | Slope of regression line |
| $f_o$ | Observed frequency (Hz) |
| $\dot{f}_o$ | Observed rate of change of frequency (Hz/sec) |
| $f_s$ | Source Frequency (Hz) |
| IAS | Indicated airspeed |
| $L_o$ | Mean zero crossing point (dB) |
| M | Source Mach No. |
| $\bar{r}$ | Correlation coefficient |
| $R_c$ | Correlation coefficient |
| $R_o$ | Minimum source to observer distance (ft) |
| $S_{x/y}$ | Mean standard error about regression line |
| $S^2$ | Variance of data about regression line |
| t | Time (seconds) |
| TAS | True airspeed |
| $T_{10}$ | "10 dB-down duration" (seconds) |
| V | Source velocity (ft/sec) |
| $\Delta$ | Increment to rating scale |

GLOSSARY (Continued)

| Symbol/Term | Definition |
|---|---|
| $\theta$ | Angle between source velocity vector and source/observer line |
| $\sigma(L_o)$ | Standard deviation of zero crossing points |

Abbreviated notations for noise rating scales

| | |
|---|---|
| PNdB | Perceived Noise Level |
| PNdBF | Perceived Noise Level corrected for pure tone content |
| S-Phon | Stevens' Phons, MK VI |
| Z-Phon | Zwickers' Phons |
| OASPL | Overall Sound Pressure Level |
| dB(A) | A-Weighted Sound Pressure Level |
| dB(B) | B-Weighted Sound Pressure Level |
| dB(C) | C-Weighted Sound Pressure Level |
| dB(J), dB(K) | Weighted Sound Pressure Level from Reference 11. |
| dB(N) | N-Weighted Sound Pressure Level |
| dB(NN) | Weighted Sound Pressure Level based on the results of this study. |

Subscripts Used With Noise Rating Scales

| | |
|---|---|
| D | Duration corrected [correction = $10 \log_{10} (T_{10}/15)$] |
| DD | Duration and Doppler corrected (Using preliminary Doppler correction $10 \log_{10} (0.1\, R_o/T^2_{10})$ ) |
| R | Range corrected [correction = $10 \log_{10} (R_o/2200)$ and is equivalent to combination of duration and Doppler corrections.] |

# I. INTRODUCTION

Predicting likely community response to noise intrusion is becoming an increasingly important factor in many areas of industrial design and town planning. It is a statistical problem since individuals differ widely in their attitudes toward noise. Factors of known importance which affect community reaction to a particular noise intrusion include location, background noise, temporal and seasonal factors, previous exposure to noise, and general community attitude (which varies between sociologically different groups of people). Consequently, considerable effort has been directed at the development of comprehensive methods for the prediction of community reaction to noise which take these factors into account. The fundamental procedure in these methods is to estimate the subjective magnitude of the noise intrusion itself, whether it be defined as loudness, noisiness, unwantedness or any other descriptor, and to modify this estimate by the likely effects of the factors described above.

Aircraft noise has been an increasingly serious community problem for many years, especially since the advent of the jet transport airplane. Now, on the threshold of a new era in aerial transportation, we are faced with the prospect of a new wave of community dissatisfaction with the noise of "Jumbo Jets" and SST's. Of immediate concern is the need to determine an acceptable scale for the rating of aircraft noise so that realistic limits can be specified during the primary stages of aircraft design. A concept which has been gaining popularity for some years is that of the Perceived Noise Level, initially proposed by Kryter (Reference 1) and stemming from earlier work by Stevens (Reference 2). This method was originally developed and validated for the sound of turbojet aircraft and calculates the noisiness or "unwantedness" rating from the energy spectrum of the sound. More recently, modifications have been proposed to take account of the duration and pure tone content of the sound, (for example References 3,4,5) being specifically addressed at the compressor whine components of turbojet, and in particular, turbofan engine noise.

The primary objective of the present study was to determine the applicability of the perceived noise level concept in rating the relative noisiness of sounds produced by representative types of general aviation aircraft. This category covers a wide range of aircraft types whose sound often has a very different character to that of larger commercial jet aircraft, so that the noise rating techniques which are satisfactory for the latter are not necessarily applicable. In general, it can be expected that general aviation aircraft, with orders-of-magnitude less installed power, will not constitute a serious noise problem in the vicinity of busy metropolitan airports. But it is important to realize that the smaller aircraft operate from many small airfields which do not serve heavy commercial traffic. Such fields are small and often located in close proximity to urban and residential areas. For this reason, it is desirable to have at hand a method by which community reaction to general aviation aircraft noise can be predicted, to at least the same level of confidence as it can for jet airliners.

The perceived noise level technique is by no means the only method available for objective evaluation of aircraft noise and it has been tested by comparing it with various other noise rating methods, through an experimental program in which a group of subjects gave their opinions on the relative acceptability of a number of recorded aircraft sounds. The method used for these tests was the well-tried "paired-comparison" technique, whereby the subjects listen to two sounds in rapid succession and judge their relative noisiness. Noisiness was specifically defined as the degree to which the sound was considered objectionable.

Prior to the execution of the main experiment, two pilot tests were carried out to evaluate a variety of possible methods by which the subjects could perform their judgments. On the basis of these tests, experimental and analytical procedures were optimized before the main experimental design was finalized.

The main experiment was performed in an acoustic facility which was originally designed for high intensity acoustic testing of hardware but which is also an ideal environment, at lower sound levels, for subjective judgment experiments in simulated free field conditions. For this purpose, sound is generated by loudspeakers in an exponential horn which expands to a working cross section measuring approximately 13 feet wide by 10 feet high. Five seated subjects are accommodated in front of a set of sound absorbant wedges, and exposed to essentially plane progressive sound waves which propagate from the horn to be absorbed by the acoustic termination behind them. The environment is thus equivalent to the free field in which the sound was originally recorded. Also, since the entire system, from original recording to the sound field at the seating positions in the facility, has a practically flat response from 25 to 10000 Hz very little sound energy is lost through poor reproduction. This is an important point which has often been overlooked in previous experiments.

Following the subjective evaluation experiments the basic aircraft sounds were analyzed, as a function of time, into 12 different noise rating units of which the perceived noise level concept formed the basis for two. Other methods which were examined for comparison included Stevens' Phons, Zwicker's Phons and a variety of weighted sound pressure levels. The correlations of these results with the judged noisiness values were then tested by statistical methods.

In all, 35 aircraft sounds were considered, representing nine aircraft types operating in a wide range of flight conditions, so that the sounds evaluated exhibited an equally wide range of spectral and temporal characteristics. The aircraft ranged from a single engined sports airplane to a four-jet transport. A significant finding which the analysis revealed is that the sound duration, defined conventionally as the interval between the "10 dB-down" points, has very little effect on the judged noisiness of the flyover sounds studied. Duration corrections when applied to five different rating methods, substantially degraded the performance of these methods as noise predictors. The study suggests that an explanation for this may lie in the effects of the Doppler frequency shift which hitherto has not been accounted for in any accepted rating scheme. A proposed modification which does account for this phenomenon, and results in a dependency on aircraft distance alone, improves the overall correlation in many cases.

Ignoring this Doppler shift correction, it was found that the pure tone corrected perceived noise level, PNdBF, is currently the most satisfactory general purpose predictor of subjective noise evaluation. Although statistical differences between the best six or so methods are small, the proposed Effective Perceived Noise Level concepr (EPNL) (References 5 and 6) is significantly inferior due to the application of a duration correction without inclusion of the Doppler correction. Slightly superior to PNdBF was a duration and Doppler corrected weighted sound pressure level, introduced in the analysis, and denoted $dB(NN)_R$. This method is based on a weighted sound pressure level scale and thus offers practical advantages over the more complex loudness/noisiness summation techniques.

## II EXPERIMENTAL DESIGN

A preliminary set of subjective judgment experiments were performed with several objectives in view. These were:

(1) To optimize the subjects' instructional set for accuracy and convenience of analysis.

(2) To ensure that useful and consistent results could be achieved.

(3) To compare results using different numbers of subjects.

(4) To develop and evaluate tape preparation and replay systems.

(5) To evaluate the progressive-wave acoustic facility for use in subjective judgment experiments.

(6) To study and develop data analysis procedures for both the subjective judgment results and the acoustic signals.

Two basic pilot experiments were conducted. The first, conducted in a semi-reverberant room, was specifically directed at the examination of various instructional sets and subjective scoring techniques. The second, based on the experience gained from the first, was performed in the progressive-wave facility with emphasis on quality of sound reproduction and consistency of results.

### A. Experiment 1

A tape was made consisting of 40 sound pairs. Five basic sound signals were used, such that four comparison signals were played at various levels in comparison with a reference signal of one octave band of pink noise centered at 1000 Hz. The latter appeared at a constant level in each of the 40 sound pairs. The four comparison signals were:

(1) An octave band of pink noise centered on 125 Hz.

(2) Pink noise, 20 - 10,000 Hz (i.e., equal energy per 1/3-octave)

(3) Simulated jet exhaust noise peaking around 200 Hz.

(4) Simulated jet noise including "compressor whine" at 4000 Hz.

The mean comparison levels were chosen by the method of individual adjustment using the average judgments of 8 subjects. The comparison sounds were then varied from relative levels of +10 to -10 dB about the mean level in increments of 5 dB. All pairs were played in both forward and reverse sequence, and the pairs were presented in random order.

A series of judgment experiments were conducted in a semi-reverberant room and Figure 1 shows a test in progress. No attempt was made to measure the acoustic characteristics of the room at this point since only the relative merits of the various scoring methods were sought. The sound pairs were amplified and reproduced by a large speaker cabinet using three loudspeakers covering the low, mid and high frequency ranges. A voice channel which announced the sound pair numbers was recorded on a second tape channel and played through an auxiliary speaker. The amplifier gain was set to give a reference signal sound pressure level of 90 dB at the subjects seating position. Initially the sounds were recorded for a duration of 4 seconds, with one second between sounds and fifteen seconds between pairs. It was found, however, that 15 seconds was far too long and for all runs after the first, the interval between pairs was reduced to 6 seconds, which proved to be adequate.

The purpose of these experiments was to evaluate four different methods for rating noise differences. These four methods are best described by the subjects' instruction and score sheets which are presented as Figures 2, 3 and 4. Twenty-three subjects took part and at least nineteen of these participated in each test.

1. <u>Method 1</u>

Each subject's score took one of three values for each sound pair; + if the second sound was judged to be noisier, - if it was judged to be less noisy, and 0 if the signals appeared equally noisy. For each comparison sound level, the number of "more noisy" votes were counted, scoring a 1/2 for an "equally noisy" result, and making the appropriate sign correction for the reverse sequency sound pairs.

The results of this procedure are shown in Figure 5. Straight lines were fitted to each set of points by the method of least squares and the relative sound pressure level corresponding to the point on the line at which 50 percent of judgments were "more noisy" is taken to be the relative level for equal noisiness.

These plots clearly indicate that it is very important to choose the correct range of relative sound pressure level for the comparison sound when this type of scoring procedure is used. In each case the line through the points has crossed the 50 percent axis at positive values of the relative sound pressure level. In at least two cases it is clear that the minimum possible percentage value of zero has caused discontinuities in in what might otherwise have been true straight lines. A quadratic curve would have been a better fit although it would probably be more correct to fit straight lines to the three right hand points.

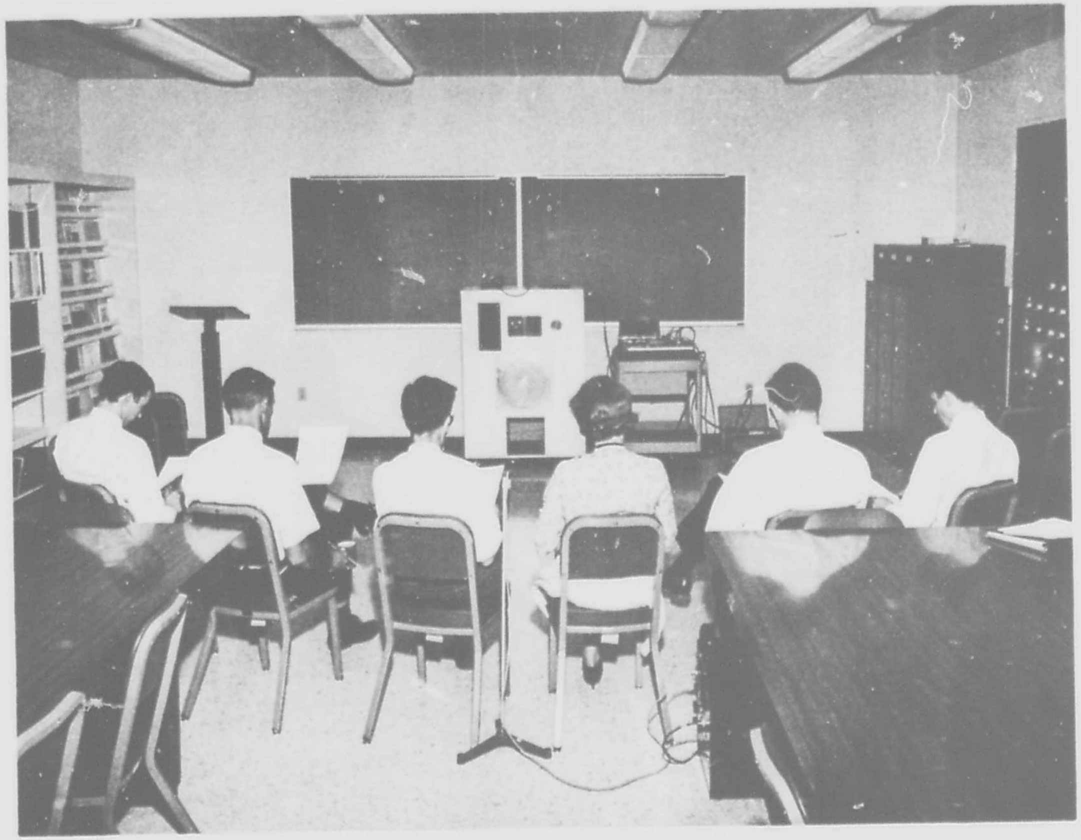


FIGURE 1 - Arrangement for Pilot Subjective Judgment Experiment in Semi-Reverberant Room

INSTRUCTIONS

NOISE RATING METHODS

The purpose of these tests is to compare the merits of various methods for judging the noisiness of sounds. Four tests will be conducted, during each of which you will hear a certain set of sounds. In each test you will be asked to judge their relative noisiness in a different way.

Method 1

When the test starts you will hear a spoken number followed by two sounds in quick succession. The number represents that pair of sounds. Decide whether the second sound is noisier (that is, more objectionable) than the first. If so enter a + sign in the box adjacent to the appropriate number on your score sheet. If you consider that it is less noisy, enter a - sign.

In making your judgment, please take account of all the effects the sound has upon you and try to imagine that the noise would occur in your home many times during the day and night. Always make a judgment for each sound pair, even though you feel you may be guessing. Remember, there is no right or wrong answer; All that is required is your own personal opinion.

Method 2

The same recording will be replayed. In this case your job, in addition to deciding whether the second sound is more or less noisy than the first, is to decide how much. To do this you must choose a scale of numbers which shows the degree of noise difference and follow your + or - sign with a number in each case. For example you may choose a range of numbers between 1 and 7 such that 1 represents a very slight difference, 4 a moderate difference, 7 a very great difference, and so on. If you think the sounds are equally noisy, enter 0.

Please feel free to choose any number scale you wish and again make a judgment in every case, though it may be a guess.

Method 3

For this test you will use a different type of score sheet in which you make a selection from nine different answers in each case. Here you must decide which of the following ratings applies to the second sound when compared with the first:

Very much noisier
Considerably noisier
Moderately noisier
Slightly noisier
Equally noisy
Slightly less noisy
Moderately less noisy
Considerably less noisy
Very much less noisy

Method 4

This test is similar to Method 2, in which you were asked to rate the noise difference as a positive or negative number. In this case, restrict your answers to the range of numbers between +5 and -5.

FIGURE 2 - Subjects' Instruction for Experiment 1

Test Number ____________ Name ______________________

| Sound Pair Number | Judgment of Second Sound |
|---|---|
| 1 | |
| 2 | |
| 3 | |
| 4 | |
| 5 | |
| 6 | |
| 7 | |
| 8 | |
| 9 | |
| 10 | |
| 11 | |
| 12 | |
| 13 | |
| 14 | |
| 15 | |
| 16 | |
| 17 | |
| 18 | |
| 19 | |
| 20 | |

| Sound Pair Number | Judgment of Second Sound |
|---|---|
| 21 | |
| 22 | |
| 23 | |
| 24 | |
| 25 | |
| 26 | |
| 27 | |
| 28 | |
| 29 | |
| 30 | |
| 31 | |
| 32 | |
| 33 | |
| 34 | |
| 35 | |
| 36 | |
| 37 | |
| 38 | |
| 39 | |
| 40 | |

FIGURE 3 - Subjects' Score Sheet for Methods 1, 2, and 4 - Experiment 1

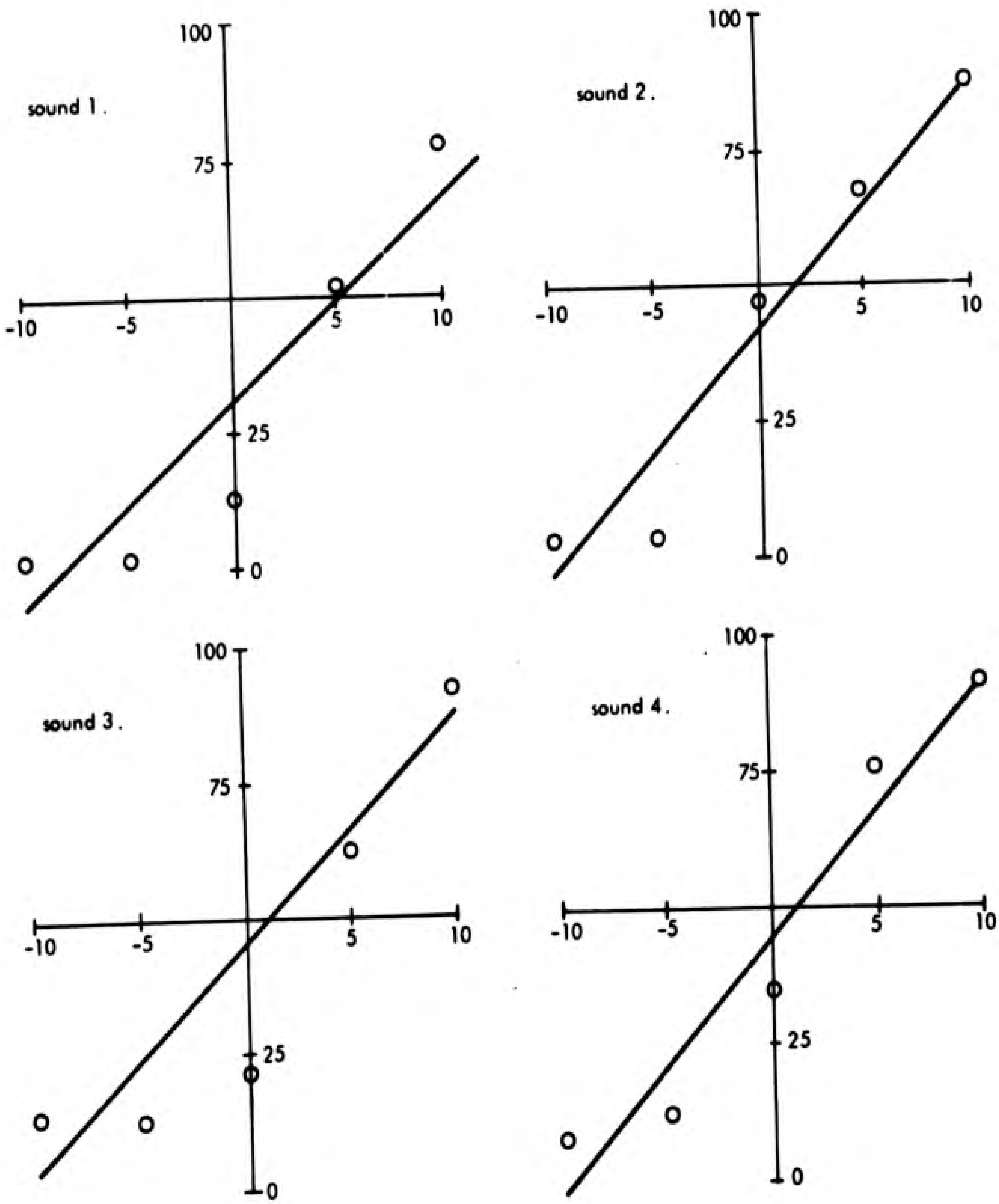
Test Number ______________ Name ____________________

Please check appropriate box in each case

| | Sound Pair Number | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Rating of Second Sound | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Very much noisier | | | | | | | | | | | | | | | | | | | | |
| Considerably noisier | | | | | | | | | | | | | | | | | | | | |
| Moderately noisier | | | | | | | | | | | | | | | | | | | | |
| Slightly noisier | | | | | | | | | | | | | | | | | | | | |
| Equally noisy | | | | | | | | | | | | | | | | | | | | |
| Slightly less noisy | | | | | | | | | | | | | | | | | | | | |
| Moderately less noisy | | | | | | | | | | | | | | | | | | | | |
| Considerably less noisy | | | | | | | | | | | | | | | | | | | | |
| Very much less noisy | | | | | | | | | | | | | | | | | | | | |

| Rating of Second Sound | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Very much noisier | | | | | | | | | | | | | | | | | | | | |
| Considerably noisier | | | | | | | | | | | | | | | | | | | | |
| Moderately noisier | | | | | | | | | | | | | | | | | | | | |
| Slightly noisier | | | | | | | | | | | | | | | | | | | | |
| Equally noisy | | | | | | | | | | | | | | | | | | | | |
| Slightly less noisy | | | | | | | | | | | | | | | | | | | | |
| Moderately less noisy | | | | | | | | | | | | | | | | | | | | |
| Considerably less noisy | | | | | | | | | | | | | | | | | | | | |
| Very much less noisy | | | | | | | | | | | | | | | | | | | | |

FIGURE 4 - Subjects' Score Sheet for Method 3 - Experiment 1

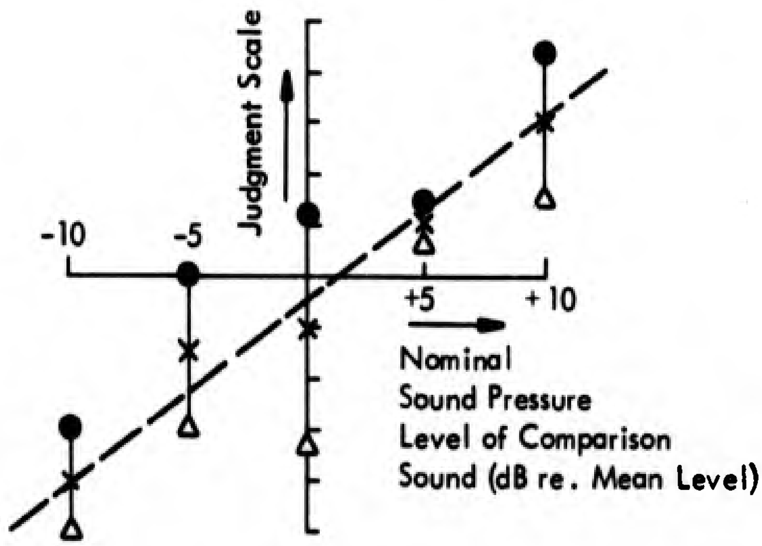



In each graph, ordinate is percentage of positive judgments, abscissa is nominal sound pressure level of comparison sound (in dB.)

FIGURE 5 - Results of Method 1 - Experiment 1

2. Methods 2, 3 and 4

The results of each of these tests were analyzed in the same way, with the "box check" results of method 3 being converted to numbers in the range +4 to -4. A typical set of results is shown in Figure 6, which represents one subject's judgments for one complete comparison (involving ten sound pairs). The triangles are the subject's judgments for the pairs in which the reference sound appears second, corrected for sign. The wide variation in one individuals' reaction to sound should be noted.




● Score for Reference as 1st of Pair
Δ Score for Reference as 2nd of Pair (with sign changed)
X Mean Score
— — Regression Line Fitted to Mean Scores

FIGURE 6- Typical Result For Methods 2, 3 or 4, For One Subject, One Comparison

The results were first analyzed on an individual basis with a linear regression line being fitted to the points representing the mean of the "forward" and "reverse" scores for each subject and each comparison. This enabled the performance of individuals to be compared. The zero crossing point of each regression line gave the relative sound pressure level of the two sounds at which they may be expected to be judged equally noisy. The findings of this analysis are summarized in Table I where the results of Methods 2, 3 and 4 are compared with Test 1. For each test several statistical parameters are given; the mean zero crossing point $L_o$, the mean standard error about the regression lines $\bar{S}_{x/y}$, the standard deviation of the zero crossing points $\sigma(L_o)$ and the mean correlation coefficient $\bar{r}$.

As in previous investigations a bias towards the second sound of each pair was found; that is, any sound was judged to be noisier when heard second than then heard first. (e.g. Figure 6). The average bias was equivalent to 2.7 dB in sound pressure level, although the scatter about this figure was considerable, corresponding to a standard deviation of 5.3 dB.

TABLE I

COMPARISON OF AVERAGE RESULTS FOR FOUR METHODS
USING LINEAR REGRESSION ANALYSIS

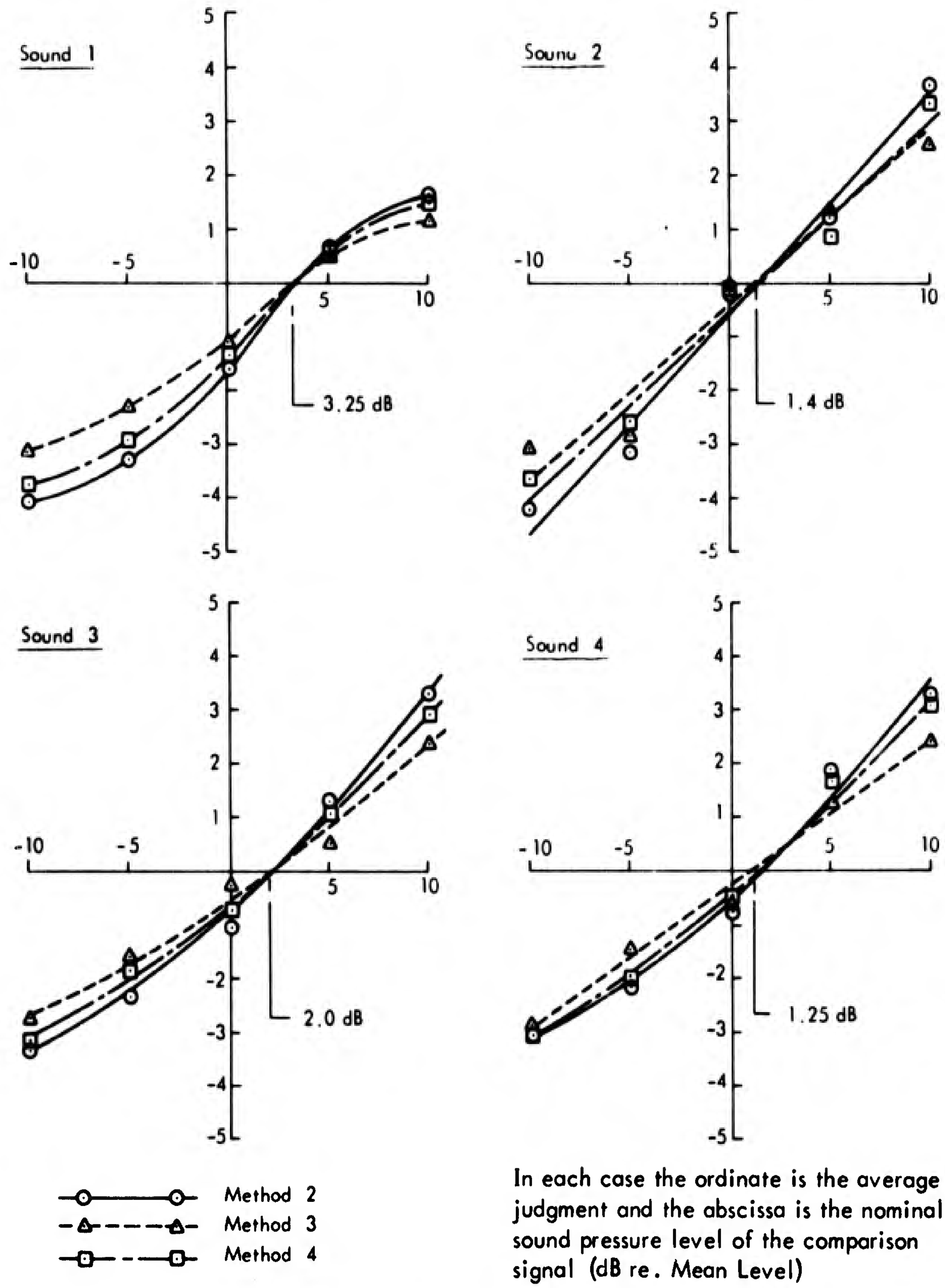
| Sound No. | Test Method | $\bar{L}_o$ | $\bar{S}_{x/y}$ | $\sigma(L_o)$ | $\bar{r}$ |
|---|---|---|---|---|---|
| 1 | 1 | 5.1 | 2.80 | | .921 |
| | 2 | 4.62 | 2.59 | 4.87 | .931 |
| | 3 | 4.63 | 2.81 | 4.23 | .921 |
| | 4 | 5.32 | 2.79 | 8.10 | .920 |
| 2 | 1 | 1.75 | 1.76 | | .970 |
| | 2 | 1.89 | 2.31 | 5.35 | .942 |
| | 3 | 2.98 | 3.57 | 5.79 | .890 |
| | 4 | 2.26 | 3.05 | 5.56 | .908 |
| 3 | 1 | 2.18 | 2.93 | | .921 |
| | 2 | 1.64 | 2.58 | 5.07 | .933 |
| | 3 | 2.92 | 3.28 | 8.22 | .899 |
| | 4 | 2.55 | 1.83 | 7.73 | .963 |
| 4 | 1 | 1.18 | 2.90 | | .925 |
| | 2 | 1.01 | 2.15 | 6.40 | .950 |
| | 3 | -.669 | 3.60 | 8.71 | .868 |
| | 4 | 0.798 | 2.10 | 6.04 | .953 |

It can be seen from Table I that the agreemr .t between the results for the various methods, as reflected in the zero crossing points, is fairly close, the largest difference being approximately 1.8 dB. However, fitting higher order curves through the mean experimental points by eye, as shown in Figure 7, suggests that the agreement should be closer. The most probable explanation for the differences observed in Table I is the use of linear regression lines. Accordingly, a second analysis was conducted on the results of Tests 2, 3 and 4 using quadratic regression lines. The results are given in Table II, and clearly show that the differences between the results for the different methods have been reduced. The standard errors associated with the regression lines- and the zero crossing points are also significantly smaller.

