JIMMY M. CAWTHORN

NASA Contractor Report 159132

NASA-CR-159132

1980 0002614

THE EFFECT OF THE DURATION OF JET AIRCRAFT FLYOVER SOUNDS ON JUDGED ANNOYANCE

Kevin P. Shepherd

UNIVERSITY OF UTAH
Salt Lake City, Utah 84112

NASA Grant NSG-1272
September 1979

NASA
National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665

THE EFFECT OF THE DURATION OF JET AIRCRAFT FLYOVER SOUNDS ON JUDGED ANNOYANCE

by

Kevin P. Shepherd *

University of Utah

SUMMARY

Two laboratory studies were conducted to examine the effect of the duration of jet aircraft flyover sounds on annoyance. Both studies utilized a nine point numerical category scaling technique. The first study revealed various methodological problems which prevented the precise measurement of the effect of duration; the duration of aircraft sounds tends to be negatively correlated with the high frequency content of the sounds which results in each noise rating scale having a different optimum duration correction. It was also found that flyover sounds which exhibit a rapid change in spectrum were judged to be particularly annoying. This effect was subjectively equivalent to a maximum of approximately 5 dB, and was capable of completely negating the effect of duration.

The second study, which was designed to overcome these various problems, confirmed the "equal energy" hypothesis. The duration correction used in the EPNL procedure is therefore considered to be appropriate.

* Presently with The Bionetics Corporation, Hampton, Virginia

INTRODUCTION

Several laboratory studies have examined the effect of the duration of aircraft flyover sounds on perceived annoyance. These studies using bands of noise or synthesized flyover sounds have generally shown that duration is of importance, although the reported magnitude of the duration correction has varied considerably, from one to six dB per doubling of duration. (1-7). Of the studies using real aircraft flyover sounds, some reported no effect of duration (8, 9, 10), while others generally reported a small effect (11-14).

An explanation of some of these differences in results is provided by Little and Mabry (15) who have shown that the magnitude of the observed duration effect can be influenced by the instructions given to the test subjects. Parry (16) claims that a duration effect is not observed unless subjects are specifically asked to rate duration i.e., a duration cue is given. However, some of the previously quoted references do not support such a conclusion. It is believed that the conflicting results can be explained, at least in part, by other differences in methodology and the choice of noise stimuli.

For laboratory experiments involving recordings of aircraft presented at their real-life noise levels, there is a high negative correlation between their peak noise level and duration, with the result that it is difficult to statistically separate their independent effects. This is particularly true if the experiment utilizes just one aircraft type. This problem can be partially overcome by using a large number of flyovers of different aircraft types, or by presenting each of the sounds at several different peak noise levels. It is interesting to note that of the two studies using real aircraft sounds and reporting large duration corrections, one (13) used the first approach and the other (14) used the latter approach. There is also some evidence from these two studies that the duration dependency is different for different noise sources.

This report describes two studies which examined the subjective effect of the duration of aircraft flyover sounds. The first study, which was concerned with possible differences between types of aircraft, had rather unexpected results and the analyses highlighted various methodological problems which prevented an accurate assessment of the effect of duration. The second study, which was designed to overcome these problems, enabled a precise estimate of the subjective effect of duration to be obtained.

ABBREVIATIONS AND SYMBOLS

EPNL	Effective Perceived Noise Level
PNLT	Tone-corrected Perceived Noise Level
PNL	Perceived Noise Level
dB(A)	
dB(B)	
dB(C)	Maximum weighted sound pressure level.
dB(D_1)	Weights may be found in reference 21.
dB(D_2)	
dB(D_3)	
C dB(A)	Composite dB(A), maximum 1/3-octave band levels are weighted regardless of time when they occur.
MK.VII	Stevens Mark VII weighting (ref. 22)
dB(NN)	Ollerhead's weighting (ref. 9)
D_5	5 dB-down duration
D_{10}	10 dB-down duration
D_{20}	20 dB-down duration
T_{60}	Time (seconds) for which signal exceeds 60 dB(A)
T_{70}	Time (seconds) for which signal exceeds 70 dB(A)
T_{80}	Time (seconds) for which signal exceeds 80 dB(A)
T_1	Integrated duration correction (EPNL procedure)
Dur	Total audible duration (seconds)
ONS	Time between onset of audibility and peak (seconds)

ABBREVIATIONS AND SYMBOLS

HF	Proportion of A-weighted energy above 800 Hz
SC	Spectral change occurring between points either side of time of peak noise level
SSV	Subjective scale value
$\overline{SSV}$	SSV averaged across subjects
r	Pearson product moment correlation coefficient
R	Multiple correlation coefficient

STUDY I

EXPERIMENTAL DESIGN

Three aircraft types were chosen for this study. Narrow-bodied jets with low bypass ratio engines were represented by the Boeing 737 and wide-bodied jets with high bypass ratios were represented by the McDonnell Douglas DC-10. Also included was the only passenger-carrying supersonic aircraft, Concorde.

Five recordings, having a wide range of durations, were selected for each aircraft type. Both approaches and departures were included and attempts were made to select the flyovers so that each source was represented by sounds having the same range of durations. In order to reduce the negative correlation between peak level and duration, Matrix 1 was formed, consisting of five peak noise levels (L_1 - L_5) and five durations (D_1 - D_5).

MATRIX 1

	D_1	D_2	D_3	D_4	D_5
L_1	1	2	3	4	5
L_2	6	7	8	9	10
L_3	11	12	13	14	15
L_4	16	17	18	19	20
L_5	21	22	23	24	25

A similar matrix was constructed for each aircraft type, giving a total of 75 sounds. It is desirable that all subjects judge all sounds and that listening sessions be of a reasonable length. To meet this objective the 75 sounds were randomly assigned to three sessions, each session consisting of five tapes of five sounds each. Matrix 2 shows the arrangement of the 75 sounds among the sessions and tapes.

MATRIX 2

Tape	Listening Session 1				
A_1	54	15	61	05	41
B_1	28	17	40	34	06
C_1	62	47	24	55	75
D_1	16	46	42	68	20
E_1	58	70	08	59	45
Tape	**Listening Session 2**				
A_2	50	31	44	07	22
B_2	69	64	37	52	09
C_2	03	12	21	43	66
D_2	01	67	13	74	33
E_2	27	04	53	38	56
Tape	**Listening Session 3**				
A_3	65	18	11	10	29
B_3	49	39	51	71	57
C_3	63	02	35	36	72
D_3	30	60	23	25	14
E_3	32	26	73	48	19

In order to minimize the effects of subject learning and fatigue, the position of each tape in each session was systematically varied for each subject group. Let the five tapes in session 1 be represented by the letters A_1 - E_1. The order of presentation of tapes in session 1 for each subject group is given by Matrix 3.

MATRIX 3

Subject Groups	Order of Presentation of Tapes				
1	A_1	B_1	E_1	C_1	D_1
2	B_1	C_1	A_1	D_1	E_1
3	C_1	D_1	B_1	E_1	A_1
4.	D_1	E_1	C_1	A_1	B_1
5	E_1	A_1	D_1	B_1	C_1

This matrix was repeated for the tapes of the other two listening sessions (A_2 - E_2 and A_3 - E_3). Thus, each tape occupied each position in a session just once.

Again, in order to minimize the effects of subject learning and fatigue, the order of presentation of the three listening sessions (S_1 - S_3) was systematically varied as follows:

MATRIX 4

Subject Group	Session Order		
1	S_1	S_2	S_3
2	S_2	S_3	S_1
3	S_3	S_1	S_2
4	S_3	S_2	S_1
5	S_1	S_3	S_2

Thirty subjects, divided into five groups of six subjects each, were used in this experiment.

EXPERIMENTAL EQUIPMENT AND PROCEDURES

Stimuli

All recordings were made directly under aircraft flight tracks in areas where background noise levels were typically 35 dB(A). The recording sites were chosen under both landing and take off routes and were at various distances from the airport to ensure a wide range of noise durations.

The test sounds were chosen according to their measured duration and their signal to noise ratio. The test tapes were made from the master recordings using high grade tape recorders and attenuators which were adjusted to achieve the desired peak noise levels in the testing room. Each of the fifteen recordings were presented at five peak noise levels, nominally 65 - 85 dB(A) in 5dB steps.

Experimental Facility

The test was conducted in the Exterior Effects Room of the Aircraft Noise Reduction Laboratory of the National Aeronautics and Space Administration Langley Research Center. This is an auditorium-like room having a volume of 330 m^3 and a reverberation time of approximately 0.5 seconds at 1k Hz. There are six two-way coaxial speakers mounted in the ceiling by which the test sounds were presented to the subjects. The seating locations of the subjects were chosen to minimize the variation in noise levels between the seats. A microphone was positioned in the center of the room with the subjects and was used to monitor the test stimuli.

Subjects

Thirty human subjects were hired for the experiment. Participation was voluntary and each subject was screened to meet minimum audiological standards, 20 dB hearing level. The subjects were randomly divided into five groups of six subjects each. One half of the subjects were male and had an age range of 18 - 32 with a median of 26 years. The female age range was 18 - 56 with a median of 29 years.

Testing Procedure

Upon arrival at the test facility, subjects were asked to read the instructions (Appendix A), and were introduced to the rating sheet (Appendix B) which they would use to record their annoyance rating for each noise. A numerical category scaling technique was used, with the ends of the nine point scale labelled "not at all annoying" and "extremely annoying". Prior to the test the subjects were asked to listen to three sounds in order to give them an indication of the kinds of sounds they were to judge.

Six subjects participated in each testing session, which lasted approximately two hours. The session was composed of three segments separated by 10 minute rest periods during which the subjects left the testing room. Each test sound was verbally identified with a number corresponding to its position on the response sheet and there were intervals of approximately six seconds between sounds, during which the number of the next sound was given. When the subjects had completed the test they were dismissed and given post-test audiograms.

DATA ANALYSIS AND RESULTS

Acoustical Analysis

The noise stimuli were monitored throughout the experimental program. The means and standard deviations of the peak noise levels measured across test sessions showed that the sounds were presented to the subject groups extremely uniformly. (Table 1).

The stimuli were analyzed into various composite and maximum frequency weighted units using the one-half second one-third-octave band time histories. In addition, various duration measures were calculated. These included the 5, 10, 15 and 20 dB(A)-down durations, found using a graphic level recorder, and an "effective" duration which was calculated from the integrated duration correction employed in the EPNL procedure (Table 2). The times (in seconds) for which the signal exceeded 60, 70 and 80 dB(A) and an "onset duration" correction were also measured. This latter correction was suggested by Nixon et al (17) who showed that the approaching (increasing sound level) part of a flyover is more annoying than the receding part.

Analysis of Variance

The effectiveness of the experimental design was examined by the use of analysis of variance. The design, in part, consisted of five tapes of five sounds each, the order of presentation of tapes to subject groups being determined by a Latin square. The analysis of variance is presented in Table 3. The significant differences between tapes is due to the experimental design, which, although randomly assigning sounds to tapes, made no attempt to balance their content. The significance of subject groups simply reflects the inter-subject variability. Matrix 4 of the experimental design was also examined; subject groups, sessions and session order were all found to be non-significant factors (Table 4).

The experiment was based upon a factorial design of five peak noise levels, five durations, and three noise sources. These factors were investigated by an analysis of variance in which all the main effects and some of the interaction terms were found to be significant at the 1% level (Table 5). As expected, the dominant variable was the peak noise level. The strength of the source-duration interaction term implies that either the durations were not the same for all noise sources or that the effect of duration is different for each aircraft type. The first explanation is possibly correct since although the flyover sounds were selected on the basis of their 10 dB(A)-down durations, considerable differences exist between their 5 and 20 dB(A) - down durations (Table 2). This will be further examined by regression analysis in a later section. Some of the other interaction terms were statistically significant but explained little of the total variance and were, therefore, considered to be of little importance.

Similar analyses were performed for each noise source separately (Table 6). All the main effects were found to be significant, with the peak noise level the dominant factor. The interaction terms were of little importance relative to the main effects.

Regression Analyses

Thirty subjects judged 25 sounds from each of three aircraft types. The arithmetic mean of the subjective scale values (SSV) was found for each of the sounds and, when plotted against peak dB(A), yielded least-squares regression lines (Figure 1) having correlation coefficients of 0.92 to 0.95. Tests were made between slopes and intercepts of these three lines and no statistical differences were found at the 5% level.

The subjective scale values were examined in terms of their relationship with several of the common noise rating scales. The three noise sources were analyzed both separately and in combination (Table 7). Many of the scales performed well, with the notable exception of EPNL which performed poorly when the aircraft were analyzed either separately or in combination.

There is an alternative method by which the noise ratings may be compared. An "equal annoyance" level was found for each sound and each rating scale using the method illustrated in Figure 2. For each sound the mean SSV is plotted against a particular rating scale and the equal annoyance level corresponding to the mean SSV of all the judgements (all sounds, all aircraft), is found from the least-squares regression line. If the subjective data were errorless, the perfect noise rating scale would result in identical equal annoyance levels for all sounds. The various noise rating scales can, therefore, be compared by examination of the distributions of the equal annoyance levels. Table 8 presents these results and the rank order of the rating scales based simply on their correlation with the subjective scale values. The rank orders based on the two procedures are in complete agreement.

Effects of Noise Duration. Various duration corrections including the logarithms of the 5, 10 and 20 dB - down durations were added to the regression equations as independent variables. The regression coefficients of the duration corrections were typically small and negative (Table 9). Since the regression coefficients for the three sources were not significantly different they have been combined.

These resulting negative duration corrections were rather unexpected, although they do explain the poor performance of EPNL in the previous regression analyses. It was possible to gain further insight into the effect of duration on perceived annoyance by plotting SSV's against duration with peak noise level held constant. (Figure 3 - 5) There is apparently no simple relationship between annoyance and duration.

There remains, however, the need to explain the observed subjective differences between flyovers presented at the same peak noise level. First it is necessary to determine whether these differences are real or whether they are simply experimental error. From the standard deviations of the subjective judgements, it was concluded that differences of approximately 0.5 SSV units between stimuli judgements were statistically significant ($p = 0.05$). Therefore, some of the observed differences are indeed real.

There are various possible explanations for the observed differences. The first is that dB(A) is a poor rating scale. Similar analyses were conducted for other rating scales with similar results. Another possibility is that different duration measures may give more reasonable results. This was attempted, with no real improvement (Table 9).

