N78-21870

REPRODUCIBLE COPY
(FACILITY CASEFILE COPY)

REPRODUCIBLE COPY
(FACILITY CASEFILE COPY)

NASA
FORMAL
REPORT

FFNo 665 Aug 65

Technical Report Documentation Page

1. Report No.
NASA CR 145318

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle
Response to actual and simulated recordings of conventional takeoff and landing jet aircraft

5. Report Date
January, 1978

6. Performing Organization Code

7. Author(s)
J. E. Mabry and B. M. Sullivan

8. Performing Organization Report No.
MAN-1023-B

9. Performing Organization Name and Address

MAN-Acoustics and Noise, Inc.
2105 North 45th Street
Seattle, Washington 98103

10. Work Unit No. (TRAIS)

11. Contract or Grant No.
NAS1-14404

12. Sponsoring Agency Name and Address

National Aeronautics and Space Administration
Langley Research Center
Hampton, Virginia 23665

13. Type of Report and Period Covered
Final Report

14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The aim of the study was to investigate comparability between noise characteristics of synthesized recordings of aircraft in flight and actual recordings.

Sixty persons made magnitude estimation judgments to 15 synthesized and 12 actual recordings, each presented at four different levels. Nominal peak presentation levels ranged from 68 to 86 dB (OASPL) in 6 dB increments with a standard signal based on USASI noise presented at 77 dB. Judgment data were obtained in the NASA Langley Exterior Effects Room.

Although the synthesized recordings were more smoothly time-varying than the actual recordings and the synthesizer could not produce a "comb-filter" effect that was present in the actual recordings, results supported the conclusion that annoyance response is comparable to the synthesized and actual recordings.

A correction for duration markedly improved the validity of engineering calculation procedures designed to measure noise annoyance, while the FAR-36 correction for tone was not effective.

Results led to the conclusion that the magnitude estimation psychophysical method is a highly reliable approach for evaluating engineering calculation procedures designed to measure noise annoyance. For repeated presentations of pairs of actual recordings, differences between judgment results for identical signals ranged from 0.0 to 0.5 dB.

17. Key Words
Noise annoyance, Psychoacoustics, Psychophysics, CTOL aircraft noise, synthesized noise signals.

18. Distribution Statement
Unclassified - unlimited

19. Security Classif. (of this report)
Unclassified

20. Security Classif. (of this page)
Unclassified

21. No. of Pages
77

22. Price

Form DOT F 1700.7 (8-72)

Reproduction of completed page authorized

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

This study was completed in coordination with Noise Effects Branch personnel of the NASA Langley Research Center. The synthesized flyovers were manufactured at NASA Langley by B. M. Sullivan of MAN-Acoustics and Noise and the judgment data were collected in the NASA Exterior Effects Room. Physical acoustical analyses of the noise signals were completed at MAN-Acoustics and Noise's laboratory as were the statistical analyses of the relationships between the acoustic data and response data. C. A. Powell and J. Cawthorn represented the NASA, in respect to coordination of contract technical deliberations and study management. We want to thank them for their approach to making the work both pleasant and interesting. Also, the authors wish to thank R. Shields of MAN-Acoustics and Noise who performed the computer programming so that numerous linear regressions and analyses of variance could be completed.

SUMMARY

The aim of the study was to investigate comparability between noise characteristics of synthesized recordings of aircraft in flight and actual recordings.

Sixty persons made magnitude estimation judgments to 15 synthesized and 12 actual recordings, each presented at four different levels. Nominal peak presentation levels ranged from 68 to 86 dB (OASPL) in 6 dB increments with a standard signal based on USASI noise presented at 77 dB. Judgment data were obtained in the NASA Langley Exterior Effects Room.

Although the synthesized recordings were more smoothly time-varying than the actual recordings and the synthesizer could not produce a "comb-filter" effect that was present in the actual recordings, results supported the conclusion that annoyance response is comparable to the synthesized and actual recordings.

A correction for duration markedly improved the validity of engineering calculation procedures designed to measure noise annoyance, while the FAR-36 correction for tone was not effective.

Results led to the conclusion that the magnitude estimation psychophysical method is a highly reliable approach for evaluating engineering calculation procedures designed to measure noise annoyance. For repeated presentations of actual recordings, differences between judgment results for identical signals ranged from 0.0 to 0.5 dB.

RESPONSE TO ACTUAL AND SIMULATED
RECORDINGS OF CONVENTIONAL TAKEOFF
AND LANDING JET AIRCRAFT

INTRODUCTION AND PURPOSE

As advanced technology commercial aircraft are in design and investigative stages, the question concerning community noise effects of these aircraft is a consideration. Utilizing engine, airframe, operational and aerodynamic parameters, there is an expectation that the significant, in terms of community noise effects, acoustic characteristics of these advanced technology aircraft can be predicted. Utilizing these predictions of the acoustic characteristics of the noise signatures, can aircraft noise simulations comparable to those obtained from actual flight be achieved? As a means of providing simulations of aircraft flyover signals, NASA Langley Noise Effects Branch has developed a synthesizer and computer program approach. The aim of the present research program is to synthesize flyover recordings using the NASA equipment and approach, and to determine their comparability to actual recordings of aircraft in flight. Two main approaches are utilized to investigate comparability between synthesized and actual recordings. The first approach involves a comparison based on acoustic parameters while the second utilizes human response data to the two sets of recordings, i.e., "Do persons perceive and respond to the synthesized recordings in the same manner as to actual recordings?"

SYNTHESIS

Three recordings of actual flyover events were selected for high quality of recording and low background noise level. The recordings were made at SEATAC International Airport, approximately 5.63 Km. (3.04 n. miles) south of the north end of the main runway. The events used were a 707 take-off, chosen for the strong tonal components, a 727 take-off with little tone and intense low frequency components and a 727 landing, a very short event with strong tones.

The synthesis program requires as input, data on the flyover broad band noise, narrow band components, pure tone components, static directivity patterns and flight data (such as speed, height, angle of flight path). Pure tones were found to sound too unreal so narrow bands of noise were used to generate "tonal" sounds. The synthesizer would form the flyover event, using Doppler shift, varying air attenuation and randomised ground reflection effects. These effects could not be manipulated to produce the comb filter effect that is audible in actual recordings; however this effect does not seem to be so audible when the event is heard binaurally.

Initially, theoretical data on the tonal components of compressor noise, the broad-band jet noise, and directivity patterns were used as input parameters but the results did not sound like the actual events. Thus, the actual recordings were analyzed, using a Spectral Dynamics Ubiquitous analyzer, to study the time-varying nature of the events, and the input data to the program was varied to approach more nearly the parameters in the recordings. The resulting syntheses were changed by altering a number of the parameters until the results sounded close to the originals. This was done strictly by listening and no attempt was made to get results which when physically analyzed matched the originals.

To the experienced ear, there were two very clear differences between the real and the synthesized flyovers: the simulations were more smoothly time-varying than the real events, in which atmospheric and other effects caused a very unsteady time history, and the real recordings, being made monophonically with a fixed microphone position, contained a very marked comb-filter effect sounding like a low-frequency "hollow" resonance, which dropped in pitch as the aircraft approached and then rose again as the aircraft flew away. This is pictured in Figure 1 which is a series of successive spectral analyses of an actual 707 take-off recording. The solid vertical line marks a constant frequency throughout the traces. Close to it, a prominent peak can be seen, shifting in frequency due to the Doppler effect. The solid curved line has been drawn in to emphasize the frequency shift in one of the troughs of the comb filter.

Figure 2 is a similar series of traces of a synthesized flyover, and the regular comb filter effect is absent.

The comb filter effect is not apparent naturally, where binaural listening and head movements radically alter the nature of the comb filter effects.

Neither the comb filter nor the unsteady time history effect could be modelled with the program and, as less experienced listeners accepted the syntheses as real events when heard under informal test conditions, it was decided to continue despite these differences.