TABLE II

QUADRATIC REGRESSION ANALYSIS OF RESULTS FROM METHODS 2,3 AND 4 - EXPERIMENT 1

| Sound No. | Test Method | $\bar{L}_o$ | $\bar{S}_{x/y}$ | $\sigma(L_o)$ |
|---|---|---|---|---|
| 1 | 2 | 3.88 | 2.31 | 4.34 |
| | 3 | 3.81 | 2.16 | 3.48 |
| | 4 | 3.52 | 2.28 | 4.62 |
| 2 | 2 | 1.64 | 1.94 | 4.26 |
| | 3 | 1.17 | 2.35 | 1.79 |
| | 4 | 1.12 | 2.32 | 2.12 |
| 3 | 2 | 1.29 | 2.28 | 4.25 |
| | 3 | 1.86 | 2.26 | 5.87 |
| | 4 | 2.64 | 1.72 | 7.86 |
| 4 | 2 | 0.68 | 1.95 | 5.56 |
| | 3 | 0.62 | 2.35 | 2.92 |
| | 4 | 0.71 | 1.95 | 5.50 |




In each case the ordinate is the average judgment and the abscissa is the nominal sound pressure level of the comparison signal (dB re. Mean Level)

FIGURE 7 - Average Judgment Results Using Scoring Methods 2, 3 and 4. Curves Fitted by Eye.

It was also concluded on the basis of this initial experiment that

(1) Subjects are strongly influenced by the phrasing of instructions. For example, the answers submitted for Test 2, for which the instructions arbitrarily quoted a number 7, included no numbers greater than 7.

(2) There is no significant difference in the various rating methods examined, although Method 3, where the score was indicated by checking a box on the answer sheet, produced the lowest standard deviation of the zero crossing point. The most noticeable difference between the three tests can be seen in Figure 7 where it is apparent that the mean slope of the curves through the experimental points increases as the range of possible scores available to the subject increases. Method 1 was rejected on the basis that its accuracy is very sensitive to the chosen range of relative sound pressure level of the two sounds. In practice it is not possible to predetermine the ideal balance since this is the objective of the test. The final choice of method, which is discussed in the next section, was therefore made on the basis of convenience for computer analysis and on the observation that a semi-pictorial scoring method gave the lowest variance between subjects.

(3) The scatter of results is very high.

(4) Technically oriented subjects give more consistent results than non-technical subjects.

(5) The reoccurence of the same reference sound in each pair appeared to have a profound effect upon the subjects' judgments. Many clear errors of sequence were noted, that is of sign, and, subjectively, there was a definite tendency to anticipate the reference sound when it appeared second in the pair.

(6) Some form of semi-automatic method is necessary to produce accurate tapes of good quality in a reasonable time.

(7) The dynamic range of signals on the tape needs to be reduced to a minimum to maintain the greatest possible signal-to-noise ratio for the recorded sounds.

### B. Experiment 2

Following the experience with Experiment 1, a more elaborate test was designed to further investigate some of the problems and to check out the entire experimental technique to be used in the main experiment. 49 subjects took part in Experiment 2 which was conducted in the progressive-wave facility, with 5 subjects at each sitting. Of these subjects, 14 were female with an average age of 29.8 and a standard deviation of 6.8 years and 35 males with an average age of 33.4 and a standard deviation of 6.4 years. The total duration of a single run was approximately 30 minutes, with a short break at the half-way point, during which 120 sound pairs were played. The major advances from the initial experiment were the use of a single instructional set/ testing technique, a reduction of the total dynamic range of the recorded signals ensuring an improved signal to noise ratio, and the use of several reference sounds as opposed to one. One of the main objectives was to ascertain the reliability of the results by the cross-comparison of several sounds, and the repeated occurrence of a single sound pair (at random intervals).

Six basic signals were used, each of 4-seconds duration. They comprised the six octave bands of pink noise centered at frequencies of 125, 250, 500, 1000, 2000 and 8000 Hz. Ten comparisons were made as depicted in Figure 8. The circles denote the basic sounds which were octave bands of the pink noise with the center frequencies marked, and the squares denote the comparison numbers.

Each line represents a comparison, involving 10 sound pairs in all, the comparison sound being played at 5 different levels both preceding and following the reference sound. The arrowhead on the line points at the reference signal in each case. The broken line denotes a single sound pair which was repeated at random intervals 20 times throughout the test, including forward and reverse sequences. This particular network was selected so that any two sounds were compared in two ways, directly and indirectly through an intermediate comparison.

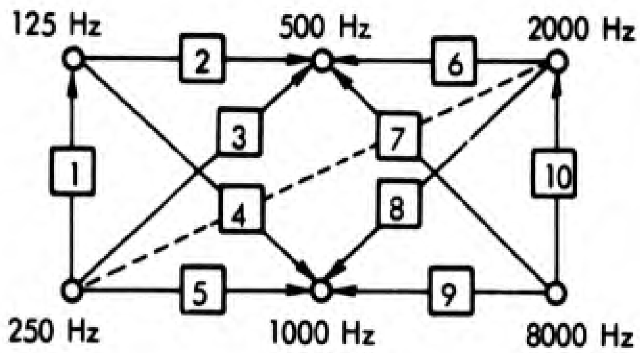



FIGURE 8- Sound Comparisons Made in Experiment 2

The sounds were played to groups of five subjects at a time, seated in the 1500 cubic feet acoustic facility in its progressive wave configuration. A description of the facility, the sound reproduction systems used and the frequency response calibrations are given in Appendix A. The five seats were aligned, side by side in front of the termination wedges and Figure 9 shows 5 subjects in attendance. Replay system 1 was used for this experiment utilizing 5 loudspeakers with the appropriate dividing network. The frequency response was essentially flat within the frequency range 25 to 10,000 Hz

FIGURE 9 - Arrangement for Subjective Tests in Progressive Wave Acoustic Facility

although there were some slight differences between the five seat positions at certain frequencies.

The instructions given to the subjects together with a sample completed score sheet are shown in Figures 10 and 11. The score sheet design is based on the results of Experiment 1 and represents a compromise between the numerical scales of Methods 2 and 4 and the pictorial layout of the Method 3 sheets which subjects claimed were "easiest" to use and which showed the least scatter of results. This format is also easy for translation by the key punch operator who prepared the results for computer analysis.

Figure 12 shows a set of histograms printed by the computer program which was written for the CDC 3300 computer to analyze the results. Each asterisk represents a single judgment for one individual subject and is the sum of his scores for that sound pair played in forward and reverse sequence (with sign correction). The top row represents the complete set of results for comparison 8. These clearly show the degree of scatter associated with collective judgment of sounds but more importantly the skewed distributions at the extreme sound pressure level differences are evident. Because this effect occurs very frequently it was concluded that a more meaningful measure of central location for each distribution is the median, or 50th percentile, rather than the mean. In the case of a symmetrical distribution, median and mean coincide, but for the skewed distributions, which in this instance largely result from the finite range of possible scores available to the subjects, the median more closely approximates the mean which would have been determined had an infinite range of numbers been available. The median is the halfway point in the scores such that half the total lie above it and the other half below it. It was further observed that in general, a skewed distribution was only found at one end of the relative sound pressure level scale since the total range is normally insufficient to cause the subjects to "run out of numbers" at both ends. This tends to confirm the conclusions that a quadratic curve fit would be adequate in the majority of cases. With a large range of relative sound level it would no doubt be necessary to proceed to a cubic fit since the curve may be expected to be S-shaped, leveling out asymptotically to the mean scores $\pm 5$. Thus a computer program was written to perform a complete analysis of the subjective response data, fitting a quadratic to the median scores for each sound pair in a set, by the method of least squares. The complete set of results was now averaged over all subjects since we were no longer interested in comparing the zero crossing points for each individual.

The error involved, in terms of the scatter about the fitted curve, is expressed in terms of percentiles as opposed to standard deviations. Therefore, in addition to the median, the 10th and 90th percentiles judgments were computed for each sound pair and further quadratic curves fitted to these points. The results of this procedure are presented in Figure 13 where the comparisons are numbered as shown in Figure 8.

It may be noted from these plots that the intervals between successive comparison signal levels are not uniform. This was due to errors in recording level during the

INSTRUCTIONS - Please Read Very Carefully

The purpose of these tests is to evaluate the noisiness of certain sounds and to determine how consistently their noisiness can be judged. The tests are part of a research program to study the noisiness of General Aviation Aircraft.

When the test starts, you will hear a sequence of sounds, grouped in pairs. Each pair is preceded by a spoken number which identifies that pair on your answer sheet. Listen to the two sounds, decide how much the second sound is more or less noisy, or objectionable than the first, and circle the appropriate number in the correct column on your answer sheet. The positive numbers, from +1 through +5 indicate that the second sound is more noisy, or objectionable. The bigger the number, the greater the difference, so that +1 represents a very slight noise difference, +2 a greater one, and so on. The negative numbers, from -1 through -5, indicate that the second sound is less noisy, or objectionable, and again, the larger the number the greater the difference. Circle the number zero if you feel that the two sounds are equally disturbing.

Remember that you are judging the second of the sounds with respect to the first. You may think that neither of the noises is objectionable or that both are objectionable. However, you should judge whether the second sound would disturb you more or less than the first sound, if heard in your home many times during the day and night.

In making your judgment, please take account of all the effects the sounds have upon you. Please record an answer in every case, even though you feel you may be guessing. Remember, there is no right or wrong answer. All that is required is your own personal opinion.

FIGURE 10 - Instructions given to Subjects for Experiment 2

Test Number 8-22-P4-C-4 NAME [illegible] Number 40

PLEASE JUDGE THE SECOND SOUND WITH RESPECT TO THE FIRST

Group A

| Sound Pair No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| More Noisy + | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 |
| | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 |
| | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 |
| | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 |
| | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 |
| Equally Noisy | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Less Noisy - | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |
| | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 |
| | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 |
| | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 |
| | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 |

Group B

| Sound Pair No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| More Noisy + | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 |
| | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 |
| | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 |
| | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 |
| | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 |
| Equally Noisy | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Less Noisy - | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |
| | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 |
| | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 |
| | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 |
| | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 |

Group C

| Sound Pair No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| More Noisy + | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 | +5 |
| | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 | +4 |
| | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 | +3 |
| | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 | +2 |
| | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | +1 |
| Equally Noisy | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Less Noisy - | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |
| | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | -2 |
| | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 | -3 |
| | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 | -4 |
| | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 | -5 |

FIGURE 11 - Example Completed Score Sheet from Experiment 2

SOUND PAIR 8 — frequency (FR.) by score (SC) for LEVEL 1 to LEVEL 5

| SC | LEVEL 1 FR. | LEVEL 2 FR. | LEVEL 3 FR. | LEVEL 4 FR. | LEVEL 5 FR. |
|---|---|---|---|---|---|
| -10 | 0 | 0 | 0 | 0 | 5 |
| -9 | 0 | 0 | 2 | 1 | 6 |
| -8 | 0 | 0 | 0 | 4 | 10 |
| -7 | 1 | 3 | 2 | 6 | 5 |
| -6 | 0 | 1 | 7 | 7 | 9 |
| -5 | 0 | 2 | 9 | 10 | 6 |
| -4 | 0 | 3 | 7 | 4 | 4 |
| -3 | 0 | 2 | 7 | 6 | 0 |
| -2 | 1 | 9 | 6 | 3 | 1 |
| -1 | 1 | 4 | 4 | 6 | 0 |
| 0 | 2 | 6 | 3 | 0 | 2 |
| 1 | 1 | 3 | 0 | 1 | 1 |
| 2 | 0 | 9 | 1 | 0 | 0 |
| 3 | 3 | 2 | 1 | 1 | 0 |
| 4 | 3 | 4 | 0 | 0 | 0 |
| 5 | 2 | 1 | 0 | 0 | 0 |
| 6 | 8 | 0 | 0 | 0 | 0 |
| 7 | 9 | 0 | 0 | 0 | 0 |
| 8 | 7 | 0 | 0 | 0 | 0 |
| 9 | 10 | 0 | 0 | 0 | 0 |
| 10 | 1 | 0 | 0 | 0 | 0 |

49 SUBJECTS

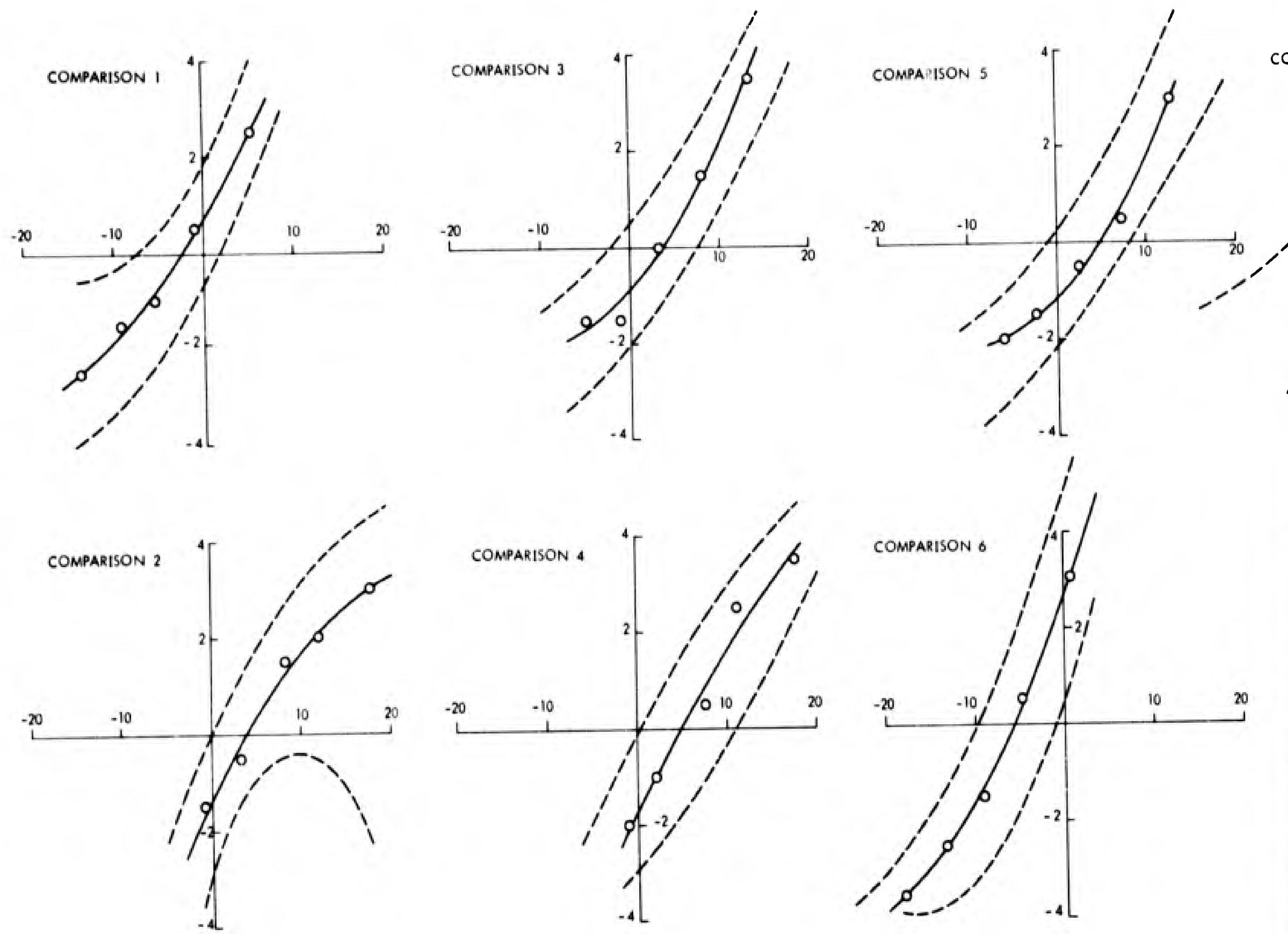
| SC | LEVEL 1 FR. | LEVEL 2 FR. | LEVEL 3 FR. | LEVEL 4 FR. | LEVEL 5 FR. |
|---|---|---|---|---|---|
| -10 | 0 | 0 | 0 | 0 | 0 |
| -9 | 0 | 0 | 0 | 0 | 3 |
| -8 | 0 | 0 | 0 | 2 | 4 |
| -7 | 0 | 1 | 0 | 3 | 3 |
| -6 | 0 | 0 | 2 | 3 | 6 |
| -5 | 0 | 1 | 4 | 2 | 2 |
| -4 | 0 | 1 | 6 | 2 | 1 |
| -3 | 0 | 1 | 3 | 1 | 0 |
| -2 | 0 | 3 | 2 | 2 | 0 |
| -1 | 1 | 2 | 2 | 5 | 0 |
| 0 | 0 | 3 | 1 | 0 | 1 |
| 1 | 0 | 2 | 0 | 0 | 0 |
| 2 | 0 | 2 | 0 | 0 | 0 |
| 3 | 1 | 1 | 0 | 0 | 0 |
| 4 | 0 | 2 | 0 | 0 | 0 |
| 5 | 0 | 1 | 0 | 0 | 0 |
| 6 | 2 | 0 | 0 | 0 | 0 |
| 7 | 5 | 0 | 0 | 0 | 0 |
| 8 | 5 | 0 | 0 | 0 | 0 |
| 9 | 6 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 |

"BEST" 20 SUBJECTS

| SC | LEVEL 1 FR. | LEVEL 2 FR. | LEVEL 3 FR. | LEVEL 4 FR. | LEVEL 5 FR. |
|---|---|---|---|---|---|
| -10 | 0 | 0 | 0 | 0 | 5 |
| -9 | 0 | 0 | 2 | 1 | 2 |
| -8 | 0 | 0 | 0 | 2 | 2 |
| -7 | 1 | 2 | 1 | 3 | 1 |
| -6 | 0 | 0 | 2 | 3 | 2 |
| -5 | 0 | 0 | 5 | 5 | 3 |
| -4 | 0 | 2 | 0 | 1 | 3 |
| -3 | 0 | 1 | 2 | 4 | 0 |
| -2 | 1 | 5 | 3 | 0 | 0 |
| -1 | 0 | 0 | 1 | 1 | 0 |
| 0 | 2 | 2 | 2 | 0 | 1 |
| 1 | 1 | 0 | 0 | 0 | 1 |
| 2 | 0 | 5 | 1 | 0 | 0 |
| 3 | 1 | 1 | 1 | 0 | 0 |
| 4 | 3 | 2 | 0 | 0 | 0 |
| 5 | 1 | 0 | 0 | 0 | 0 |
| 6 | 2 | 0 | 0 | 0 | 0 |
| 7 | 2 | 0 | 0 | 0 | 0 |
| 8 | 1 | 0 | 0 | 0 | 0 |
| 9 | 4 | 0 | 0 | 0 | 0 |
| 10 | 1 | 0 | 0 | 0 | 0 |

"WORST" 20 SUBJECTS

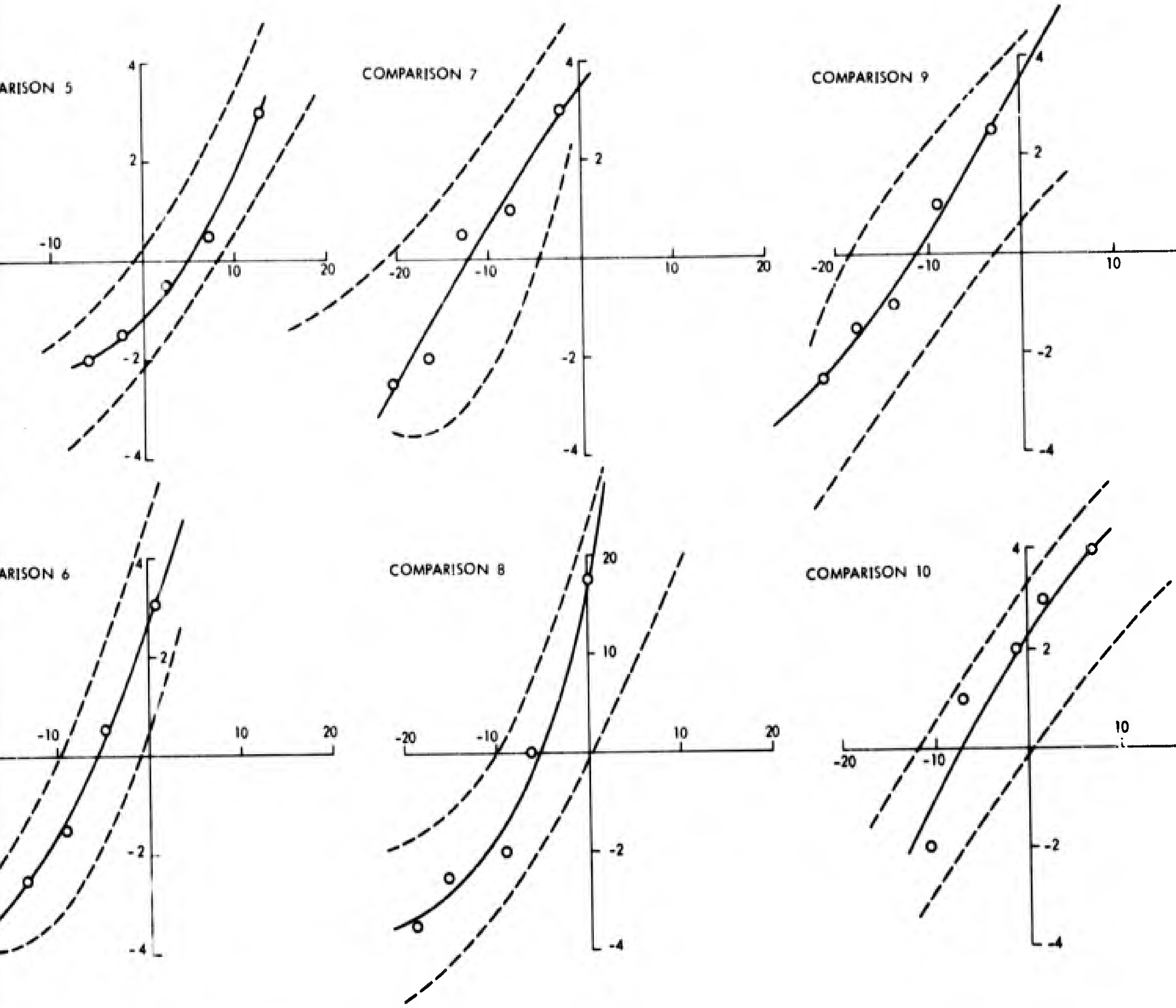
FIGURE 12 - Histograms of Judgment Results for a Particular Comparison in Experiment 2




In each case ordinate is subjective judgment scale, is relative SPL of comparison sound, continuous line square quadratic fit to median points (plotted) and b lines fit 10th and 90th percentiles.

FIGURE 13 - Quadratic Curve Fit to Percentile Analysis of Subjective Ju

A




inate is subjective judgment scale, abcissa
comparison sound, continuous line is least
it to median points (plotted) and broken
90th percentiles.

ercentile Analysis of Subjective Judgment Results for Experiment 2

BLANK PAGE

manufacture of the tape; it was intended to utilize 5 dB steps, as previously. The levels given are the true sound pressure levels, in dB, of the comparison sounds relative to the reference in each case. The 10th and 90th percentile curve intersection points give an estimate of the scatter associated with the calculated zero crossing, since approximately 80 percent of the actual scores lie within these two boundaries. This range is somewhat larger than the width of two standard deviations which would bound approximately 63 percent of the results. The 10th percentile curve for sound pair 2 shows an unfortunate trend which illustrates one of the penalties of automating the analysis of these results. Due to a small number of excessively dispersed points and the constraints of a quadratic curve fit, the line has failed to cross the zero axis so that the required scatter band could not be computed.

The 49 subjects who participated in this experiment were selected from Wyle Laboratories personnel to give a reasonable mix of sex, age, and background. The group consisted of 35 males and 14 females. The performance of each individual was assessed by calculating the variance of his bias, i.e., the additional weighting he gave to the noisiness of a particular sound because it was the second of the pair. There was considerable variation in both the bias and the standard deviation of the bias. The former had values ranging from -0.241 to +1.129 judgment units and the standard deviation from 0.595 to 1.711 units. Examination of the results for the 20 repetitions of the same sound pair revealed adequate consistency in all but a very few cases. The standard deviation of the scores for these sound pairs averaged over all subjects was 0.86 judgment units. Only in about 5 cases was there any clear evidence of an "adjustment period" where the subject was obviously in the process of adjusting his personal rating scale during the first few sound pairs.

In order to examine the consequences of using a limited number of subjects, the results of the "best" and "worst" twenty subjects were analyzed independently, where "best" and "worst" were determined from a rank listing of the standard deviation of bias as discussed above. The second and third rows of histograms presented in Figure 12 show the results of making this separation. The overall spread of results is obviously greater for the worst 20 subjects. Comparison of the results using all 49 subjects, the 20 best subjects and the 20 worst subjects are made in Table III. The most noticeable difference apparent in this table is the increase in the 80 percent confidence interval for the worst 20 subjects. Standard t-tests on these comparisons for both the medians and the 80 percent confidence intervals showed that

(1) There is no difference between the results for all 49 subjects and those for the best twenty at the 5 percent level of significance.

(2) There is a difference between all 49 and the worst 20 and between the best and worst 20 at the 5 percent level of significance.

TABLE III

EFFECTS OF SUBJECT GROUP SELECTION
(Zero Crossings are Relative to Mean Comparison Level)

| Comparison No. | Median Zero Crossing (dB) | | | 80% Confidence Interval (dB) | | |
|---|---|---|---|---|---|---|
| | 49 Subs | Best 20 | Worst 20 | 49 Subs | Best 20 | Worst 20 |
| 1 | - 2.4 | - 2.0 | - 2.5 | 9.0 | 8.0 | 9.4 |
| 2 | 3.6 | 3.4 | 4.9 | 5.3 | 7.5 | 4.1 |
| 3 | 3.8 | 3.0 | 4.9 | 9.7 | 8.9 | 14.1 |
| 4 | 4.7 | 5.0 | 5.4 | 10.6 | 9.7 | 15.9 |
| 5 | 4.9 | 4.9 | 4.4 | 9.3 | 9.0 | 10.7 |
| 6 | - 5.7 | - 5.9 | - 6.2 | 8.8 | 11.8 | 12.4 |
| 7 | -12.2 | -13.1 | - 9.6 | 16.8 | 15.7 | 17.4 |
| 8 | - 6.1 | - 6.2 | - 4.2 | 10.7 | 8.9 | 12.7 |
| 9 | -11.5 | -10.7 | -10.7 | 15.2 | 15.2 | 18.3 |
| 10 | - 7.3 | - 8.3 | - 6.2 | 12.2 | 10.2 | 14.8 |
| | | | AVERAGE | 10.8 | 10.5 | 13.0 |

However, since the mean differences between the median results are only 1.4 dB between all 49 and the worst 20 subjects and 1.6 dB between the best 20 and the worst 20 subjects it is apparent that the differences are not too serious even in this extreme comparison.

In order to examine these differences in a little more detail, the 49 subjects were subdivided in two further ways and the performances of the various sub-groups, again rated by the deviations of the bias, were compared. The first division was into male and female sub-groups and no differences were found at the 5 percent significance level. Secondly, the male subjects were divided into

(a) University graduates in science and engineering
(b) Other white collar workers
(c) Blue collar workers
(d) Female employees.

This separation showed the existence of a significant difference between the consistency of group (a) and each of the other three groups between which there was no difference at the 5 percent level of significance.

Figure 8, showing the network of comparisons used in Experiment 2, illustrates the manner in which all comparisons are repeated - directly and indirectly. For example, the 250 Hz sound is compared directly with the 500 Hz band in comparison number 3 and indirectly through comparisons 1 and 2. In fact it is easy to see that these two sounds are compared by 4 indirect routes. A measure of the consistency of the results, therefore, is to calculate and compare the relative sound pressure level of these two sounds as judged equally noisy via the five independent routes. An equivalent step is to examine the results for 5 sound pairs, each compared via two possible routes. This has been done in Table IV, for the three subject groups studied previously, using the zero crossing data from Table III. It can be seen that the errors vary considerably depending on the subjects selected; another indication of the large variance involved in subjective judgment. It is ironical that the largest error, 3.8 dB arises from the use of the "best" 20 subjects' results.

TABLE IV

NETWORK ERRORS FOR EACH SUBJECT GROUP

| Equivalent Routes | Magnitude of Error Between Routes - (dB) | | |
|---|---|---|---|
| | 49 Subjects | Best 20 | Worst 20 |
| 3 / 1 - 2 | 2.2 | 1.6 | 2.4 |
| 4 / 1 - 5 | 2.6 | 2.1 | 1.5 |
| 7 / 6 - 10 | 0.8 | 1.1 | 2.7 |
| 8 / 9 - 10 | 1.9 | 3.8 | 0.4 |
| 5-1-2 / 9-10-6 | 2.2 | 0.0 | 0.4 |

The actual relative noisiness results of this experiment are presented and discussed in Section IV.

The main conclusions from this second pilot experiment can be summarized as follows:

(1) The optimum method for automatic analysis of the subjective results is to fit quadratic curves to the median values.

(2) In view of the curve fitting problem which can be caused by a few stray points it is probably safer to use the 20th and 80th percentiles for confidence limit estimation.

(3) The scatter of judgments is very high in every case but can be reduced by careful selection of subjects. However, this almost certainly narrows the range of subject type since it was shown that the most consistent subjects were University graduates with technical and scientific backgrounds. Fortunately, although a choice of less consistent subjects increases the variance of results, the median results do not vary substantially. The differences are certainly no more than the network errors listed in Table IV. It is thus considered that 20 subjects chosen at random will give meaningful results.

(4) 30 minutes is a little excessive for one sitting. Several subjects complained of fatigue which affected their concentration. 20 minutes per session is probably more acceptable.

## III MAIN EXPERIMENT

### A. Sound Recording

The choice of aircraft sounds studied was dictated to some extent by what was available at the time of the various recording sessions. The final selection, for which sound recordings were made under a variety of conditions, is listed in Table V.