In listening to the sounds which were judged most annoying it was felt that they gave an impression of close proximity. There are various characteristics of the sounds which might produce this impression, including the duration, the rise time and the rate of change of sound level. These latter two statistics are found to be highly related to duration and offered no significant improvement. Another possible explanation is Doppler shift, due to an aircraft being a moving source. It is clearly impossible to extract Doppler shift information from the spectral data, but an estimate was made on the basis that when an aircraft is overhead, the Doppler shift is inversely proportional to its altitude. Estimates of the altitude of the flyovers were made and added to the regression equations as independent variables, but no significant improvement was found.

The spectra at the peak dB(A) level of the sounds were examined and it appeared that the most annoying sounds contained the most high frequency energy. This is investigated further in the following section.

Effects of High-Frequency Energy. It is not immediately obvious how to quantify the amount of high frequency energy present in a flyover sound. The following simple approach was taken. The 1/2 second, 1/3 octave band spectra for the period of 4 seconds either side of the peak dB(A) point were tabulated for each flyover. The band levels were A-weighted and the proportion of energy above various high frequency/low frequency boundaries was calculated. The proportions derived from the different high/low boundaries were found to be highly correlated and so the boundary was arbitrarily selected to be 800 Hz. The maximum proportion of A-weighted acoustic energy above 800 Hz in any 1/2 second time interval was designated HF.

The regression analyses were rerun with the inclusion of this new variable and a significant improvement was found:

$$
\begin{aligned}
\overline{SSV} \;&\propto\; dB(D_3) \; -53.15 && R = 0.936 \\
&\propto\; dB(D_3) \; -4.26 \log (D_{10}) \; -51.11 && R = 0.948 \\
&\propto\; dB(D_3) \; +0.05\ HF \; -57.16 && R = 0.958 \\
&\propto\; dB(D_3) \; -0.43 \log D_{10} + 0.05\ HF \; -56.74 && R = 0.958 \\
&\propto\; dB(A) \; -53.53 && R = .918 \\
&\propto\; dB(A) \; -5.37 \log D_{10} \; -48.16 && R = .937 \\
&\propto\; dB(A) \; +0.10\ HF \; -54.55 && R = .960 \\
&\propto\; dB(A) \; +0.10\ HF \; +0.45 \log D_{10} \; -55.15 && R = .961
\end{aligned}
$$

The addition of the variable HF resulted in an improvement for all of the noise rating scales, both with and without the inclusion of duration corrections. This effect is illustrated by Figure 6 in which equal annoyance levels are shown as a function of HF. There is clearly a strong linear relationship ($r = -0.79$) between these two variables such that increasing high frequency content results in increased noisiness.

The above equations yielded another particularly interesting result; the addition of HF caused a large change in the coefficient of duration. A high correlation ($r = -0.7$) was found between the two variables HF and duration, but this is not unexpected since, for any aircraft, increasing the altitude will always result in longer durations and reduced high-frequency content due to atmospheric attenuation. The implication of this result is that experiments using real aircraft sounds will display a duration correction which will vary according to the frequency weighting of the noise rating scale. This variation will be such that those rating scales having the most high-frequency emphasis will require the largest duration corrections. This is demonstrated by the following:

	dB per Doubling of Duration
dB(A)	-1.5
dB(D_3)	-1.3
dB(NN)	-0.7
dB(A) + HF	0.1

dB(NN) is a weighting proposed by Ollerhead (9) which has extreme high frequency emphasis.

Other studies have found similar results. For example Kryter (18) found that the addition of a duration correction to dB(A) had a very adverse effect, whereas the addition to the high-frequency emphasizing units (dBD, PNL) had a less adverse effect. Ollerhead (13) reported that a duration correction adversely affected dB(A), but was of positive benefit to PNL and the D-weighted units. Powell (19) found that PNL required a larger duration correction than dB(A).

It is still necessary to explain why this variable HF is apparently of subjective importance. Initially it was thought that the inadequate high-frequency emphasis of the A-weighting network was the cause, but upon closer examination it was found that the optimum high frequency emphasis would have to be enormous (more than 15 dB relative to dB(A)), and it is unlikely that equal noisiness contours could be so much in error. An alternative explanation may lie with some variable highly correlated with the high-frequency content. A possible clue may be found in a study conducted by Ollerhead (9) in which he made recordings of a single aircraft at several altitudes to achieve a range of durations and showed that duration did not influence annoyance. He also made a long duration recording directly under a circling aircraft and shaped the recording to give similar durations to those of the flyovers used in the first experiment. This time a duration effect was observed. Ollerhead noted the absence of Doppler shift in the second experiment and concluded that this was probably responsible for the results. He attempted to confirm this in a later experiment, but failed (13). There is clearly some characteristic of the real flyovers that counteracted the effect of duration. It is worth noting that the Doppler shift was not the only difference between the two experiments; the sounds in the first experiment had spectral variations, the second group of experimental sounds did not. There is a possibility, therefore, that the change in the spectrum during a flyover is of importance since these changes would generally be highest for the short duration flyovers and thus counteract the effect of duration. Also the spectrum change is probably correlated with the high frequency content and may explain the good performance of this latter variable in the previous regression analyses.

The Effects of Spectral Change. A method had to be found to quantify the spectral change that occurs during a flyover. It was felt that the spectral changes close to the peak were likely to be of most importance so the time intervals 1 1/2, 2 1/2 and 5 seconds either side of the peak were examined. It was apparent that the changes in each 1/3-octave band were due more to overall level changes rather than changes in the spectrum. A measure of spectral change that was independent of overall level changes was obtained by calculating dB(A) values at the peak and at the points 1 1/2, 2 1/2 and 5 seconds either side of the peak. These dB(A) changes were subtracted from the changes in each 1/3 octave band which were than logarithmetically averaged to give a measure of spectral change that was independent of overall level changes. An example of the computation of spectral change is given in Table 10, where, for reasons of ease of understanding, octave band rather than 1/3 octave band data have been used.

The spectral changes occurring during the three time intervals were found to be highly correlated with each other, and negatively correlated with duration, and each had a similar effect when added to the regression equations. These results are presented in Table 11. The addition of this variable (SC) was statistically significant for the three aircraft, both singularly and in combination. The relationship between noisiness and spectral change is illustrated in Figure 7, in which equal annoyance levels are plotted against the spectral change occurring between the points 2 1/2 seconds either side of the peak noise level. There is clearly a strong linear relationship between these two variables ($r = -0.8$) such that increasing spectral change results in increased noisiness.

The addition of the spectral change statistic to the regression equations resulted in large changes in the coefficient of duration. These two variables, duration and spectral change, are negatively correlated (typically $r = -0.7$) so that flyovers having the shortest duration generally have the most spectral change and those having the longest durations have the least spectral change. There is clearly a possibility, therefore, that the effects of these two variables will tend to counteract one another. This can be seen in Table 10, where, for the DC-10 stimuli, the coefficient of duration changed from -7 to +12 when the variable SC was added to the equation. In other words, when no account was taken of spectral change, duration appeared to have no positive effect on noisiness. However, when spectral change was included, the duration had a strong positive effect on annoyance.

The range of values of spectral change was approximately 10 dB and given a regression coefficient of 0.6 (Table 11, all aircraft), this implies that this variable is subjectively equivalent to peak level changes of approximately 6 dB. This figure varies a little depending upon the choice of rating scale, but is generally in the range of 5 - 6 dB.

The high correlation between duration and spectral change results in wide confidence intervals for the regression coefficient of duration. For example, for the B 727 stimuli, the coefficient of duration has a 95% confidence interval of approximately ±5. This illustrates that for any set of stimuli having differing spectral changes, precise estimates of the subjective effect of duration can not be determined.

STUDY II

EXPERIMENTAL DESIGN

The primary aim of this study is to reduce the negative correlation between duration and spectral change, thus enabling the precise measurement of the subjective effect of the duration of aircraft sounds. It is also intended that the correlation between duration and high frequency content be reduced so that the measured duration correction will be independent of the choice of noise rating scale.

The previous study showed that the spectral change occurring near the peak of the flyover was of subjective importance. This spectral change is due to the directivity characteristics of a jet engine and should be best studied by the use of synthesized flyover sounds. However, a similar phenonemon, Doppler shift may be investigated by using real flyover sounds and the present study is aimed at this.

The little previous research on the subjective importance of Doppler shift has produced conflicting results. Ollerhead (a) produced results which implied that Doppler shift was of importance and unsuccessfully attempted to confirm this in a later study (13). Pearsons et al (20) showed that Doppler shift had a rather weak effect.

Ideally the experiment should be designed so that the variables peak noise level, duration, Doppler shift and spectrum are orthogonal to one another. Achieving such a design using real aircraft flyovers presents several difficulties, the major ones being the relationships between the variables duration, specturm, and Doppler shift. There is clearly a tendency for duraation to be negatively correlated with both Doppler shift and high frequency content. This problem was partially overcome by manually altering the duration of flyovers when rerecording, thus enabling duration to be varied independently of Doppler shift and spectrum. In order to achieve a range of Doppler shifts, flyovers at several altitudes would be required. However, this would result in the confounding of Doppler shift and spectrum since the higher altitude flyovers would have less high-frequency content due to atmospheric attenuation. Consequently it was necessary to perform some spectral filtering in order to eliminate this confounding. In order to further extend the range of Doppler shifts, some of the flyover recordings were played at one-half normal speed, thus halving the maximum Doppler shift.

A Boeing 727 was chosen for this study due to its wide availability for recording purposes and because its spectrum has minimal pure tone components, thus simplifying the necessary frequency filtering. The three basic test sounds were recordings made of take offs at altitudes of approximately 120, 240 and 480 meters. The 10 dB(A) - down durations were found to be approximately 5, 10, and 20 seconds. These sounds were played in the testing room and their spectra measured using a real time analyzer. The spectra averaged over the time period defined by the 5 dB-down points were found, and, as expected, there was a decrease in the high frequency content for the higher altitude flyovers (Figure 8). These curves were used to determine the high frequency attenuation required to achieve equalized spectra. Both the filtered and unfiltered sounds were then manually adjusted to give other durations when rerecorded. A description of the test sounds is presented in Table 12. The set of stimuli may also be described by the following matrix which presents the combinations of spectrum and Doppler shift that were included in the study. All of the sounds were presented with three different durations except those having a Doppler shift of 4B, which were presented with only two durations.

Spectrum			
$f_{1/2}$	f_1	f_2	
B/2	B	2B	
B	2B	4B	B = Doppler shift
2B	4B		

Each of the sounds was presented at four peak levels of nominally 65, 72, 79 and 86 dB(A), giving a total of 88 test stimuli.

In order to simplify the experimental design, two of the sounds were repeated at each of the four peak noise levels. Hence there were twenty-four sounds each presented at four peak levels. The design was composed of four sessions, each session containing four tapes of six sounds each. The distribution of the sounds on the tapes is given by Table 13. Each peak level occurs in each possible position on a tape an equal number of times and all twenty-four stimuli occur in each session.

The order of presentation of tapes within each session was systematically varied for each subject group as follows:

Order of Tapes Within Each Session

Subject Group I	A	B	D	C
II	B	C	A	D
III	C	D	B	A
IV	D	A	C	B

The order of presentation of sessions for each subject group was as follows:

Order of Sessions

Subject Group					
I	S_1	S_2	S_4	S_3	
II	S_2	S_3	S_1	S_4	
III	S_3	S_4	S_2	S_1	
IV	S_4	S_1	S_3	S_2	S_1 - S_4 are sessions 1 - 4

Four groups of ten subjects were used in the experiment. Half of the subjects were male and had an age range of 18 - 59 years with a median of 28 years. The female age range was 18 - 64 with a median of 28 years.

The experimental facilities and procedures were identical to those employed in the first study.

DATA ANALYSIS AND RESULTS

Acoustical Analysis

The stimuli were monitored throughout the experimental program. These monitored levels are presented in Table 14, from which it is clear that the sounds were presented uniformly to the subject groups.

The stimuli were analyzed into various composite and maximum frequency weighted units using the one-third-octave time histories. Typical peak level spectra are presented in Figure 9. Various duration measures were calculated, including an "effective" duration which was found from the integrated correction used in the computation of EPNL. These duration measures are presented in Table 15.

Analysis of Variance

The order of presentation of sessions and tapes within sessions was determined by Latin squares. The analyses of variance of these are presented in Tables 16 and 17. Differences between sessions, session order and tape order were not statistically significant. There were significant differences between subject groups and tapes which reflect the intersubject variability and the unequal distribution of the sounds and their peak levels amongst the tapes.

The design of the experiment with regard to duration, Doppler shift and spectra is rather complex, but can be briefly summarized as follows:

		Spectrum		
		$f_{1/2}$	f_1	f_2
	B/2	4, 5, 6		
	B	13, 14, 15	1, 2, 3	
Doppler Shift	2B	20, 21, 22	10, 11, 12	7, 8, 9
	4B		18, 19	16, 17

The numbers refer to the sounds (see Table 12), and indicate that each occupied cell of the matrix contains sounds of three durations except those having a Doppler shift of 4B, which contain only two. Analysis of variance was used to investigate duration by examination of all the sounds except those numbered 16 - 19. The experiment may then be considered to consist of six stimuli, each being presented with three durations and four peak noise levels. Table 18 shows all the main effects to be significant and the interaction effects to be relatively weak. The peak level of the sound is clearly the dominant variable. The strength of the "sounds X duration" interaction is probably due to sounds of nominally the same duration displaying differences in their measured durations (Table 15). The "sounds X peak noise level" interaction term is probably due to the peak levels being set in terms of dB(A) which may not sufficiently account for the spectral differences between the sounds.

Analysis of variance was used to examine the effect of the differences in the spectral content of the sounds, independently of duration and Doppler shift, by careful selection of the sounds. For example, sounds 7 - 12 and 20 - 22 have a range of spectra but the same durations and Doppler shift. These results are presented in Table 19, in which all the main effects were found to be significant. The same procedure was applied to Doppler shift (Table 20) and again the main effects were found to be significant. The peak noise level is clearly the dominant variable; the other variables will be further examined in the following sections.

Regression Analysis

The mean annoyance response was found for each stimulus and correlated with various noise measures (Table 21). All of the scales performed well, but examination of the correlation coefficients and the standard errors of estimate showed EPNL to be statistically superior to all other rating scales. This is presumably due to EPNL being the only rating scale which employs a duration correction.

The effect of duration was further investigated by adding the various duration measures to the regression equations as independent variables (Table 22). In all cases the addition of the duration measures was statistically beneficial and, in contrast to the previous study, their regression coefficients were positive. The "equal energy" hypothesis predicts that the integrated duration correction (log D_1) would have a coefficient of value 10. From examination of the standard errors associated with the coefficients of this variable, it was concluded that the results were in complete statistical agreement with the hypothesis.