The best simulation for each original was chosen and two variations were generated by the synthesizer, one with the tonal components raised by 10dB relative to the broad-band noise and one with the flyover velocity

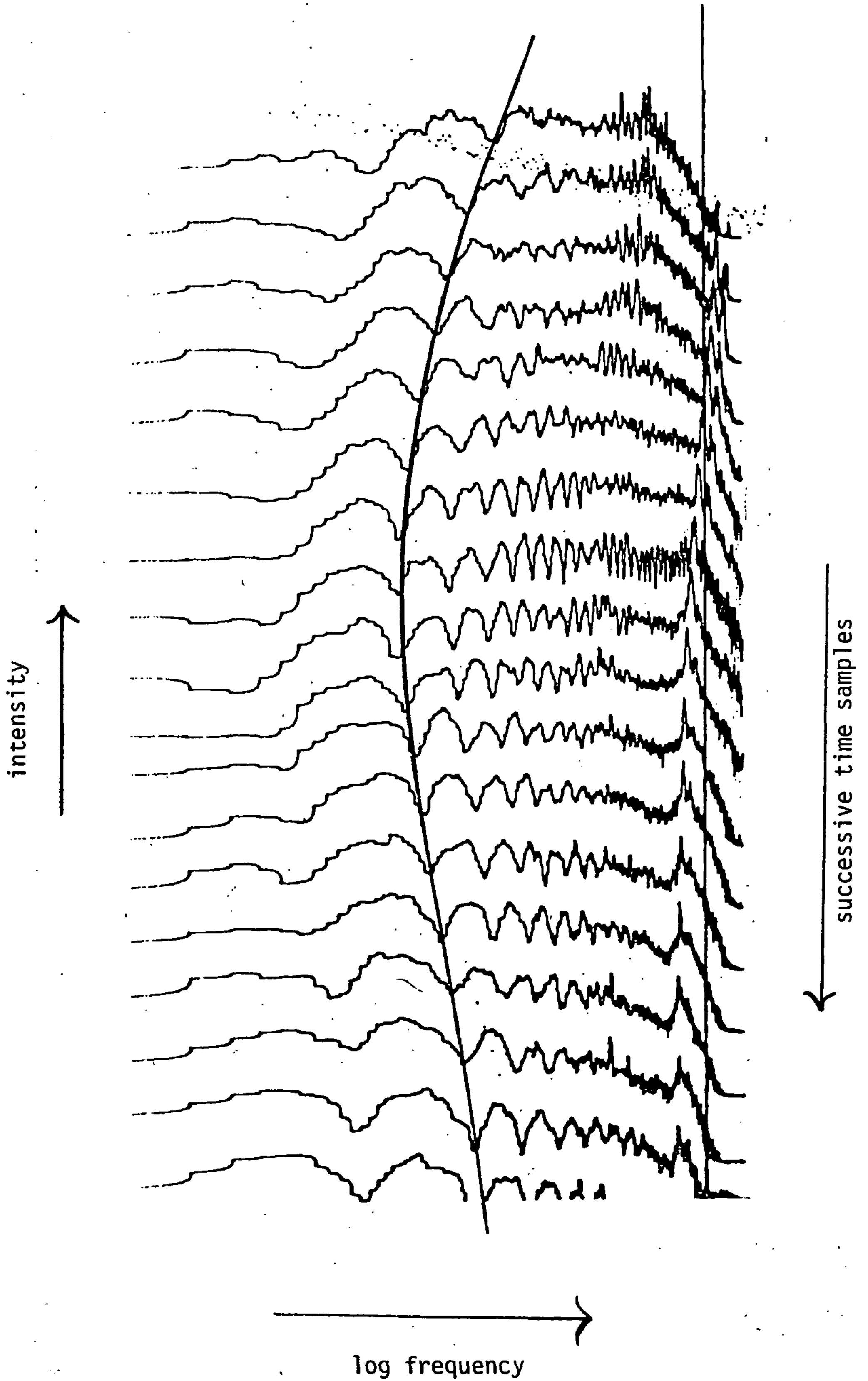

Figure 1: Successive spectral analyses of 707 take-off actual recording, showing ground reflection comb-filter effect

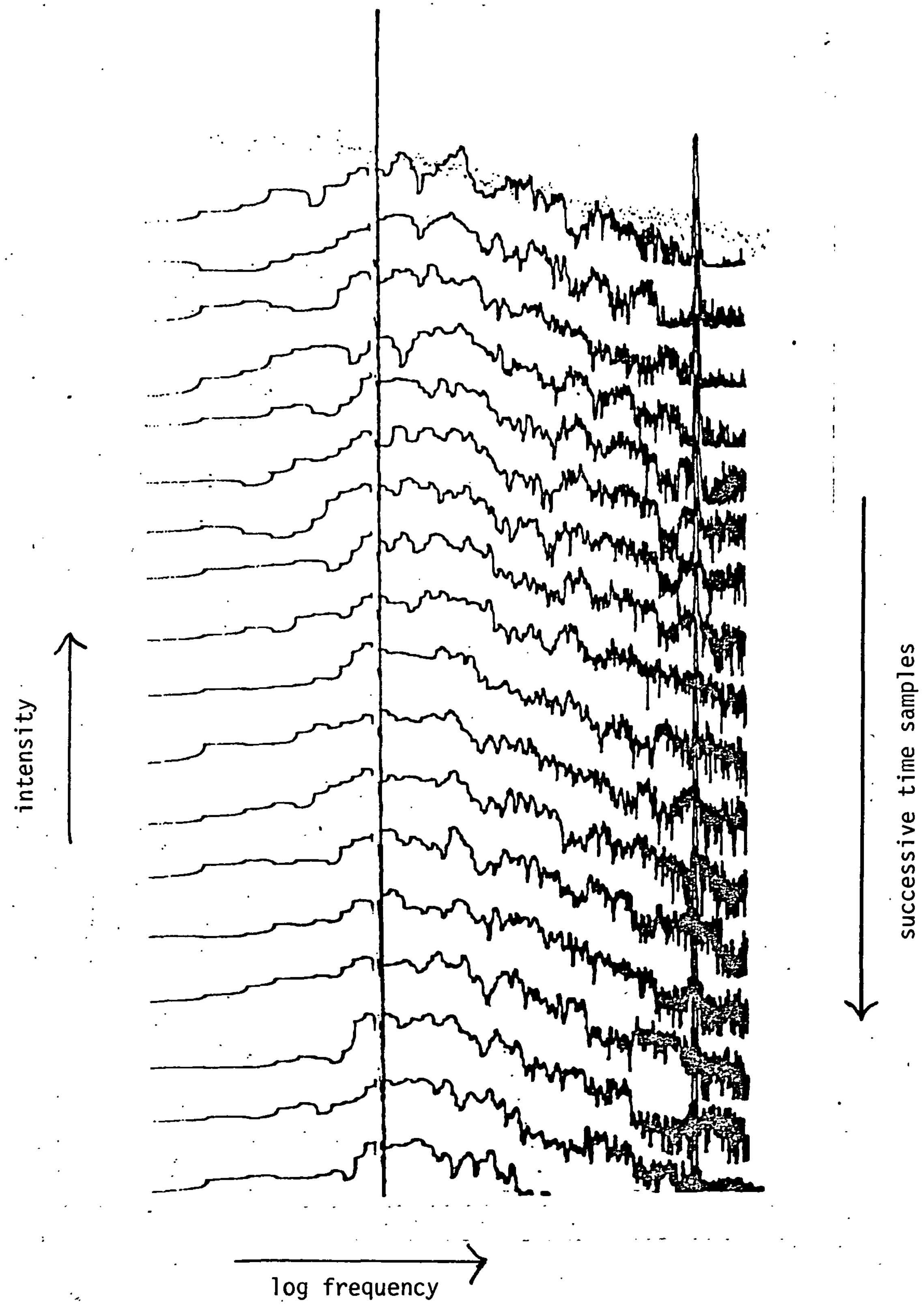

Figure 2: Successive spectral analyses of a synthesized flyover event

decreased by a factor of 2.5. The "best-effort" synthesis was designated v_1t_1; the one with increased tone was v_1t_2 and the long duration, slow velocity was v_2t_1.