TABLE V

AIRCRAFT WHOSE SOUND RECORDINGS WERE USED IN THE MAIN EXPERIMENT

| Aircraft Type/Model | Classification | All-Up-Weight | Installed H.P./Thrust |
|---|---|---|---|
| Piper Cherokee 140 | Single Piston Engine, 4-Place Sport/Business Aircraft | 2150 lbs | 150 H.P. |
| Piper Cherokee 6 | Single Piston Engine, 6-Place Utility Aircraft | 3400 lbs | 260 H.P. |
| Piper Aztec | Twin-Piston Engine, 6-Place Executive Transport | 5200 lbs | 500 H.P. |
| Turbo-Commander | Twin Turboprop 7-9 Seat Executive Transport | 8950 lbs | 1300 S.H.P. |
| Lear Model 23 | Twin Turbojet 8-Seat Executive Transport | 12,500 lbs | 5700 lbs |
| Douglas DC-9/30 | Twin Turbofan 115 Seat Short-Medium haul Transport | 98,000 lbs | 28000 lbs |
| Boeing 707/120B | 4 Turbofan Long Haul Transport | 257,000 lbs | 72000 lbs |

Details of the methods by which sound recordings were obtained are confined to Appendix A. The first three aircraft in Table V, which are illustrated in Figure 14, were hired for short periods and flown in a variety of conditions in the vicinity of the recording microphone. Permission was granted by the Huntsville-Madison County Airport Authority to make these flights at the new Huntsville Airport while it was still under construction. This airport is located in open country some eight miles from Huntsville and the very low


Piper Cherokee 140


Piper Aztec

Piper Cherokee 6

FIGURE 14 – Piston Engined Aircraft Whose Sounds were Evaluated

background noise permitted the acquisition of interference free recordings of good quality. Accurate flight information was recorded by an observer in each aircraft who was in radio communication with the recording engineer on the ground. Sound recordings of the remaining aircraft were obtained at Memphis Metropolitan Airport with the cooperation of the Executive Director and his staff. These aircraft were performing normal operations at the airfield and the recording equipment was simply moved from place to place within the airport environs to obtain the necessary variety of recordings. The various recordings for each aircraft model were not necessarily obtained from the same aircraft. Table VI gives a complete description of the flyover sounds used in the main experiment. Each signal has been allocated an identification code for ease of description which is included in the table. All signals were obtained in conditions of low windspeed (in all cases, except one, less than 5 mph) and, where possible, in the early or late hours of daylight.

B. Preparation of Tapes

In order to assess the direct effects of signal duration on judged noisiness, long duration sound recordings were made of the three Piper aircraft in circling flight over the microphone. The recordings were made in each case with the aircraft performing (approximately) a 2g turn. The Cherokee 140 and Aztec were flown at an altitude of 500 ft over the microphone at velocities of 90 mph IAS* and 155 mph TAS**, respectively, and the Cherokee 6 at 250 ft at a velocity of 125 mph TAS. The recorded sounds were not exactly stationary due to perturbations from the ideal flight path but the level remained constant to within 2 dB. Using a variable gain amplification circuit, five signals were made from each original recording. These had realistic amplitude time histories, with their overall sound pressure level rising and falling 20 dB at a uniform rate, but with total durations ranging from 2 to 32 seconds. The results sounded like aircraft flyovers but with the curious effect of no frequency change. The shaped signals are given identification codes with the duration, in seconds, following a vertical stroke; e.g., 140/4, 6/8, A/16.

The pilot studies showed the desirability of using a variety of reference signals and the main experiment was designed so that, although the noisiness of each aircraft sound could be related to that of a standard reference sound, the majority of comparisons were made indirectly through an intermediate reference signal. The standard reference, as in the pilot experiments, was an octave band of pink noise, centered on 1000 Hz, with a duration of 4 seconds. To provide some intermediate references the same shaping functions used for the Piper Aircraft sounds were applied to a shaped wideband noise spectrum which simulated jet exhaust noise. These signals, four in number, having total durations of 4, 8, 16 and 32 seconds and identified by the codes JN4, JN8, JN16 and JN32, were shaped through a set of 1/3 octave filters to peak at 250 Hz and fall away at 2 dB per 1/3 octave on either side.

*IAS = Indicated Airspeed　　**TAS = True Airspeed

## TABLE VI

## FLYOVER SIGNALS EVALUATED

| Code | Aircraft | Flight Condition | Measurement Point | Altitude (ft) | Velocity (mph) | Peak* SPL |
|---|---|---|---|---|---|---|
| 140 F | Piper Cherokee 140 | Max. level speed | Vertically below | 250 | 120 TAS | 85.0 |
| 140 T1 | Piper Cherokee 140 | Take-off comm. 2500' before mic. | Vertically below | 200 | 70 IAS | 88.0 |
| 140 T2 | Piper Cherokee 140 | Take-off comm. 5000' before mic. | Vertically below | 400 | 85 IAS | 78.0 |
| 140 T3 | Piper Cherokee 140 | Take-off comm. 7500' before mic. | Vertically below | 600 | 85 IAS | 76.5 |
| 6 T | Piper Cherokee 6 | Take-off comm. 2500' before mic. | Vertically below | 300 | 90 IAS | 94.5 |
| 6 F1 | Piper Cherokee 6 | Level cruise at 75% power | Vertically below | 250 | 160 TAS | 96.0 |
| 6 F2 | Piper Cherokee 6 | Level cruise at 75% power | Vertically below | 500 | 140 TAS | 86.0 |
| 6 F3 | Piper Cherokee 6 | Level cruise at 75% power | Vertically below | 1000 | 145 TAS | 80.5 |
| AT | Piper Aztec | Take-off | Vertically below | 100 | 135 TAS | 90.5 |
| AF1 | Piper Aztec | Level cruise at 75% power | Vertically below | 250 | 175 TAS | 98.5 |
| AF2 | Piper Aztec | Level cruise at 75% power | Vertically below | 500 | 165 TAS | 90.0 |
| AF3 | Piper Aztec | Level cruise at 75% power | Vertically below | 1000 | 170 TAS | 82.5 |
| DCT | DC-9 Series 30 | Take-off comm. 10500' before mic. | 1020' sideline | 565 | -- | 98.0 |
| DCL | DC-9 Series 30 | Landing, 1/2 mi. before touch down | Vertically below | 220 | -- | 102.5 |
| 707T | B707-120B | Take-off comm. 10500' before mic. | 1020' sideline | 575 | -- | 102.0 |
| 707L | B707-120B | Landing, 1/2 mi. before touch down | Vertically below | 200 | -- | 114.0 |
| LT1 | Lear Model 23 | Take-off comm. 12800' before mic. | Vertically below | 1090 | -- | 100.0 |
| LT2 | Lear Model 23 | Take-off comm. 10500' before mic. | 1020' sideline | 435 | -- | 99.0 |
| LL | Lear Model 23 | Landing, 1/2 mi. before touch down | Vertically below | 110 | -- | 103.5 |
| TC | Turbo-Commander | Landing, 1/2 mi. before touch down | Vertically below | 220 | -- | 89.0 |

*Maximum measured level, dB re. 0002 μ Bar

A secondary objective of the main experiment was to develop an equal noisiness contour for use as a sound pressure level weighting function in the analysis of the final results, since it was thought that such a function, derived under the same experimental conditions, would provide an interesting comparison with other standard weighting functions. Accordingly, nine 1/3 octave band wide noise signals, each of four seconds duration were included in the set of signals for evaluation. Their frequencies were centered on all octave intervals between and including 31.5 Hz and 8 KHz. For reference, they have been assigned the codes 3OB31 through 3OB8K.

All sounds presented to the subjects had a rise and decay time of approximately 20 milliseconds to avoid any extraneous effects due to transients. These were generated by passing each signal through a 60 dB dynamic range variable gain amplifier whose bias voltage was controlled by a discharging, or charging, capacitor. So that tapes of good quality could be produced with a minimum of transient signals due to tape recorder operation or other causes, and to maintain rigid control of the duration of the signals, an automatic recording system was developed for dubbing the original signals onto the master tapes for the subjective experiments. This system is described in some detail in Appendix A.

The paired comparison sounds which are listed in Table VII in the Analysis of Results, and will therefore not be repeated here, were recorded, as before, in random sequence; this time with the reference sound varied at 5 dB intervals about a balanced mean level. The latter was estimated for each comparison using various available noise rating techniques since a preliminary "individual adjustment" experiment would have been a major task with so many sounds involved. Each pair was recorded ten times, five times each in forward and reverse sequence. In all, 480 sound pairs were recorded in sets of 40 per tape which played for approximately 20 minutes each. The tapes were arranged so that a set of three comprised a completely self-contained test, the results of which could be analyzed by computer. Each test comprised 120 sound pairs and there were 4 tests in all. The actual record and replay levels of each aircraft sound had to be altered from the levels measured at the recording site in order to restrict the total dynamic range to a minimum. As stated previously, it is desirable to minimize the total range in order to maintain the maximum possible signal to noise ratio. In addition to the tape noise, most elements in the replay circuit give rise to some electronic noise, which can be troublesome in a wide frequency range system.

C. Experiment

The experiments were conducted in the progressive wave facility; this time using replay system 2, as described in Appendix A. It was found that two speakers gave an equal performance to the five in Experiment 2, and since these were hidden well down in the primary horn of the facility, obscured from the view of the subjects, any risk of distraction by the speakers was removed. Following audiometric screening tests, 28 subjects took part in the experiment, 13 females with average age and standard deviation of 29.9 and 10.5 years, respectively, and 15 males of average age and standard deviation of 31.8 and 5.9 years. Of these, 22 subjects had experience in the previous experiments. Unfortunately, due to scheduling problems, only 20 complete sets of data were obtained for each test. The instruction and score sheets were identical to those used in Experiment 2. The sound was monitored continuously during the tests by a microphone located alongside the central seat position, and recorded for possible reference at a later time.

D. Analysis of Aircraft Signals

A computer program was written to calculate a series of noisiness or loudness ratings as a function of the 1/3 octave band spectrum levels. The input routine was designed so that a time history of sound pressure levels could be read for each 1/3 octave band until a complete array, fully describing the time history of a particular aircraft flyover sound was read. The program computes, for each time interval, the following ratings which again, for convenience, are given abbreviations. These are, for the most part, accepted terminology:

OASPL - Overall Sound Pressure Level in dB re: .0002 μbar.
dB(A) - A-weighted sound pressure level.
dB(B) - B-weighted sound pressure level.
dB(C) - C-weighted sound pressure level.
dB(N) - N-weighted sound pressure level.
PNdB - Perceived Noise Level (according to the method and tables in Reference 7).
PNdBF - Pure tone corrected Perceived Noise Level (Reference 5).
S Phon - Phons following Stevens' MK VI method (Reference 8).
Z Phon - Phons following Zwicker for free field conditions. (The method is based on the work presented in Reference 9 but uses a chart adapted by Kryter in Reference 10).

In addition, the program includes provision to calculate a number of arbitrarily weighted sound pressure levels according to weighting functions which can be read prior to the computation. For all sounds, three further weighted levels were computed, defined by:

dB(J) }
dB(K) } - Weighted sound pressure levels which reportedly showed good correlation with subjective test data during recent research at NASA, Marshall Space Flight Center (Reference 11).
dB(NN) - A weighted sound pressure level based on the function generated during this study (see p. 38).

All sound pressure level weighting functions are shown in Figure 15.

Using a graphic level recorder as described in Appendix A, 1/3 octave band levels in the range 50 to 10,000 Hz were plotted as a continuous function of time for all recorded aircraft signals evaluated during this study. The actual signals analyzed were the final master recordings which were replayed to the subjects. The resulting traces were read at 1/2 second intervals, punching the data onto cards for computer input. Care was exer-

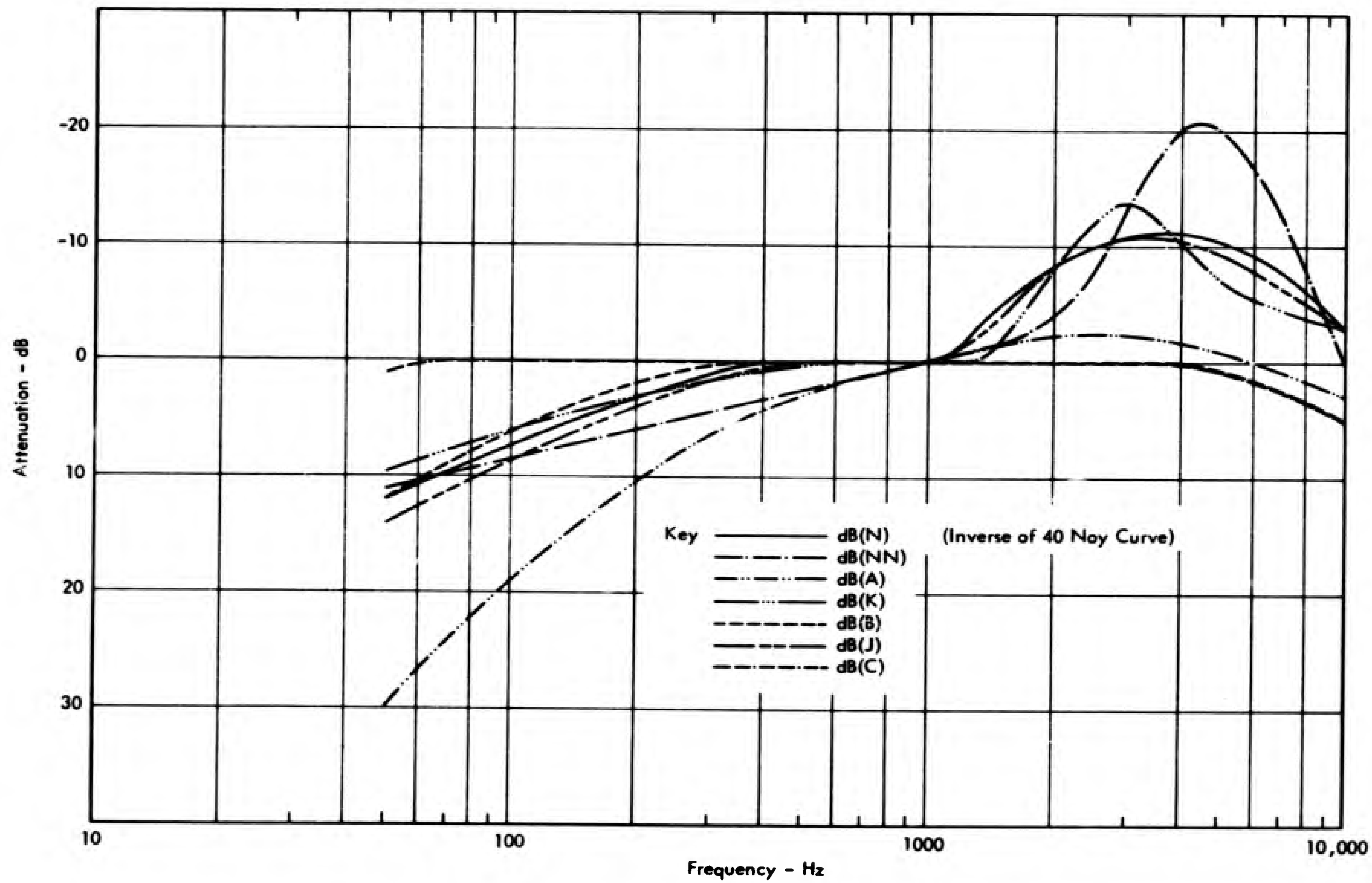



FIGURE 15 - Weighting Functions used by Sound Analysis Program

cised to ensure that the same reference instant, at which readings commenced, was common to each 1/3 octave band trace for each signal. From this data, 1/2-second time histories of each rating scale were computed for each signal.

In order to illustrate the complexity of the actual sound signal reaching an observer as a piston engined aircraft flies overhead, a recording of the Cherokee 6 cruising 1000 ft above the microphone, which was not included in the evaluation experiment, was analyzed using a Federal Scientific "Ubiquitous" real time spectrum analyzer. This instrument performs narrow band spectrum analysis at a succession of very short time intervals so that the results may be viewed as a continuously varying signal on an oscilloscope. By photographing successive sweeps on the oscilloscope and subsequently reading the "instantaneous" harmonic amplitudes, Figure 16 was constructed which shows the time and frequency variation of various identifiable sound harmonics recorded as the aircraft flew by.

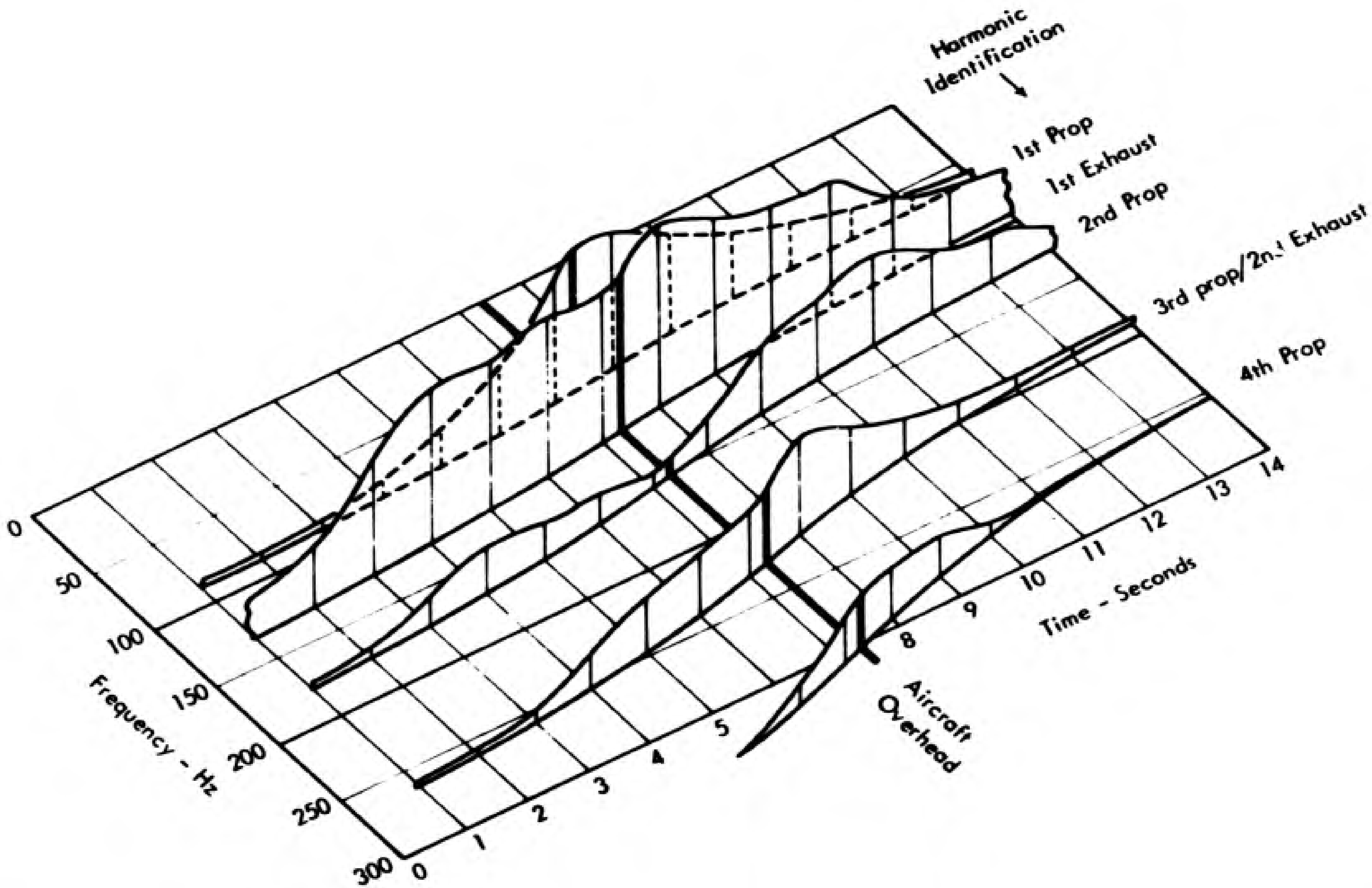



FIGURE 16 - Graphical Representation of Flyover Noise Spectrum History for Piper Cherokee 6 Aircraft Cruising 1000 ft. Over Microphone at 145 m.p.h. and 2300 r.p.m. The individual curves show the time variation of each sound harmonic. Vertical displacement is linearly proportional to the instantaneous sound energy in that harmonic. (The "Aircraft Overhead" marker denotes the time of arrival of the sound generated when the aircraft was overhead - approximately one second earlier).

## IV. ANALYSIS OF RESULTS

Table VII contains the complete set of the results from the paired comparison experiment. Both the nominal and the achieved peak levels of each signal are given; the latter being measured during the experiment by a microphone located alongside the central seat location. The differences can be attributed mainly to errors in recording level during the manufacture of the master tape loops but the slight spectrum errors due to the acoustic characteristics of the room must also contribute. The computed zero crossing points, the intersections of the quadratic curves fitted to the median judgments with the zero axis (see previous section), are listed in the eighth column. On either side of this are the estimated 20th and 80th percentiles about these points. The mean difference between these limits, averaged over all sound pairs and which is the average range within which 60 percent of judgments lie, is 7.2 dB. The locations filled with an asterisk represent indeterminate solutions which arise when, due to an unusual distribution of the actual percentile points, the fitted quadratic curve fails to intersect the zero axis. Figure 17 is a set of fitted curves for six typical comparisons and Figure 18 shows a selection of histograms representing the distribution of subjective results for three comparisons. Comparing these results with Figures 12 and 13 it is clear that, despite the use of narrower confidence limits, the scatter of results in this experiment is considerably less than obtained previously in Pilot Experiment 2. The reasons for this are obscure although the subjects possibly became more consistent with practice. Also the fact that the subjects found aircraft sounds more interesting than the artificial noise signals of previous experiments may have improved their concentration.

By making the appropriate comparisons between references and between each signal and its own reference, each sound is related to the sound pressure level of the Standard Reference (an octave band of pink noise centered on 1000 Hz with a duration of 4 seconds) which is directly or indirectly judged to be equally noisy. The relative sound pressure levels of the Standard Reference (SRL) are given in the last column of Table VII.

### A. Equal Noise Contour for Narrow Band of Noise

Figure 19 shows the results of the paired comparisons of narrow bands of noise. The continuous line is the 21 Noy contour adapted from Reference 14 which is included for comparison. Twenty-one Noys correspond to the noisiness of the octave band at 1000 Hz with a sound pressure level of 84 dB re: .0002 μbar which corresponds to the average level of the current experiment. Two sets of experimental data are plotted. The first six points were obtained, for octave bands of pink noise, during Pilot Experiment 2 reported in Section II. The second nine points are taken from Table VII and are the results of the main experiment for 1/3 octave bands of noise. It may be noted that in both cases the results are not exactly equal noisiness contours since they represent comparisons made at different levels. They have been normalized for the present

BLANK PAGE

TABLE VII

## RESULTS OF PAIRED COMPARISON EXPERIMENT

| Comparison Sound Code | Reference Sound Code | Nominal SPL of Comparison (dB) | (A) Measured SPL of Comparison (dB) | Nominal Mean SPL of Reference (dB) | (B) Measured SPL of Reference (dB) | Paired Comparison Results (C): Estimate of 20th Percentile (dB) | Zero Crossing Point of Median Judgment (dB) | Estimate of 80th Percentile (dB) | (D) Mean SPL of Reference Relative to Comparison \|B - A\| dB | (E) Mean SPL of Standard Reference Judged Equally Noisy to Reference (dB) | (F) Zero Crossing Point = Rounded Value of C | Relative SPL of Standard Reference for Equal Noisiness of Comparison = D + E + F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 140F | A/8 | 75 | 76 | 80 | 81 | 9.0 | 6.7 | 4.1 | +5 | -10 | +7 | +2 |
| 140T1 | 6/8 | 85 | 84 | 80 | 78 | 25.0 | 13.0 | 9.5 | -6 | -11 | +13 | -4 |
| 140T2 | 6/8 | 80 | 77 | 80 | 78 | 12.5 | 9.8 | 6.2 | +1 | -11 | +10 | 0 |
| 140T3 | 6/8 | 80 | 79 | 80 | 78 | 24.3 | 7.7 | 5.9 | -1 | -11 | +8 | -4 |
| 6T | A/8 | 85 | 85 | 80 | 81 | 17.9 | 12.2 | 8.3 | -4 | -10 | +12 | -2 |
| 6F1 | A/8 | 85 | 85 | 80 | 81 | * | 11.2 | * | -4 | -10 | +11 | -3 |
| 6F2 | A/8 | 85 | 86 | 80 | 81 | 18.3 | 11.0 | 10.0 | -5 | -10 | +11 | -4 |
| 6F3 | A/8 | 85 | 86 | 80 | 81 | 14.4 | 10.5 | 8.5 | -5 | -10 | +11 | -4 |
| AT | JN8 | 85 | 86 | 80 | 81 | 7.6 | 5.0 | 2.4 | -5 | -6 | +5 | -6 |
| AF1 | 140/8 | 85 | 85 | 80 | 76 | * | 10.7 | +6.2 | -9 | -3 | +11 | -1 |
| AF2 | 140/8 | 80 | 79 | 80 | 76 | 7.7 | 5.6 | 1.6 | -3 | -3 | +6 | 0 |
| AF3 | 140/8 | 80 | 80 | 80 | 76 | 7.0 | 5.2 | 1.1 | -4 | -3 | +5 | -2 |
| DCT | JN8 | 80 | 82 | 85 | 86 | 2.4 | -0.3 | -3.0 | +4 | -6 | 0 | -2 |
| DCL | JN4 | 85 | 88 | 85 | 84 | 21.1 | 15.5 | 13.0 | -4 | -4 | +16 | +8 |
| 707T | JN8 | 80 | 83 | 85 | 86 | 2.9 | -2.6 | -6.1 | +3 | -6 | -3 | -6 |
| 707L | JN4 | 85 | 86 | 85 | 84 | * | 12.1 | 8.4 | -2 | -4 | +12 | +6 |
| LT1 | JN8 | 80 | 80 | 85 | 81 | 3.6 | 0.5 | -2.0 | +1 | -6 | +1 | -4 |
| LT2 | JN8 | 80 | 79 | 85 | 81 | 3.2 | -2.4 | -3.3 | +2 | -6 | -2 | -6 |
| LL | JN4 | 85 | 89 | 85 | 84 | 10.0 | 7.4 | 2.8 | -5 | -4 | +7 | -2 |
| TC | JN4 | 90 | 91 | 85 | 84 | 8.5 | 6.9 | 1.5 | -7 | -4 | +7 | -4 |
| 140/2 | JN8 | 85 | 80 | 80 | 81 | -3.4 | -6.0 | -7.7 | +1 | -6 | -6 | -11 |
| 140/4 | JN8 | 85 | 83 | 80 | 82 | -0.2 | -3.1 | -5.7 | -1 | -6 | -3 | -10 |
| 140/8 | JN8 | 80 | 76 | 80 | 81 | -0.8 | -2.3 | -5.2 | +5 | -6 | -2 | -3 |
| 140/16 | JN8 | 80 | 78 | 80 | 82 | 3.9 | 1.0 | -0.7 | +4 | -6 | +1 | -1 |

D

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| LT1 | JN8 | 80 | 80 | 85 | 81 | 3.6 | 0.5 | -2.0 | +1 | -6 | +1 | -4 |
| LT2 | JN8 | 80 | 79 | 85 | 81 | 3.2 | -2.4 | -3.3 | +2 | -6 | -2 | -6 |
| LL | JN4 | 85 | 89 | 85 | 84 | 10.0 | 7.4 | 2.8 | -5 | -4 | +7 | -2 |
| TC | JN4 | 90 | 91 | 85 | 84 | 8.5 | 6.9 | 1.5 | -7 | -4 | +7 | -4 |
| 140/2 | JN8 | 85 | 80 | 80 | 81 | -3.4 | -6.0 | -7.7 | +1 | -6 | -6 | -11 |
| 140/4 | JN8 | 85 | 83 | 80 | 82 | -0.2 | -3.1 | -5.7 | -1 | -6 | -3 | -10 |
| 140/8 | JN8 | 80 | 76 | 80 | 81 | -0.8 | -2.3 | -5.2 | +5 | -6 | -2 | -3 |
| 140/16 | JN8 | 80 | 78 | 80 | 82 | 3.9 | 1.0 | -0.7 | +4 | -6 | +1 | -1 |
| 140/32 | JN8 | 75 | 73 | 80 | 81 | +2.9 | -2.0 | * | +8 | -6 | -2 | 0 |
| 6/2 | JN8 | 85 | 78 | 80 | 80 | -3.1 | -8.1 | -11.2 | +2 | -6 | -8 | -12 |
| 6/4 | JN8 | 85 | 87 | 80 | 81 | -0.2 | -3.2 | -6.5 | -6 | -6 | -3 | -15 |
| 6/8 | JN8 | 80 | 73 | 80 | 80 | -3.3 | -6.6 | -8.7 | +2 | -6 | -7 | -11 |
| 6/16 | JN8 | 80 | 80 | 80 | 81 | +0.2 | -3.3 | -7.7 | +1 | -6 | -3 | -8 |
| 6/32 | JN8 | 75 | 75 | 80 | 80 | 1.4 | -2.3 | -8.3 | +5 | -6 | -2 | -3 |
| A/2 | JN8 | 80 | 79 | 80 | 82 | -5.8 | -8.4 | -9.7 | +3 | -6 | -8 | -11 |
| A/4 | JN8 | 85 | 80 | 80 | 80 | 5.1 | 0.9 | -2.2 | 0 | -6 | +1 | -5 |
| A/8 | JN8 | 80 | 81 | 80 | 82 | -2.9 | -6.2 | -9.2 | +1 | -6 | -5 | -10 |
| A/16 | JN8 | 80 | 81 | 80 | 80 | -0.8 | -3.8 | -9.4 | -1 | -6 | -3 | -10 |
| A/32 | JN8 | 75 | 75 | 80 | 82 | -0.8 | -5.2 | -10.2 | +7 | -6 | -5 | -4 |
| JN4 | STD+ | 85 | 84 | 85 | 85 | -2.5 | -5.0 | -7.6 | +1 | 0 | -5 | -4 |
| JN8 | STD | 80 | 81 | 85 | 85 | -5.0 | -9.7 | -12.2 | +4 | 0 | -10 | -6 |
| JN16 | STD | 75 | 76 | 85 | 85 | -7.7 | -9.6 | -13.9 | +9 | 0 | -10 | -1 |
| JN32 | STD | 75 | 76 | 85 | 85 | -6.0 | -13.0 | * | +9 | 0 | -13 | -4 |
| STD | 30B63 | 75 | 75 | 85 | 85 | 5.7 | 1.9 | -1.5 | +10 | 0 | +2 | -12 |
| 30B31 | 30B63 | 85 | 85 | 80 | 80 | 6.8 | 3.6 | -0.4 | -5 | -12 | +4 | -13 |
| 30B125 | 30B63 | 75 | 75 | 85 | 85 | 2.2 | -2.5 | -6.6 | +10 | -12 | -3 | -5 |
| 30B250 | STD | 85 | 85 | 85 | 85 | -2.7 | -6.4 | -7.5 | 0 | 0 | -6 | -6 |
| 30B500 | 30B63 | 75 | 74 | 85 | 85 | 4.4 | -1.7 | -4.5 | +11 | -12 | -2 | -3 |
| 30B1K | STD | 85 | 85 | 85 | 85 | 2.4 | -1.1 | -4.0 | 0 | 0 | -1 | -1 |
| 30B2K | STD | 75 | 77 | 85 | 85 | +0.8 | -3.7 | -6.4 | +8 | 0 | -4 | +4 |
| 30B4K | STD | 75 | 78 | 85 | 85 | 20.4 | 12.9 | 9.2 | +7 | 0 | +13 | +20 |
| 30B8K | STD | 80 | 80 | 85 | 85 | 5.9 | 2.4 | -1.2 | +5 | 0 | +2 | +7 |