Equal annoyance levels were calculated for each of the 22 sounds and for each rating scale using the procedure described earlier (Figure 2). These levels may be used to illustrate the effect of duration. Figure 10 shows lines having constant Doppler shift and spectrum and clearly illustrates the relationship between duration and annoyance.

Doppler shift was also examined by using the equal annoyance levels. Figure 11 shows lines of constant duration and spectrum, but no clear effect of Doppler shift is apparent. When added to the regression equations as an independent variable no significant improvements were found.

A possible reason for Doppler shift having no clear effect may be due to the choice of stimuli, since a recent study by McCurdy (7) showed that synthesized flyover sounds having significant pure tone content require a small correction for Doppler shift, whereas sounds with no tones do not. The stimuli used in the present study were chosen for their lack of tones.

CONCLUSIONS

Two studies were conducted concerning the subjective reaction to the duration of aircraft flyover sounds. The first study highlighted several methodological problems which prevented the precise measurement of the effect of duration. Changes in the spectral characteristics of aircraft sounds caused by atmospheric attenuation were found to have a rather unexpected effect on the measured duration correction. As the altitude of an aircraft flyover is increased, the high-frequency attenuation is also increased, so that the high frequency content and the duration become negatively correlated. This results in each noise rating scale having a different optimum duration correction. It was also found in this study that flyovers which exhibit a rapid change in spectrum were found to be particularly annoying. Differences in the spectral change characteristics of the stimuli in this study were subjectively equivalent to peak level changes of approximately 5 dB, and this variable is therefore considered to be of importance. The effect of spectral change was sufficient to completely negate any effect of duration.

It is believed that these spectral phenomena are responsible for much of the disagreement found in the results of previous studies. Synthesized flyovers and bands of noise do not exhibit such spectral variations and therefore enable reasonable estimates of the subjective effect of duration to be made. On the other hand experiments using real aircraft sounds have reported a wide range of duration corrections, which, it is believed, is due to the spectral characteristics of the particular test stimuli.

The second study, which was designed to overcome the major problems encountered in the earlier study, confirmed the "equal energy" hypothesis. The duration correction used in the calculation of EPNL is therefore considered appropriate. This study also examined the effect of Doppler shift using aircraft noises with minimal pure-tone content. It was concluded that Doppler shift was not related to annoyance.

APPENDIX A

Instructions for Subjects

We are asking you to help us solve a problem concerned with noise; how annoying are various kinds of sounds? First we will ask you to listen to some of the sounds you will be judging so you will have some familiarity with them.

The sounds you are to rate will be presented to you one at a time. We would like you to try to imagine that you are hearing these sounds while out of doors. Please consider both the peak noise level and the duration of the noise when making your judgments. Listen to all of the sound before making your judgment. Notice that on your answer sheet each sound has nine possible ratings. "0" is for no annoyance while "8" is for extremely annoying. You should place the sounds on the scale according to their degree of annoyance. For example, a sound causing a small amount of annoyance may be scored a "2" or a "3", a sound causing a high amount of annoyance may be scored a "6" or "7", and so on.

Your ratings should reflect only your own opinion of the noise, that is what we want.

APPENDIX B

Typical Answer Sheet

NAME ______________________ DATE ______________

NOISE NUMBER	NOT AT ALL ANNOYING								EXTREMELY ANNOYING
1	0	1	2	3	4	5	6	7	8
2	0	1	2	3	4	5	6	7	8
3	0	1	2	3	4	5	6	7	8
4	0	1	2	3	4	5	6	7	8
5	0	1	2	3	4	5	6	7	8
1	0	1	2	3	4	5	6	7	8
2	0	1	2	3	4	5	6	7	8
3	0	1	2	3	4	5	6	7	8
4	0	1	2	3	4	5	6	7	8
5	0	1	2	3	4	5	6	7	8
1	0	1	2	3	4	5	6	7	8
2	0	1	2	3	4	5	6	7	8
3	0	1	2	3	4	5	6	7	8
4	0	1	2	3	4	5	6	7	8
5	0	1	2	3	4	5	6	7	8
1	0	1	2	3	4	5	6	7	8
2	0	1	2	3	4	5	6	7	8
3	0	1	2	3	4	5	6	7	8
4	0	1	2	3	4	5	6	7	8
5	0	1	2	3	4	5	6	7	8
1	0	1	2	3	4	5	6	7	8
2	0	1	2	3	4	5	6	7	8
3	0	1	2	3	4	5	6	7	8
4	0	1	2	3	4	5	6	7	8
5	0	1	2	3	4	5	6	7	8

TABLE I

MEAN AND STANDARD DEVIATION OF PEAK dB(A) VALUES OF EXPERIMENTAL SOUNDS

		D_1	D_2	D_3	D_4	D_5
B737	L_1	84.8 (0.40)	85.6 (0.49)	85.0 (0.30)	85.8 (0.40	85.0 (0.63)
	L_2	79.8 (0.40)	80.0 (0.30)	80.8 (0.40)	81.0 (0.30)	79.6 (0.49)
	L_3	73.6 (0.49)	74.0 (0.63)	75.4 (0.49)	74.6 (0.49)	75.2 (0.40)
	L_4	69.8 (0.40)	69.4 (0.49)	70.6 (0.49)	70.4 (0.49)	70.0 (0.49)
	L_5	64.4 (0.49)	64.0 (0.30)	64.8 (0.75)	66.0 (0.63)	65.4 (0.80)
DC10	L_1	84.4 (0.49)	84.6 (0.49)	85.2 (0.40)	85.0 (0.40)	85.2 (0.75)
	L_2	80.8 (0.40)	80.4 (0.49)	80.4 (0.40)	79.8 (0.40)	81.2 (0.40)
	L_3	74.2 (0.75)	75.4 (0.49)	75.2 (.79)	75.2 (0.75)	74.8 (0.40)
	L_4	70.4 (0.49)	69.8 (0.40)	70.6 (0.49)	70.2 (0.40)	70.0 (0.30)
	L_5	64.6 (0.49)	64.4 (0.49)	65.4 (0.49)	65.2 (0.74)	65.8 (0.40)
Concorde	L_1	85.2 (0.75)	85.4 (0.49)	85.0 (0.30)	84.0 (0.63)	86.0 (0.63)
	L_2	79.0 (0.63)	80.8 (0.40)	79.8 (0.40)	80.2 (0.40)	80.2 (0.75)
	L_3	75.2 (0.40)	74.6 (0.49)	75.4 (0.49)	74.6 (0.49)	76.0 (0.63)
	L_4	69.8 (0.40)	69.8 (0.40)	69.6 (0.49)	70.4 (0.49)	70.8 (0.40)
	L_5	65.4 (0.80)	65.0 (0.89)	65.4 (0.49)	65.0 (0.30)	66.0 (0.30)

Each cell contains the peak dB(A) for each sound averaged over the five presentations and the standard deviation (in parentheses) of the peak level. L_1 - L_5 and D_1 - D_5 refer to the peak levels and durations (see Matrix 1, experimental design).

TABLE 2

DURATIONS (IN SECONDS) OF THE SOUND STIMULI

Noise Source		5 dB(A) down duration	10 dB(A) down duration	15 dB(A) down duration	20 dB(A) down duration	Effective duration *
B737	D_1 **	3.0	5.0	8.0	13.0	2.39
	D_2	4.5	6.5	11.5	14.0	2.82
	D_3	8.0	14.0	25.0	35.0	5.46
	D_4	6.5	16.0	26.0	31.0	4.39
	D_5	14.0	28.0	40.0	53.0	5.44
DC10	D_1	2.5	5.0	8.0	13.5	2.09
	D_2	4.5	8.0	12.0	17.0	3.88
	D_3	10.0	14.0	17.0	25.0	4.98
	D_4	12.0	20.0	28.0	40.0	10.79
	D_5	14.0	27.0	39.0	50.0	6.98
Concorde	D_1	3.0	4.5	6.0	9.5	2.09
	D_2	5.0	7.5	9.0	11.0	2.60
	D_3	11.0	15.0	19.0	24.0	4.50
	D_4	12.0	18.5	23.0	32.0	8.48
	D_5	17.0	29.0	32.0	36.0	9.45

* The "effective duration" was calculated from the integrated duration correction used in the calculation of EPNL

** D_1 - D_5 refer to the durations of the flyovers as given in Experimental Design

TABLE 3

ANALYSIS OF VARIANCE OF MATRIX 3

	Source of Variation	Degrees of Freedom	Sums of Squares	Means Squares	F
Session 1	Subject Groups	4	73.46	18.36	9.53 *
	Tapes	4	267.67	66.92	34.64 *
	Tape Order	4	6.58	1.64	0.85^{ns}
	Residual	12	23.10	1.93	
	Total	24	370.81		
Session 2	Subject Groups	4	12.06	3.01	1.79^{ns}
	Tapes	4	159.27	39.82	23.66 *
	Tape Order	4	5.96	1.49	0.89^{ns}
	Residual	12	20.15	1.68	
	Total	24	197.44		
Session 3	Subject Groups	4	7.78	1.94	0.85^{ns}
	Tapes	4	195.85	48.96	21.45 *
	Tape Order	4	14.90	3.72	1.63^{ns}
	Residual	12	27.22	2.28	
	Total	24	245.75		

* significant at 1% level

ns - not significant

TABLE 4

ANALYSIS OF VARIANCE OF MATRIX 4

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F
Subject Groups	4	244.74	61.19	2.03^{ns}
Sessions	2	103.01	51.51	2.05^{ns}
Session Order	2	103.89	51.95	2.41^{ns}
Residual	6	152.27	25.38	
Total	14	603.91		

ns - not significant

TABLE 5

ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F
1. Subjects	29	2388.39	82.36	68.00 *
2. Source (B737, etc)	2	31.48	15.74	7.54 *
3. Peak Noise Level	4	4298.65	1074.66	324.00 *
4. Duration	4	219.21	54.80	27.50 *
1 X 2	58	121.18	2.09	1.73 *
1 X 3	116	385.69	3.32	2.74 *
1 X 4	116	231.66	1.99	1.64 *
2 X 3	8	38.75	4.84	4.56 *
2 X 4	8	307.52	38.44	26.20 *
3 X 4	16	40.76	2.55	1.96 *
1 X 2 X 3	232	246.58	1.06	0.88
1 X 2 X 4	232	341.68	1.47	1.21
1 X 3 X 4	464	604.35	1.30	1.17
2 X 3 X 4	32	118.18	3.69	3.05 *
Residual	928	1122.56	1.21	
Total	2249	10496.67		

* significant at 1% level

TABLE 6

ANALYSIS OF VARIANCE FOR EACH AIRCRAFT TYPE

	Source of Variation	Degrees of Freedom	Sums of Squares	Means Squares	F
B737	1. Subjects	29	881.63	30.40	28.33 *
	2. Peak Level	4	1574.88	393.72	263.58 *
	3. Duration	4	212.99	53.25	32.58 *
	1 X 2	116	173.27	1.49	1.39 *
	1 X 3	116	189.57	1.63	1.52 *
	2 X 3	16	43.18	1.70	2.52 *
	Residual	474	497.85	1.07	
	Total	749	3573.38		
DC10	1. Subjects	29	739.48	25.50	20.08 *
	2. Peak Level	4	1561.91	390.48	209.93 *
	3. Duration	4	227.64	56.91	31.67 *
	1 X 2	116	215.77	1.86	1.47 *
	1 X 3	116	208.43	1.80	1.42 *
	2 X 3	16	32.03	2.00	1.58
	Residual	464	589.08	1.27	
	Total	749	3574.35		
Concorde	1. Subjects	29	888.46	30.64	22.21 *
	2. Peak Level	4	1200.60	300.15	143.14 *
	3. Duration	4	86.10	21.52	14.24 *
	1 X 2	116	243,23	2.10	1.52 *
	1 X 3	116	175.34	1.51	1.10
	2 X 3	16	83.73	5.23	3.79 *
	Residual	464	640.02	1.37	
	Total	749	3317.49		

* significant at 1% level

TABLE 7

REGRESSION ANALYSIS OF SUBJECTIVE SCALE VALUES AND

VARIOUS RATING SCALE UNITS - EXPERIMENT 2

		dBA	PNL	PNLT	EPNL	dB(B)	dB(C)	dB(D_1)	dB(D_2)	dB(D_3)	MKVII	CdB(A)
B737	$\overline{SSV}$	.91	.90	.87	.79	.81	.81	.87	.89	.94	.89	.90
DC10	$\overline{SSV}$	.92	.92	.91	.78	.77	.72	.89	.91	.95	.90	.88
Concorde	$\overline{SSV}$	.94	.92	.92	.86	.91	.89	.91	.92	.93	.92	.93
All Sources	$\overline{SSV}$	.92	.90	.89	.80	.82	.78	.87	.89	.94	.89	.90

$\overline{SSV}$ Mean subjective scale value

TABLE 8

RANK ORDER OF RATING SCALES

Rank Order	Standard Deviation of Equal Annoyance Levels		Correlation With Raw SSV	
1	$dB(D_3)$	2.04	$dB(D_3)$	0.94
2	dB(A)	2.57	dB(A)	0.92
3	PNL	2.72	PNL	0.90
4	PNLT	2.90	PNLT	0.89
5	$dB(D_2)$	2.91	$dB(D_2)$	0.89
6	MKVII	2.98	MKVII	0.89
7	$dB(D_1)$	3.22	$dB(D_1)$	0.87
8	dB(B)	3.96	dB(B)	0.82
9	EPNL	4.35	EPNL	0.80
10	dB(C)	4.57	dB(C)	0.78

TABLE 9

REGRESSIONS OF SSV's AND VARIOUS RATING SCALES AND DURATION CORRECTIONS

B737

$\overline{SSV} \propto$ dB(A) - 9.15 $\log_{10}$ (D_5) - 45.20 (R = .048)

$\propto$ dB(A) - 9.25 $\log_{10}$ (D_{10}) - 43.00 (R = .966)

$\propto$ dB(A) - 10.30 $\log_{10}$ (D_{20}) - 38.35 (R = .962)

$\propto$ dB(A) - 6.36 $\log_{10}$ (ONS) - 48.82 (R = .941)

$\propto$ dB(A) - 7.08 $\log_{10}$ (T_{60}) - 48.46 (R = .952)

$\propto$ dB(A) - 7.14 $\log_{10}$ (T_{70}) - 83.14 (R = .940)