Additionally, as it was found that even the best synthesis sounded less complex than the real event, a fourth synthesis for each original was generated, formed by combining two separate simulations. This allowed for more complex tone variation with different tones before and after overhead, and for some low-frequency interference effects. This variation was designated C.

Thusly each of the three flyover events were represented by four simulations which are:

Description	Designation
"Best effort"	v_1t_1
Increased tone	v_1t_2
Long duration	v_2t_1
Combination synthesis	C

PHYSICAL ACOUSTICAL ANALYSES

As indicated above, comparisons between the synthesized and actual recordings were to be based on both physical acoustical analyses and an evaluation by a group of persons. Since an evaluation of the various signals by persons is a function of what they perceive in a particular listening environment, the acoustical analyses are based on the signals as they were presented to persons in the test listening environment.

Figures 3,5,6,8,9,10,11,13 and 14 show the spectra (using 1/2 second integration time) found to give peak dBA at Seat #1 for the top presentation level. The actual recording and the v_1t_1 and C syntheses for all three flyover events are shown. For most of these events, peak levels for dBA and dBAT (dBA tone corrected according to FAR 36) occured at identical times so spectra are identical; in those cases where the spectra for peak dBA and peak dBAT differed, the peak dBAT spectra are given in Figures 4, 7 and 12.

It is clear that for all three cases, the simulations have lower high-frequency (>500 Hz) and higher low-frequency (<80 Hz) components than their real counterparts.

In the case of the 727 landing (Figure 5), the high-frequency lack in the synthesis v_1t_1 is very noticeable compared with the actual event (Figures 3 and 4). It is also apparent that the "tonal" components introduced in v_1t_1 at 2675 Hz and 2725 Hz are too strong relative to the higher frequency "tonal" components (at 3320, 3760, 3950 and 4050 Hz), which produces the sharp peak in the spectrum at 2.5K Hz. The actual-event spectrum shows a much broader peak from 2.5 K Hz to 6.3 K Hz.

However, the combination simulation (C) (Figures 6 and 7) has more high frequency energy and a somewhat broader peak (from 2.5 KHz to 4 KHz). The actual spectrum shows two low frequency dips, one in the 80Hz band and the other in the 200 Hz band, which are probably due to ground reflection effects.

The spectrum for the 727 take-off (Figures 8, 9 and 10) show less difference between the real flyover and v_1t_1 at the high frequency end. Again the C simulation is closer to the actual event, though the peak in the 630 Hz band is too prominent.

For the 707 take-off (Figures 11 to 14) there is again a large difference between v_1t_1 and actual at the high frequency end. The lack of broad-band energy in v_1t_1 makes the tonal components, introduced at 2950 and 3230 Hz, form a very prominent peak in the 3.15 KHz band. As before, the combination synthesis has increased broad-band high frequency energy but the "tone" is also raised significantly. The actual spectrum is displayed for two 1/2 seconds, the one giving peak dBA (Figure 11) and the other peak dBAT (Figure 12). It is clear that the peak band at 3.15 KHz is very variable, but even in the peak dBAT spectrum the presence of broadband noise in the surrounding bands should mask some of this peak, making it less subjectively apparent.

Looking at Table I and II, which calculate the tone and duration corrections (averaged over all seats), the masking effect is clear from the tone correction data for the 707 T/O. The correction for the actual flyover is 3.4 dB, whereas for the v_1t_1 simulation it is 4.2 dB and for the C version it is 5.2 dB. The tone correction for the 727 L is also too high for the simulations, though that for the take-off is somewhat low.

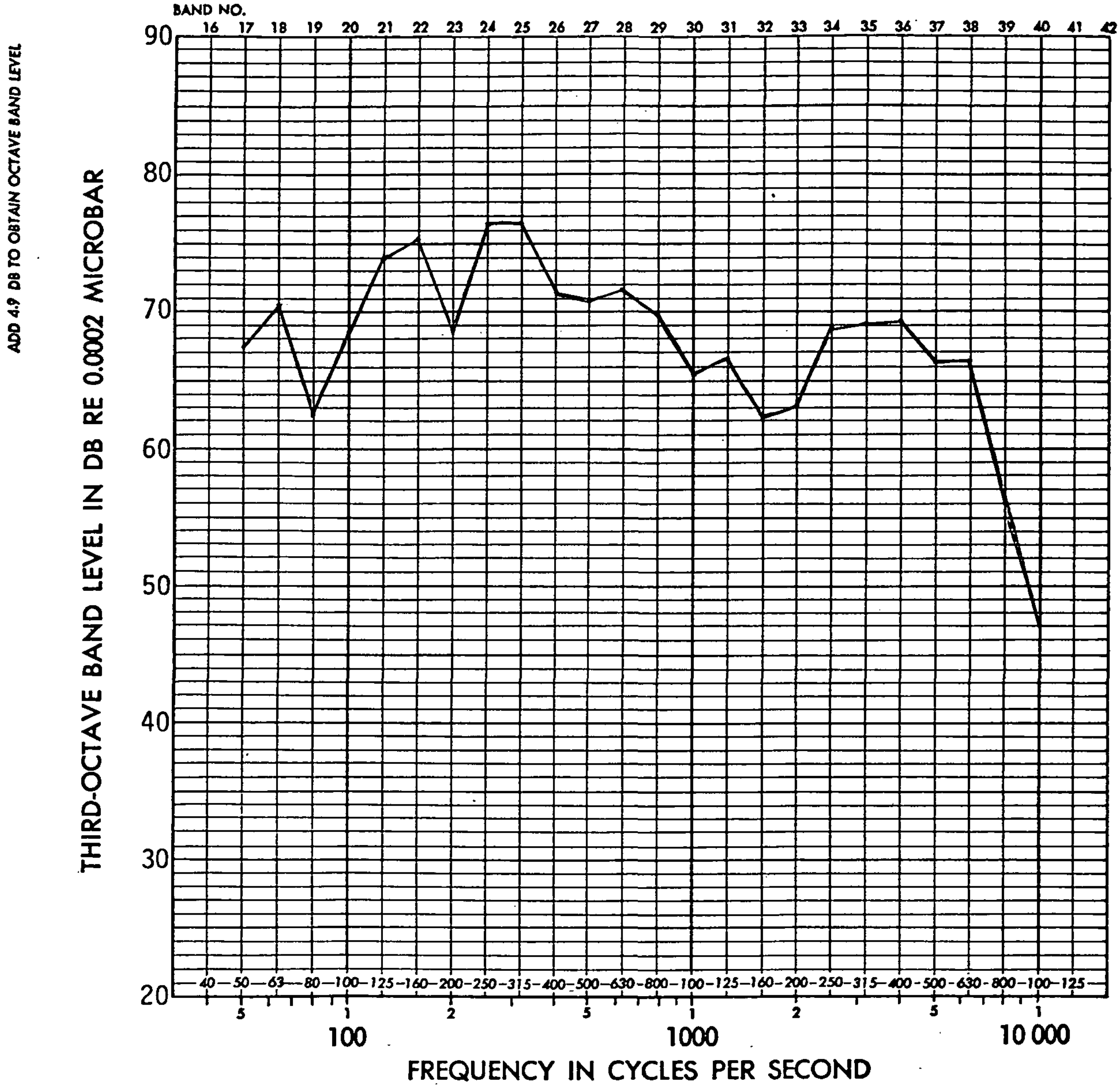

Figure 3: 727 Landing - Actual Recording
1/3 octave x 1/2 second spectrum giving peak dBA