+STD = Standard Reference Sound

B

BLANK PAGE

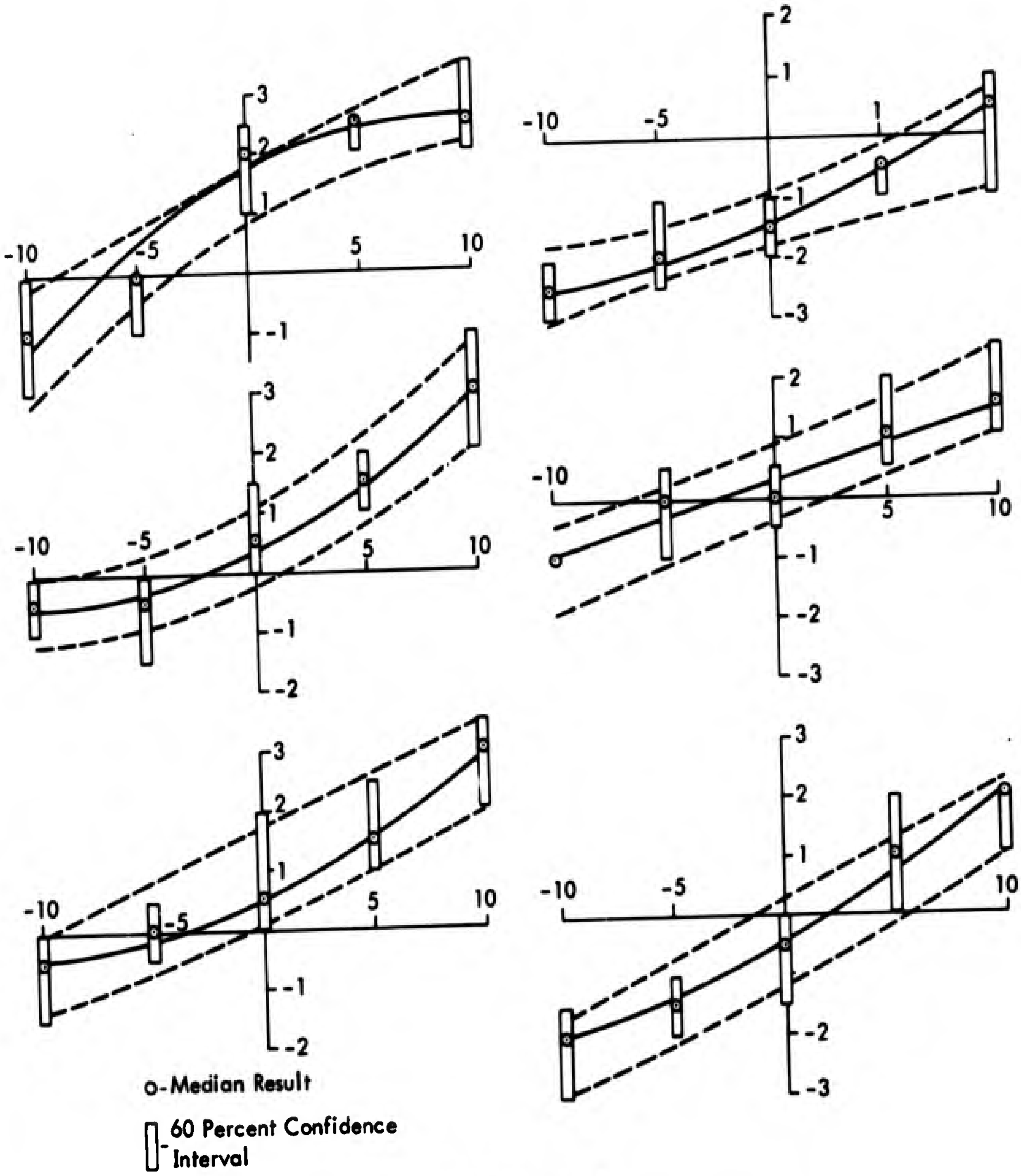



FIGURE 17 - Computed Curve-Fits For Some Typical Sound Comparisons from the Main Experiment

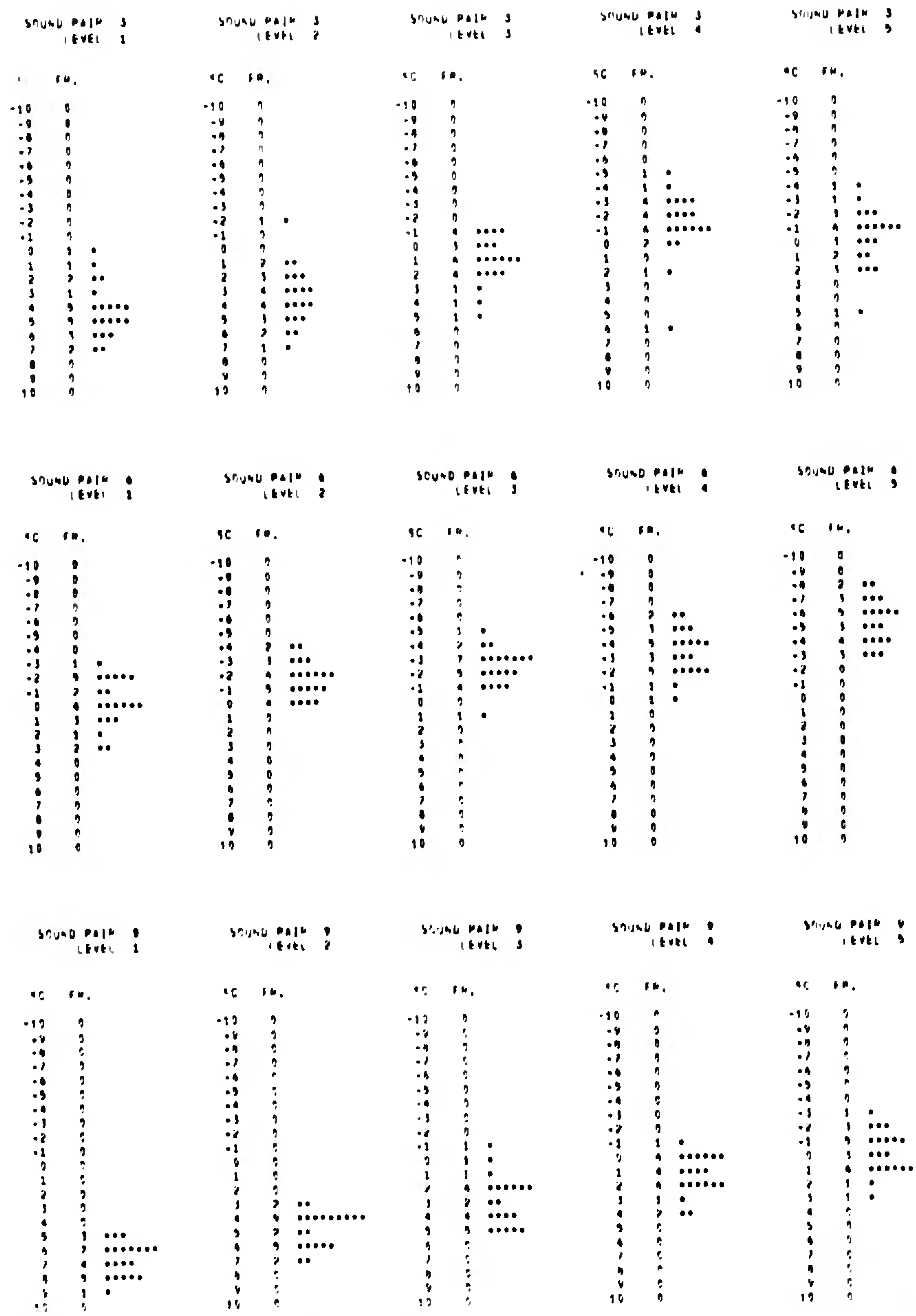


FIGURE 18 - Histograms of Judgment Results for 3 Typical Comparisons From Main Experiment

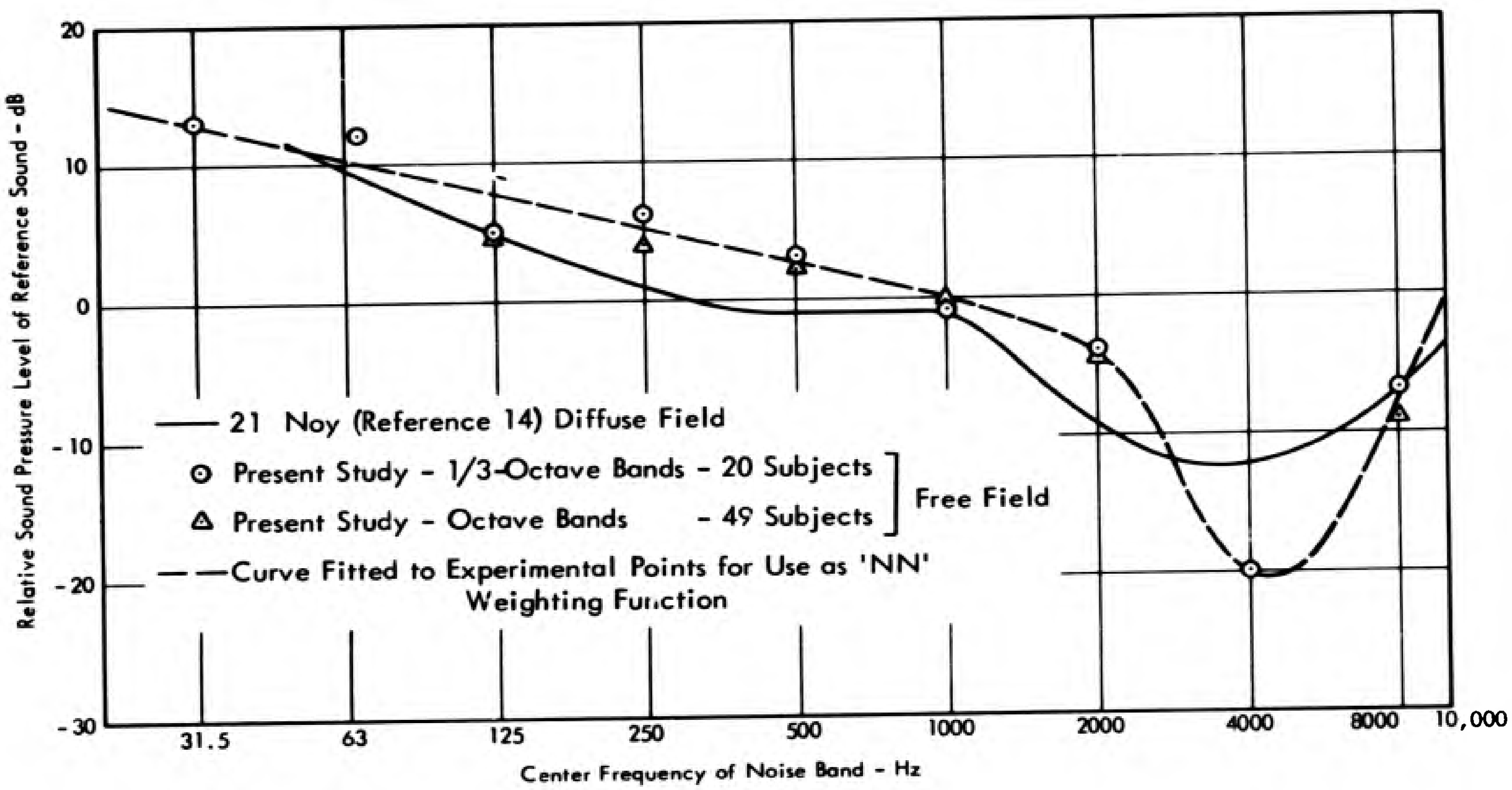



FIGURE 19 - Equal Noisiness Contours for Narrow Bands of Noise

purposes by plotting them as sound pressure levels relative to that of the Standard Reference Sound. As explained previously, this variation was necessary to maintain a reasonable signal-to-noise ratio in the tape recording system. Nevertheless, it is encouraging to note the close agreement of the two sets of data at the coincident frequencies which confirms the repeatability of results. The two experiments were conducted on two completely separate occasions using different numbers of subjects and different sound reproduction equipment. The differences between the Noy contour and the results of the present study are substantial although the result for the 1/3 octave at 4000 Hz is the major difference at high frequencies. The broken line has been fitted to the data points for use as a weighting function in the calculation of a weighted sound pressure level referred to as dB(NN). This noise rating method is included among those studied in the following sections.

## B. Subjective Evaluation of the Aircraft Sounds

The levels at which the recorded aircraft signals were played to the subjects differed from the actual levels at the recording site because of the necessity to compress their total dynamic range. Making the assumption that the Standard Reference Level (SRL) for judged equal noisiness varies linearly with the aircraft signal level, Figure 20 has been prepared which shows the equivalent judged noisiness of all 20 real flyover sounds in terms of the SRL. These correspond to the actual noise leveis heard, out of doors, for the flight condition and observer locations noted.

### 1. Correlation of Noise Rating Scales and Subjective Evaluation

The maximum values of the noise ratings under consideration were obtained from the computed 1/2-second time histories by inspection and are recorded in Table VIII. The slight differences between the computed Overall Sound Pressure Levels and the measured peak values listed in Table VII are probably explained by the digitization of the 1/3 octave band level histories at 1/2 second time intervals, together with rounding errors. In order to minimize the effect of these minor differences the final "equal noisiness" level of the Standard Reference Signal given in the final column of Table VIII is obtained by referring the judged relative levels given in the final column of Table VII to the computed, rather than the measured, overall sound pressure levels of the comparison signals.

Using each of the various noise ratings as the independent variable, linear regression lines were fitted through the corresponding values of the Standard Reference Level (SRL) as shown in Figures 21 through 25. Given an ideal rating method and a "perfect" set of experimental results all the points would fall on a continuous curve, which, because the scales, both ordinate and abscissa, are logarithmic, might be expected to be a straight line. The scatter of the actual data about the best fitting straight line can therefore be attributed to a combination of experimental error and the inadequacy of the rating method. The same set of dependent variables are common to each regression so that the experimental errors remain constant. The accuracy of the

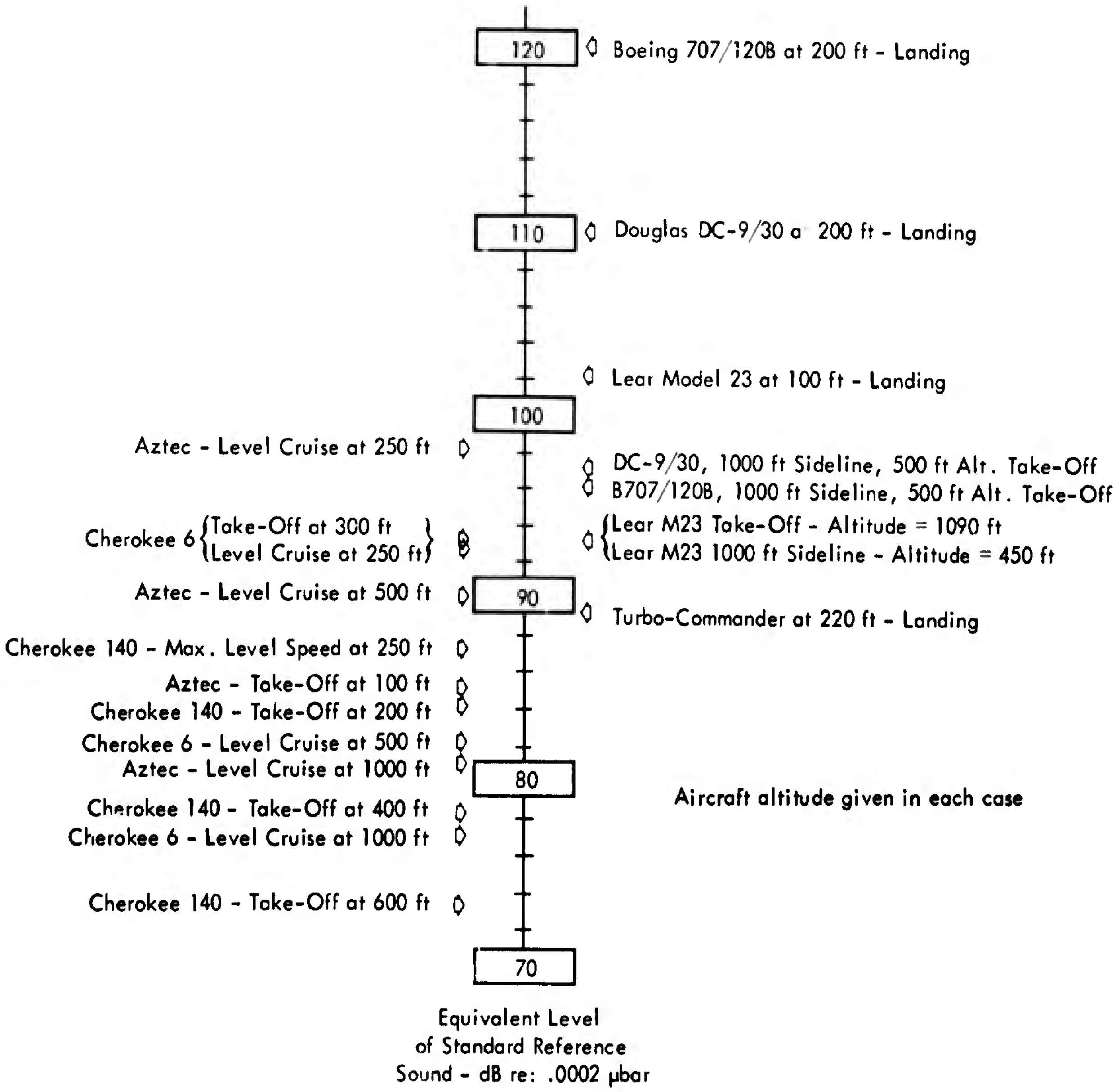



FIGURE 20 - Judged Noisiness of Aircraft Flyover Sounds Extrapolated from Subjective Test Data

## TABLE VIII

## MAXIMUM VALUES OF NOISE RATINGS FROM COMPUTED 1/2 SECOND TIME HISTORIES

| | Aircraft Sound Code | OASPL | dB(A) | dB(B) | dB(C) | dB(N) | dB(NN) | dB(J) | dB(K) | PNdB | PNdBF | S-Phon | Z-Phon | SRL |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Piston-Engined Aircraft Flyovers | 140F | 78 | 72 | 76 | 78 | 76 | 74 | 75 | 75 | 83 | 87 | 82 | 86 | 80 |
| | 140T1 | 85 | 75 | 81 | 85 | 81 | 79 | 80 | 81 | 87 | 93 | 86 | 89 | 81 |
| | 140T2 | 78 | 70 | 76 | 78 | 75 | 73 | 75 | 75 | 82 | 87 | 81 | 82 | 78 |
| | 140T3 | 77 | 70 | 75 | 80 | 74 | 74 | 73 | 74 | 81 | 87 | 80 | 83 | 73 |
| | 6T | 86 | 79 | 84 | 86 | 84 | 81 | 83 | 83 | 89 | 93 | 88 | 91 | 84 |
| | 6F1 | 86 | 78 | 83 | 86 | 83 | 81 | 82 | 83 | 90 | 94 | 88 | 91 | 83 |
| | 6F2 | 87 | 75 | 83 | 87 | 82 | 80 | 81 | 83 | 88 | 91 | 87 | 88 | 83 |
| | 6F3 | 88 | 76 | 84 | 88 | 83 | 81 | 82 | 83 | 89 | 92 | 87 | 88 | 84 |
| | AT | 87 | 78 | 83 | 87 | 84 | 85 | 84 | 85 | 91 | 93 | 89 | 94 | 81 |
| | AF1 | 85 | 75 | 81 | 85 | 82 | 86 | 82 | 83 | 90 | 91 | 89 | 93 | 84 |
| | AF2 | 80 | 71 | 76 | 80 | 77 | 78 | 76 | 77 | 84 | 86 | 84 | 87 | 80 |
| | AF3 | 82 | 70 | 77 | 82 | 77 | 77 | 76 | 78 | 84 | 89 | 83 | 87 | 80 |
| Jet Aircraft | DCT | 81 | 78 | 81 | 81 | 81 | 79 | 81 | 81 | 87 | 89 | 86 | 90 | 80 |
| | DCL | 89 | 87 | 87 | 88 | 94 | 99 | 94 | 95 | 101 | 102 | 98 | 102 | 96 |
| | 707T | 83 | 77 | 81 | 83 | 81 | 80 | 81 | 81 | 88 | 89 | 87 | 91 | 77 |
| | 707L | 84 | 85 | 84 | 84 | 93 | 96 | 93 | 94 | 99 | 104 | 96 | 97 | 90 |
| | LT1 | 78 | 74 | 77 | 78 | 78 | 75 | 77 | 77 | 84 | 84 | 83 | 87 | 74 |
| | LT2 | 79 | 76 | 78 | 79 | 79 | 77 | 79 | 79 | 86 | 90 | 84 | 88 | 73 |
| | LL | 88 | 84 | 86 | 88 | 89 | 91 | 88 | 89 | 96 | 99 | 94 | 99 | 86 |
| * | TC | 92 | 81 | 87 | 92 | 86 | 87 | 85 | 87 | 95 | 99 | 93 | 94 | 88 |
| 'Shaped' Flyover Signals | 140/2 | 80 | 70 | 77 | 80 | 76 | 74 | 76 | 76 | 83 | 85 | 84 | 84 | 69 |
| | 140/4 | 83 | 73 | 80 | 83 | 79 | 77 | 79 | 80 | 86 | 89 | 86 | 87 | 73 |
| | 140/8 | 76 | 67 | 74 | 76 | 73 | 71 | 72 | 73 | 79 | 82 | 80 | 81 | 73 |
| | 140/16 | 78 | 68 | 75 | 78 | 74 | 72 | 73 | 74 | 80 | 83 | 80 | 81 | 77 |
| | 140/32 | 73 | 64 | 70 | 73 | 70 | 67 | 69 | 70 | 76 | 78 | 77 | 78 | 73 |
| | 6/2 | 78 | 64 | 72 | 78 | 71 | 70 | 70 | 72 | 77 | 77 | 77 | 77 | 66 |
| | 6/4 | 87 | 72 | 81 | 86 | 80 | 79 | 79 | 81 | 86 | 86 | 85 | 85 | 72 |
| | 6/8 | 78 | 66 | 74 | 78 | 73 | 71 | 72 | 73 | 79 | 79 | 79 | 79 | 67 |
| | 6/16 | 80 | 69 | 76 | 80 | 75 | 73 | 74 | 75 | 81 | 82 | 81 | 82 | 72 |
| | 6/32 | 75 | 65 | 71 | 75 | 71 | 69 | 70 | 71 | 77 | 78 | 77 | 79 | 72 |
| | A/2 | 79 | 65 | 73 | 79 | 72 | 71 | 71 | 73 | 70 | 79 | 78 | 78 | 68 |
| | A/4 | 80 | 70 | 77 | 80 | 76 | 74 | 76 | 76 | 82 | 84 | 83 | 83 | 75 |
| | A/8 | 81 | 67 | 76 | 81 | 75 | 73 | 74 | 76 | 81 | 82 | 81 | 80 | 71 |
| | A/16 | 81 | 64 | 74 | 81 | 73 | 72 | 72 | 74 | 70 | 79 | 79 | 78 | 71 |
| | A/32 | 75 | 65 | 72 | 75 | 71 | 69 | 71 | 71 | 77 | 78 | 78 | 79 | 71 |

* Turbo-Prop

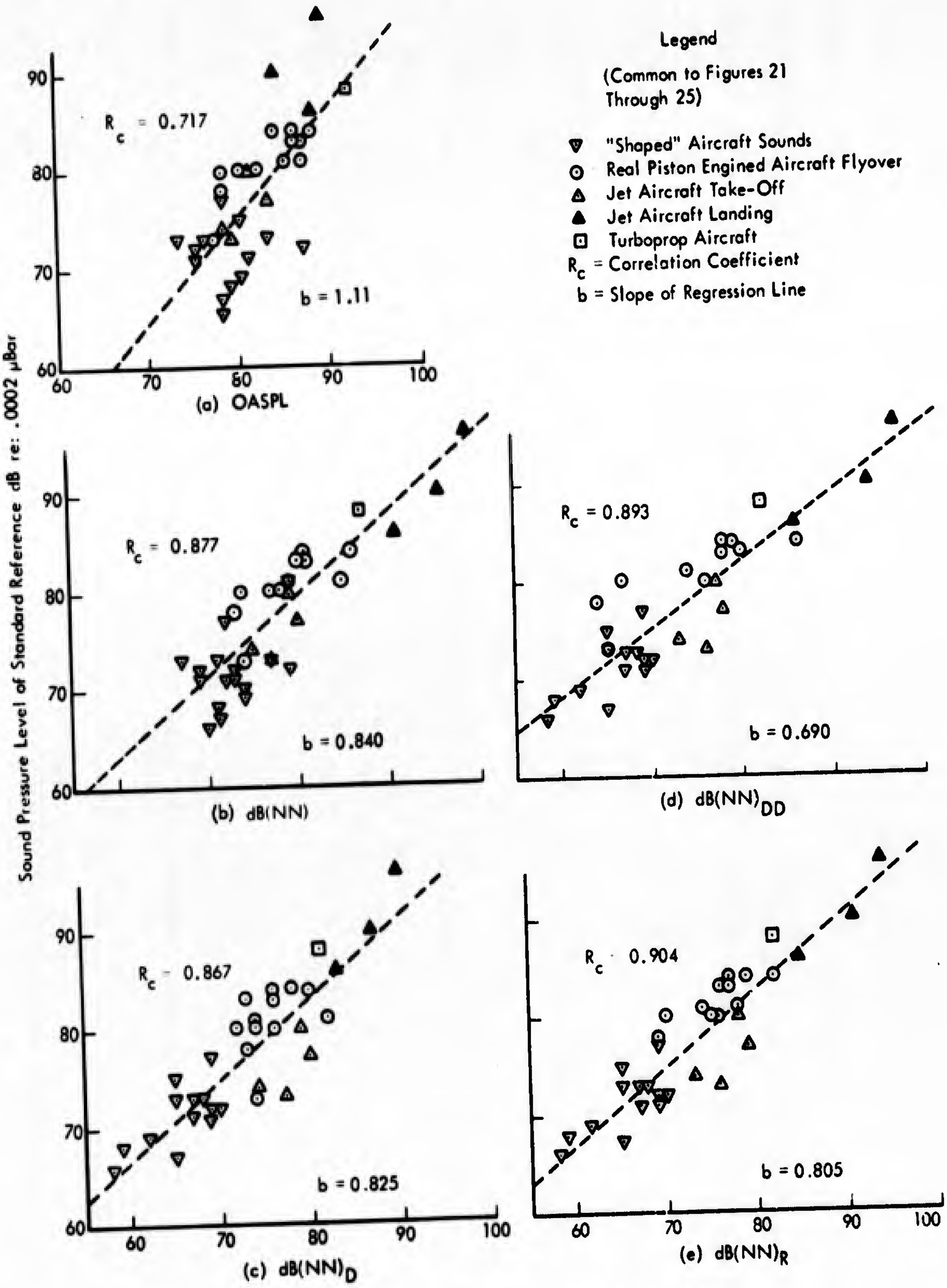



FIGURE 21 - Regression Analyses of Various Ratings. Abscissa Definitions in Glossary