$\propto$ dB(A) - 11.71 $\log_{10}$ (Dur) - 38.95 (R = .934)

$\propto$ dB(A) - 14.40 $\log_{10}$ (T_1) - 44.35 (R = .950)

$\propto$ PNL - 10.25 $\log_{10}$ (D_{10}) - 59.30 (R = .970)

$\propto$ PNLT - 11.85 $\log_{10}$ (D_{10}) - 59.40 (R - .965)

DC10

$\overline{SSV} \propto$ dB(A) - 6.80 log (D_5) - 49.70 (R - .955)

$\propto$ dB(A) - 7.20 log (D_{10}) - 47.65 (R = .954)

$\propto$ dB(A) - 7.95 log (D_{20}) - 43.90 (R = .947)

$\propto$ dB(A) - 6.65 log (ONS) - 55.00 (R - .933)

$\propto$ dB(A) - 6.28 $\log_{10}$ (T_{60}) - 51.76 (R = .953)

$\propto$ dB(A) - 3.12 $\log_{10}$ (T_{70}) - 57.19 (R = .932)

$\propto$ dB(A) - 8.81 $\log_{10}$ (Dur) - 43.57 (R = .939)

$\propto$ dB(A) - 8.75 $\log_{10}$ (T_1) - 49.30 (R = .961)

$\propto$ PNL - 7.90 $\log_{10}$ (D_{10}) - 63.30 (R = .956)

$\propto$ PNLT - 7.70 $\log_{10}$ (D_{10}) - 63.10 (R = .947)

TABLE 9 - Concluded

Concorde		
$\overline{SSV}$	$\propto$	$dB(A) + 1.33 \log (D_5) - 51.72$ (R = .941)
	$\propto$	$dB(A) + 0.61 \log (D_{10}) - 51.22$ (R = .940)
	$\propto$	$dB(A) + 1.11 \log (D_{20}) - 51.83$ (R = .940)
	$\propto$	$dB(A) + 2.11 \log (ONS) - 49.83$ (R = .942)
	$\propto$	$dB(A) - 0.68 \log (T_{60}) - 47.53$ (R = .940)
	$\propto$	$dB(A) - 0.63 \log (T_{70}) - 50.39$ (R = .940)
	$\propto$	$dB(A) + 0.78 \log (Dur) - 57.17$ (R = .939)
	$\propto$	$dB(A) - 0.39 \log (T_1) - 50.39$ (R = .939)
	$\propto$	$PNL - 0.83 \log (D_{10}) - 64.00$ (R = .924)
	$\propto$	$PNLT - 0.35 \log (D_{10}) - 65.76$ (R = .921)
All Sources		
$\overline{SSV}$	$\propto$	$dB(A) - 4.40 \log (D_5) - 47.40$ (R = .932)
	$\propto$	$dB(A) - 5.37 \log (D_{10}) - 45.75$ (R = .937)
	$\propto$	$dB(A) - 5.58 \log (D_{20}) - 46.79$ (R = .934)
	$\propto$	$dB(A) - 3.57 \log (ONS) - 49.10$ (R = .926)
	$\propto$	$dB(A) - 4.50 \log (T_{60}) - 50.77$ (R = .933)
	$\propto$	$dB(A) - 3.04 \log (T_{70}) - 56.71$ (R = .928)
	$\propto$	$dB(A) - 5.45 \log (Dur) - 46.05$ (R = .926)
	$\propto$	$dB(A) - 5.85 \log (T_1) - 47.40$ (R = .935)
	$\propto$	$PNL - 6.47 \log (D_{10}) - 62.95$ (R = .933)
	$\propto$	$PNLT - 6.58 \log (D_{10}) - 61.68$ (R = .923)

D_x = 'x' dB - down duration

T_x = Time for which signal exceeded 'x' dB(A)

T_1 = Integrated duration correction

ONS = Onset duration correction

Dur = Total audible duration

TABLE 10

THE COMPUTATION OF SPECTRAL CHANGE

	Octave Band Centre Frequency								
	63	125	250	500	1k	2k	4k	8k	dB(A)
1 Spectrum 1 1/2 secs. before peak	78	80	81	72	81	75	72	70	83.46
2 Spectrum at peak	82	84	85	83	79	78	75	70	85.53
3 Spectrum 1 1/2 secs. after peak	84	85	88	80	74	73	70	68	83.27
4 Change between times 1 and 2	4	4	4	11	2	3	3	0	
5 Change between times 2 and 3	2	1	3	3	5	5	5	2	
Total Change (4 + 5)	6	5	7	14	7	8	8	2	
Change exceeding dB(A) level change.	1.67	0.67	2.67	9.67	2.67	3.67	3.67	0.00	
	Average Change = 4.24								

** dB(A) level change = (85.53 - 83.46) + (85.53 - 83.27) = 4.33

TABLE 11

THE EFFECT OF SPECTRAL CHANGE

<u>B727</u>		
$\overline{SSV}$	$\propto$ dB(A) - 9.2 log D_{10} - 43.0	R = .966
	$\propto$ PNL - 10.2 log D_{10} - 59.3	R = .970
	$\propto$ dB(A) - 1.3 log D_{10} + 0.7 sc - 57.0	R = .983
	$\propto$ PNL - 4.4 log D_{10} + 0.5 sc - 68.7	R = .979
<u>DC10</u>		
$\overline{SSV}$	$\propto$ dB(A) - 7.2 log D_{10} - 47.6	R = .954
	$\propto$ PNL - 7.9 log D_{10} - 63.3	R = .956
	$\propto$ dB(A) + 12.2 log D_{10} + 2.7 sc - 76.3	R = .983
	$\propto$ PNL + 11.0 log D_{10} + 2.6 sc - 91.3	R = .984
<u>Concorde</u>		
$\overline{SSV}$	$\propto$ dB(A) + 0.6 log D_{10} - 51.2	R = .940
	$\propto$ PNL - .08 log D_{10} - 64.0	R = .924
	$\propto$ dB(A) + 1.4 log D_{10} + 1.0 sc - 56.0	R = .974
	$\propto$ PNL - .10 log D_{10} + 1.0 sc - 67.0	R = .957
<u>All Aircraft</u>		
$\overline{SSV}$	$\propto$ dB(A) - 5.37 log D_{10} - 48.2	R = .937
	$\propto$ PNL - 6.47 log D_{10} - 62.9	R = .933
	$\propto$ dB(A) - 1.47 log D_{10} + 0.6 sc - 55.3	R = .961
	$\propto$ PNL - 3.11 log D_{10} + 0.5 sc - 68.5	R = .950

sc = spectral change between points 2 1/2 seconds either side of peak noise

TABLE 12

TEST STIMULI - STUDY 2

Sound No.	Description	Duration	Doppler Shift	Spectrum
1	480m flyover	2d	β	f_1
2	480m flyover	d	β	f_1
3	480m flyover	d/2	B	f_1
4	480m flyover at 1/2 speed	2d	β/2	$f_{1/2}$
5	480m flyover at 1/2 speed	d	β/2	$f_{1/2}$
6	480m flyover at 1/2 speed	d/2	β/2	$f_{1/2}$
7	240m flyover	2d	2β	f_2
8	240m flyover	d	2β	f_2
9	240m flyover	d/2	2β	f_2
10	240m flyover (filtered)	2d	2β	f_1
11	240m flyover (filtered)	d	2β	f_1
12	240m flyover (filtered)	d/2	2β	f_1
13	240m flyover (filtered) at 1/2 speed	2d	β	$f_{1/2}$
14	240m flyover (filtered) at 1/2 speed	d	β	$f_{1/2}$
15	240m flyover (filtered) at 1/2 speed	d/2	β	$f_{1/2}$
16	120m flyover (filtered	d	4β	f_2
17	120m flyover (filtered)	d/2	4β	f_2
18	120m flyover (filtered	d	4β	f_1
19	120m flyover (filtered)	d/2	4β	f_1
20	120m flyover (filtered) at 1/2 speed	2D	2β	$f_{1/2}$
21	120m flyover (filtered) at 1/2 speed	d	2β	$f_{1/2}$
22	120m flyover (filtered) at 1/2 speed	d/2	2β	$f_{1/2}$

TABLE 13

ARRANGEMENT OF SOUNDS WITHIN TAPES AND SESSIONS

	Tape	Order of Sounds					
Session 1	A	17^1	20^2	18^4	5^3	1^1	19^2
	B	16^2	7^3	22^1	2^4	11^2	8^3
	C	6^3	15^4	9^2	12^1	3^3	13^4
	D	20^4	21^1	4^3	14^2	10^4	8^1
Session 2	B	2^1	20^2	7^4	8^3	6^4	16^3
	C	1^2	14^3	13^1	8^4	20^1	3^4
	D	19^3	4^4	21^2	15^1	12^2	10^1
	A	5^4	18^1	11^3	17^2	9^3	22^2
Session 3	C	7^1	2^2	16^4	20^3	10^2	6^1
	D	18^2	21^3	5^1	22^4	20^3	13^2
	A	12^3	9^4	15^2	3^1	19^4	17^3
	B	11^4	8^1	1^3	8^2	4^1	14^4
Session 4	D	9^1	8^2	17^4	2^3	12^4	15^3
	A	3^2	22^3	19^1	21^4	11^1	1^4
	B	13^3	8^4	4^2	20^1	7^2	16^1
	C	20^4	14^1	18^3	6^2	10^3	5^2

X^y where X is the sound number (1-22) and y is the peak noise level (1-4).

TABLE 14

MEAN (AND STANDARD DEVIATION) OF PEAK dB(A) VALUES

OF SOUNDS PRESENTED TO SUBJECT GROUPS

Sound	L_1	L_2	L_3	L_4
1	86.00(0.20)	79.25(0.43)	72.50(0.50)	64.75(0.43)
2	85.25(0.43)	79.50(0.50)	73.25(0.43)	65.25(0.43)
3	87.00(0.20)	80.25(0.43)	72.75(0.69)	65.25(0.43)
4	87.75(0.69)	80.25(0.43)	73.25(1.09)	65.25(0.43)
5	87.75(0.69)	79.50(0.50)	73.00(0.20)	65.25(0.43)
6	86.75(0.43)	80.00(0.71)	72.75(0.69)	65.50(0.50)
7	86.00(0.20)	79.25(0.43)	72.00(0.20)	64.25(0.43)
8	86.00(0.50)	79.50(0.61)	71.75(0.97)	64.87(0.33)
9	81.00(0.20)	79.25(1.09)	71.25(0.43)	65.75(0.43)
10	85.25(0.43)	80.00(0.20)	72.25(0.43)	66.50(0.87)
11	85.25(0.43)	78.50(0.50)	70.75(0.43)	65.00(0.20)
12	85.25(0.43)	80.00(0.20)	72.00(0.20)	65.25(0.43)
13	85.25(0.43)	80.25(0.43)	72.25(0.43)	65.25(0.43)
14	87.50(0.50)	80.25(1.09)	71.75(0.43)	66.50(0.87)
15	85.25(0.43)	79.25(0.43)	73.25(0.43)	65.50(1.14)
16	86.00(0.20)	78.75(0.43)	70.25(0.43)	65.00(0.71)
17	86.00(0.20)	78.00(0.20)	71.50(0.50)	65.25(0.43)
18	84.00(0.20)	78.75(0.69)	70.75(0.43)	64.25(0.43)
19	84.75(0.43)	78.00(0.20)	70.00(0.20)	64.75(1.30)
20	86.25(0.34)	79.25(0.34)	73.12(0.61)	66.50(0.50)
21	87.00(1.22)	80.25(0.43)	73.00(0.69)	65.25(0.43)
22	86.75(0.43)	79.25(0.43)	72.25(0.43)	66.50(0.50)

TABLE 15

DURATIONS OF STIMULI (SECONDS)

Sound No.	5 dB(A) down	10 dB(A) down	20 dB(A) down	Effective Duration, (D_1)	Estimated Duration, (D_2)
1	12.0	19.5	37	7.5	16.0
2	3.5	9.5	20	2.7	7.5
3	3.0	6.5	15	1.8	5.0
4	13.0	19.0	42	7.6	19.0
5	6.0	10.0	23	3.5	8.0
6	4.0	6.5	17	2.6	6.5
7	11.0	20.0	37	6.6	16.0
8	7.0	11.0	20	4.3	9.0
9	3.0	6.0	15	2.8	6.0
10	11.0	20.0	30	5.1	15.5
11	6.0	11.0	19	4.0	8.5
12	4.0	6.0	15	3.1	6.5
13	14.0	20.0	30	7.8	14.5
14	4.5	10.0	19	2.8	8.0
15	2.5	5.0	12	1.8	4.0
16	8.0	11.0	20	5.5	9.5
17	3.5	6.0	15	2.6	5.5
18	7.5	11.0	19	4.9	8.5
19	2.5	6.5	15	1.8	5.0
20	8.0	18.0	31	5.1	13.0
21	5.5	10.0	25	4.3	9.0
22	3.5	6.0	13	2.4	4.5

D_5, D_{10}, D_{20} were read from graphic level recorder.

D_1 is the "effective" duration found from EPNL correction.

D_2 is the digitally-derived 10 PNdBT-down duration.