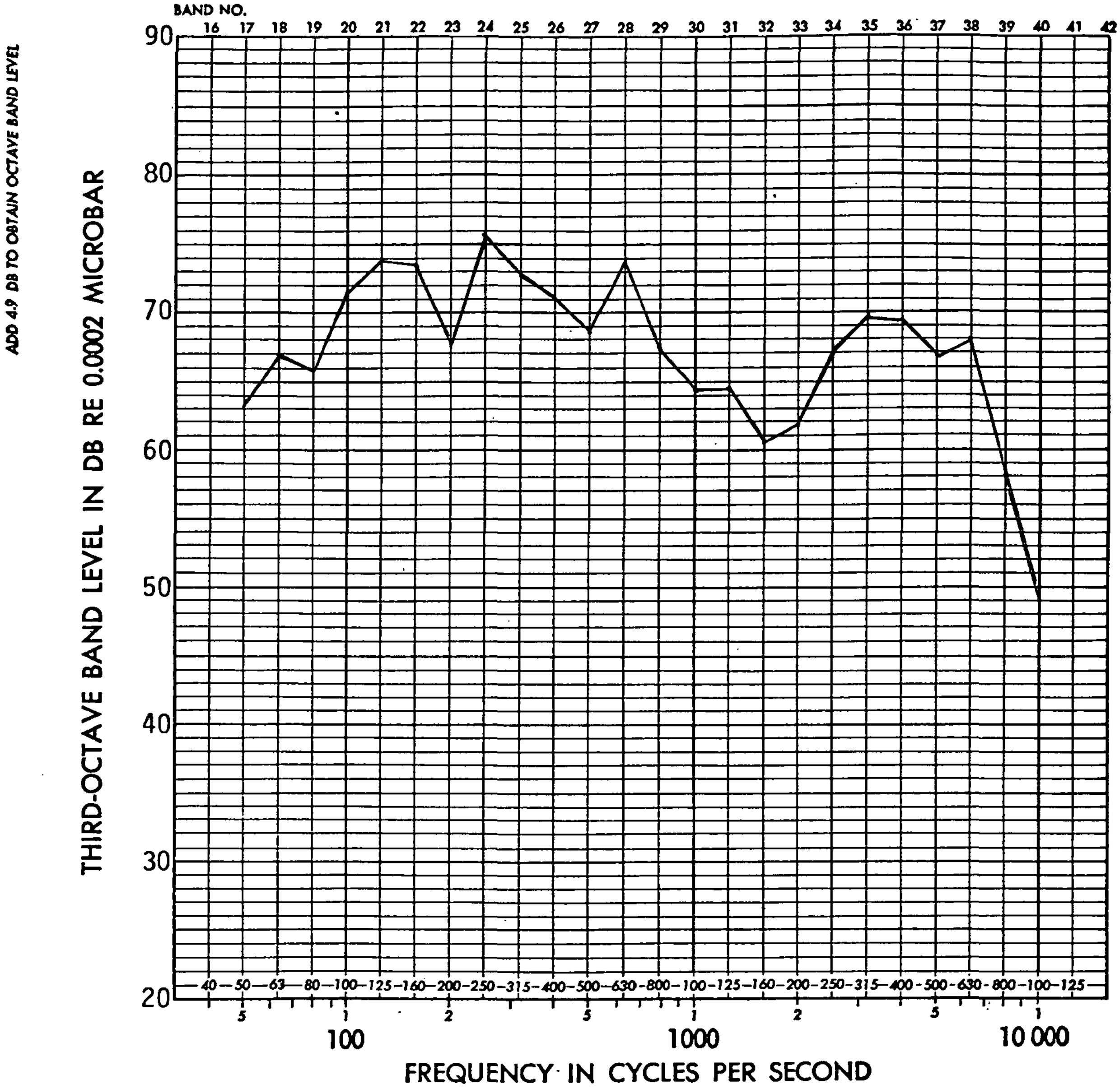

Figure 4: 727 Landing - Actual Recording
1/3 octave x 1/2 second spectrum giving peak dBAT
(dBA tone-corrected according to FAR 36)

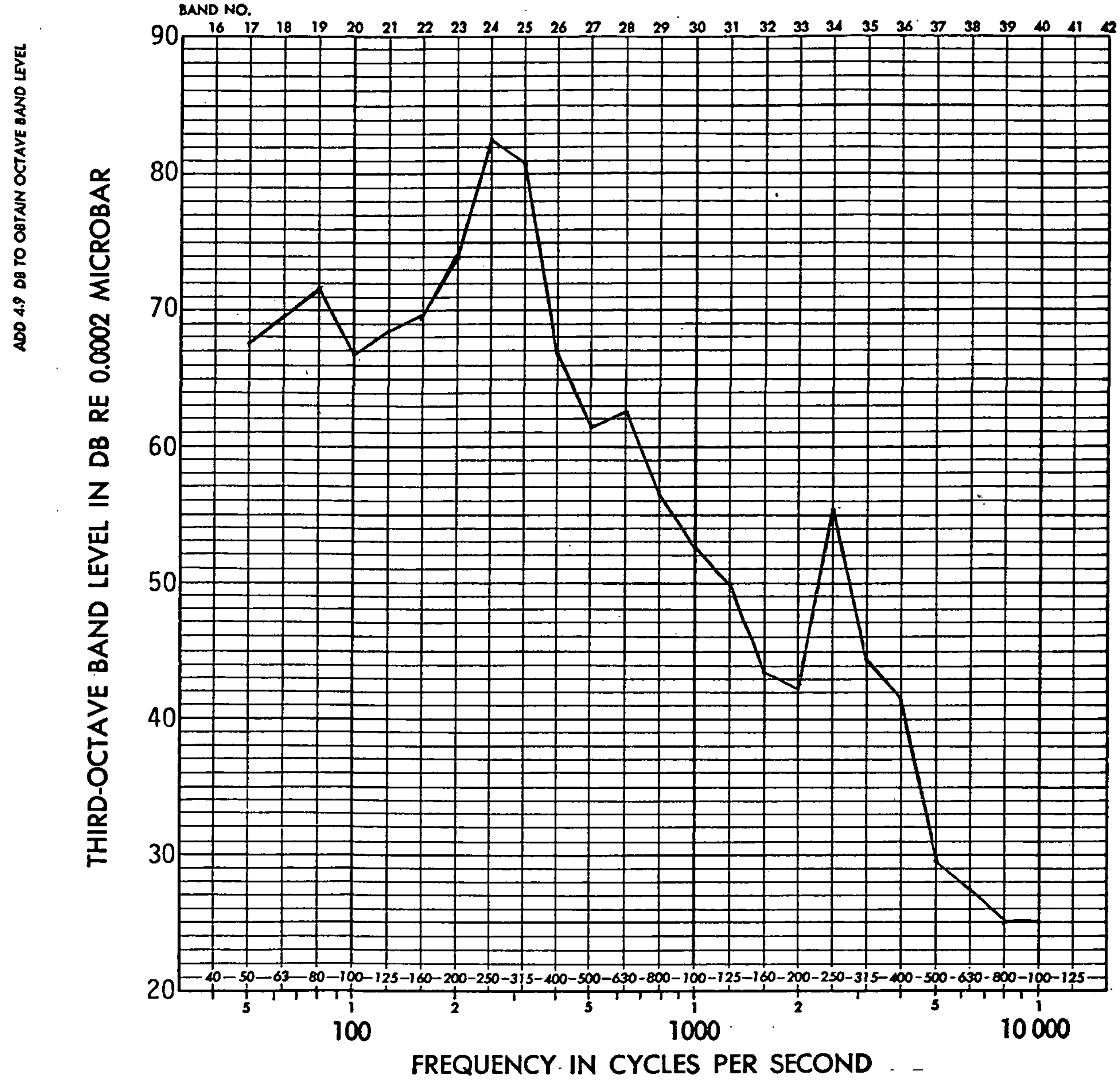

Figure 5: 727 Landing - v_1t_1 Synthesis
1/3 octave x 1/2 second spectrum giving peak dBA and dBAT
(tone-corrected according to FAR 36)