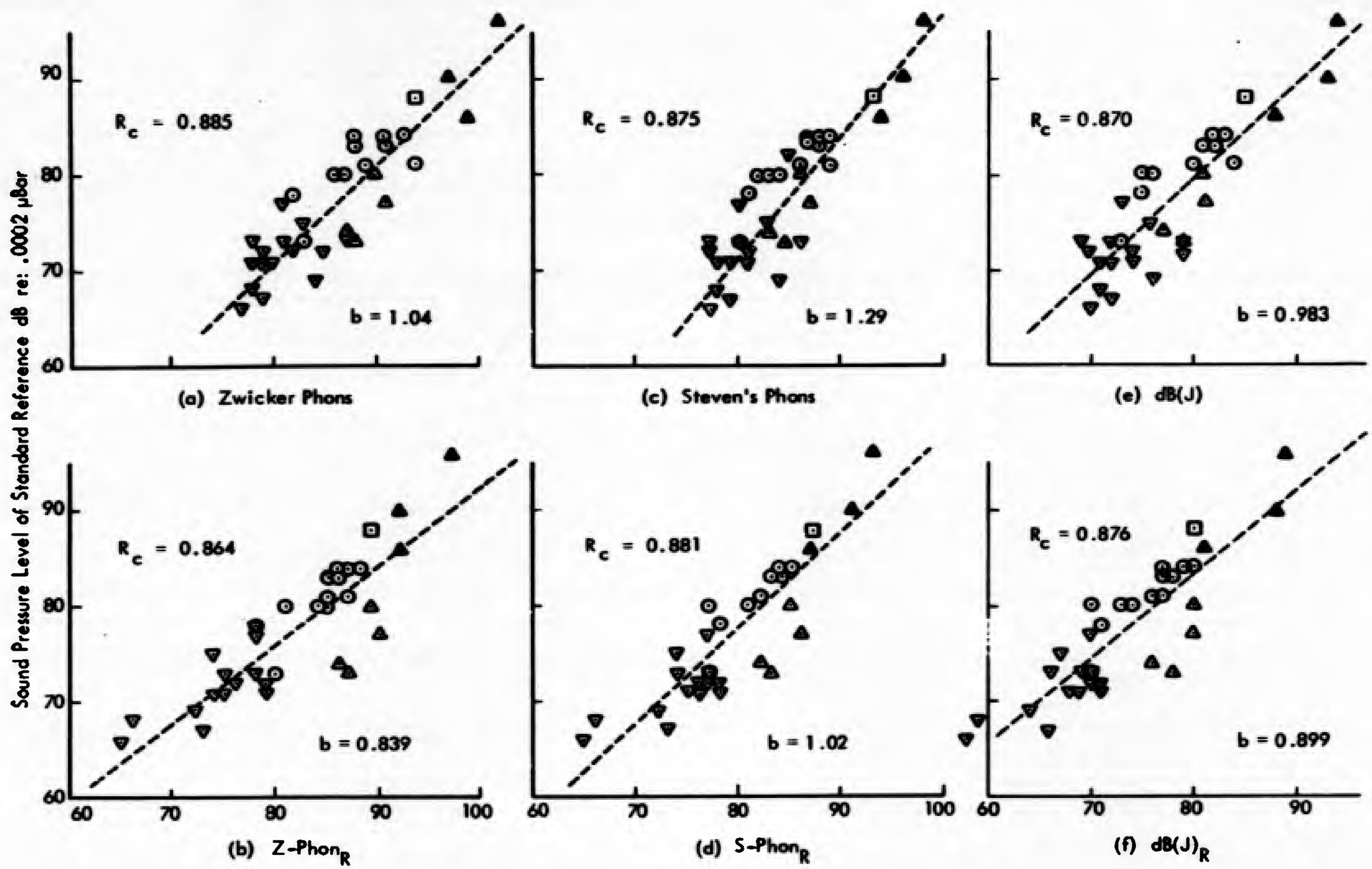



FIGURE 22 – Regression Analyses of Various Ratings. Abscissa Definitions in Glossary

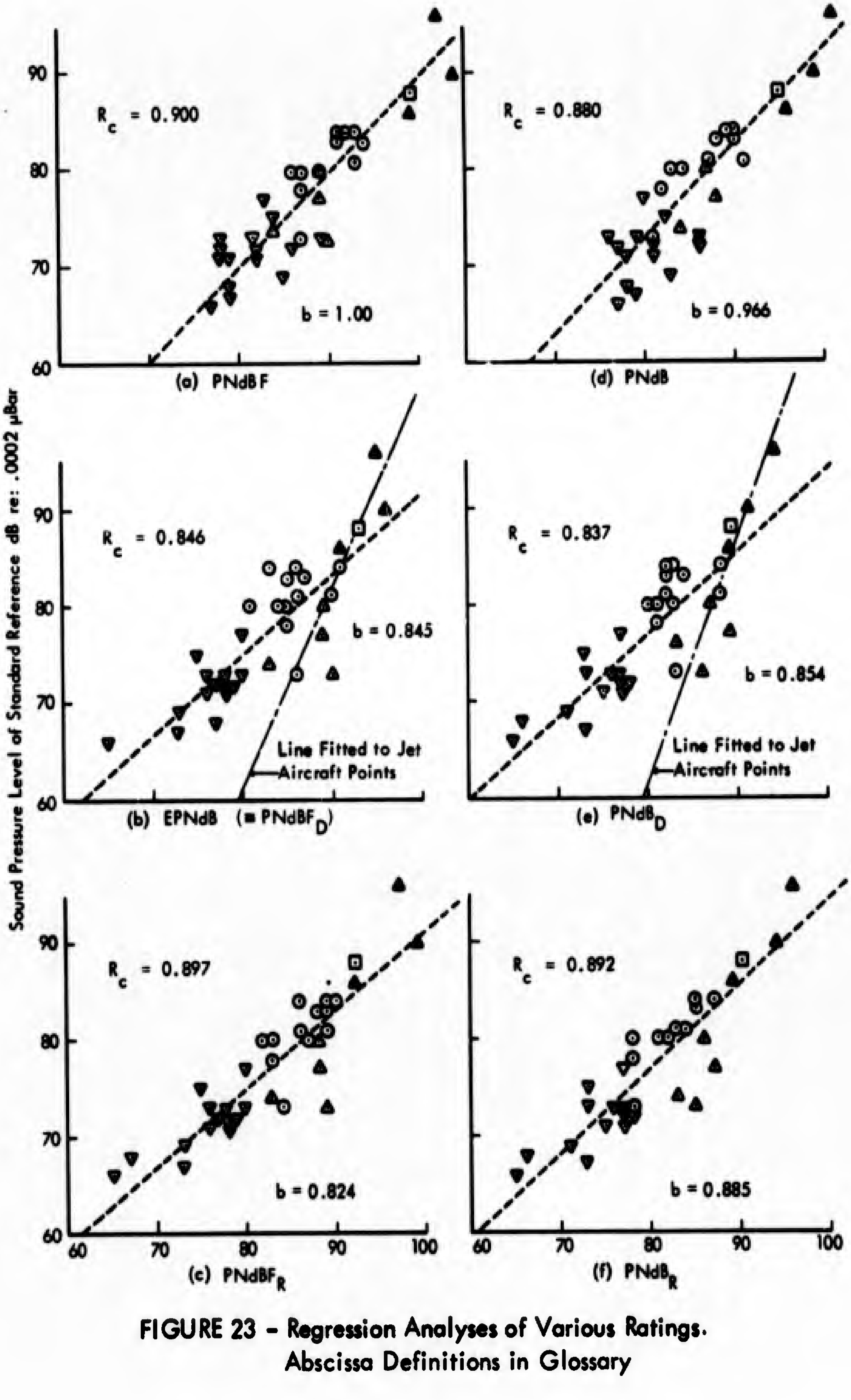



FIGURE 23 – Regression Analyses of Various Ratings. Abscissa Definitions in Glossary

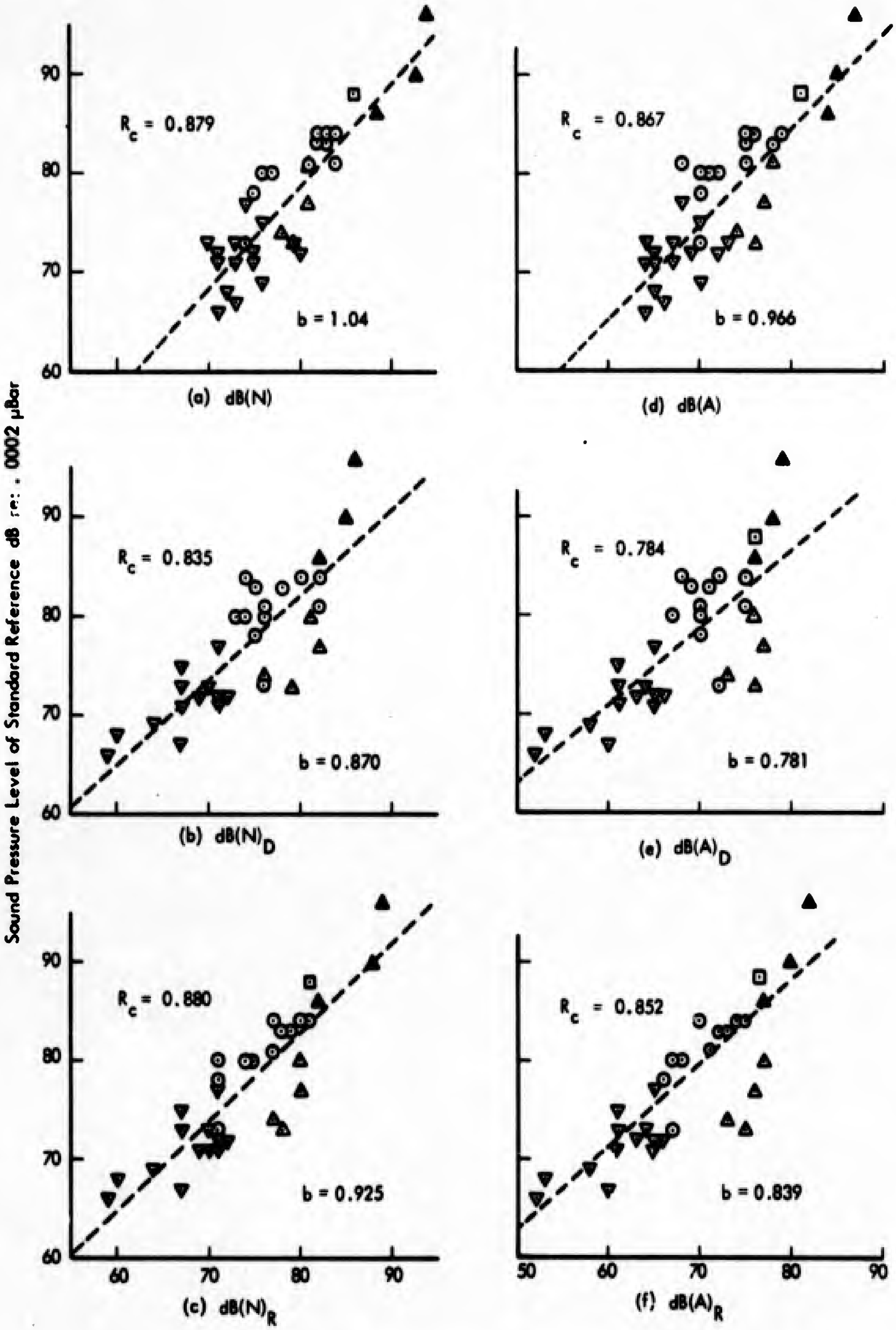



FIGURE 24 – Regression Analyses of Various Ratings. Abscissa Definitions in Glossary

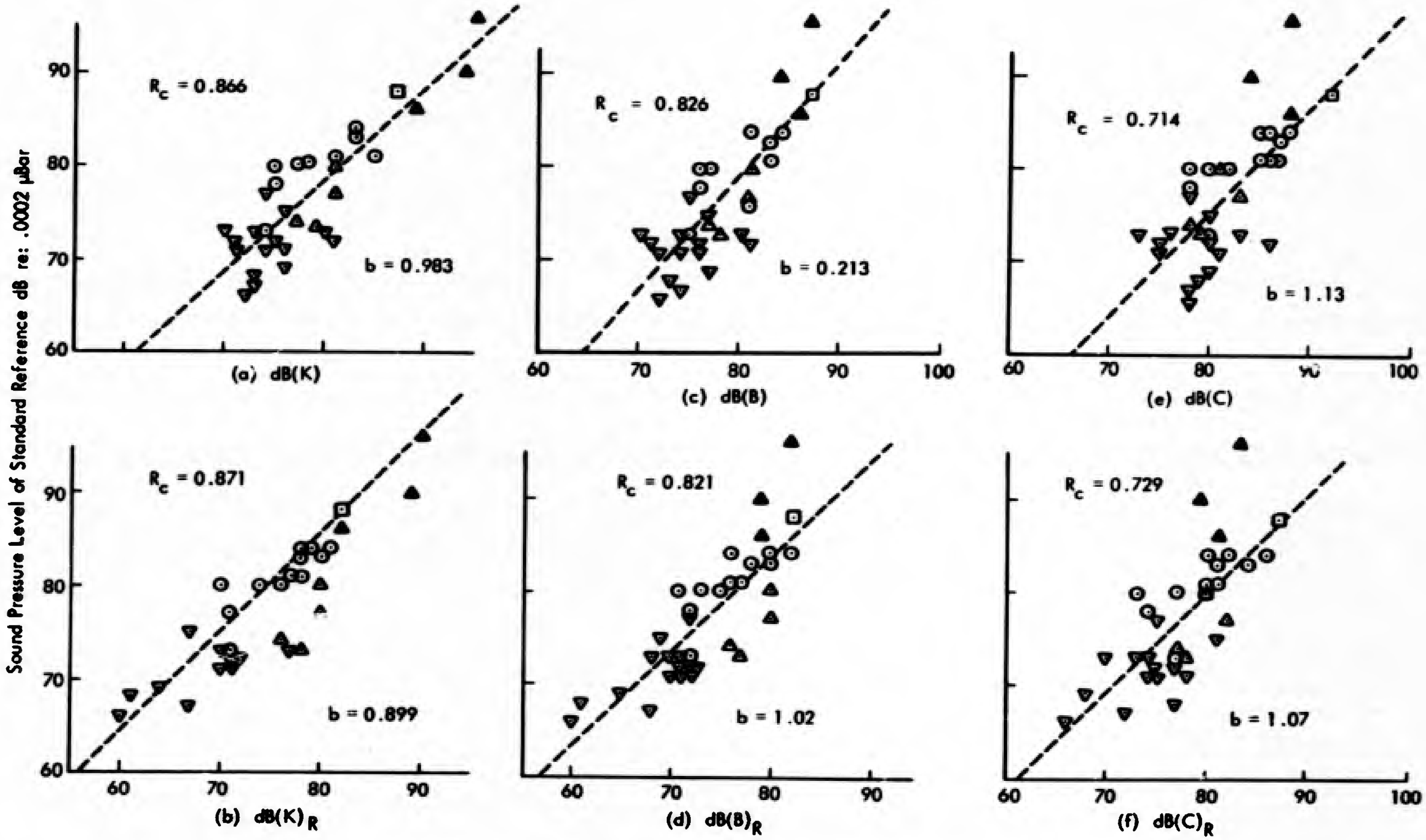



FIGURE 25 – Regression Analyses of Various Ratings. Abscissa Definitions in Glossary

various rating methods can thus be directly compared by comparing the scatter of the points about the regression lines. Standard methods of statistical analysis are applied to estimate the "goodness of fit" of the regression lines. Two parameters are calculated; the correlation coefficient $R_c$ which has a value of unity if the scatter is zero, and the variance about the regression line $S^2$. Mathematically these are expressed as follows:

$$R_c = \frac{N\Sigma x_i y_i - \Sigma x_i \Sigma y_i}{\sqrt{[N\Sigma x_i^2 - (\Sigma x_i)^2]\ [N\Sigma y_i^2 - (\Sigma y_i)^2]}}$$

$$S^2 = \frac{\Sigma (y_i - a - b x_i)^2}{N - 2}$$

N is the number of data points, $x_i$ are the values of the independent variable (the rating unit), and $y_i$ are the values of the dependent variable (the judged results). The summations are carred out over all the i's, and a and b are the coefficients of the regression line ($y' = a + b x$) which are obtained by the method of least squares. Physically, the square of the correlation coefficient is the fraction of the total variance of the judged results accounted for by the regression on the rating method, so that the residue is a measure of the dispersion about the regression line. The variance $S^2$ is simply the mean square error of the data about the line in the y direction.

Table IX shows the results of the analysis of variance of regression for the 12 basic noise rating methods. Regression lines have been fitted to all 35 data points corresponding to 20 real flyover signals and 15 artifically shaped steady aircraft signals. The rating scales are listed in rank order of $R_c$ and the table indicates the relative performance of these methods in predicting the subjective noisiness of this particular set of signals which includes the sounds of turbojet, turbofan, turboprop and piston engined aircraft in various flight configurations. Essentially we are examining the usefulness of the methods as general purpose estimators. This point is further discussed in the first paragraph of Section IVC, (p. 72).

For 35 data points the correlation coefficient must exceed a critical value of 0.430 if we are to conclude that there is, in fact, some positive correlation between the judged results and the noise ratings at the 5 percent level of significance. This is the statement of the fact that even if the correlation were zero, implying a random set of results, there is still a 5 percent probability that the correlation would exceed 0.430. Since all the coefficients in Table IX substantially exceed this value, all ratings are statistically related to the measured subjective reaction.

TABLE IX

RELATIVE PERFORMANCE OF PEAK RATINGS IN PREDICTING JUDGED NOISINESS OF AIRCRAFT SOUNDS

| Rank | Correlation Coefficient $R_c$ | $S^2$** | Method | Percentage Probability That Methods are Significantly Different* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.900 | 9.21 | PNdBF | 66 | 69 | 70 | 72 | 72 | 76 | 78 | 78 | 93 | 99 | 99 |
| 2 | 0.885 | 10.7 | Z-Phons | | 53 | 54 | 56 | 57 | 60 | 63 | 64 | 85 | 98 | 98 |
| 3 | 0.880 | 11.0 | PNdB | | | 51 | 53 | 53 | 57 | 60 | 61 | 83 | 98 | 98 |
| 4 | 0.879 | 11.1 | dB(N) | | | | 52 | 53 | 56 | 59 | 60 | 83 | 98 | 98 |
| 5 | 0.877 | 11.3 | dB(NN) | | | | | 51 | 55 | 57 | 58 | 81 | 98 | 98 |
| 6 | 0.875 | 11.4 | S-Phons | | | | | | 53 | 56 | 57 | 81 | 98 | 98 |
| 7 | 0.870 | 11.8 | dB(J) | | | | | | | 52 | 53 | 78 | 97 | 97 |
| 8 | 0.867 | 12.1 | dB(A) | | | | | | | | 50 | 76 | 97 | 97 |
| 9 | 0.866 | 12.2 | dB(K) | | | | | | | | | 75 | 97 | 97 |
| 10 | 0.826 | 15.5 | dB(B) | | | | | | | | | | 89 | 89 |
| 11 | 0.717 | 25.7 | OASPL | | | | | | | | | | | 51 |
| 12 | 0.714 | 23.9 | dB(C) | | | | | | | | | | | |

For example we may be 85% confident that the Zwicker Phon method is a superior predictor to dB(B)

* Where 50% probability implies that the two methods are equally good.

** Regression Line Variance

(These results are selected from the comprehensive tabulations in Table XVII.)

To make a comparison between any two methods, it is necessary to conduct an "F-test" of their respective variances about regression. A description of the theory behind the F-test can be found in most texts on statistics and it suffices to state here that it is designed to determine the probability that the two variances are significantly different given a normal distribution of errors in each case. The probability that we can reject the hypothesis that the two variances are equivalent is (Reference 12)

$$P(f) = \int_F^{\infty} \frac{(f-1)\,!}{\left(\frac{f-2}{2}\,!\right)} f^f F^{f-2/2} f(1+F)^{-f} \, dF$$

where F is the F-ratio $S_2^2/S_1^2$, the ratio of the two variances and f is the number of degrees of freedom which must be equal for both cases. In fact, f = N-2 where N is the number of data points. Alternatively, the probability that the two variances are statistically different is 1 - P(F). The above relationship has been calculated as a function of F for f = 33 and is plotted in Figure 26. Using this curve the probability that any rating method is a better predictor than any other has been calculated and listed in Table IX. It should be remembered that a probability of 50 percent implies that the two methods are equivalent and a probability of 100 percent means that one method is superior to the other beyond any possible doubt. Expressed in another way, 90 percent probability implies favorable odds of 9 to 1.

It is as well to discuss the significance of these probabilities at this point, since considerable use is made of them in subsequent paragraphs. It was mentioned earlier that the displacement of any data point from the fitted regression line is the sum of two errors; the experimental error and the error due to the inadequacy of the rating method. This is based on the assumption that the true relationship between the two variables, of which the regression line is an estimate, is in fact linear. Accepting this, the error of the ith point can be written

$$\epsilon_i = E_i + M_i$$

where $E_i$ is the "experimental error" and $M_i$ is the rating error. Thus, in comparing two rating methods, we are forming the F-ratio

$$F_{12} = \frac{\Sigma(E1_i + M1_i)^2}{\Sigma(E2_i + M2_i)^2}$$

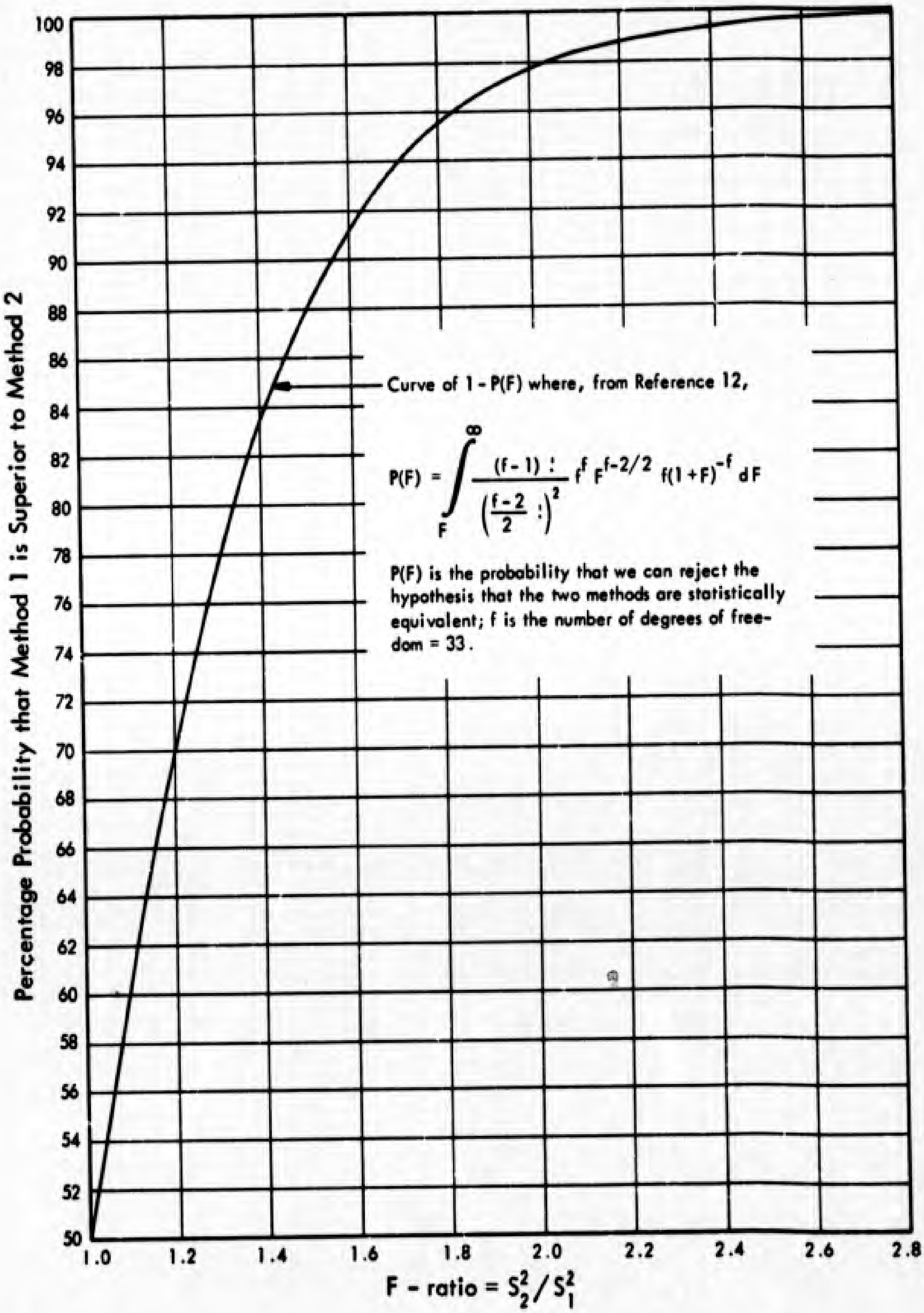



FIGURE 26 - Probability of Significant Differences for 35 Data Points as a Function of F-Ratio

If the experiment is "perfect" so that the errors $E_i$ are zero, or at least small compared to the $M_i$, then the F-ratio expresses the ratio of the errors for which we are looking, namely, the errors in the rating methods. However, the "experimental" error $E_i$ is the combination of two types of variance; the true experimental error due to measurement deficiencies and the very large errors due to human psychological and physiological variations. The latter are reflected in the large deviations found in the recorded subjective judgment data. It is thus probable that the $E_i$ are in fact large compared to the $M_i$ so that the F ratios are considerably smaller than the actual ratios of the rating method variances. The consequence of this is that the confidence with which we can discriminate between the performance of any two methods is low. For this reason, the usual practice of only drawing statistical conclusions at the 95 percent, or greater, confidence level must be rejected if any useful conclusions are to be drawn at all.

For example, referring to Table IX we see that only in relatively few cases does the probability exceed 95 percent. This is only found in fact when comparing various methods with the worst two, OASPL and dB(C). These two methods are so poor that, as shown by the correlation coefficients, their regression lines only account for approximately 50 percent of their variance. This does not mean that we cannot attach significance to the lower probabilities. It is obviously significant that the pure tone correction applied to PNdBF makes it 69 percent probable that this corrected method is an improvement over PNdB. This is equivalent to saying that, on the average, we can expect PNdBF to give us better results than PNdB twice out of every three comparisons; a worthwhile improvement.

2. Effects of Duration

Previous work has shown that signal duration has a definite influence on the subjective noisiness of sound. Kryter and Pearson (Reference 3) and Pearson (Reference 4) conducted experiments using broadband sound which suggested that Perceived Noise Levels should be modified by a correction which varied between 2 dB and 6 dB per duration doubling, depending on the absolute range of duration concerned. The latest duration correction in the proposed FAA Certification Procedures (Reference 5) is

$$\Delta PNdB = 10 \log_{10} \left( \frac{T_{10}}{15} \right)$$

where $\Delta$PNdB is an increment to be added to the uncorrected peak PNdBF, and $T_{10}$ is the time interval between the 10 dB-down points in the PNdBF time history of the aircraft flyby noise. This corresponds to a duration correction of 3 dB per duration doubling and represents a good average of experimental results.

Figure 27(a) shows how the evaluated noisiness of the shaped aircraft signals of the present study varies with duration. The total duration of the signals, during which the overall sound pressure level was varied linearly with time from -20 dB to 0 to -20 dB from the peak, varied between 2 and 32 seconds so that the corresponding values of $T_{10}$ range from 1 to 16 seconds. The ordinate scale in this figure is SRL minus dB(NN) where dB(NN) is selected as a representative noise rating unit. In this case, any rating method would produce the same result since the frequency spectrum of the sound is constant. The increase of judged noisiness with duration is apparent, although the increase ranges from 3 dB per doubling for the Cherokee 140 sound to 1.75 dB per doubling for the Aztec sound.

Figure 27(b) shows the corresponding result for real flyover signals recorded for the same set of aircraft. Admittedly, the scatter of the points about the fitted lines is greater in two cases, and the significance of the particular choice of dB(NN) is now more important since due to different velocities, there are some changes in the frequency spectrum. Also the $T_{10}$ values now correspond to the 10 dB down points of the NN-weighted SPL histories of the various flyovers which are tabulated in Table X.

TABLE X

DATA FOR FIGURE 27(b)

| Aircraft | Code | Velocity (mph) | Altitude Over Mic | Flight Configuration | $T_{10}$ (sec) |
|---|---|---|---|---|---|
| Cherokee 140 | 140F | 120 TAS | 250 | Level Max. Speed | 8.7 |
| | 140T1 | 70 TAS | 200 | Take-Off | 4.7 |
| | 140T2 | 85 TAS | 400 | Take-Off | 16.9 |
| | 140T3 | 85 TAS | 600 | Take-Off | 23.0 |
| Cherokee 6 | 6T | 90 TAS | 300 | Take-Off | 4.2 |
| | 6F1 | 160 TAS | 250 | Level Cruise | 2.2 |
| | 6F2 | 140 TAS | 500 | Level Cruise | 5.7 |
| | 6F3 | 145 TAS | 1000 | Level Cruise | 11.5 |
| Aztec | AT | 135 TAS | 100 | Take-Off | 7.8 |
| | AF1 | 175 TAS | 250 | Level Cruise | 2.0 |
| | AF2 | 165 TAS | 500 | Level Cruise | 5.6 |
| | AF3 | 170 TAS | 1000 | Level Cruise | 10.5 |

It should be noted that three signals for each aircraft type were recorded for approximately equal flight conditions so that duration differences are mainly due to aircraft altitude. However, the most noticable feature of these results is that there is no

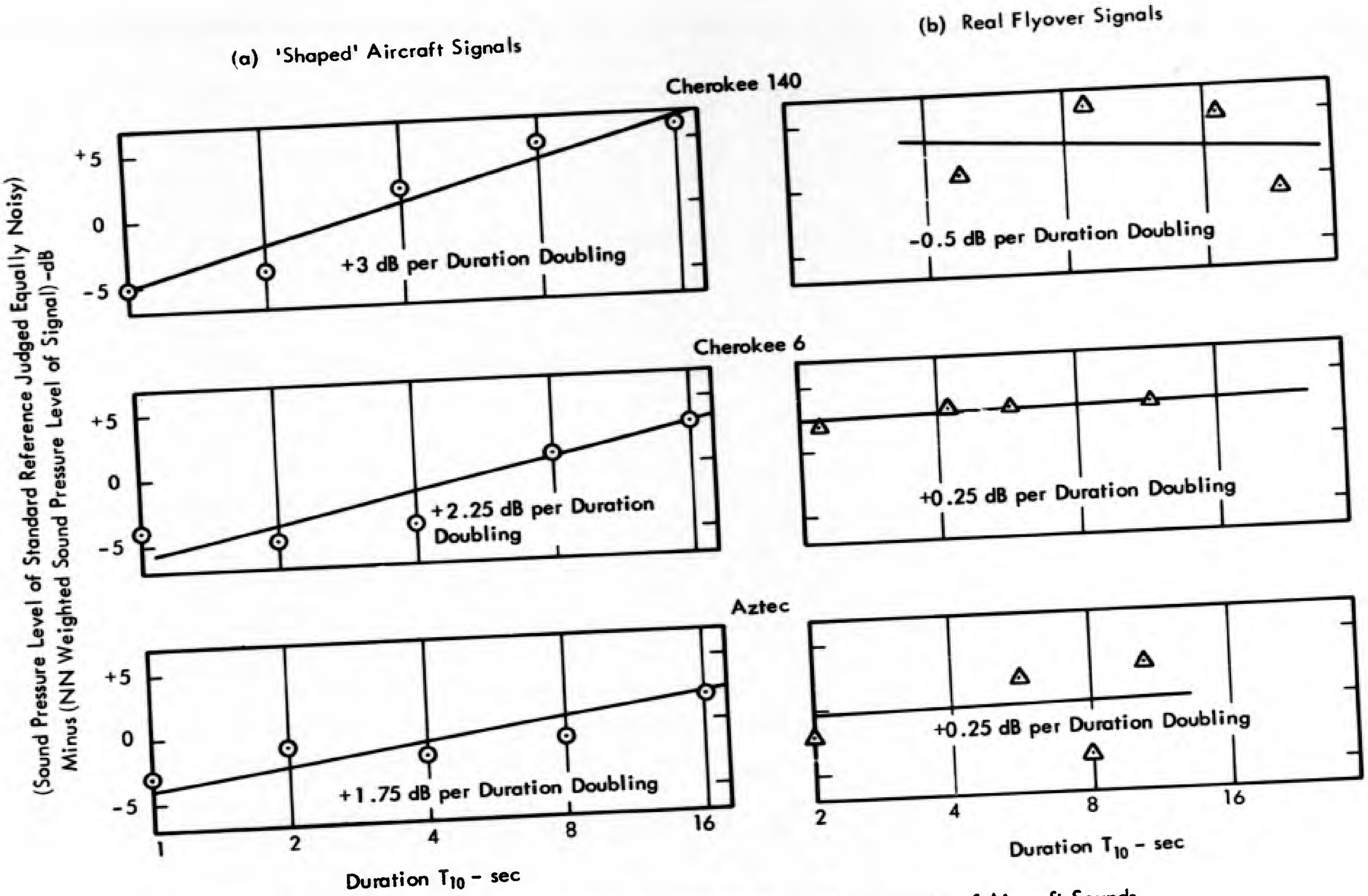



FIGURE 27 – Effect of Signal Duration on Judged Noisiness of Aircraft Sounds
($T_{10}$ = 10 dB Down Duration)

obvious effect of duration. The fitted lines have slopes of -0.5 dB, 0.25 dB and 0.25 dB per duration doubling.