TABLE 16

ANALYSIS OF VARIANCE: TAPE AND TAPE ORDER EFFECTS

	Source of Variation	Degrees of Freedom	Sums of Squares	Means Squares	F
Session 1	Subject Groups	3	4.779	1.593	22.4 *
	Tapes	3	3.343	1.114	15.7 *
	Tape Order	3	0.265	0.088	1.24
	Residual	6	0.426	0.071	
	Total	15	8.813		
Session 2	Subject Groups	3	5.829	1.943	97.1 *
	Tapes	3	1.443	0.481	24.0 *
	Tape Order	3	0.369	0.123	6.1
	Residual	6	0.118	0.020	
	Total	15	7.759		
Session 3	Subject Groups	3	4.005	1.335	78.5 *
	Tapes	3	4.304	1.435	84.4 *
	Tape Order	3	0.120	0.040	2.35
	Residual	6	0.103	0.017	
	Total	15	8.532		
Session 4	Subject Groups	3	4.106	1.369	29.1 *
	Tapes	3	4.730	1.577	33.6 *
	Tape Order	3	0.214	0.071	1.51
	Residual	6	0.280	0.047	
	Total	15	0.331		

* significant at 1% level

TABLE 17

ANALYSIS OF VARIANCE: SESSION AND SESSION ORDER EFFECTS

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	*F*
Subject Groups	3	7178.12	2392.71	81.6 *
Sessions	3	358.85	119.62	4.1
Session Order	3	151.63	50.54	1.72
Residual	6	176.10	29.35	
Total	15	7864.70		

* significant at 1% level

TABLE 18

ANALYSIS OF VARIANCE OF RESPONSES TO SOUNDS HAVING THREE DURATIONS

(ALL SOUNDS EXCEPT NUMBERS 16-19: SEE TABLE 12)

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F
1. Sounds	5	362.43	72.48	66.5 *
2. Duration	2	269.24	134.62	115.8 *
3. Peak Noise Level	3	5800.89	1933.63	442.0 *
4. Subjects	39	2756.12	70.67	68.6 *
1 X 2	10	42.51	4.25	4.54 *
1 X 3	15	140.70	9.38	8.20 *
1 X 4	195	213.03	1.09	1.06
2 X 3	6	16.89	2.81	2.90 *
2 X 4	78	90.76	1.16	1.13
3 X 4	117	513.04	4.38	4.25 *
1 X 2 X 3	30	61.62	3.05	1.99 *
1 X 2 X 4	390	365.32	0.94	0.91
1 X 3 X 4	585	669.12	1.14	1.11
2 X 3 X 4	234	226.55	0.97	0.94
Residual	1170	1208.34	1.03	
Total	2879	12736.56		

* significant at 1% level

TABLE 19

ANALYSIS OF VARIANCE INVESTIGATING THE EFFECTS

OF THE SPECTRAL CONTENT OF THE SOUNDS

(a) Sounds 1, 2, 3, 13, 14, 15 (Doppler Shift β)

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F
1. Spectrum	1	115.51	115.51	101.3 *
2. Duration	2	95.63	47.82	47.8 *
3. Peak Level	3	1625.43	541.81	206.3 *
4. Subjects	39	921.83	23.64	20.6 *
1 X 2	2	17.77	8.89	10.8 *
1 X 3	3	10.01	3.34	2.82
1 X 4	39	44.53	1.14	0.99
2 X 3	6	16.77	2.79	2.58
2 X 4	78	78.28	1.00	0.87
3 X 4	117	307.53	2.63	2.29 *
1 X 2 X 3	6	15.02	2.50	2.18
1 X 2 X 4	78	63.81	0.82	0.72
1 X 3 X 4	117	138.78	1.19	1.03
2 X 3 X 4	234	254.64	1.09	0.95
Residual	234	268.06	1.14	
Total	956	3973.61		

* significant at 1% level

TABLE 19 - Continued

(b) Sounds 7-12, 20-22 (Doppler Shift 2β)

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F
1. Spectrum	2	75.09	37.54	35.6 *
2. Duration	2	100.11	50.05	58.3 *
3. Peak Level	3	2699.56	899.85	342.0 *
4. Subjects	39	1366.58	35.04	37.8 *
1 X 2	4	12.61	3.15	3.32
1 X 3	6	31.32	5.22	5.66 *
1 X 4	78	82.13	1.05	1.13
2 X 3	6	15.02	2.50	2.61
2 X 4	78	66.95	0.86	0.92
3 X 4	117	308.08	2.63	2.82 *
1 X 2 X 3	12	29.12	2.43	2.62 *
1 X 2 X 4	156	148.00	0.95	1.02
1 X 3 X 4	234	216.12	0.92	0.99
2 X 3 X 4	234	224.59	0.96	1.03
Residual	668	434.90	0.93	
Total	1439	5810.19		

* significant at 1% level

TABLE 19 - Continued

(c) Sounds 16-19 (Doppler Shift 4β)

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F
1. Spectrum	1	73.58	73.58	73.6 *
2. Duration	1	25.20	25.20	54.0 *
3. Peak Level	3	1222.12	407.34	219.0 *
4. Subjects	39	573.84	14.71	21.2 *
1 X 2	1	9.75	9.75	8.6 *
1 X 3	3	18.09	6.03	6.7 *
1 X 4	39	39.36	1.00	1.4
2 X 3	3	8.79	2.93	3.0
2 X 4	39	18.23	0.47	0.7
3 X 4	117	217.94	1.86	2.7 *
1 X 2 X 3	3	10.24	3.41	4.9 *
1 X 2 X 4	39	44.43	1.13	1.6
1 X 3 X 4	117	100.72	0.86	1.2
2 X 3 X 4	117	113.52	0.97	1.4
Residual	117	81.32	0.70	
Total	639	2557.15		

* significant at 1% level

TABLE 20

ANALYSIS OF VARIANCE INVESTIGATING THE EFFECTS

OF DOPPLER SHIFT

(a) Sounds 4, 5, 6, 13, 14, 15, 20, 21, 22 (Spectrum $f_{1/2}$)

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F
1. Doppler Shift	2	310.38	155.19	138.5 *
2. Duration	2	184.20	92.10	93.4 *
3. Peak Level	3	3398.05	1132.68	420.0 *
4. Subjects	39	1397.33	35.83	36.7 *
1 X 2	4	11.76	2.94	2.7
1 X 3	6	78.67	13.11	9.9 *
1 X 4	78	87.56	1.12	1.2
2 X 3	6	13.52	2.25	2.3
2 X 4	78	77.08	0.99	1.0
3 X 4	117	316.39	2.70	2.8 *
1 X 2 X 3	12	24.84	2.07	2.1
1 X 2 X 4	156	172.46	1.10	1.1
1 X 3 X 4	234	310.72	1.33	1.4 *
2 X 3 X 4	234	227.87	0.97	1.0
Residual	486	456.91		
Total	1439	7076.75		

* significant at 1% level

TABLE 20 - Continued

(b) Sounds 2, 3, 11, 12, 18, 19 (Spectrum f_1)

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F
1. Doppler Shift	2	41.76	20.88	25.6 *
2. Duration	1	11.05	11.05	9.9 *
3. Level	3	1486.86	495.62	263.6 *
4. Subjects	39	868.37	22.27	20.7 *
1 X 2	2	23.91	11.95	1.7
1 X 3	6	20.29	3.38	3.4 *
1 X 4	78	63.49	0.81	0.7
2 X 3	3	5.43	1.81	1.8
2 X 4	39	43.66	1.12	1.0
3 X 4	117	220.60	1.88	1.8 *
1 X 2 X 3	6	10.77	1.80	1.7
1 X 2 X 4	78	52.51	0.67	0.6
1 X 3 X 4	234	232.12	0.99	0.9
2 X 3 X 4	117	118.03	1.01	0.9
Residual	234	251.13	1.07	
Total	956	3449.99		

* significant at 1% level

TABLE 20 - Continued

(c) Sounds 8, 9, 16, 17 (Spectrum f_2)

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F
1. Doppler Shift	1	8.56	8.56	13.4 *
2. Duration	1	12.10	12.10	11.4 *
3. Peak Level	3	1292.00	430.67	222.3 *
4. Subjects	39	604.37	15.50	21.1 *
1 X 2	1	2.50	2.50	3.9
1 X 3	3	13.60	4.53	4.4 *
1 X 4	39	24.82	6.64	0.9 *
2 X 3	3	21.16	7.05	6.5 *
2 X 4	39	41.27	1.06	1.4
3 X 4	117	226.62	1.94	2.6 *
1 X 2 X 3	3	1.51	0.50	0.7
1 X 2 X 4	39	24.87	0.64	0.9
1 X 3 X 4	117	120.02	1.02	1.4
2 X 3 X 4	117	126.46	1.08	1.5
Residual	117	86.11	0.74	
Total	639	2605.99		

* significant at 1% level

TABLE 21

CORRELATION COEFFICIENTS OF MEAN SUBJECTIVE SCALE VALUES AND THE RATING SCALES

Correlation Coefficient

	dB(A)	dB(B)	dB(C)	dB(D)	PNL	MKVII	PNLT	EPNL
$\overline{SSV}$	.933	.941	.953	.938	.942	.945	.949	.965

Standard Errors of Estimate

Rating Scale	Standard Error
dB(A)	3.008
dB(B)	2.854
dB(C)	2.572
dB(D_2)	2.925
PNL	2.930
MKVII	2.672
PNLT	2.759
EPNL	2.354

TABLE 22

REGRESSIONS OF SSV's AND VARIOUS RATING SCALES AND DURATION CORRECTIONS

SSV $= .18(dB(A) + 10.99 \log D_1) - 12.0$	$R = .963$
$= .18(dB(A) + 10.18 \log D_2) - 11.3$	$R = .961$
$= .18(dB(A) + 8.83 \log D_5) - 11.4$	$R = .962$
$= .18(dB(A) + 9.50 \log D_{10}) - 11.9$	$R = .957$
$= .18(dB(A) + 13.5 \log D_{20}) - 13.4$	$R = .962$
$= .18(dB(D_2) + 8.86 \log D_1) - 12.6$	$R = .959$
$= .18(dB(D_2) + 8.03 \log D_2) - 12.0$	$R = .957$
$= .18(dB(D_2) + 6.91 \log D_5) - 12.0$	$R = .957$
$= .18(dB(D_2) + 7.64 \log D_{10}) - 12.6$	$R = .955$
$= .18(dB(D_2) + 11.03 \log D_{20}) - 13.76$	$R = .959$
$= .17(PNL + 9.08 \log D_1) - 13.1$	$R = .963$
$= .17\ (PNL + 8.30 \log D_2) - 12.5$	$R = .961$
$= .17\ (PNL + 7.14 \log D_5) - 12.6$	$R = .961$
$= .17\ (PNL + 7.90 \log D_{10}) - 13.1$	$R = .959$
$= .17\ (PNL + 11.36 \log D_{20}) - 14.2$	$R = .963$

D_2 is an estimated 10-dB down duration.

D_1 is the "effective" duration found by integration

D_5, D_{10}, D_{20} are the 5, 10 and 20 dB(A) - down durations measured from a graphic level recorder.

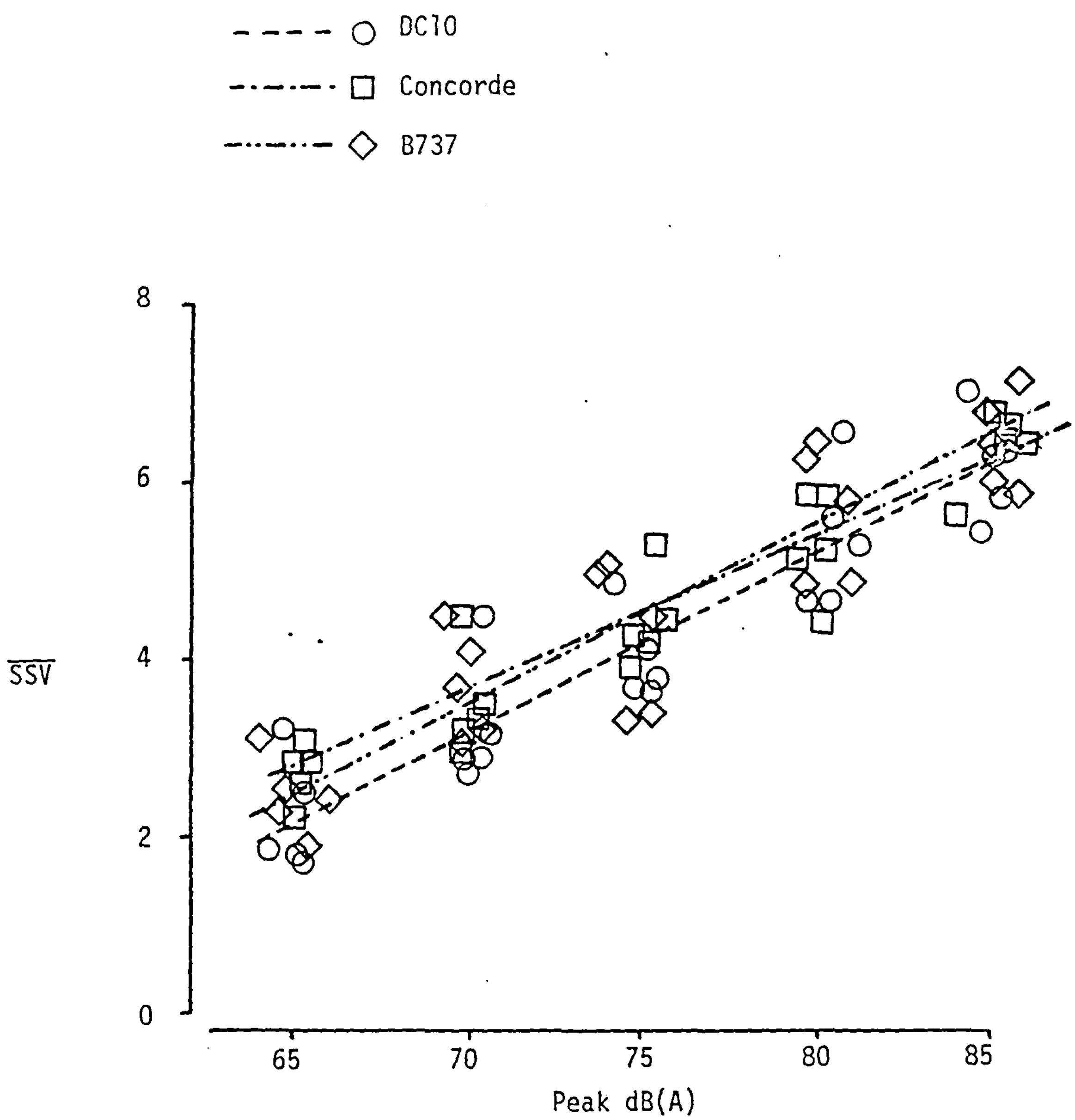

Figure 1.- Least-squares regression lines of $\overline{SSV}$ on peak dB(A) for each aircraft type.