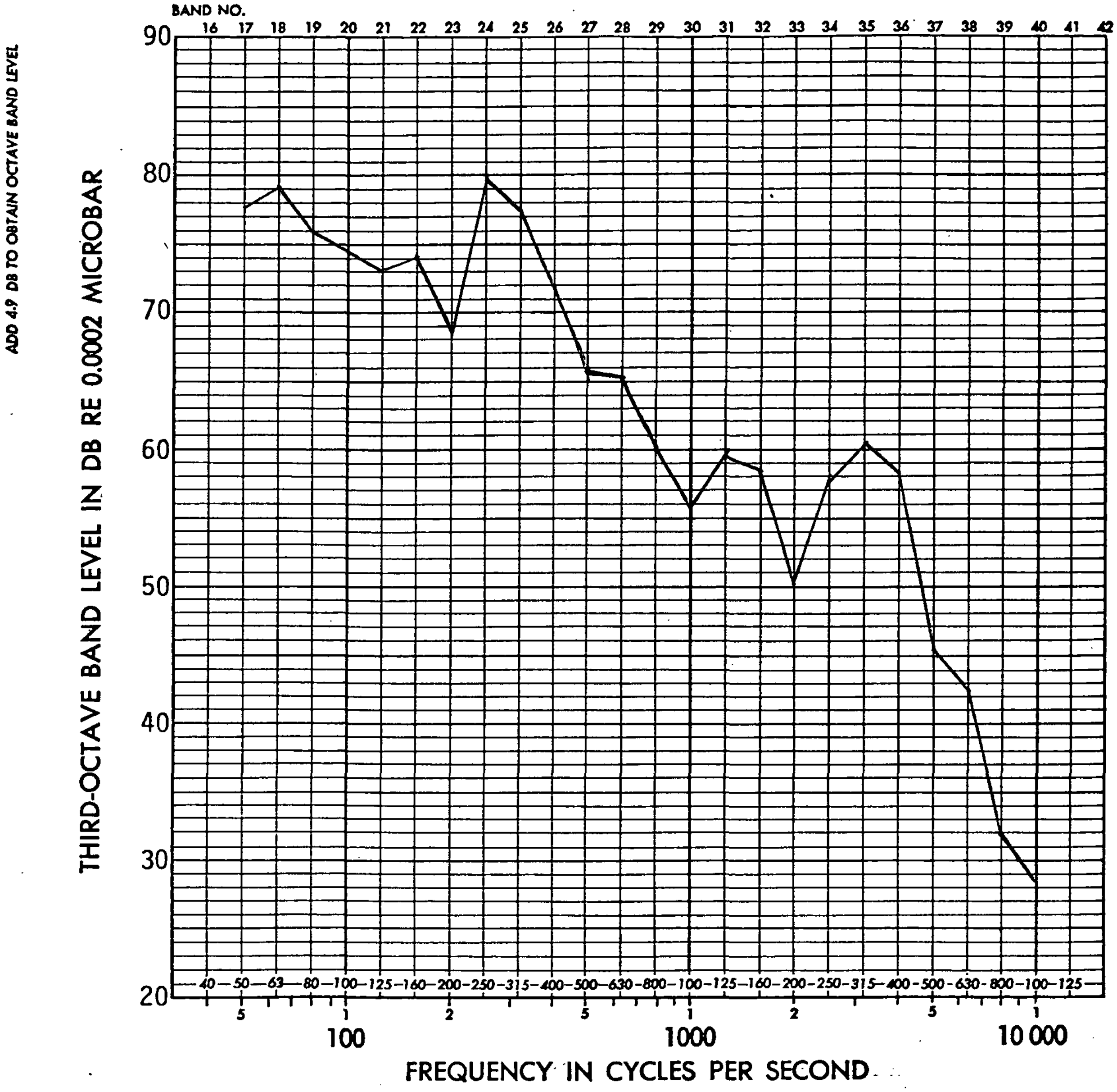

Figure 6: 727 Landing - C Synthesis
1/3 octave x 1/2 second spectrum giving peak dBA

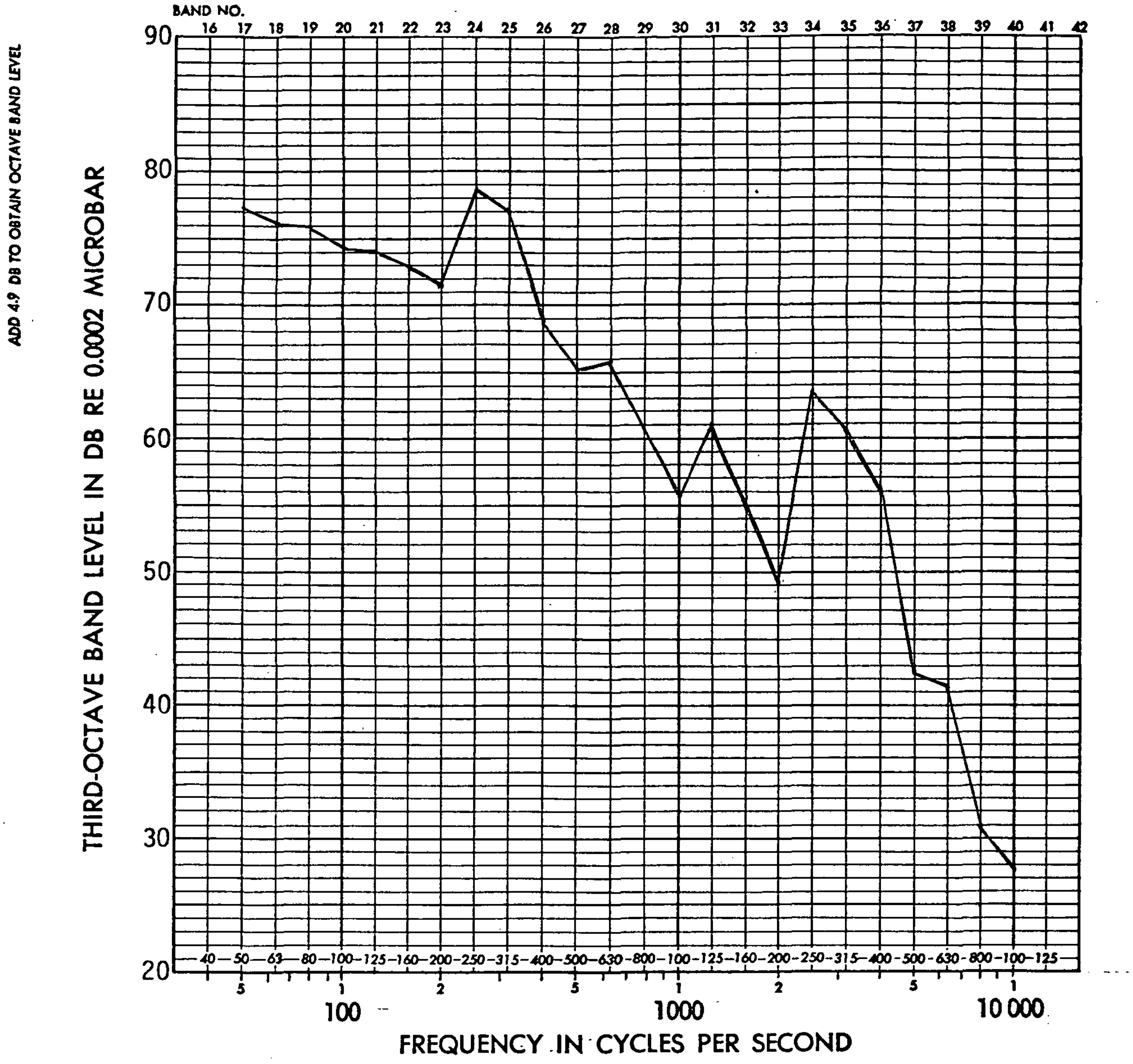

Figure 7: 727 Landing - C Synthesis
1/3 octave x 1/2 second spectrum giving peak dBAT
(dBA tone-corrected according to FAR 36)