To further investigate the significance of duration in the subjective evaluation of all the aircraft sounds under study, corrections equivalent to that proposed in Reference 5, namely, 3 dB/duration doubling, was also applied to the units PNdB, PNdBf, dB(N) and dB(A). The actual correction used in all cases is

$$\Delta = 10 \log_{10} \left(\frac{T_{10}}{15}\right)$$

where $T_{10}$ is now the 10 dB-down duration of the particular rating unit to which it is being applied. Figure 28 shows that differences in the time histories of the various scales do cause differences in the appropriate values of $T_{10}$. The values of $T_{10}$ for each method used, and each signal, are tabulated in Table XI. Only the real flyover signals are included; the durations of the 15 shaped sounds are equal to 0.5 times the specified durations for all methods. The duration corrected rating units for all signals and each method are listed in Table XII. The duration corrected units are identified by the subscript D.

A regression analysis on each set of D-corrected data revealed that in all five cases the performance of the rating method was degraded. The correlation coefficients computed for the uncorrected and corrected data are shown in Table XIII together with the probabilities of degradation. This is significant at the 88 percent confidence level in the case of PNdBF and at the 90 percent level for dB(A). The effects are illustrated in Figures 21, 23, and 24 which compares the distribution of the data points for both uncorrected and corrected cases. It is clear from these plots that the variance of the results for the special shaped sound has generally been reduced by the duration correction but that the increased scatter of the real flyover results outweighs this gain to give an overall deterioration.

### 3. Effects of Doppler Frequency Shift

The characteristics of the sound signal so far considered by the various rating methods are its spectral energy distribution, its peak overall level and its duration. It is clear that all these quantities do have a strong influence upon the subjective "magnitude" of the sound. However, in spite of previous results already mentioned and the evidence of Figure 27(a), the expected effects of duration in the case of real aircraft flyover noise have not been observed in the present study.

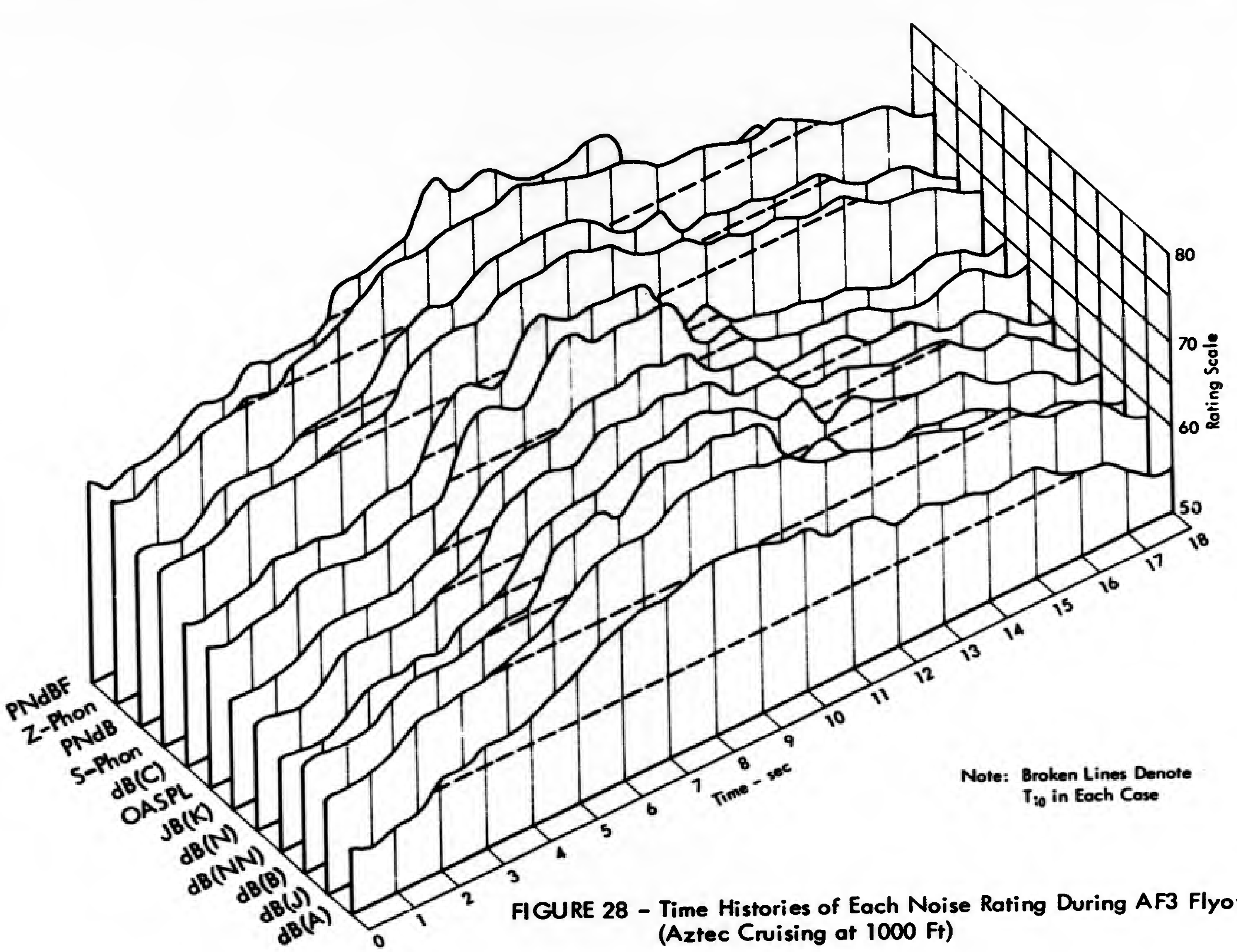



FIGURE 28 – Time Histories of Each Noise Rating During AF3 Flyover
(Aztec Cruising at 1000 Ft)

## TABLE XI

## $T_{10}$ DURATIONS FOR REAL FLYOVER SIGNALS

| Signal Code | $T_{10}$ (Seconds) | | | | |
|---|---|---|---|---|---|
| | PNdBF* | PNdB | dB(N) | dB(NN) | dB(A) |
| 140F | 8.2 | 9.1 | 9.0 | 8.7 | 9.5 |
| 140T1 | 3.0 | 5.0 | 4.6 | 4.7 | 4.7 |
| 140T2 | 10.2 | 12.2 | 16.0 | 16.9 | 15.3 |
| 140T3 | 25.0 | 25.0 | 25.5 | 23.0 | 22.0 |
| 6T | 2.9 | 3.8 | 3.6 | 4.2 | 3.0 |
| 6F1 | 1.9 | 2.2 | 2.3 | 2.2 | 1.8 |
| 6F2 | 5.7 | 6.0 | 5.7 | 5.7 | 5.9 |
| 6F3 | 11.3 | 12.0 | 11.1 | 11.5 | 11.1 |
| AT | 7.7 | 8.2 | 8.4 | 7.8 | 8.2 |
| AF1 | 2.1 | 2.2 | 2.5 | 2.0 | 2.8 |
| AF2 | 4.8 | 5.5 | 5.5 | 5.6 | 6.5 |
| AF3 | 13.2 | 11.7 | 11.2 | 10.5 | 14.6 |
| DCT | 15.0 | 16.0 | 15.5 | 14.4 | 12.7 |
| DCL | 3.1 | 2.7 | 2.5 | 2.0 | 2.5 |
| 707T | 13.8 | 17.5 | 18.0 | 13.1 | 14.0 |
| 707L | 2.3 | 2.5 | 2.4 | 1.9 | 3.2 |
| LT1 | 11.5 | 11.1 | 10.1 | 11.7 | 11.3 |
| LT2 | 14.4 | 14.7 | 14.9 | 14.6 | 14.7 |
| LL | 2.3 | 2.9 | 2.8 | 2.3 | 2.3 |
| TC | 3.7 | 4.1 | 4.8 | 4.1 | 4.8 |

* Duration corrected PNdBF = EPNdB.

TABLE XII

DURATION CORRECTED NOISE RATINGS
COMPARED WITH JUDGED REFERENCE LEVEL SRL (dB)

| Aircraft Sound Code | EPNdB | $PNdB_D$ | $dB(N)_D$ | $dB(NN)_D$ | $dB(A)_D$ | SRL |
|---|---|---|---|---|---|---|
| 140F | 84 | 81 | 74 | 72 | 70 | 80 |
| 140T1 | 86 | 82 | 76 | 74 | 70 | 81 |
| 140TL | 85 | 81 | 75 | 73 | 70 | 78 |
| 140T3 | 86 | 83 | 76 | 74 | 72 | 73 |
| 6T | 86 | 83 | 80 | 76 | 72 | 84 |
| 6F1 | 85 | 82 | 75 | 73 | 69 | 83 |
| 6F2 | 87 | 84 | 78 | 76 | 71 | 83 |
| 6F3 | 91 | 88 | 82 | 80 | 75 | 84 |
| AT | 90 | 88 | 82 | 82 | 75 | 81 |
| AF1 | 83 | 82 | 74 | 78 | 68 | 84 |
| AF2 | 81 | 80 | 73 | 74 | 67 | 80 |
| AF3 | 85 | 83 | 76 | 76 | 70 | 80 |
| DCT | 89 | 87 | 81 | 79 | 76 | 80 |
| DCL | 95 | 94 | 86 | 90 | 79 | 96 |
| 707T | 89 | 89 | 82 | 80 | 77 | 77 |
| 707L | 96 | 91 | 85 | 87 | 78 | 90 |
| LT1 | 83 | 83 | 76 | 74 | 73 | 74 |
| LT2 | 90 | 86 | 79 | 79 | 76 | 73 |
| LL | 91 | 89 | 82 | 83 | 76 | 86 |
| TC | 93 | 89 | 84 | 81 | 76 | 88 |

| Aircraft Sound Code | EPNdB | $PNdB_D$ | $dB(N)_D$ | $dB(NN)_D$ | $dB(A)_D$ | SRL |
|---|---|---|---|---|---|---|
| 140/2 | 73 | 71 | 64 | 62 | 58 | 69 |
| 140/4 | 80 | 77 | 70 | 68 | 64 | 73 |
| 140/8 | 76 | 73 | 67 | 65 | 61 | 73 |
| 140/16 | 80 | 77 | 71 | 69 | 65 | 77 |
| 140/32 | 78 | 76 | 70 | 67 | 64 | 73 |
| 6/2 | 65 | 65 | 59 | 58 | 52 | 66 |
| 6/4 | 77 | 77 | 71 | 70 | 63 | 72 |
| 6/8 | 73 | 73 | 67 | 65 | 60 | 67 |
| 6/16 | 79 | 78 | 72 | 70 | 66 | 72 |
| 6/32 | 78 | 77 | 71 | 69 | 65 | 72 |
| A/2 | 77 | 66 | 60 | 59 | 53 | 68 |
| A/4 | 75 | 73 | 67 | 65 | 61 | 75 |
| A/8 | 76 | 75 | 69 | 67 | 61 | 71 |
| A/16 | 76 | 75 | 70 | 69 | 61 | 71 |
| A/32 | 78 | 77 | 71 | 69 | 65 | 71 |

## TABLE XIII

### EFFECTS OF DURATION CORRECTION, $10 \log_{10}\left(\frac{T_{10}}{15}\right)$

| Method | Correlation Coefficient $R_c$ | | Probability of Deterioration |
|---|---|---|---|
| | Peak Value | D-Corrected Peak Value | |
| PNdBF | 0.900 | 0.846 | 88% |
| PNdB | 0.880 | 0.837 | 79% |
| dB(N) | 0.879 | 0.835 | 79% |
| dB(NN) | 0.877 | 0.867 | 57% |
| dB(A) | 0.867 | 0.784 | 90% |

The main difference between the shaped aircraft sounds, which do exhibit duration effects, and the real flyover sounds, which apparently do not, is the lack of a Doppler frequency shift in the former. Now this change of frequency observed as the aircraft flies by is very significant, and increases with aircraft velocity. The hearing mechanism is at least as sensitive to pitch changes as to small changes in sound pressure level and it seems likely that the Doppler effect would be at least as important, subjectively, as the duration effect.

Consider a simple sound source moving in a straight line at velocity V which emits a pure tone of frequency $f_s$. The frequency $f_0$ observed by a stationary observer is given by the Doppler equation

$$f_0 = f_s (1 - M \cos \theta)^{-1}$$

(See Figure 29)

where M is the source Mach number (= $V/a_0$ where $a_0$ is the speed of sound) and $\theta$ is the angle between the velocity vector and the line joining source and observer at the instant the sound was generated. The latter is somewhat earlier than the instant of observation due to the finite sound propagation time.

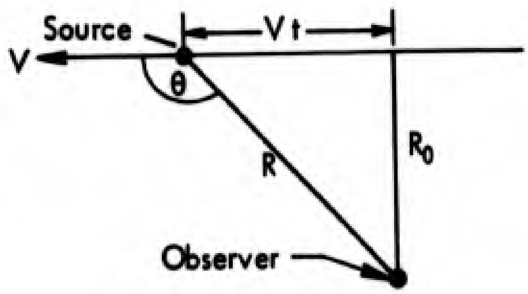


FIGURE 29 - Geometry of Source Motion

If t = 0 defines the instant of generation of the sound that travels the shortest distance $R_0$ to the observer, the angle $\theta$ can be defined as

$$\theta = \tan^{-1} \left( - \frac{R_0}{Vt} \right)$$

and by differentiation

$$\frac{d\theta}{dt} = \frac{V}{R_0} \sin^2 \theta \quad .$$

The magnitude of the Doppler shift rate $df_0/dt$ can thus be written

$$\frac{df_0}{dt} = \frac{d\theta}{dt} \cdot \frac{df_0}{d\theta}$$

$$= - \frac{MV}{R_0} \left( \frac{\sin \theta}{1 - M \cos \theta} \right)^2 \cdot f_s$$

where the negative sign indicates that the observed frequency is decreasing. The Doppler shift rate has a maximum value at the angle $\theta$ given by

$$\frac{d}{d\theta}\left(\frac{\sin\theta}{1 - M\cos\theta}\right) = 0$$

which is found to be $\theta = \cos^{-1} M$, giving a maximum of

$$\left(\frac{df_0}{dt}\right)_{max} = \frac{-M}{1 - M^2} \cdot \frac{V}{R_0} \cdot f_s$$

It is interesting to note that the sound having the maximum observed frequency shift rate is generated before the source reaches its closest point to the observer. However, by the time that sound reaches the observer the source has moved to precisely that point. Thus a person on the ground hears the maximum rate of change of frequency as he sees an aircraft overhead, or at its closest point of approach.

If M is small, which for aircraft maneuvering in the vicinity of airfields it generally is, we can write the approximation

$$\left(\frac{df_0}{dt}\right)_{max} \simeq \frac{-MV}{R_0} f_s$$

It is useful to nondimensionalize the frequency change rate as follows, putting $\dot{f}_o = \frac{df_0}{dt}$ :

$$\left(\frac{\dot{f}_0}{f_s}\right)_{max} \simeq \frac{-MV}{R_0} ,$$

since this ratio is the same for all source frequencies and provides a simple functional form for the Doppler effect.

Let us now postulate a Doppler shift correction to a subjective rating scale so that

$$N = N_0 + 10 \log_{10}\left(\frac{T_{10}}{15}\right) + G\left(\frac{\dot{f}_0}{f_s}\right)_{max}$$

where $N_0$ is the uncorrected peak value of the rating during the flyover, the second term is the duration correction discussed previously and the third term is some unknown function of the maximum frequency shift rate. For any aircraft flyover, the value of $(f_0/f_s)_{max}$ can be calculated from the aircraft velocity, minimum distance and the ambient speed of sound. However, if these values are not known for a given sound recording, the frequency shift rate could only be obtained by direct analysis of the signal using real time, narrow band spectrum analysis techniques such as those discussed in Section III. In the present case an alternative approach was followed in which the signal duration relationships were used to obtain a useful approximation to the Doppler shift term.

Referring again to Figure 29, and assuming firstly that the sound source is spherically uniform and secondly that atmospheric attenuation can be ignored, the observed sound pressure level can be written

$$\overline{p^2} = \overline{p_0^2} \left(\frac{R_0}{R}\right)^2$$

where $\overline{p_0^2}$ is the level at $t = 0$ when $R = R_0$. Substituting $R = (R_0^2 + V^2 t^2)^{\frac{1}{2}}$,

$$\overline{p^2} = \overline{p_0^2} \left(\frac{R_0^2}{R_0^2 + V^2 t^2}\right) = \overline{p_0^2} \left[\frac{1}{1 + \left(\frac{V^2 t^2}{R_0^2}\right)}\right]$$

Thus,

$$\left(\frac{Vt}{R_0}\right)^2 = \left(\frac{\overline{p_0^2}}{\overline{p^2}}\right) - 1$$

Any specific signal duration T, defined by the time interval between the instants when the sound pressure level passes through some arbitrary value of $\overline{p^2}/\overline{p_0^2}$, thus implies

$$\frac{VT}{R_0} = \text{constant} \qquad \text{or} \qquad T = k \cdot \frac{R_0}{V}$$

where k is a constant which depends on the chosen value of $\overline{p^2}/\overline{p_0^2}$.

It is interesting to pause at this point since this result now gives a further clue to the subjective importance of the Doppler shift. As discussed by Bauer and Torick (Reference 13) the loudness of a sound has two components, "sensory" loudness and "perceptual" loudness, where the former is the component which is directly attributable to auditory (or other physical) stimulation. The latter takes account of semantic content of the sound such as danger warnings. Since noisiness is defined as "unwantedness" it must include both sensory and perceptual effects. The semantic content of an aircraft sound which, as indicated above, tells the observer something about the proximity of the aircraft is likely to be very significant. Any subconscious notions of danger due to aircraft accident undoubtedly increases with reduced distance from the flight path so that, indirectly, subjective noisiness could increase with aircraft proximity, even if the overall sound pressure level remains the same. The exact manner in which a listener judges the distance of an aircraft in the absence of visual information is difficult to identify. The signal duration, as defined by $T_{10}$, is proportional to the ratio $R_0/V$ only, so that it alone tells the observer nothing about the distance between him and the aircraft, although familiarity with the sound of a particular aircraft type under various operational conditions might enable him to deduce it. However, it does seem possible that the additional Doppler shift information, namely, $- MV/R_0$, might be used subconsciously to complement the duration information in arriving at the final judgment, since the two sets of information are sufficient to identify both the velocity and distance of the aircraft. If this occurs, the ability to perform this analysis would be found in the majority of people who have a long history of exposure to aircraft, motor vehicle and other forms of transport noise. In summary, the duration effect is mostly a sensory attribute, which roughly corresponds to an integration of sound energy arriving at the ear, whereas the Doppler effect would be mostly perceptual, influencing an observer's judgment through the meaning he attaches to the sound.

Continuing the analysis, and substituting the equation $V = k R_0 T^{-1}$ into the Doppler shift relationship, we find that

$$\left(\frac{\dot{f}_0}{f_s}\right)_{max} \simeq -\frac{k^2}{a_0}\left(\frac{R_0}{T}\right)^2 \cdot \frac{1}{R_0} = -K\frac{R_0}{T^2}$$

where $K = k^2/a_0$. Thus, for a particular source distance, the peak shift rate is inversely proportional to the square of the duration.

A suitable form for the corrected rating scale is possibly

$$N = N_0 + 10 \log_{10}\left(\frac{T_{10}}{15}\right) + A \log_{10}\left(\frac{B \cdot R_0}{T_{10}^2}\right)$$

where A and B are constants to be determined. It is emphasized that the expression for the Doppler shift term is only approximate in the case of real aircraft noise since (1) the sound source is not spherically uniform, and (2) the duration $T_{10}$ applies to the particular rating scale. However, the errors introduced by this approximation would be offset to some extent by the simplicity of calculation in a practical rating scale.

The proposed correction is evaluated by applying it to one of the noise rating methods under study and applying the same regression analysis to determine whether or not the corrected scale shows any improvement. The Doppler term is only applied to those signals which contain a Doppler shift; i.e., the 20 real flyover signals. As a first attempt the correction

$$10 \log_{10} \left( \frac{0.1 \, R_0}{T_{10}^2} \right)$$

was applied to $dB(NN)_D$ where the value of $R_0$ for each case is obtained from Table VI. A plot of the result is shown in Figure 21(d) where the corrected rating is denoted $dB(NN)_{DD}$. It was evident from an inspection of this figure that this step "over corrected" the long duration signals; here the slope of the regression line is very low. Based on this observation and the fact that the results plotted in Figure 27(b) indicate that noisiness is dependent on both duration and Doppler shift, a second attempt was based in the square root of the frequency shift rate. Now

$$dB(NN) + 10 \log_{10} \left( \frac{T_{10}}{15} \right) + 10 \log_{10} \left( \frac{0.1 \, R_0}{T_{10}^2} \right)^{\frac{1}{2}}$$

$$= dB(NN) + 5 \log_{10} R_0 + 10 \log_{10} (0.1)^{\frac{1}{2}} - 10 \log_{10} 15$$

$$\simeq dB(NN) + 5 \log_{10} \left( \frac{R_0}{2200} \right) = dB(NN)_R$$

so that the total correction does not explicitly contain a duration term but appears dependent only on distance. Figure 21(e) illustrates the effect of this second correction, where the duration/Doppler corrected scale is now denoted by $dB(NN)_R$. The regression analysis of the four NN-weighted sound pressure levels, with and without corrections, gave the results listed in Table XIV which shows that this second correction is an improvement over the uncorrected rating and a substantial improvement (P = 80 percent)

TABLE XIV

EFFECT OF DURATION AND DOPPLER CORRECTIONS ON NN-WEIGHTED SOUND PRESSURE LEVEL RATINGS

| Rating Method | Correlation Coefficient $R_c$ | Variance About Regression | Probability of Improvement Over Uncorrected Method (percent) |
|---|---|---|---|
| dB(NN) | 0.877 | 11.3 | |
| $dB(NN)_D$ | 0.876 | 12.1 | 43* |
| $dB(NN)_D + 10 \log_{10}\left(\frac{0.1 R_0}{T_{10}^2}\right)$ | 0.893 | 9.90 | 64 |
| $dB(NN)_R$ | 0.904 | 8.94 | 75 |

* i.e., a deterioration.

over the duration correction above. The R-correction has been applied to each of the rating methods under study with the exception of OASPL and the corrected values are listed in Table XV. The results of a regression analysis of the R-values which are presented in Table XVI show that in seven cases out of eleven the R-correction represents a slight improvement over the uncorrected peak rating value. However, the deteriorations in the other cases are relatively small. The important finding is that the R-correction represents a very substantial improvement over the duration correction even though the net effect is essentially that the Doppler shift correction has compensated for the duration correction. Further, the particular form of the Doppler correction applied was chosen on the basis of the dB(NN) results. It is very probable that a further study of the constant terms would improve the average correlation.

It may be noted that the distance correction term tends to very large positive or negative values as R tends to infinity or zero. Whereas this might appear to constitute some limitation on its validity, it should be remembered that under these conditions the first term (which varies in similar fashion to the overall sound pressure level) tends to large negative and positive values. To appreciate this point, assume that dB(NN) varies with distance due to spherical spreading only so that

$$dB(NN)_R = dB(NN)_{REF} - 20 \log_{10} R_0 + 5 \log_{10} R_0 = dB(NN)_{REF} - 15 \log_{10} R_0$$

where $dB(NN)_{REF}$ is some reference level which is a function of spectrum shape only. Thus, the inverse square law has simply become an inverse 1.5 power law which means that the rating scale has, in fact, become less sensitive to distance. This may be a very significant result. Comparisons of subjective response to aircraft noise indoors

## TABLE XV

## DURATION AND DOPPLER CORRECTED NOISE RATINGS COMPARED WITH JUDGED REFERENCE LEVEL SRL-dB

| | Aircraft Sound Code | $dB(A)_R$ | $dB(B)_R$ | $dB(C)_R$ | $dB(N)_R$ | $dB(NN)_R$ | $dB(J)_R$ | $dB(K)_R$ | $PNdB_R$ | $PNdBF_R$ | $S\text{-}Phon_R$ | $Z\text{-}Phon_R$ | SRL |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Piston-Engined Aircraft Flyovers | 140F | 67 | 71 | 73 | 71 | 70 | 70 | 70 | 78 | 82 | 77 | 81 | 80 |
| | 140T1 | 71 | 77 | 81 | 77 | 74 | 76 | 77 | 83 | 89 | 82 | 85 | 81 |
| | 140T2 | 66 | 72 | 74 | 71 | 69 | 71 | 71 | 78 | 83 | 78 | 78 | 78 |
| | 140T3 | 67 | 72 | 77 | 71 | 69 | 70 | 71 | 78 | 84 | 77 | 80 | 73 |
| | 6T | 75 | 80 | 82 | 80 | 77 | 79 | 79 | 85 | 89 | 84 | 87 | 84 |
| | 6F1 | 73 | 78 | 81 | 78 | 76 | 77 | 78 | 85 | 89 | 83 | 86 | 83 |
| | 6F2 | 72 | 80 | 84 | 79 | 77 | 78 | 80 | 85 | 88 | 84 | 85 | 83 |
| | 6F3 | 74 | 82 | 86 | 81 | 79 | 80 | 81 | 87 | 90 | 85 | 86 | 84 |
| | AT | 71 | 76 | 80 | 77 | 78 | 77 | 78 | 84 | 86 | 82 | 87 | 81 |
| | AF1 | 70 | 76 | 80 | 77 | 82 | 77 | 78 | 85 | 86 | 84 | 88 | 84 |
| | AF2 | 68 | 73 | 77 | 74 | 75 | 73 | 74 | 81 | 83 | 81 | 84 | 80 |
| | AF3 | 68 | 75 | 80 | 75 | 76 | 74 | 76 | 82 | 87 | 81 | 85 | 80 |
| Jet Aircraft | DCT | 77 | 80 | 80 | 80 | 78 | 80 | 80 | 86 | 88 | 85 | 89 | 80 |
| | DCL | 82 | 82 | 83 | 89 | 94 | 89 | 90 | 96 | 97 | 93 | 97 | 96 |
| | 707T | 76 | 80 | 82 | 80 | 79 | 80 | 80 | 87 | 88 | 86 | 90 | 77 |
| | 707L | 80 | 79 | 79 | 88 | 91 | 88 | 89 | 94 | 99 | 91 | 92 | 90 |
| | LT1 | 73 | 76 | 77 | 77 | 73 | 76 | 76 | 83 | 83 | 82 | 86 | 74 |
| | LT2 | 75 | 77 | 78 | 78 | 78 | 78 | 78 | 85 | 89 | 83 | 87 | 73 |
| | LL | 77 | 79 | 81 | 82 | 85 | 81 | 82 | 89 | 92 | 87 | 92 | 86 |
| * | TC | 76 | 82 | 87 | 81 | 82 | 80 | 82 | 90 | 92 | 87 | 89 | 88 |
| 'Shaped' Flyover Signals | 140/2 | 58 | 65 | 68 | 64 | 62 | 64 | 64 | 71 | 73 | 72 | 72 | 69 |
| | 140/4 | 64 | 71 | 74 | 70 | 68 | 70 | 71 | 77 | 80 | 77 | 78 | 73 |
| | 140/8 | 61 | 68 | 70 | 67 | 65 | 66 | 77 | 73 | 76 | 74 | 75 | 73 |
| | 140/16 | 65 | 72 | 75 | 71 | 69 | 70 | 71 | 77 | 80 | 77 | 78 | 77 |
| | 140/32 | 64 | 70 | 73 | 70 | 67 | 69 | 70 | 76 | 78 | 77 | 78 | 73 |
| | 6/2 | 52 | 60 | 66 | 59 | 58 | 58 | 60 | 65 | 65 | 65 | 65 | 66 |
| | 6/4 | 63 | 72 | 77 | 71 | 70 | 70 | 72 | 77 | 77 | 76 | 76 | 72 |
| | 6/8 | 60 | 68 | 72 | 67 | 65 | 66 | 67 | 73 | 73 | 73 | 73 | 67 |
| | 6/16 | 66 | 73 | 77 | 72 | 70 | 71 | 72 | 78 | 79 | 78 | 79 | 72 |
| | 6/32 | 65 | 71 | 75 | 71 | 69 | 70 | 71 | 77 | 78 | 77 | 79 | 72 |
| | A/2 | 53 | 61 | 77 | 60 | 59 | 59 | 61 | 66 | 67 | 66 | 66 | 68 |
| | A/4 | 61 | 69 | 81 | 67 | 65 | 67 | 67 | 73 | 75 | 74 | 74 | 75 |
| | A/8 | 61 | 70 | 74 | 69 | 67 | 68 | 70 | 75 | 76 | 75 | 74 | 71 |
| | A/16 | 61 | 71 | 78 | 70 | 69 | 69 | 71 | 75 | 76 | 76 | 75 | 71 |
| | A/32 | 65 | 72 | 75 | 71 | 69 | 71 | 71 | 77 | 78 | 78 | 79 | 71 |

* Turbo-Prop

TABLE XVI

THE EFFECTIVENESS OF THE R-CORRECTION

| Method | Correlation Coefficient $R_c$ | | | % Probability of Improvement | |
|---|---|---|---|---|---|
| | Uncorrected | D-Corrected + | R-Corrected ++ | R-Corrected vs Uncorrected | R-Corrected vs D-Corrected |
| PNdBF | 0.900 | 0.846 | 0.896 | 46* | 86 |
| PNdB | 0.880 | 0.837 | 0.892 | 61 | 86 |
| dB(N) | 0.879 | 0.835 | 0.880 | 51 | 80 |
| dB(NN) | 0.877 | 0.867 | 0.904 | 75 | 80 |
| dB(A) | 0.867 | 0.784 | 0.845 | 34* | 80 |
| Z-Phon | 0.885 | - | 0.864 | 34* | - |
| S-Phon | 0.875 | - | 0.881 | 54 | - |
| dB(J) | 0.870 | - | 0.876 | 53 | - |
| dB(K) | 0.866 | - | 0.871 | 53 | - |
| dB(B) | 0.826 | - | 0.821 | 48* | - |
| dB(C) | 0.714 | - | 0.729 | 54 | - |

* Probabilities < 50% represent a deterioration

\+ D-Correction = $10 \log_{10}\left(\frac{T_{10}}{15}\right)$

++ R-Correction = $5 \log_{10}\left(\frac{R_0}{2200}\right)$

and outdoors (e.g., Reference 15) have shown that people are more tolerant to the noise outdoors, presumably simply because they expect to find more noise outside their homes. Such an effect could have a strong general influence in that people expect aircraft to be noisier at shorter distances and consequently make allowance for this fact when making noisiness judgments.