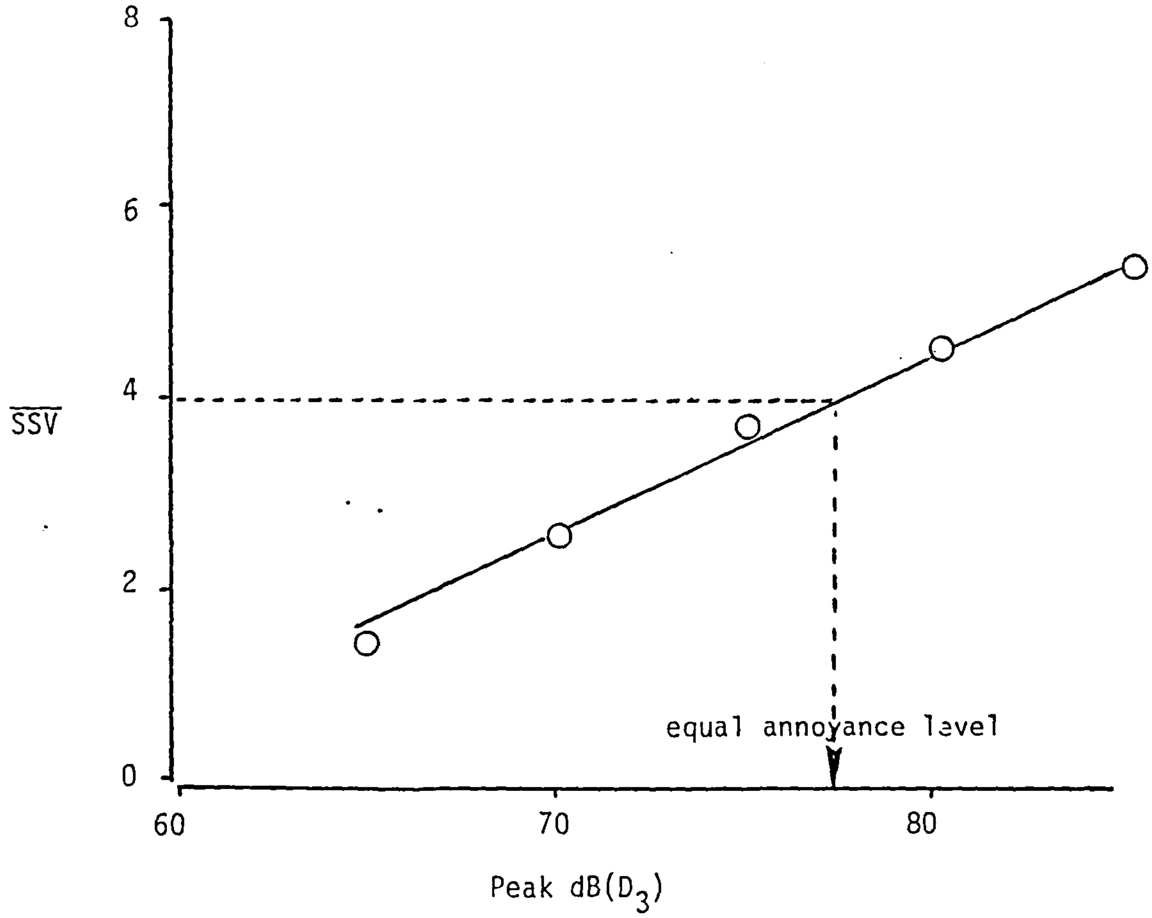

Figure 2.- Derivation of equal annoyance level.

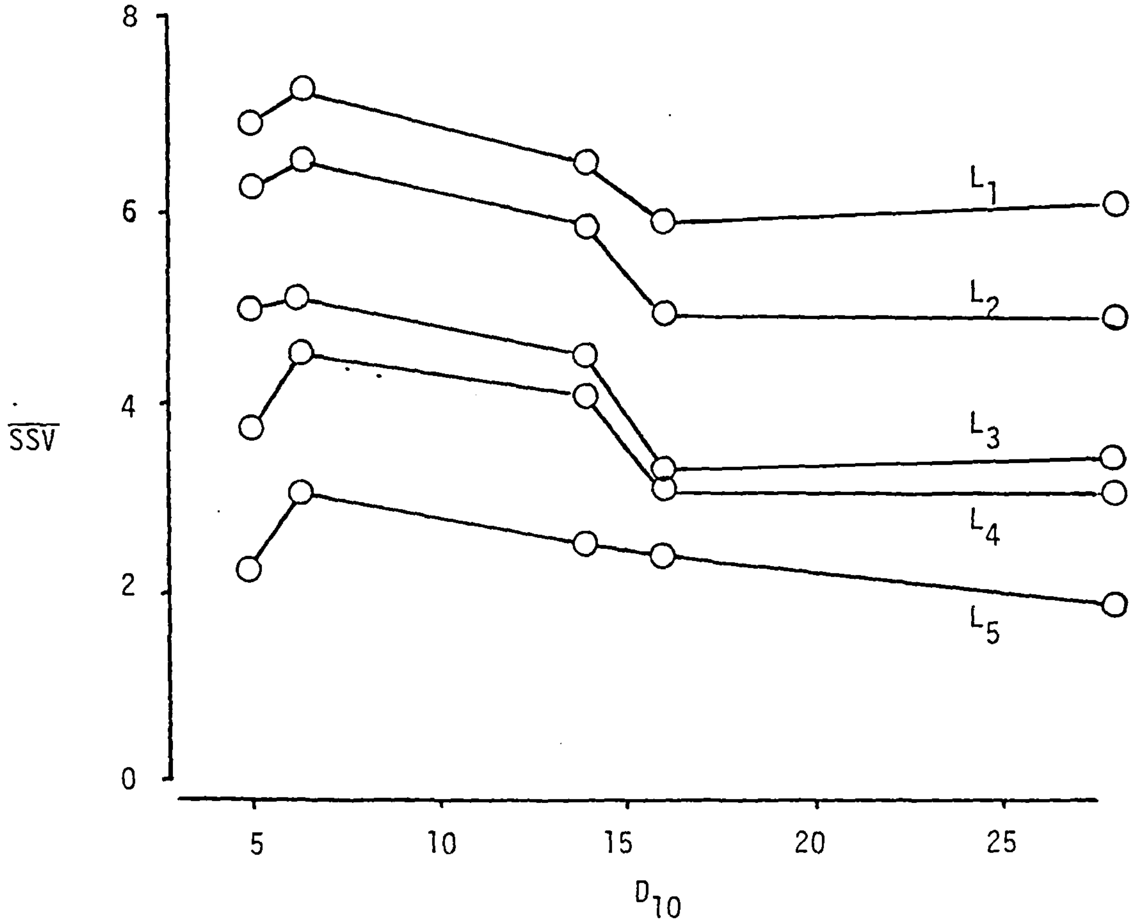

Figure 3.- Mean subjective scale values plotted against duration for each peak level (L_1-L_5) - B737.

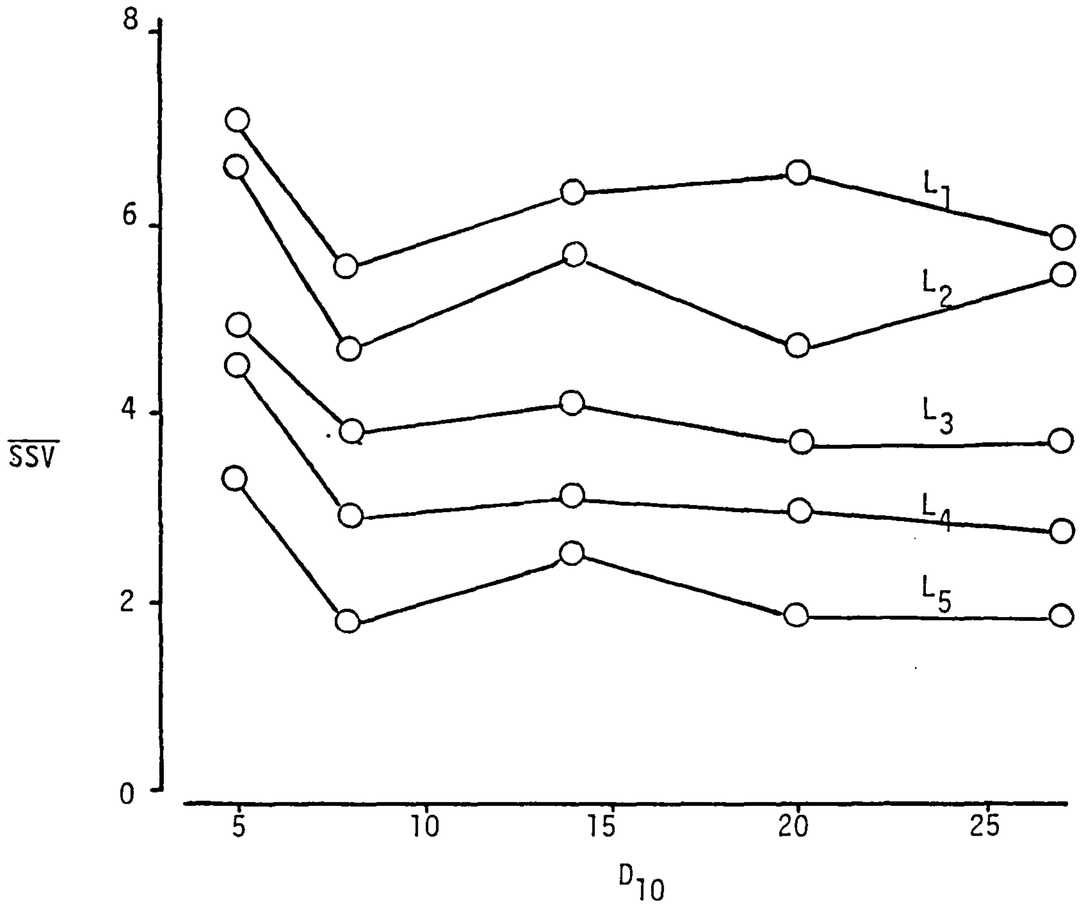

Figure 4. - Mean subjective scale values plotted against duration for each peak level (L_1-L_5) - DC10

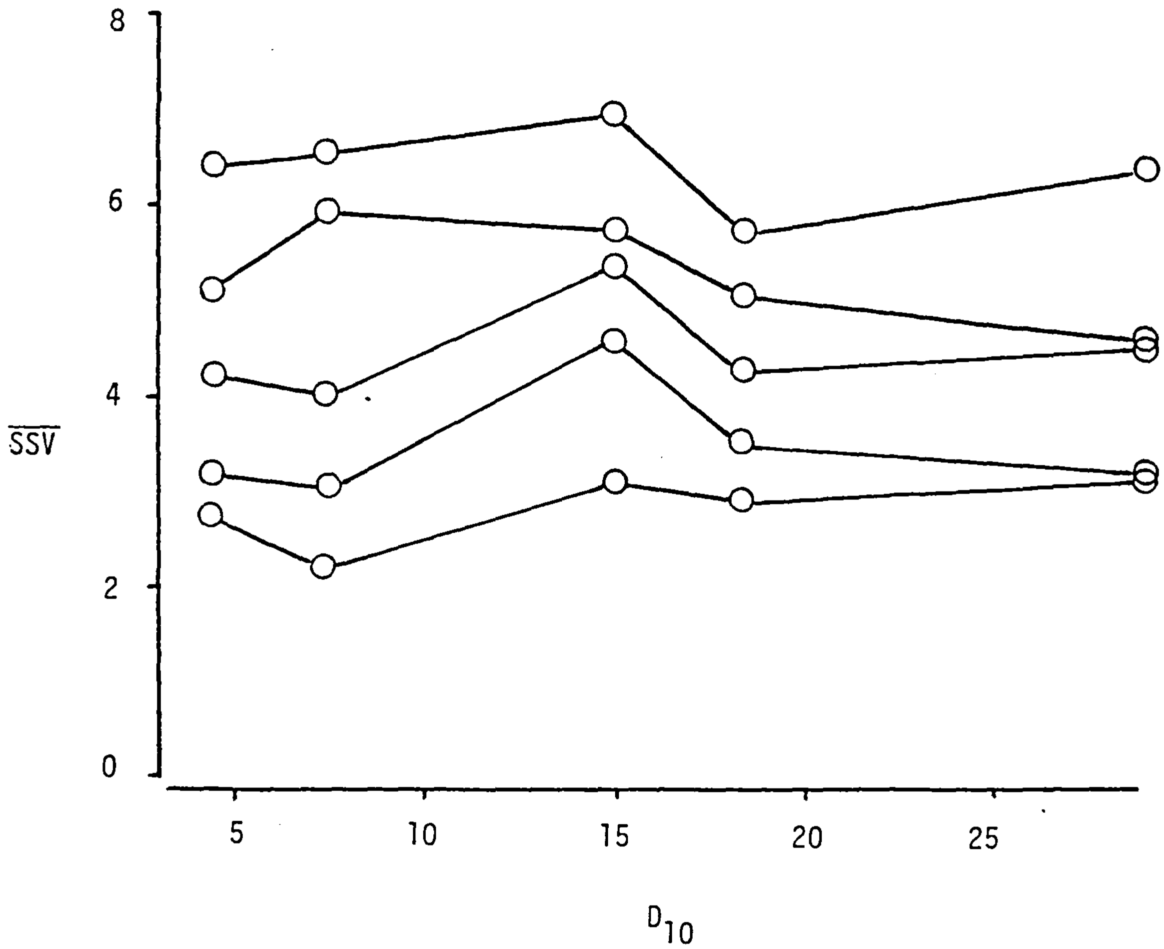

Figure 5.- Mean subjective scale values plotted against duration for each peak level (L_1-L_5) - Concorde