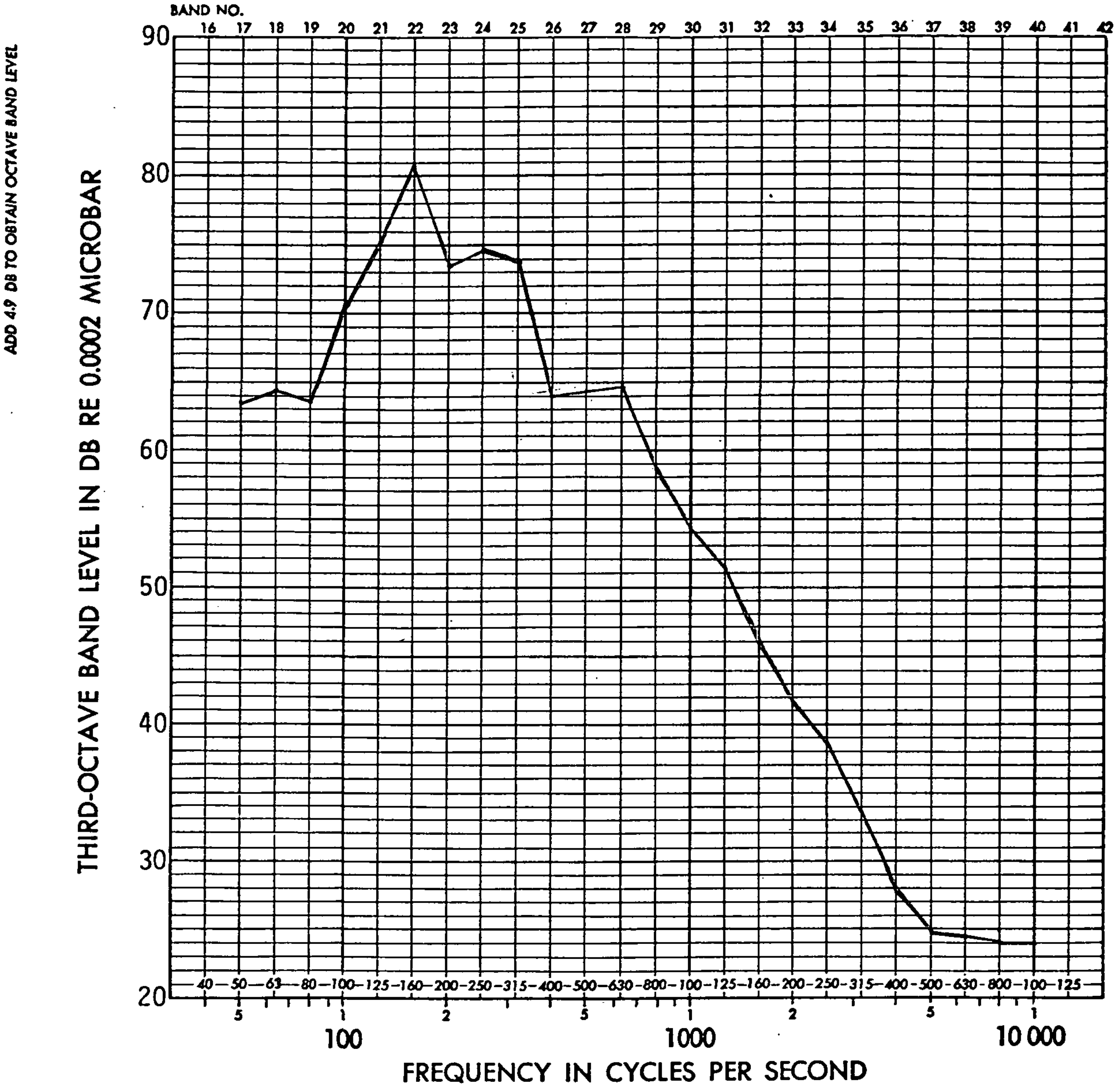

Figure 8: 727 Take-off - Actual Recording
1/3 octave x 1/2 second spectrum giving peak dBA and dBAT
(tone-Corrected according to FAR 36)

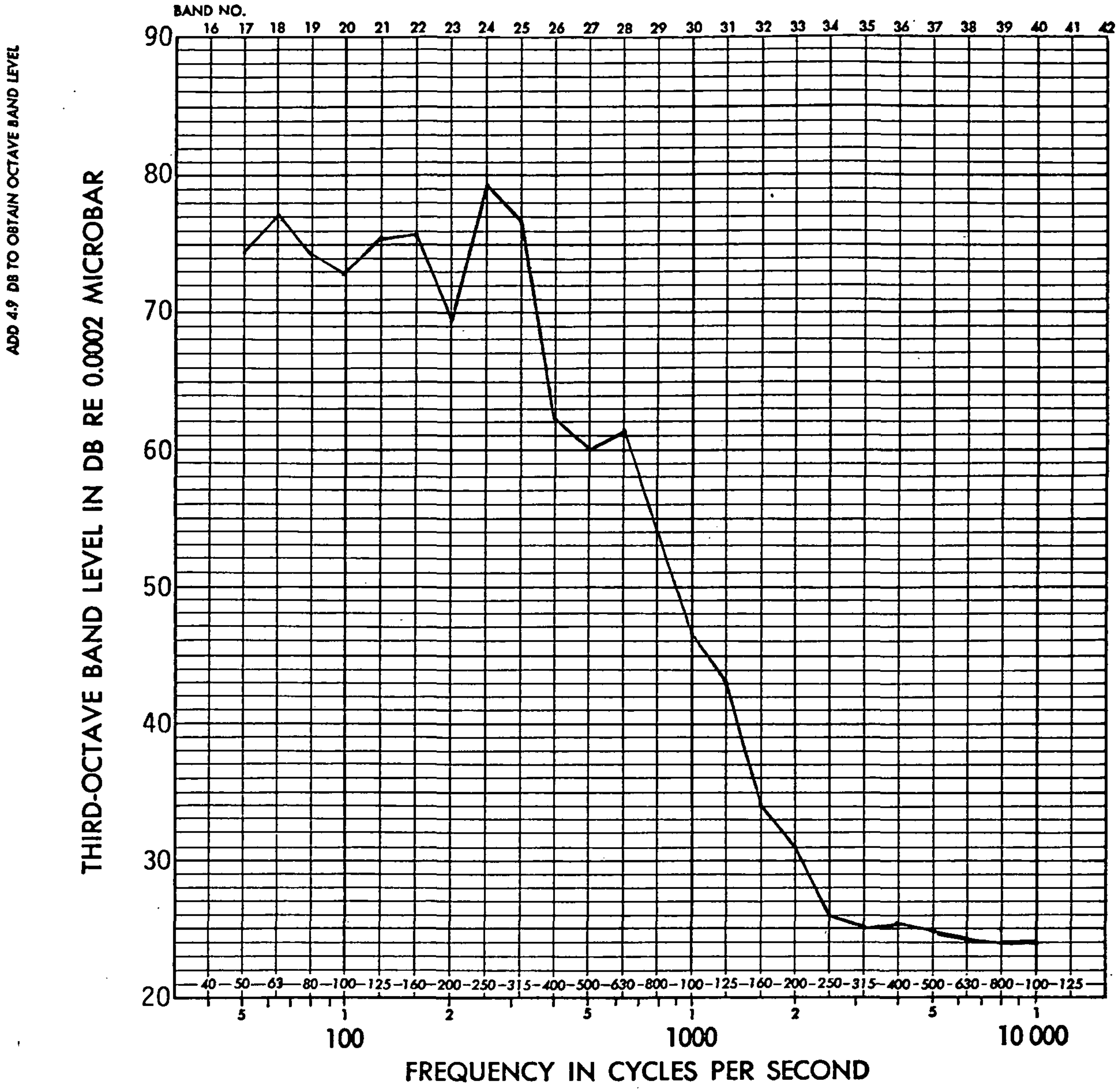

Figure 9: 727 Take-off - $v_1 t_1$ Synthesis
1/3 octave x 1/2 second spectrum giving peak dBA and dBAT
(tone-corrected according to FAR 36)