Nevertheless, it should be noted that the above results are based on a number of approximations including the neglect of atmospheric absorption. To avoid large errors which the latter effect could introduce it is recommended that the R-correction be restricted to distances less than about 2000 ft.

## C. Discussion

The main findings of the study are summarized in Table XVII which lists all corrected and uncorrected noise rating methods in rank order of correlation with the judged results. The percentage probability that a significant difference exists between any two of these methods is also tabulated for easy reference. It is clear that the differences between any two successive methods are marginal although the total range of variation is quite large. The probabilities listed are the integral values of the actual calculated values; for example, the probability that the best method is superior to the worst is actually 99.7 percent. Before commencing a discussion of some of the more important features of these results, it is as well to emphasize that we have measured the accuracy of each noise rating method as a general purpose predictor of the subjective evaluation of a wide range of aircraft sounds. The objective of this study was to evaluate the applicability of these methods to the sounds of the general aviation aircraft type defined in Section III. By including these sounds in one set with some sounds of commercial jet aircraft, for which certain rating methods are known to be satisfactory, the relative applicability of the methods can be estimated directly. Furthermore, any reliable noise rating technique should accurately predict the magnitude of any sound as a function of its characteristics. The inclusion of the artificially shaped sounds in the same set ensures a very wide range of sound characteristics upon which to base the evaluation. A final point is that the consolidation of data enables conclusions to be made with statistically greater confidence than if the data were divided into subsets corresponding to aircraft types or real and "artificial" sounds.

In addition to the value of the correlation coefficient for each rating method, we have the direct information provided by the plots of the data points in Figures 21 through 25, where the points are identified by the classifications: piston engined aircraft fly-overs, shaped piston engine aircraft sound, jet aircraft take-off, jet aircraft landing and turboprop aircraft. The results will be discussed with reference to these figures.

### 1. Loudness Summation Techniques

It is obviously significant that six of the seven best methods are complex loudness summation techniques as opposed to weighted sound pressure levels, and that

# TABLE XVII

## RANK LISTING OF ALL NOISE RATING METHODS STUDIED AND THE PERCENTAGE PROBABILITY THAT A SIGNIFICANT DIFFERENCE EXISTS BETWEEN ANY TWO.

| Rank | Method | Correlation Coefficient | Standard Deviation About Regression (dB) | dB(NN)$_R$ | PNdBF | PNdBF$_R$ | PNdB$_R$ | Z-Phon | S-Phon$_R$ | PNdB | dB(N)$_R$ | dB(N) | dB(NN) | dB(J)$_R$ | S-Phon | dB(K)$_R$ | dB(J) | dB(A) | dB(NN)$_D$ | dB(K) | Z-Phon$_R$ | EPNdB | dB(A)$_R$ | PNdB$_D$ | dB(N)$_D$ | dB(B) | dB(B)$_R$ | dB(A)$_D$ | dB(C)$_R$ | OASPL |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | dB(NN)$_R$ | .904 | 2.95 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2 | PNdBF | .900 | 3.04 | 53 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | PNdBF$_R$ | .896 | 3.09 | 57 | 54 | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4 | PNdB$_R$ | .892 | 3.16 | 62 | 58 | 54 | | | | | | | | | | | | | | | | | | | | | | | | |
| 5 | Z-Phon | .885 | 3.27 | 69 | 66 | 62 | 58 | | | | | | | | | | | | | | | | | | | | | | | |
| 6 | S-Phon$_R$ | .881 | 3.30 | 71 | 68 | 64 | 60 | 52 | | | | | | | | | | | | | | | | | | | | | | |
| 7 | PNdB | .880 | 3.32 | 72 | 69 | 65 | 61 | 53 | 51 | | | | | | | | | | | | | | | | | | | | | |
| 7 | dB(N)$_R$ | .880 | 3.32 | 72 | 69 | 65 | 61 | 53 | 51 | 50 | | | | | | | | | | | | | | | | | | | | |
| 9 | dB(N) | .879 | 3.33 | 73 | 70 | 66 | 62 | 54 | 52 | 51 | 51 | | | | | | | | | | | | | | | | | | | |
| 10 | dB(NN) | .877 | 3.36 | 75 | 72 | 68 | 64 | 56 | 54 | 53 | 53 | 52 | | | | | | | | | | | | | | | | | | |
| 11 | dB(J)$_R$ | .876 | 3.38 | 75 | 72 | 69 | 64 | 57 | 54 | 53 | 53 | 53 | 51 | | | | | | | | | | | | | | | | | |
| 12 | S-Phon | .875 | 3.38 | 75 | 72 | 69 | 64 | 57 | 54 | 53 | 53 | 53 | 51 | 50 | | | | | | | | | | | | | | | | |
| 13 | dB(K)$_R$ | .871 | 3.44 | 78 | 76 | 72 | 68 | 60 | 58 | 57 | 57 | 56 | 55 | 53 | 53 | | | | | | | | | | | | | | | |
| 14 | dB(J) | .870 | 3.44 | 78 | 76 | 72 | 68 | 60 | 58 | 57 | 57 | 56 | 55 | 53 | 53 | 50 | | | | | | | | | | | | | | |
| 15 | dB(A) | .867 | 3.48 | 80 | 78 | 75 | 71 | 63 | 61 | 60 | 60 | 59 | 57 | 56 | 56 | 52 | 52 | | | | | | | | | | | | | |
| 16 | dB(NN)$_D$ | .867 | 3.48 | 80 | 78 | 75 | 71 | 63 | 61 | 60 | 60 | 59 | 57 | 56 | 56 | 52 | 52 | 50 | | | | | | | | | | | | |
| 17 | dB(K) | .866 | 3.49 | 81 | 78 | 67 | 71 | 64 | 62 | 61 | 61 | 60 | 58 | 57 | 57 | 53 | 53 | 50 | 50 | | | | | | | | | | | |
| 18 | Z-Phon$_R$ | .864 | 3.52 | 82 | 80 | 77 | 73 | 66 | 64 | 63 | 63 | 62 | 60 | 59 | 59 | 55 | 55 | 52 | 52 | 51 | | | | | | | | | | |
| 19 | EPNdB | .846 | 3.73 | 89 | 88 | 85 | 83 | 77 | 76 | 75 | 75 | 74 | 72 | 71 | 71 | 68 | 68 | 65 | 65 | 64 | 62 | | | | | | | | | |
| 20 | dB(A)$_R$ | .845 | 3.74 | 90 | 88 | 86 | 83 | 78 | 76 | 75 | 75 | 74 | 73 | 72 | 72 | 68 | 68 | 66 | 66 | 65 | 63 | 50 | | | | | | | | |
| 21 | PNdB$_D$ | .837 | 3.82 | 92 | 90 | 88 | 86 | 81 | 80 | 79 | 79 | 78 | 76 | 76 | 76 | 73 | 73 | 70 | 70 | 69 | 68 | 55 | 54 | | | | | | | |
| 22 | dB(N)$_D$ | .835 | 3.85 | 92 | 91 | 89 | 87 | 82 | 81 | 80 | 80 | 79 | 78 | 77 | 77 | 74 | 74 | 71 | 71 | 70 | 69 | 57 | 56 | 51 | | | | | | |
| 23 | dB(B) | .826 | 3.94 | 94 | 93 | 91 | 90 | 85 | 84 | 83 | 83 | 83 | 81 | 81 | 81 | 78 | 78 | 76 | 76 | 75 | 74 | 62 | 61 | 56 | 55 | | | | | |
| 24 | dB(B)$_R$ | .821 | 3.99 | 95 | 94 | 92 | 91 | 87 | 86 | 85 | 85 | 84 | 83 | 83 | 83 | 80 | 80 | 78 | 78 | 77 | 76 | 64 | 64 | 59 | 57 | 52 | | | | |
| 25 | dB(A)$_D$ | .784 | 4.34 | 98 | 98 | 97 | 96 | 94 | 94 | 93 | 93 | 93 | 92 | 92 | 92 | 91 | 91 | 90 | 90 | 89 | 88 | 80 | 80 | 77 | 75 | 70 | 68 | | | |
| 26 | dB(C)$_R$ | .729 | 4.79 | 99 | 99 | 99 | 99 | 98 | 98 | 98 | 98 | 98 | 97 | 97 | 97 | 97 | 97 | 96 | 96 | 96 | 96 | 92 | 92 | 90 | 89 | 87 | 85 | 71 | | |
| 27 | OASPL | .717 | 4.86 | 99 | 99 | 99 | 99 | 98 | 98 | 98 | 98 | 98 | 98 | 98 | 98 | 97 | 97 | 97 | 97 | 97 | 97 | 94 | 93 | 92 | 91 | 89 | 87 | 74 | 53 | |
| 28 | dB(C) | .714 | 4.89 | 99 | 99 | 99 | 99 | 98 | 98 | 98 | 98 | 98 | 98 | 98 | 98 | 97 | 97 | 97 | 97 | 97 | 97 | 94 | 93 | 92 | 91 | 89 | 87 | 75 | 54 | 51 |

four of these six are based on the Perceived Noise Level concept. Ignoring for the present the first ranking technique, it can be seen that the six methods, PNdBF, $PNdBF_R$, $PNdB_R$, Zwicker's Phons, Stevens' $Phons_R$ and PNdB are, in fact, fairly closely matched although it could be argued that, since the probability of a difference between the best and worst of these six is 69 percent, PNdBF is twice as good a predictor as PNdB. This follows from the probability that on an average we can expect PNdBF to be more accurate 69 times out of 100. Turning to Figure 23 we see that for both PNdB and PNdBF each of the subsets of points, corresponding to the various aircraft classifications, fall about the regression lines. The main effect of the pure tone correction in PNdBF appear to have been a shift of the entire set of points about 3 dB to the right although, even though it is not particularly apparent in the figure, there is some reduction of scatter. Surprisingly, the tone correction appears to have had an adverse effect on the points corresponding to landing turbojet and turbofan aircraft, precisely those cases where an improvement would be expected. The regression slopes b are almost identical at 1.00 for PNdBF and 0.966 for PNdB.

Similar comments apply to Zwicker Phons regarding the distribution of points about the regression line which in this case has a slope of 1.04. However, the jet take-off sounds all fall below the line. Application of the duration correction to both PNdB and PNdBF greatly increases the scatter of the points, as indeed it does for all rating methods where it has been used. The plots strongly suggest that the reason for this is that the correction has significantly different effects upon piston engined and jet aircraft noise. Although the slope of the regression line in both cases has been reduced to approximately 0.85, lines fitted by eye to the jet aircraft points both have slopes which exceed 2.0. The variance of the shaped sounds is reduced, as would be expected from the results shown in Figure 27, but the real piston engine aircraft flyover sounds are more dispersed. It seems safe to conclude that the duration corrected methods EPNdB and $PNdB_D$ are inferior either for general purpose application or for jet aircraft noise. In the latter case it is possible that good correlation could be found but the reduced sensitivity implied by the large slope is undesirable. Figure 23 shows how the Doppler correction has remedied some of the faults introduced by the duration correction although the jet take-off sounds fall well below the line, by an average of around 4 dB.

Stevens' Phons fit all aircraft types reasonably well although an apparently low scatter is misleading because the slope of the regression line is high (1.29). The piston engined aircraft points tend to fall above the line and the jet aircraft below. The R-correction improves Stevens' Phons with a probability of 54 percent but an advantage which is not reflected in this probability is the increased sensitivity obtained by reducing the regression slope to 1.02.

2. Weighted Sound Pressure Levels

Some of the weighted sound pressure levels show surprisingly good correlation with the subjective results, notably dB(N) and dB(NN). The probabilities that these methods are worse than PNdB are only 51 and 53 percent, respectively. In particular, the NN-weighted results for the various categories fall very evenly about the regression methods are worse than PNdB are only 51 and 53 percent, respectively. In particular, the NN-weighted results for the various categories fall very evenly about the regression line. Also dB(NN) takes good care of the landing jet noise, with its dominating compressor whine, which the great majority of methods fail to do. This suggests that the large contribution of the NN-weighting function around 4000 Hz goes a long way towards compensating for the lack of any pure-tone correction. It is interesting to speculate that intense pure tones cause problems with other aircraft noise evaluation methods merely because these are usually compressor generated and have frequencies in this critical range. As before, the duration correction, when used without a Doppler correction, is responsible for a large deterioration in correlation, largely because of the different effects upon piston engine and jet aircraft noise. The Doppler correction has a significant effect upon both dB(N) and dB(NN) to the extent that $dB(NN)_R$ is the highest ranking method of those examined. However, it should be remembered that the constants in the Doppler term were chosen on the basis of their effect upon the NN-weighted scale so that this result is not too surprising. The prime reason for the choice of a 2200 ft reference range in the Doppler correction was the need to match the shaped signal results, which have no Doppler correction, with the remainder. It is probable that the PNdBF scale could have been optimized if it had been chosen as a base.

The weighted sound pressure level methods, dB(A), dB(J) and dB(K), fall in the mid-range of performance and little need be said about them except that they are practically equivalent statistically. The B-weighted SPL is rather worse and OASPL and dB(C) are too poor to be given any consideration as noisiness rating methods.

### 3. Summary

Some desirable features of a general purpose aircraft noise rating method are:

- Accuracy - the ability to predict subjective evaluation of aircraft noise with the smallest possible error.
- Generality - applicability to the sound of all types of aircraft operating under all conditions under which noise is a problem.
- Simplicity - preferably a method which is amenable to use in a simple measuring device for monitoring purposes.
- Sensitivity - preferably a one-to-one ratio between the scale units and the sound pressure level of a standard reference.

Of all those examined, the method which best meets all these requirements is the weighted sound pressure level dB(N); correlation coefficient 0.879; standard error 3.3 dB; slope of regression line 0.966.

Since the requirement for a slope of unity is a minor one, an attractive alternative method is $dB(NN)_R$, which appears to minimize the variance between aircraft, and, in the present study, had the lowest variance of all the methods examined. (Correlation coefficient .904, standard error 2.95 dB, slope of regression line 0.840.) Further, it is likely to be superior to dB(N) three times out of four. In common with any weighted sound pressure level technique, it has the advantage that it can in practice be implemented as a direct measuring scale using a sound level meter weighting network. Furthermore, the R-correction is extremely simple to apply; potentially it could be incorporated into the sound level meter as a scale adjustment operated by a calibrated dial.

Of the remainder, the methods which most satisfactorily meet these requirements are, in rank order,

PNdBF - Perceived Noise Level in PNdB (maximum value) corrected for pure tone content; correlation coefficient .900; standard error 3.0 dB; slope of regression line 1.00.

Phons Zwicker - (Peak value). Correlation coefficient 0.885; standard error 3.3 dB; slope of regression line 1.04.

PNdB - Perceived Noise Level in PNdB (maximum value). Correlation coefficient 0.880; standard error 3.3 dB; slope of regression line 0.966.

# V. CONCLUSIONS

Analysis of the results of an experiment in which a group of subjects evaluated the relative noisiness of a wide range of general aviation aircraft sounds in relation to the noise of commercial jet transports yielded the following findings:

1. The performances of various noise rating methods have been compared. The statistical differences between the better methods are small but this is largely due to the masking effects of the large subjective variance which is apparent in the experimental results. Consequently we must accept conclusions at a relatively low level of significance.

2. Although the accepted effects of signal duration were clearly observed for a set of artificially shaped sounds, it appeared initially that duration had little effect on the subjective evaluation of the real aircraft flyover sounds studied. Application of a duration correction to five different noise rating methods substantially degraded their performance.

3. An explanation offered for the lack of duration effects is that, although they are clearly observed for sounds which have no Doppler frequency shift, in the case of aircraft flyover noise the subjective effects of this shift tend to cancel those due to duration. A crude rating scale modification put forward to account for the Doppler phenomenon leads to a complete cancellation of the duration correction leaving a term which is dependent on aircraft distance only. Thus the combined correction is a function of aircraft range only, defined as the shortest distance between the aircraft and observer during the flyover. For a fairly narrow range of aircraft velocities this range correction is roughly equivalent to the square root of the previously proposed duration correction. In seven cases out of eleven applications it led to improvements in the correlation of the ratings with subjective results. Furthermore, a thorough study of this correction to optimize the constants would probably lead to improvements in its effectiveness. Further study is required.

4. Although the statistical differences between the various rating methods are small, the best four methods for general purpose aircraft noise rating purposes <u>of those currently in use</u> are, in rank order

Pure tone corrected perceived noise level (PNdBF)
Phons (Zwicker) (Z-Phon)
Perceived Noise Level in dB (PNdB)
N-Weighted sound pressure level (dB(N))

where the maximum instantaneous value is used in each case.

The worst two by far are:

Overall sound pressure level
and C-Weighted sound pressure level.

5. The NN-Weighting curve derived from the results of the present experiment, which shows a large peak around 4000 Hz proved equal to the N-Weighting as a prediction method and moreover, showed superior correlation for the landing jet aircraft sounds which displayed large pure tone contents around the aurally sensitive frequencies. Further, the NN-Weighted sound pressure level was less sensitive to the duration correction and when used [illegible] Doppler correction proved to be the best general method of those studied.

6. The above best four methods appear to rate general aviation aircraft noise and commercial jet transport noise equally well. Most other methods examined tended to overrate the relative magnitude of jet aircraft take-off noise, including that of the small jet aircraft.

7. In view of the small differences between the various methods the maximum N-Weighted sound pressure level which can be read conveniently from an appropriate sound level meter would appear to be the most suitable technique for general aircraft noise monitoring, at the present time.

## REFERENCES


1. Kryter, K.D., "Scaling Human Reactions to the Sound from Aircraft," J. Acoust. Soc. Am. 31, 1415-1429, No. 11., Nov. 1959.

2. Stevens, S.S., "The Calculations of the Loudness of Complex Noise," J. Acoust. Soc. Am. 28, 807-832, September 1956.

3. Kryter, K.D., and Pearsons, K.S., "Some Effects of Spectral Content and Duration on Perceived Noise Level," J. Acoust. Soc. Am. 35, 866-883, No. 6, June 1963.

4. Pearsons, K.S., "The Effects of Duration and Background Noise Level on Perceived Noisiness," FAA-ADS-78, April 1966.

5. F.A.A. Draft Proposal for Aircraft Noise Certification Criteria, December 1, 1967.

6. Society of Automotive Engineers, Research Project Committee R.2.5, "Technique for Developing Noise Exposure Forecasts," FAA-DS-67-14, August 1967.

7. Society of Automotive Engineers, Proposed Attachment for Draft SAE ARP Defining Effective Perceived Noise Level, February 1, 1968.

8. Stevens, S.S., "Procedure for Calculating Loudness: Mark VI," J. Acoust. Soc. Am. 33, 1577-1585, No. 11, November 1961.

9. Zwicker, E., "Graphic Evaluation of Loudness from 1/3 Octave Band Sound Pressure Level Curves," (In German) Frequenz, 13, 234-238, No. 8, 1959.

10. Kryter, K.D., "Review of Research and Methods for Measuring the Loudness and Noisiness of Complex Sounds," NASA, CR-422, April 1966.

11. Jones, J., NASA Marshall Space Flight Center, 1968, To be Published.

12. Crow, E.L., Davis, F.A., Maxfield, M.W., Statistics Manual, Dover Publications Inc., New York, N.Y.

13. Bauer, B.B., and Torick, E.L., "Researches in Loudness Measurement," I.E.E.E Trans. on Audio and Electro-acoustics, Vol. AU-14, No. 3., 141-151, September 1966.

14. Society of Automotive Engineers, "Definitions and Procedures for Computing the Perceived Noise Level of Aircraft Noise," ARP 865, October 1964.

15. Robinson, D.W., "The Subjective Basis For Aircraft Noise Limitation," J. Roy. Aero. Soc., No. 678, Vol. 71, 396-400, June 1967.

ACKNOWLEDGEMENTS

The author wishes to thank the following members of Wyle Laboratories Research Staff who contributed to this project:

Mr. K. McK. Eldred for his guidance and encouragement, Mr. E.N. Esslinger who designed, built and operated all the instrumentation, Mr. D.M. Lister who prepared the computer programs and assisted in the statistical analyses, and Dr. S.W. Radcliffe for his assistance in running the subjective experiments.

Thanks are also due to the Airport Authorities of Huntsville and Memphis Airports for their very willing cooperation during recording sessions.

## APPENDIX A

## INSTRUMENTATION SYSTEMS USED FOR RECORDING, DUBBING, ANALYZING AND REPLAYING SOUND RECORDINGS AND DESCRIPTION OF PROGRESSIVE WAVE FACILITY

### 1. Recording System

The instrumentation used for recording aircraft sounds in the field is shown in Figure A 1. A one inch diameter B and K Model 4131 microphone was mounted on a 5 foot high tripod with the plane of its diaphragm approximately parallel to the aircraft path during flyover. This was connected through a B and K 2613 cathode follower and 100 feet of cable to a B and K 2603 amplifier set which incorporates a 10 dB step attenuator. The amplifier output was recorded on one channel of an Ampex AG 500 twin channel, 1/4 inch tape recorder operating at 7-1/2 inches per second tape speed. This was powered from a 320 ampere hour 12 volt battery through a Terado 50-191 250 watt inverter which provided a 125 volt supply at 60 Hz. The AG-500 has an overall signal to noise ratio of approximately 60 dB and it was endeavored to adjust the attenuator setting in each case so that the peak recorded level came as close as possible to the maximum without the risk of overloading. At frequent intervals throughout the recording sessions the system was calibrated, end-to-end, using a B and K 4220 pistonphone. The entire system, from microphone to replay output, had a flat response between 20 and 15,000 Hz within ± 2 dB.

The other channel of the recorder was used for voice identification of aircraft type and other information required to define the aircraft maneuvers.

All aircraft recorded were also photographed to provide additional identification and to enable the distance of the aircraft to be calculated where necessary.

### 2. Automatic Record System for Preparation of Paired Comparison Tapes

It became clear early in the program that a form of automatic procedure was desirable for the manufacture of the tape recordings required for the paired comparison tests. Early efforts were made to perform this task manually, but the time involved was excessive and the resulting quality of the tapes poor. Consequently, the system depicted in Figure A 2 was developed, which enabled tapes of good quality to be prepared rapidly and accurately.

Basically, the method consists of two sequential dubbing processes. The original signals were transferred from the Ampex AG-500 to a tape loop on a Sangamo 14 channel F.M. recorder. Fourteen different signals were recorded side-by-side on this loop. The loop was then played continuously with the output from all channels feeding into the Automatic Record System which transferred the signals in a predetermined sequence

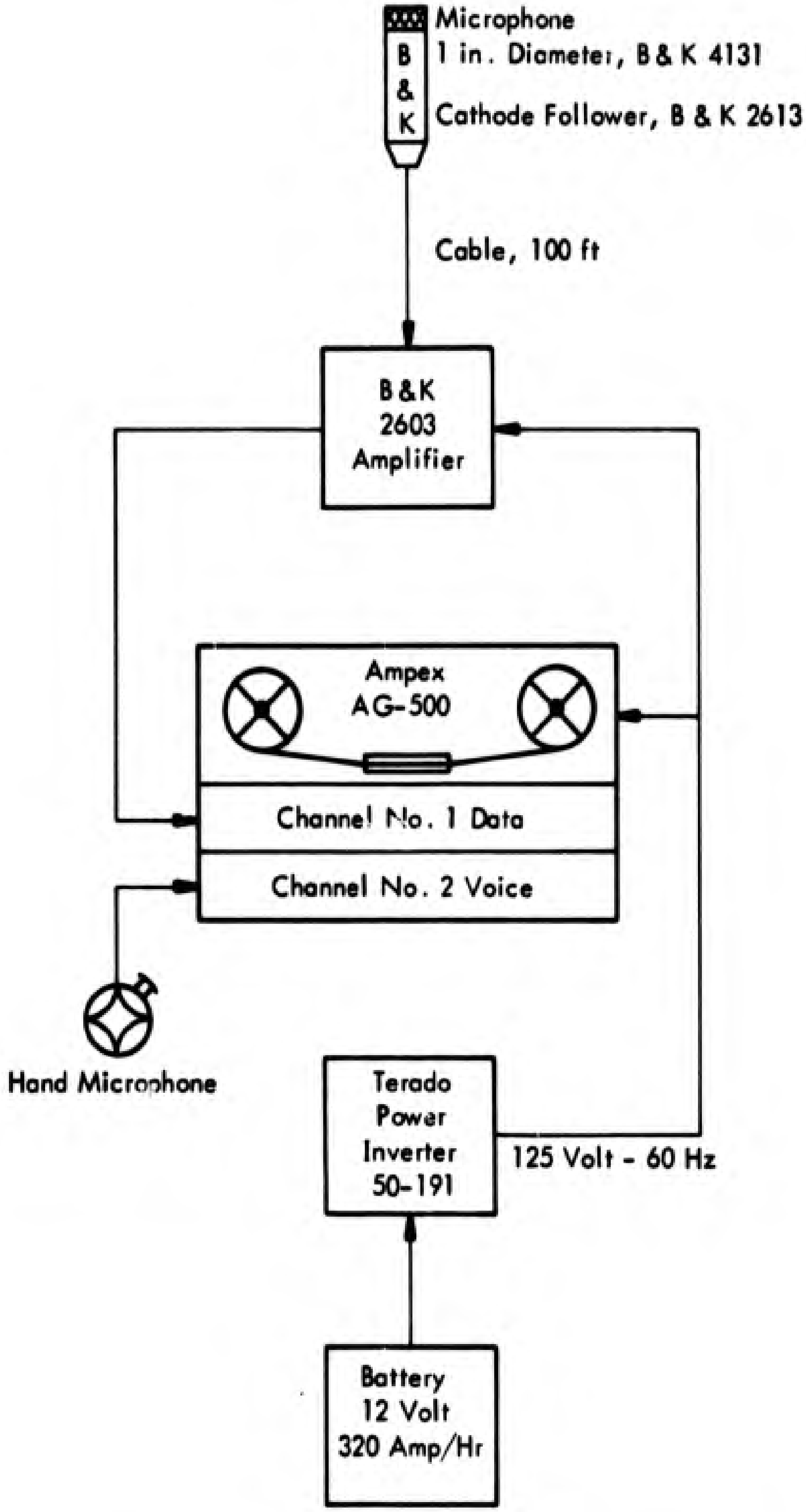



FIGURE A1 - Instrumentation Used for Recording Aircraft Noise

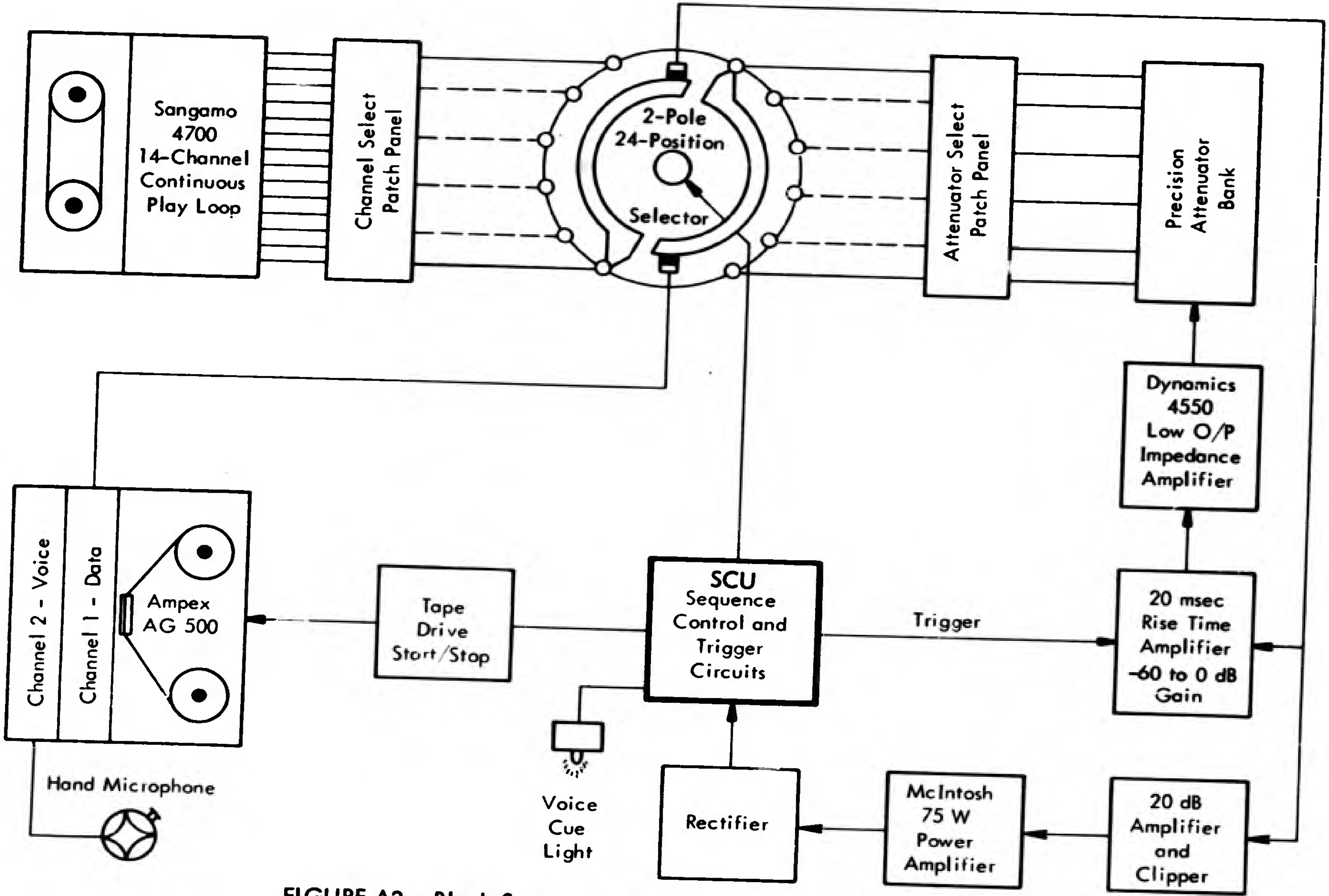



FIGURE A2 – Block Systematic Diagram of Automatic Record System For Preparation of Paired-Comparison Tapes through the required attenuator in each case, back onto the 1/4 inch Ampex recorder. The intervals between sounds and between pairs were accurately controlled by the system, which, once initiated, transferred 24 signals in succession with no manual intervention other than recording the appropriate voice identification on the second channel at the correct points. The automatic process will now be described in detail.