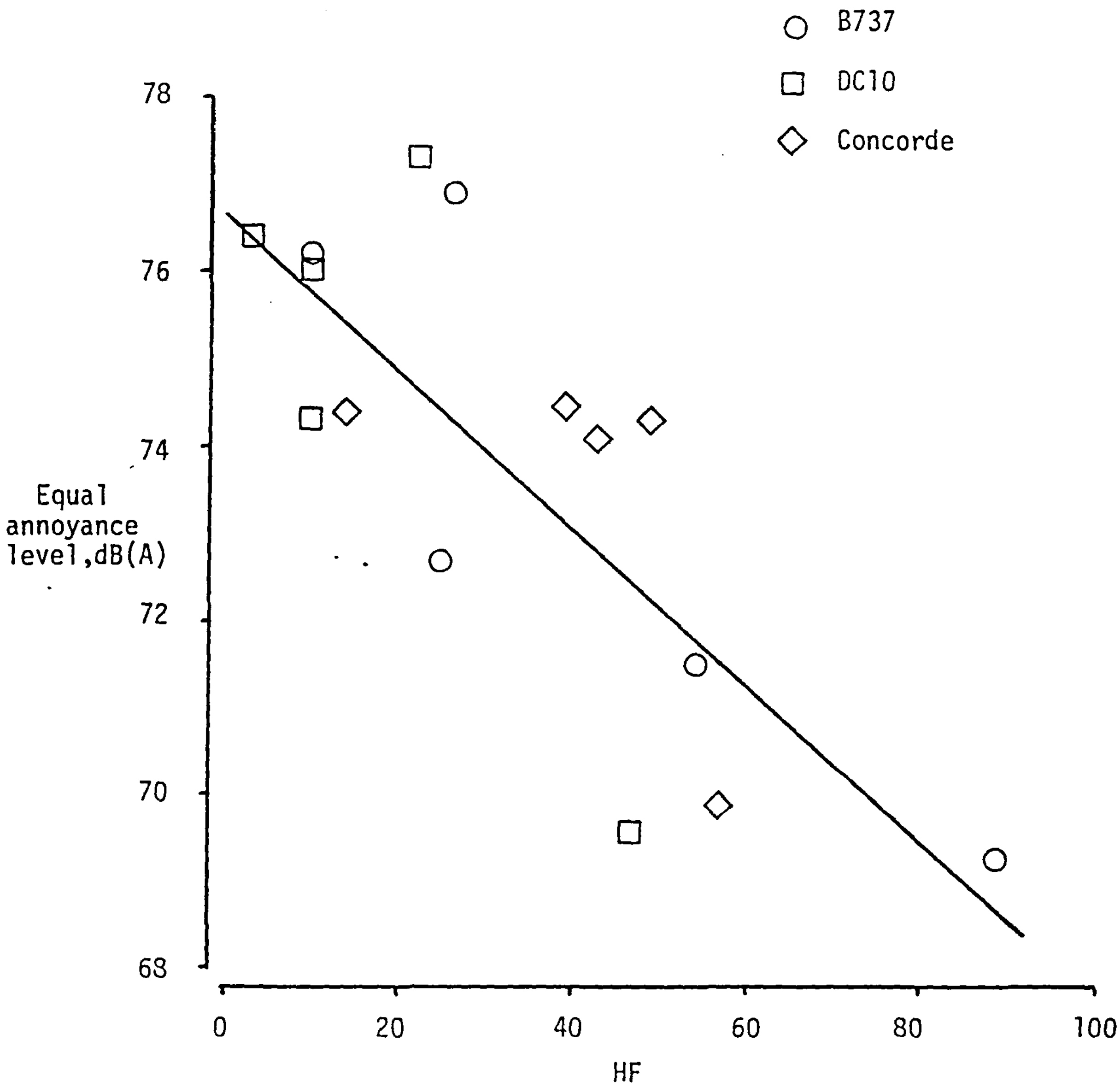

Figure 6.- The regression of equal annoyance level on the percentage of A-weighted high frequency energy (HF).

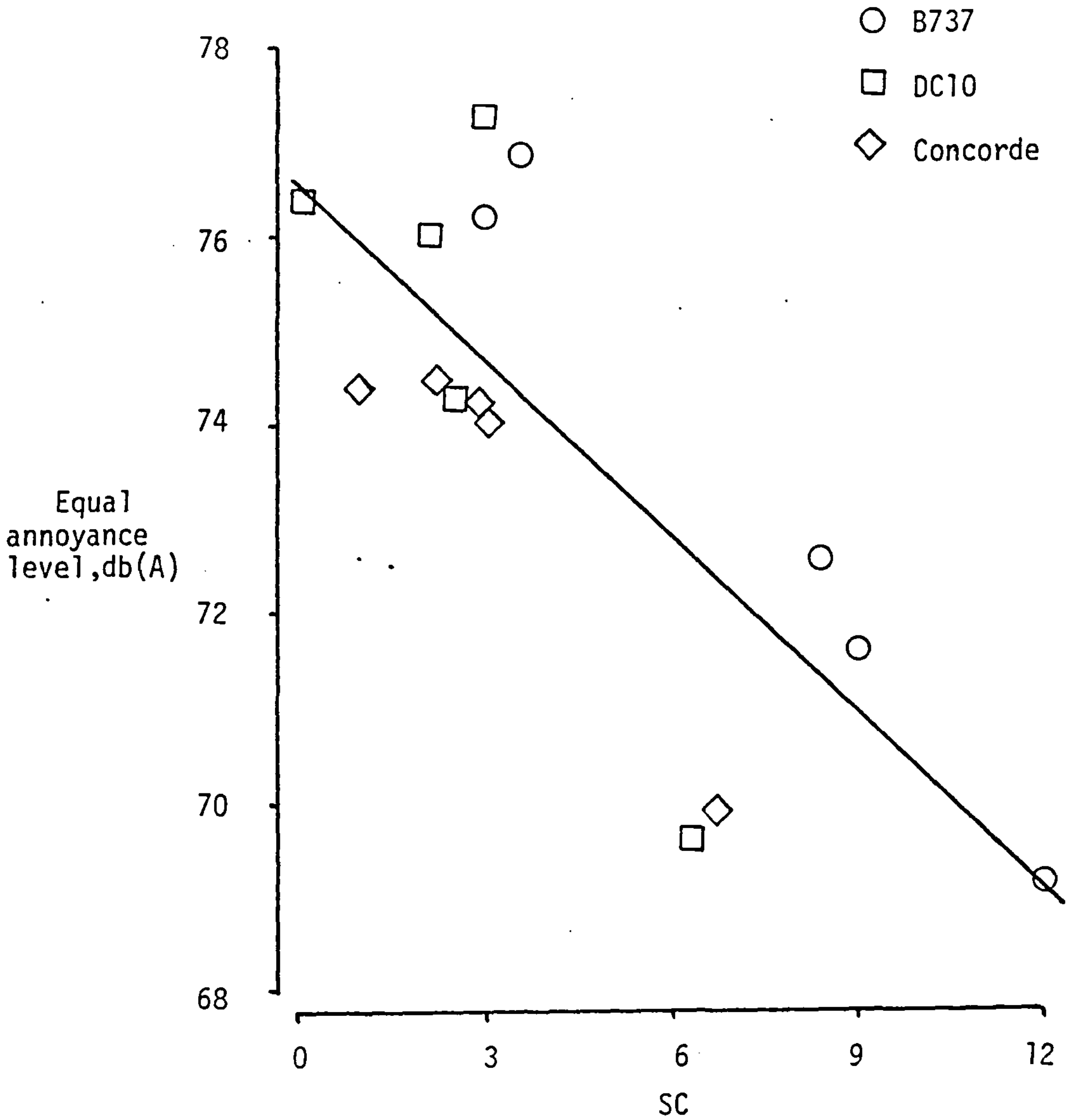

Figure 7.- The regression of equal annoyance level on spectral change.

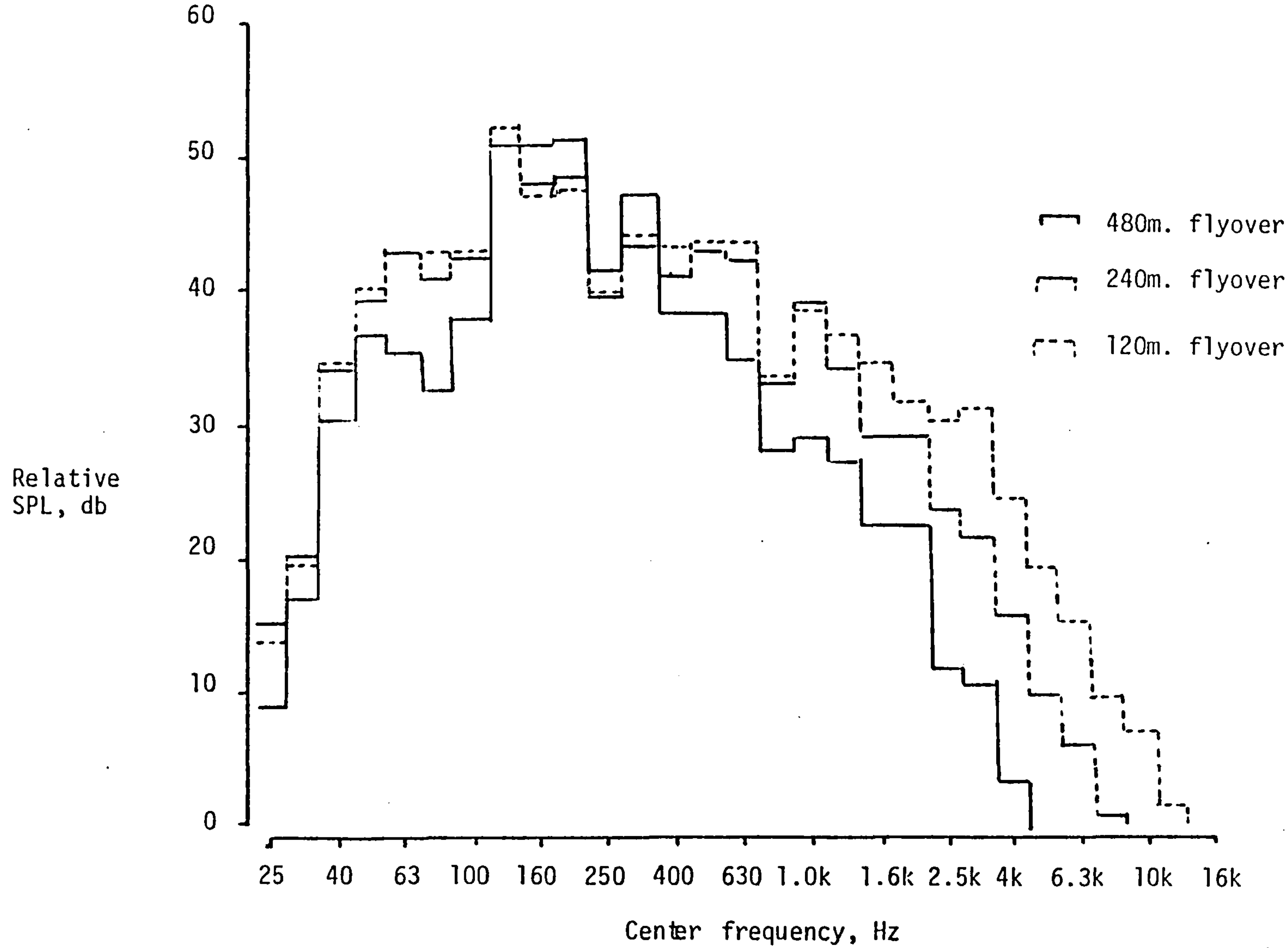

Figure 8.- One-third-octave spectra of the B727 flyovers.

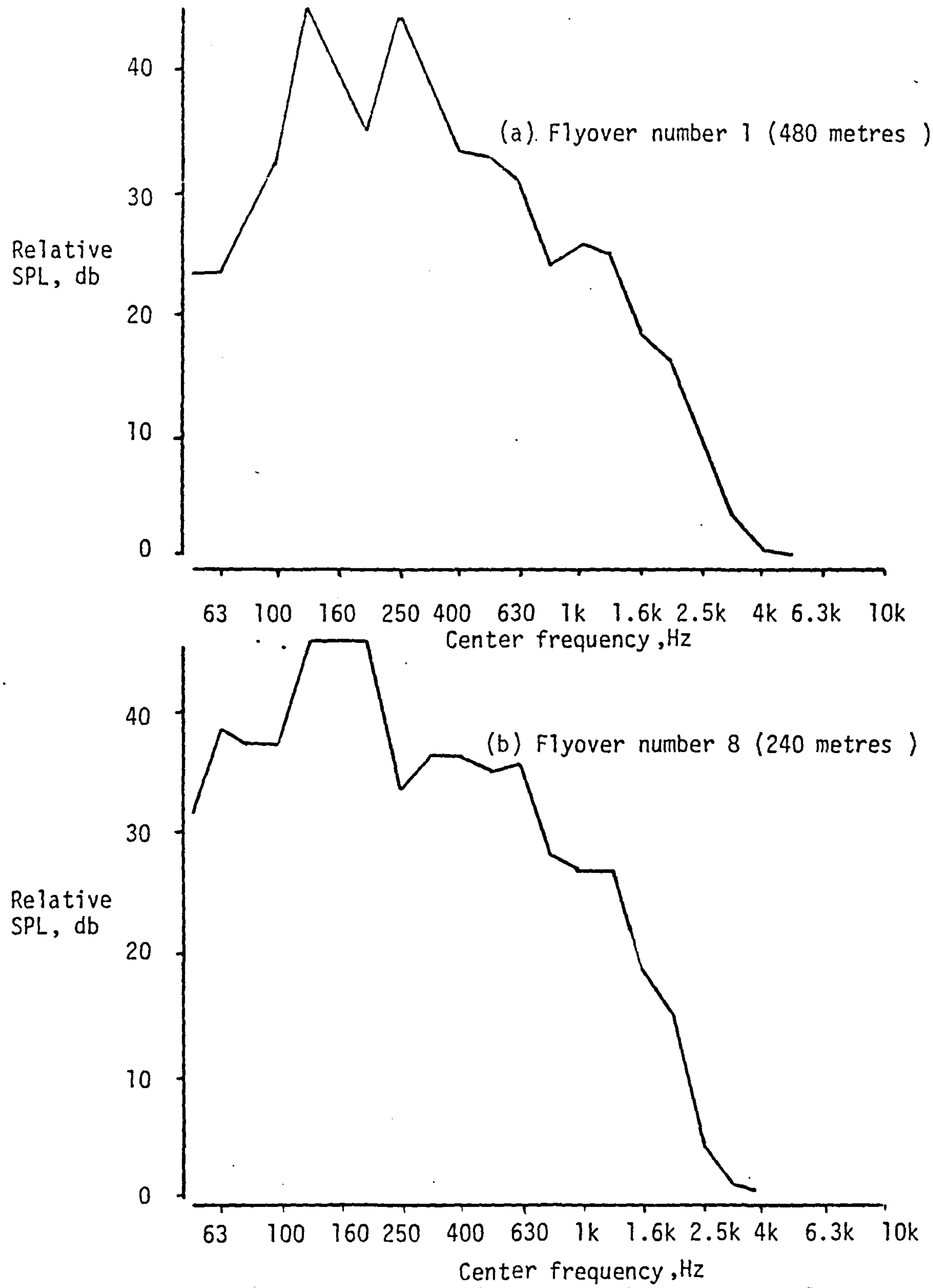

Figure 9.- Typical one-third-octave peak spectra of flyovers.