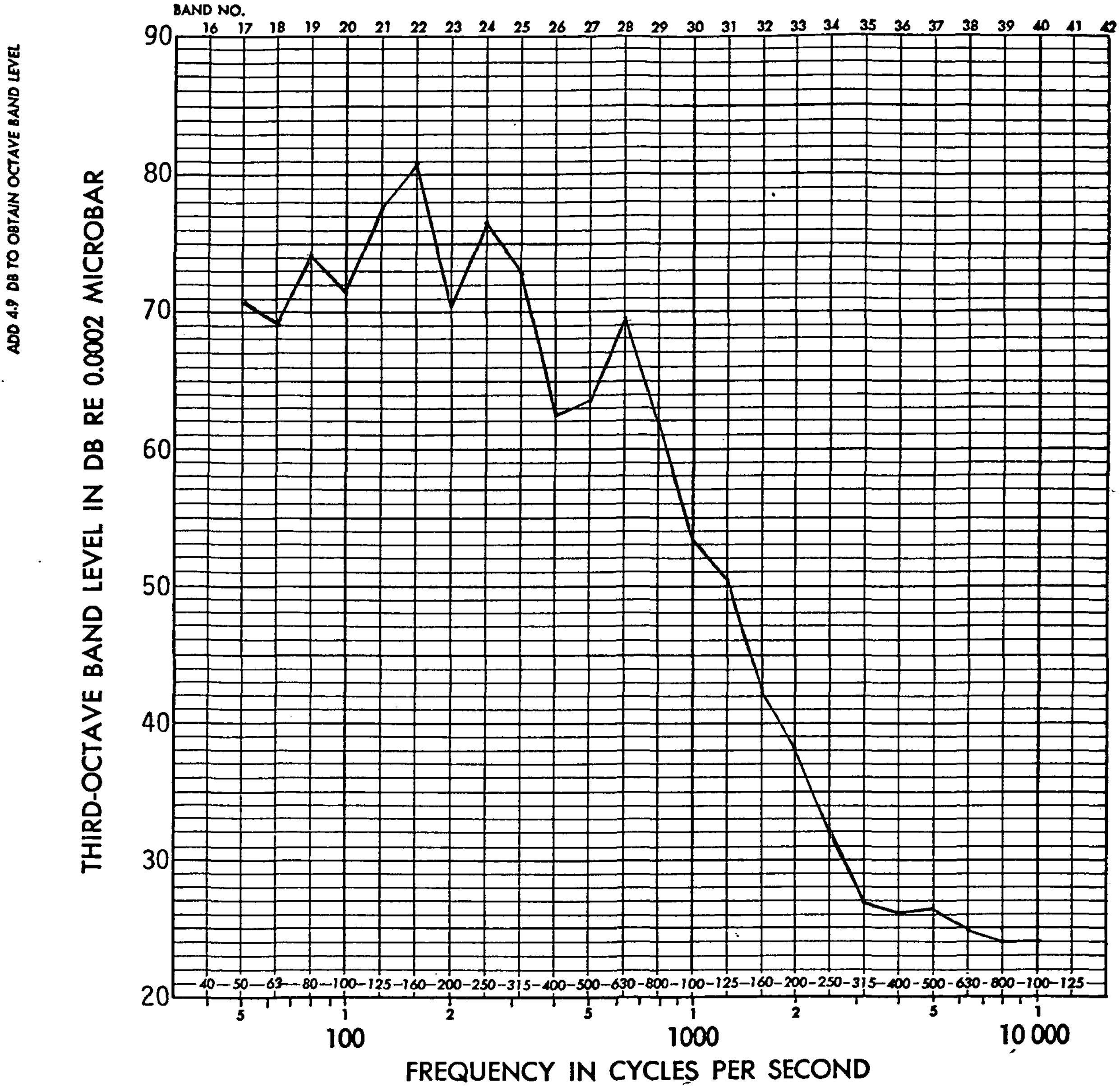

Figure 10: 727 Take-off - C Synthesis
1/3 octave x 1/2 second spectrum giving peak dBA and dBAT
(tone-corrected according to FAR 36)

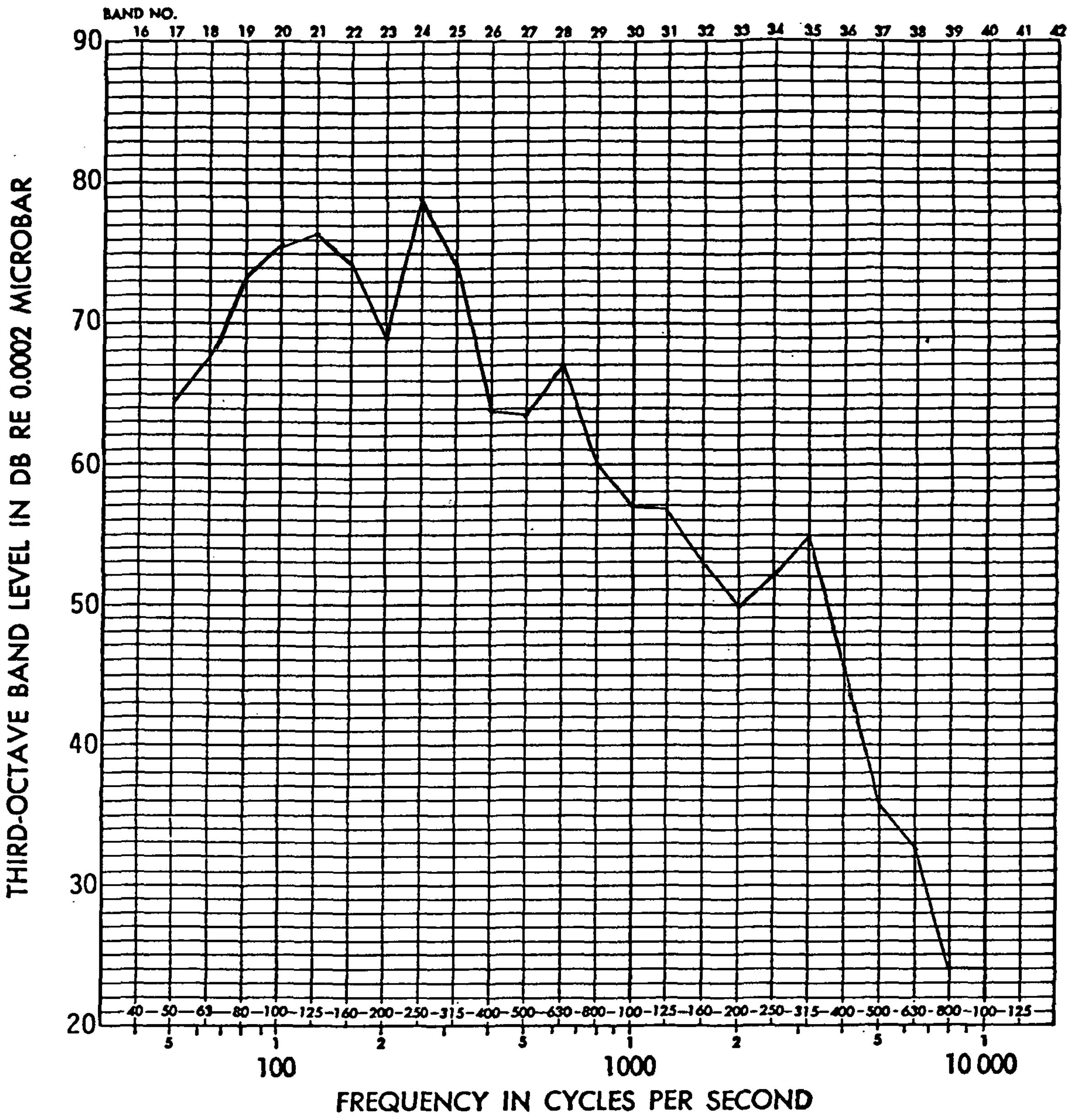

Figure 11: 707 Take-off - Actual Recording
1/3 octave x 1/2 second spectrum giving peak dBA