Graphic level time history plots of the original signals were inspected to determine the required cutoff in each case. In general these were chosen as the 20 dB down points whenever possible. The signals were then transferred onto the desired channel of the Sangamo Recorder through a variable gain amplifier whose bias voltage was controlled by a capacitor. By charging or discharging this capacitor through a suitable resistor the gain of the amplifier could be changed from 0 to -60 dB or vice-versa in approximately 20 milliseconds. This had the effect of applying a rise or decay time, of the same duration, to the transferred signals. For a reason which will be explained shortly, an additional 1/2 second duration was included at the beginning of the transferred signal.

At the heart of the Automatic Record System is the Sequence Control Unit (SCU) which uses the signal replayed from the tape loop to trigger various actions including all operations of the Ampex AG-500 recorder, the stepping of the selector switch, the control of the "rise time" amplifier and the warning light which tells the operator to record the necessary voice identification on channel 2.

The automatic sequency of operations involved in recording the paired comparison tapes is, referring to Figure A2, as follows:

(1) The two panels are patched to select the desired sequence of the first 24 sounds (comprising twelve sound pairs) and the associated attenuator setting in each case.

(2) With the loop running in the "reproduce" mode the system is energized with the 24 position selector in position 1. The Ampex tape drive is stopped and an interlock within the SCU prevents any action if the signal on the monitored channel happens to be playing. The interlock is automatically disengaged when the signal ends.

(3) The start of the signal, which is amplified, clipped and rectified through the SCU imput circuit causes the SCU to start the Ampex tape moving in the "record" mode. The input to channel 1 however, which passes from the Sangamo recorder, through the two stages of the selector and the required attenuator, also passes through the variable gain amplifier (VGA) at its -60 dB setting so that its level is essentially zero. One half second later, when the tape is up to speed, the SCU triggers the VGA causing the recorded signal to be amplified to its true level in 20 milliseconds. Thus, although the first 1/2 second of signal is lost, the required rise time is replaced.

(4) Recording continues until the end of the signal which causes two actions by the SCU. Firstly the VGA gain is dropped back to -60 dB, effectively grounding

the recorder input and secondly the selector switch is stepped to position 2, thus monitoring the second required channel. Also the interlock is reset so that if a signal is found on this channel no action is taken. After a further half second the ampex tape drive is stopped and the interlock de-energized beginning a second cycle.

(5) The second cycle is identical to the first with the exception that after the sound is recorded the tape runs for 5-1/2 seconds, instead of 1/2 second, during which time the warning light is illuminated and the voice identification for the next pair of sounds is recorded on channel 2 of the Ampex. This completes an entire cycle of operation and the cycle recommences at step 3.

(6) The process continues until all twenty-four sounds have been recorded at which point the process pauses until the next 24 sounds are pre-programmed before continuing.

3. Progressive Wave Facility

The facility is illustrated in Figure A 3. It is a dual prupose facility in that it was designed for reverberant or progressive wave high intensity acoustic testing. Basically the sound is generated at the end of a 36 feet long horn which expands into a 1500 cubic feet reverberation room. In the progressive wave configuration a portable set of absorbent wedges are installed against the wall facing the horn. The horn has a low frequency cutoff of 15 Hz and the wedges are highly absorbent down to this frequency. With the wedges installed, the subjects for the judgment experiments were seated across the mouth of the horn which is approximately 13 feet wide and 10 feet high. The walls of the room are constructed of concrete and are over 12 inches thick so that ambient sound levels within are extremely low. Details of the acoustic characteristics of the room are included in the next section.

4. Sound Reproduction Equipment

The equipment used in calibration of the progressive wave facility and to reproduce the paired-comparison tape recordings is illustrated in Figure A 4. Two loudspeaker configurations were used, denoted by Systems 1 and 2. System 1 was used in Pilot Experiment 2 and System 2 in the Main Experiment. The essential difference is that in the latter system two speakers are used as opposed to five (excluding the voice speaker). The noise generator and oscillator were used for frequency response measurements at the seating locations during loudspeaker installation and adjustment and to adjust the 1/3-octave levels of the B and K 1612-5A shaping filter network. The frequency response characteristics of Systems 1 and 2 were very similar and Figure A5 shows both random and sine sweep spectra measured at the five seat locations using System 2. The differences at high frequencies observed in the 1/3-octave plots for the various seat positions are very sensitive to speaker position and are associated with the high frequency radiation patterns. The amplitude fluctuations recorded during the constant amplitude

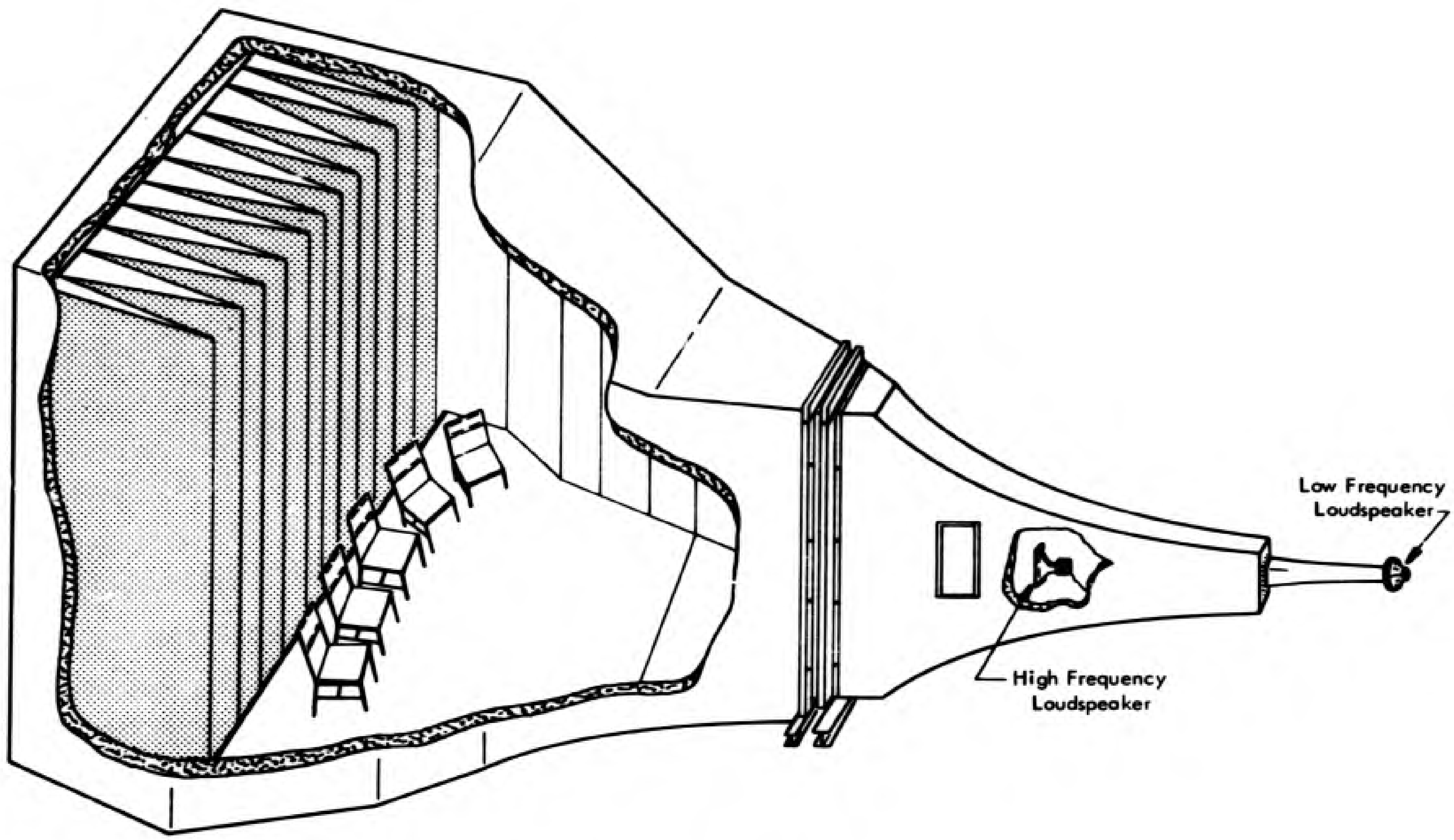
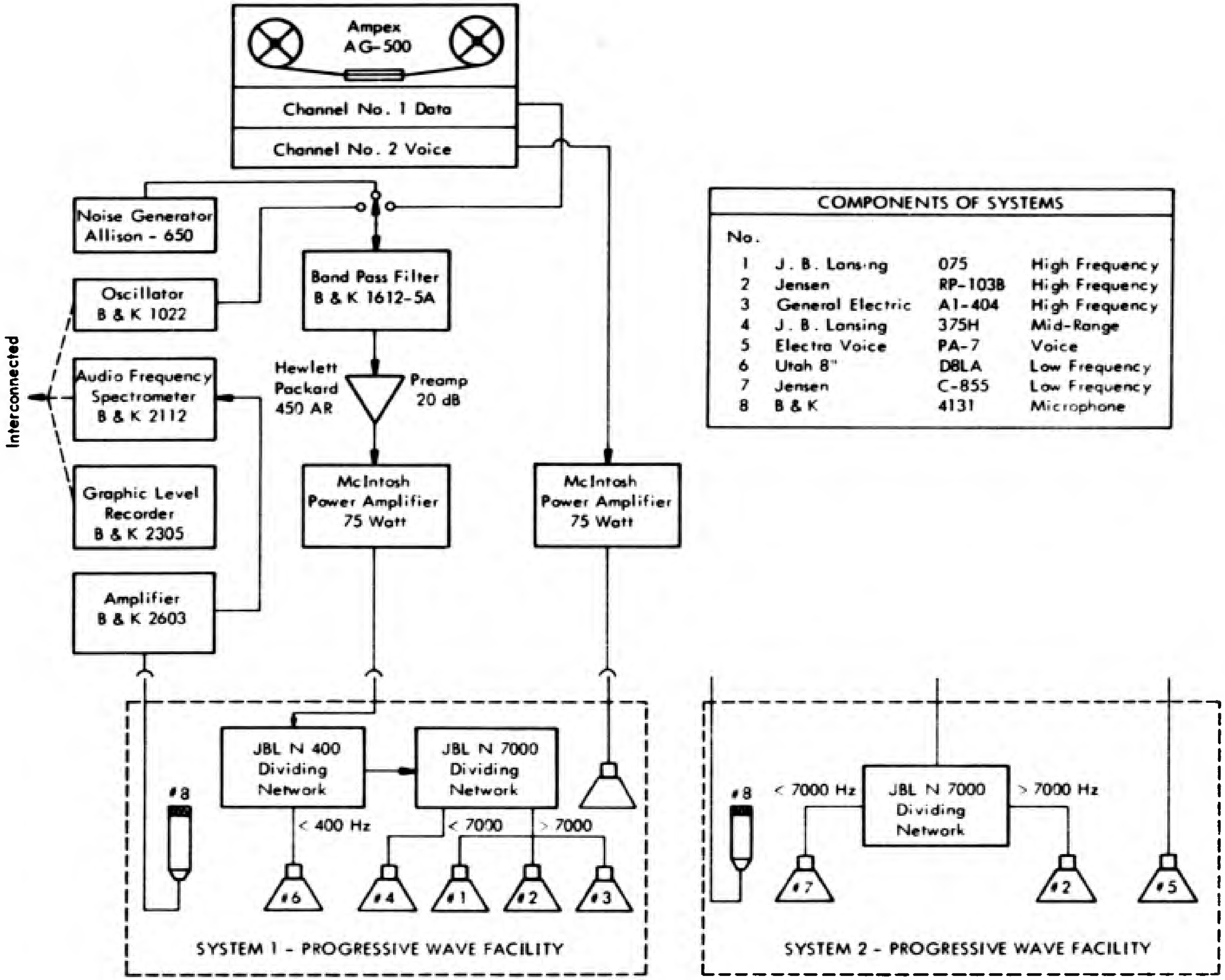



FIGURE A3 - Cutaway View of Progressive Wave Facility Showing Arrangement for Subjective Judgment Experiment sweep frequency input are the consequences of various standing wave patterns which arise due to the highly reflective hard walls. Although it is believed that these effects could be substantially reduced by appropriate treatment of the walls in the vicinity of the seating positions it was not possible to perform these modifications during the present study. In order to better evaluate t e magnitude of these fluctuations a similar analysis was made in the semi-reverberant room used for Plot Experiment 1. A wide frequency range 3-unit loudspeaker cabinet was used with the same pink noise and sine sweep input. Figure A6 shows the results. It is clear that the progressive wave facility represents an order of magnitude improvement.

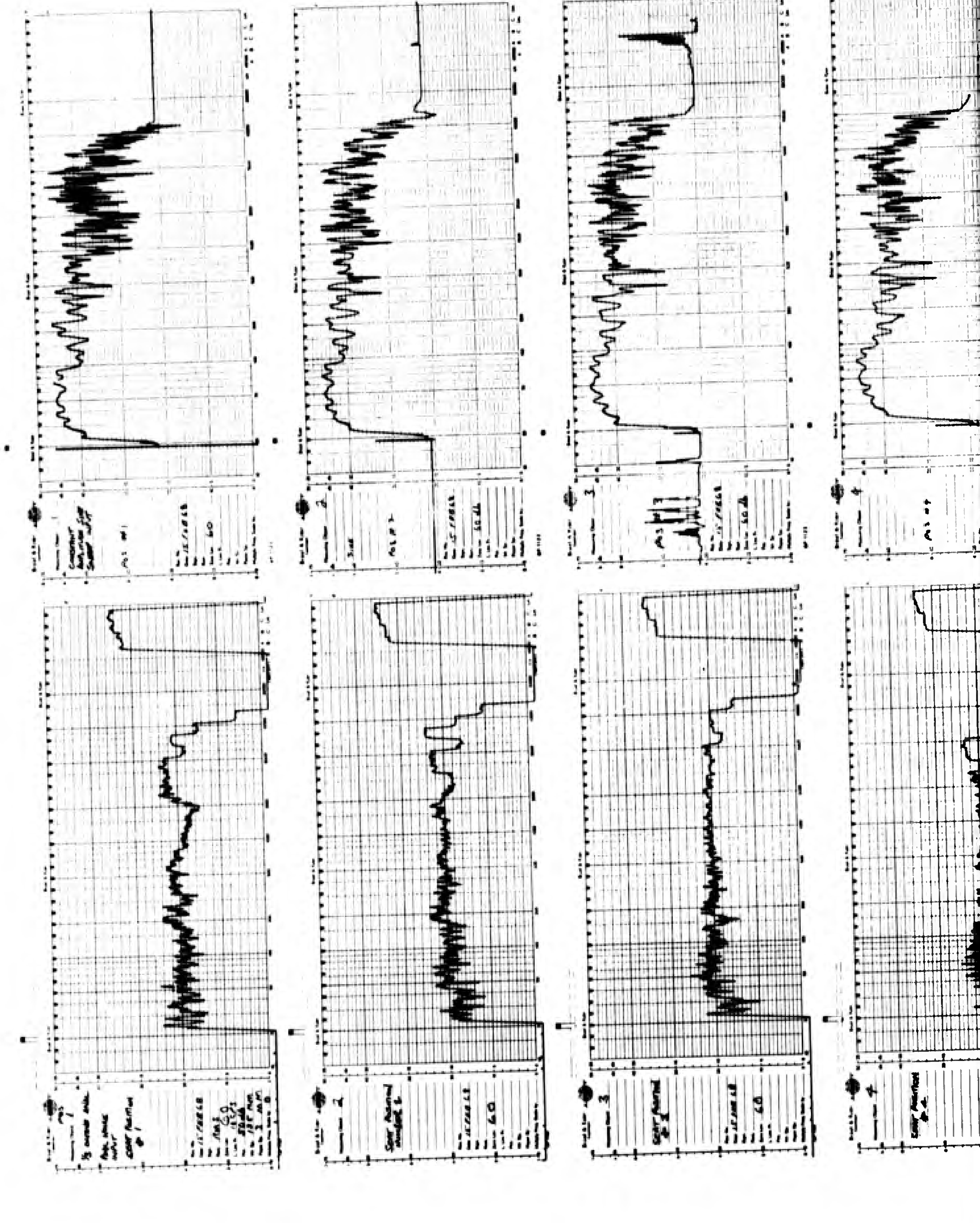



COMPONENTS OF SYSTEMS

| No. | | | |
|---|---|---|---|
| 1 | J. B. Lansing | 075 | High Frequency |
| 2 | Jensen | RP-103B | High Frequency |
| 3 | General Electric | A1-404 | High Frequency |
| 4 | J. B. Lansing | 375H | Mid-Range |
| 5 | Electro Voice | PA-7 | Voice |
| 6 | Utah 8" | D8LA | Low Frequency |
| 7 | Jensen | C-855 | Low Frequency |
| 8 | B & K | 4131 | Microphone |

FIGURE A4 – Instrumentation for Sound Reproduction and Analysis

BLANK PAGE



A

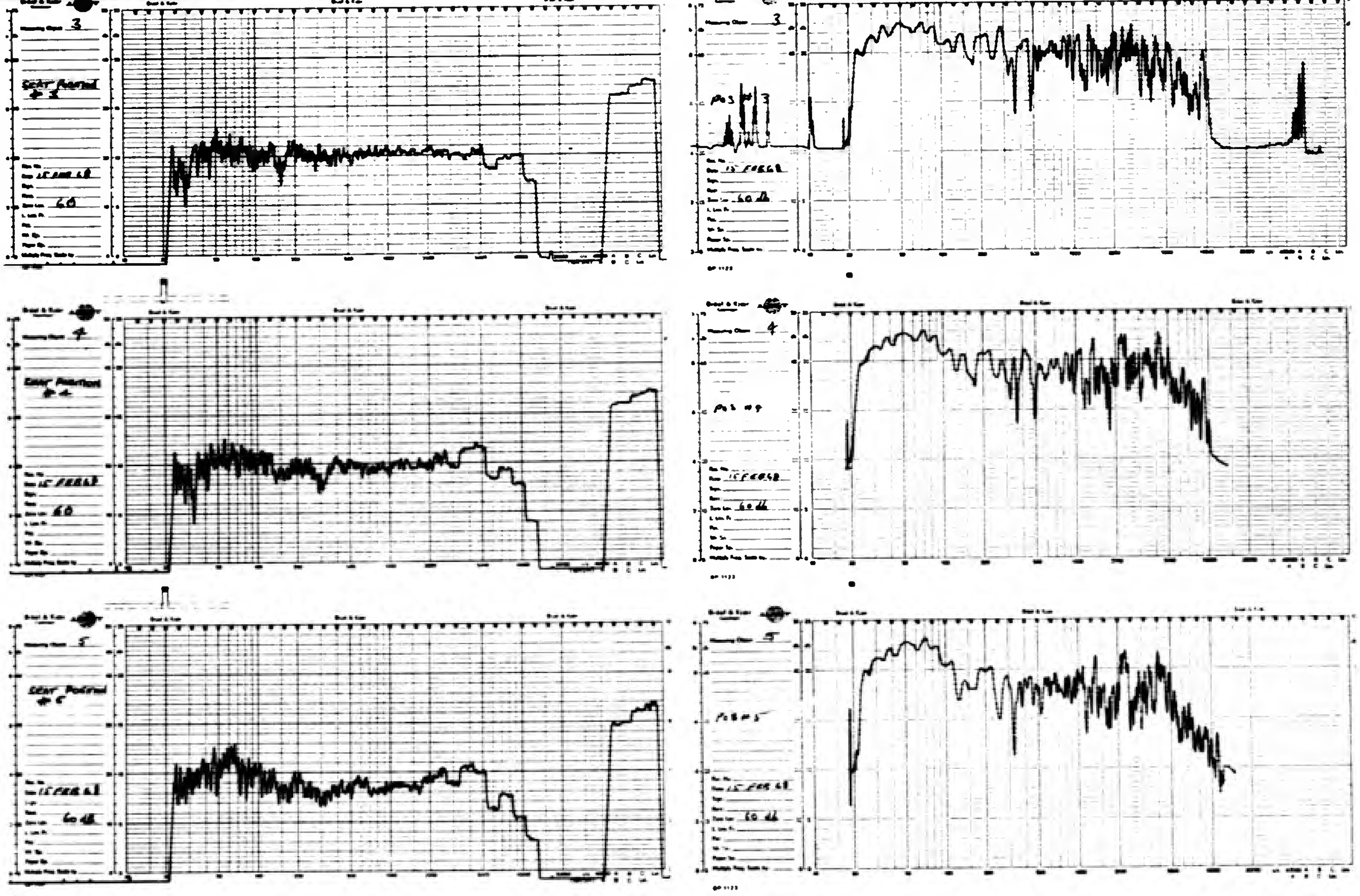
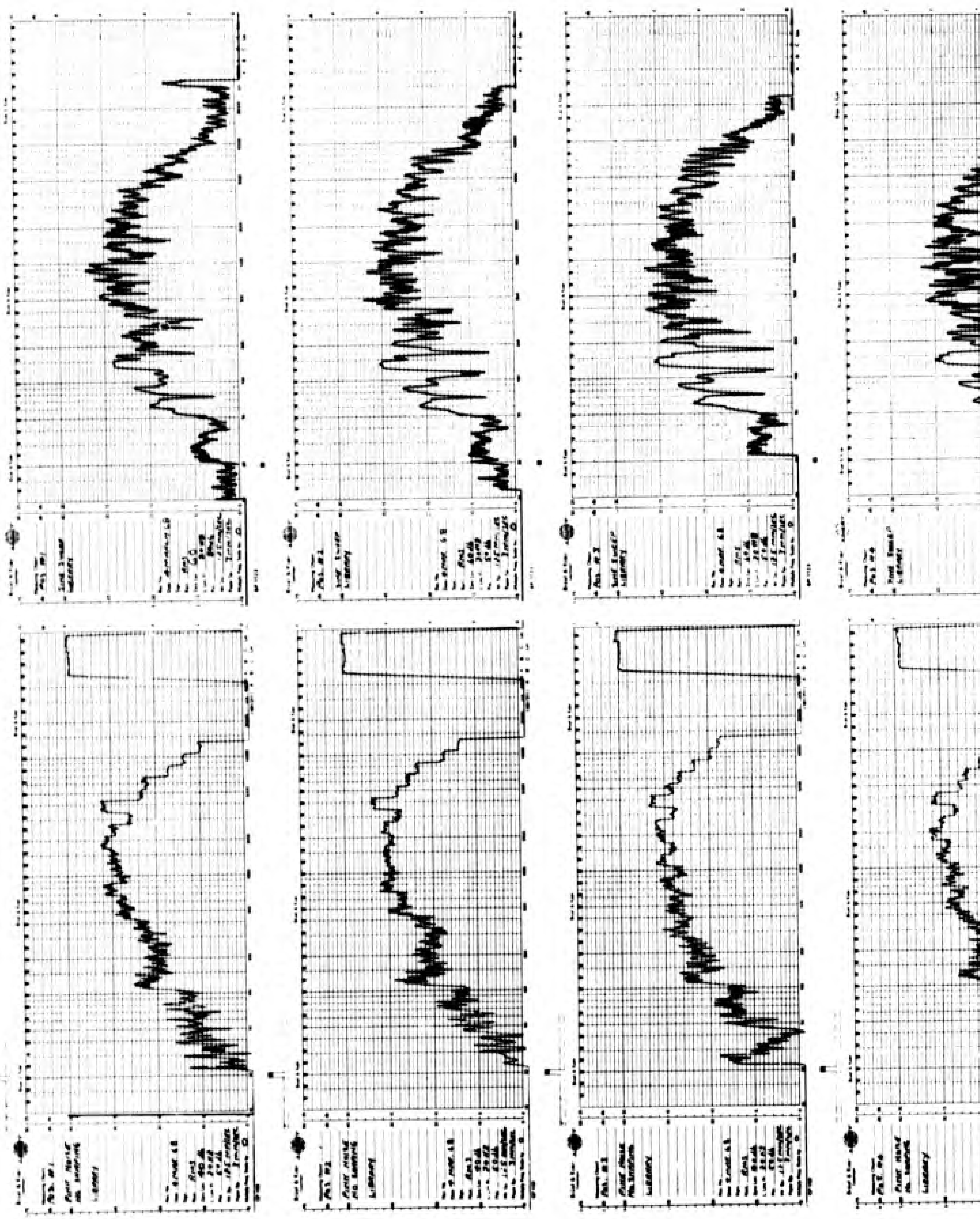


FIGURE A5 - Measured Frequency Response in 1500 ft$^3$. Acoustic Facility Configured for Subjective Judgement Experiments. Left Hand Column Shows 1/3 Octave Band Analysis of Microphone Output at Each of Five Seating Positions (With Seat 3 Located at Center of Cross-Section). Input to Shaping Filters is Pink Noise of Constant 1/3 Octave Band Level. The Right Hand Column is the Overall Sound Pressure Level, Measured at the Various Seat Positions, for a Constant Amplitude Sine Wave Input, Sweeping Through the Frequency Range.



A

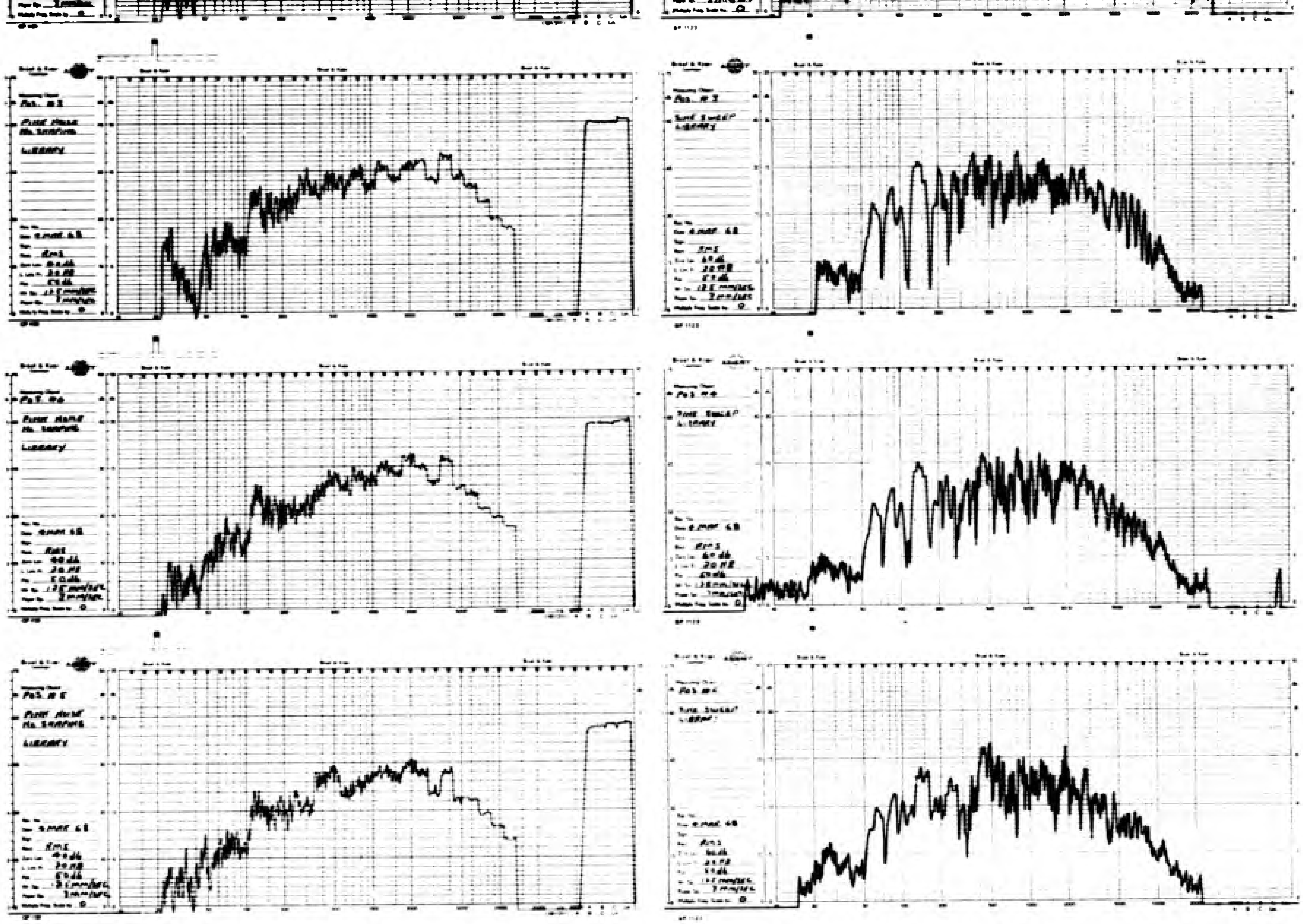


FIGURE A6 – Measured Frequency Response at 5 Typical Seat Locations in the Semi-Reverberant Room Used in Pilot Experiment 1. Left Hand Column Shows 1/3 Octave Band Analysis of Microphone Output at Each of Five Seating Positions (With Seat 3 Located at Center). Input to Speaker Network is Pink Noise of Constant 1/3 Octave Band Level. The Right Hand Column is the Overall Sound Pressure Level, Measured at the Various Seat Positions, For a Constant Amplitude Sine Wave Input, Sweeping Through the Frequency Range.



B