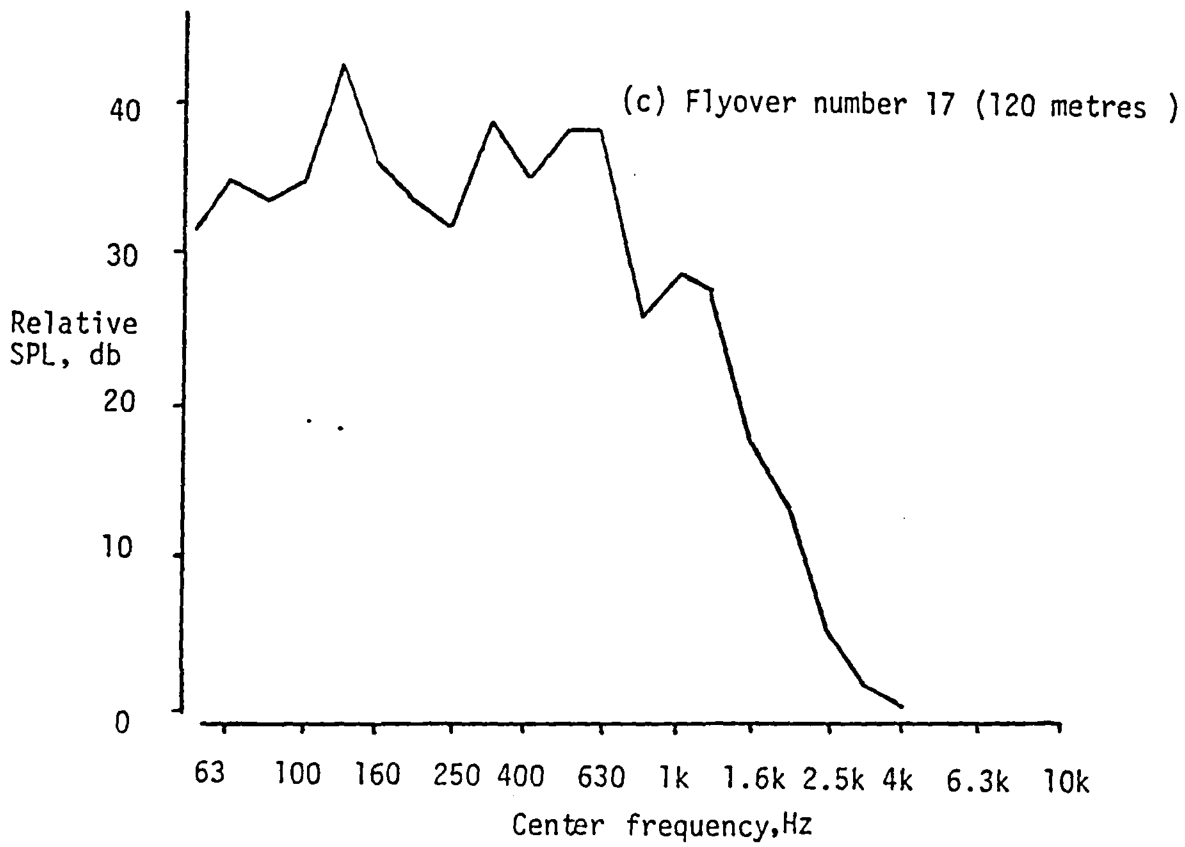

Figure 9. concluded.- Typical one-third-octave peak spectra of flyovers.

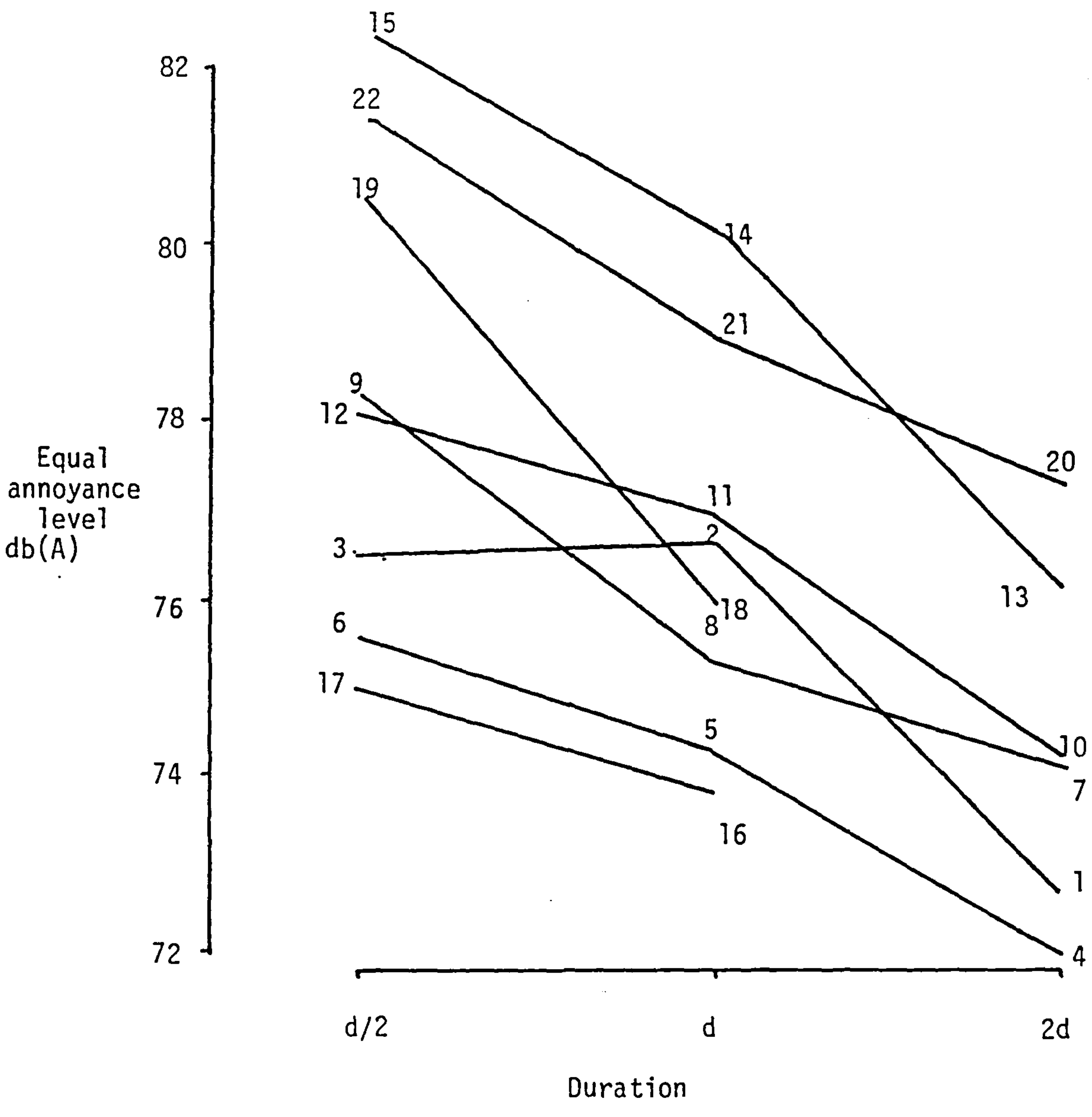

Figure 10.- The relationship between equal annoyance level and duration. (The numbers assigned to the sounds refer to the experimental design).

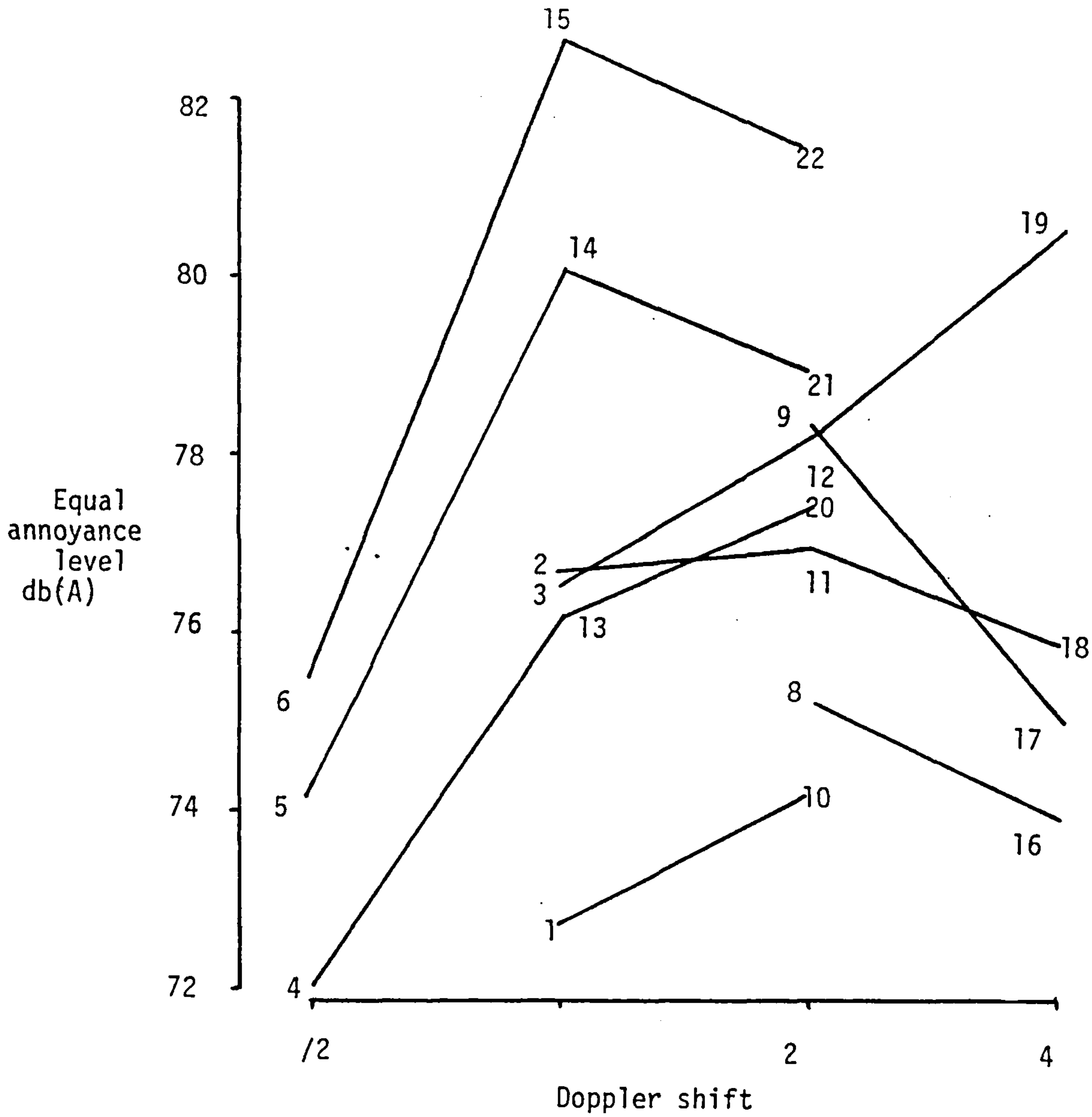

Figure 11.- The relationship between equal annoyance levels and Doppler shift (the numbers assigned to the sounds refer to the experimental design).

REFERENCES

1. Kryter, Karl D, and Pearsons, Karl S. "Some Effects of Spectral Content and Duration on Perceived Noise Level", Journal of the Acoustical Society of America, 35, (6), 866-883 (1963)

2. Pearsons, Karl S "The Effects of Duration and Background Level on Perceived Noisiness", FAA- ADS - 78 (1966)

3. Powell, Clemans A., "A Subjective Evaluation of Synthesized STOL Airplane Noises", NASA TN D-7102 (1973)

4. Little, J. W. and Mabry, J. E. "Sound Duration and Its Effect on Judged Annoyance", Journal of Sound and Vibration 9, (2), 247-262 (1969)

5. Hinterkeuser, Ernest G. and Sternfeld, Harry, Jr., "Subjective Response to Synthesisized Flight Noise Signatures of Several Types of V/STOL Aircraft", NASA CR -1118 (1968)

6. Sternfeld, Harry, Jr., HinterKeuser, Ernest G, Hackman, Roy B., and Davis, Jerry, "Acceptability of VTOL Aircraft Noise Determined by Absolute Subjective Testing", NASA CR-2043 (1972)

7. McCurdy, D.A. and Powell, C.A., "Effects of Duration and Other Noise Characteristics on the Annoyance Caused by Aircraft Flyover Noise", NASA TP-1386 (1979)

8. Little, J. W. and Mabry, J. E. "Empirical Comparisons of Calculation Procedures for Estimating Annoyance of Jet Aircraft Flyovers", Journal of Sound and Vibration 10, (1), 71-80 (1969)

9. Ollerhead, J. B., "Subjective Evaluation of General Aviation Noise", FAA-68 -35 (1968)

10. Kryter, Karl D. "Possible Modifications to the Calculation of Perceived Noisiness", NASA CR-1636 (1970)

11. Hecker, M.H.L. and Kryter, K. D. "Comparison Between Subjective Ratings of Aircraft Noise and Various Objective Measures", FAA-68-33 (1968)

12. Kryter, K. D., Johnson, P. J. and Young, J. R. "Judgment Tests of Flyover Noise from Various Aircraft", NASA CR-1635 (1970)

13. Ollerhead, J. B. "An Evaluation of Methods for Scaling Aircraft Noise Perception", NASA CR-1883 (1971)

14. Shepherd, Kevin P. "The Subjective Evaluation of Noise from Light Aircraft", NASA CR-2773 (1976)

15. Little, J. W. and Mabry, J. E. "Sound Duration and its Effect on Judged Annoyance", Journal of Sound and Vibration 9,(2), 247-262 (1969)

/

REFERENCES

16. Parry, H. J. and Parry J. K. "The Interpretation and Meaning of Laboratory Determinations of the Effect of Duration of the Judged Acceptability of Noise", Journal of Sound and Vibration 20, (1), 51-57 (1972).

17. Nixon, C. W., von Gierke, H. E., Rosinger G. "Comparative Annoyances of 'approaching' versus 'receding' sound sources", Journal of the Acoustical Society of America, 45, 330 (1969)

18. Clarke, Frank R. and Kryter, Karl D., "The Method of Paired Comparison and Magnitude Estimation in Judging the Noisiness of Aircraft", NASA CR-2107 (1972)

19. Powell, Clemans, A. "Judgments of Relative Noisiness of a Supersonic Transport and Several Commercial-Service Aircraft" NASA TN D-8434 (1977)

20. Pearsons, Karl S., Bennett, Ricarda, Fidell Sanford, "Study of the Effects of the Doppler Shift on Perceived Annoyance" NASA CR-1779 (1971)

21. Kryter, Karl D. "The Effects of Noise on Man", Academic Press 1970

22. Stevens, S. S. "Perceived Level of Noise by Mark VII and Decibels (E)" Journal of the Acoustical Society of America 51, 575-601 (1972)