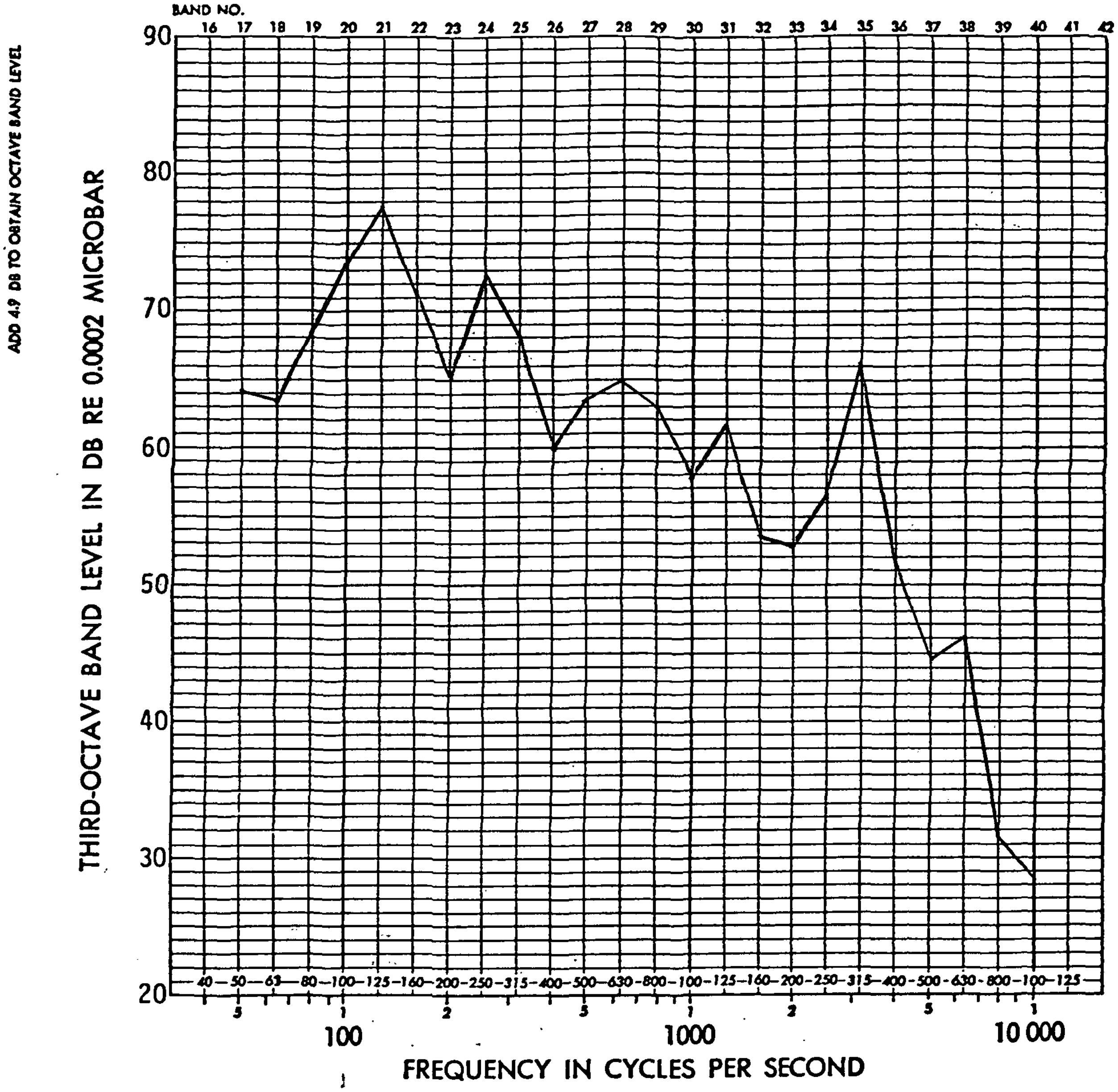

Figure 12: 707 Take-off - Actual Recording
1/3 octave x 1/2 second spectrum giving peak dBAT
(dBA tone-corrected according to FAR 36)

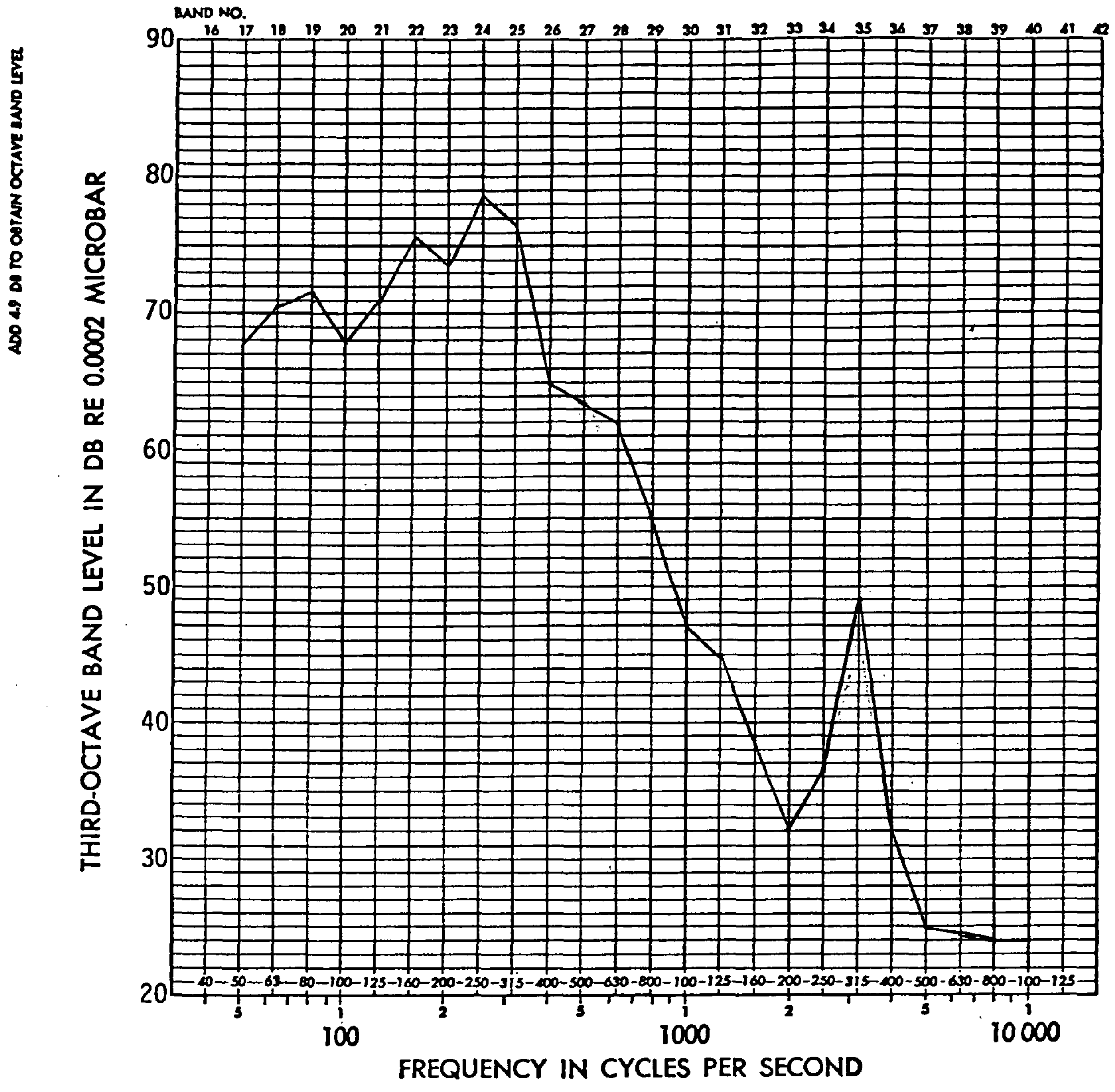

Figure 13: 707 Take-off - v_1t_1 Synthesis
1/3 octave x 1/2 second spectrum giving peak dBA and dBAT
(tone-corrected according to FAR 36)

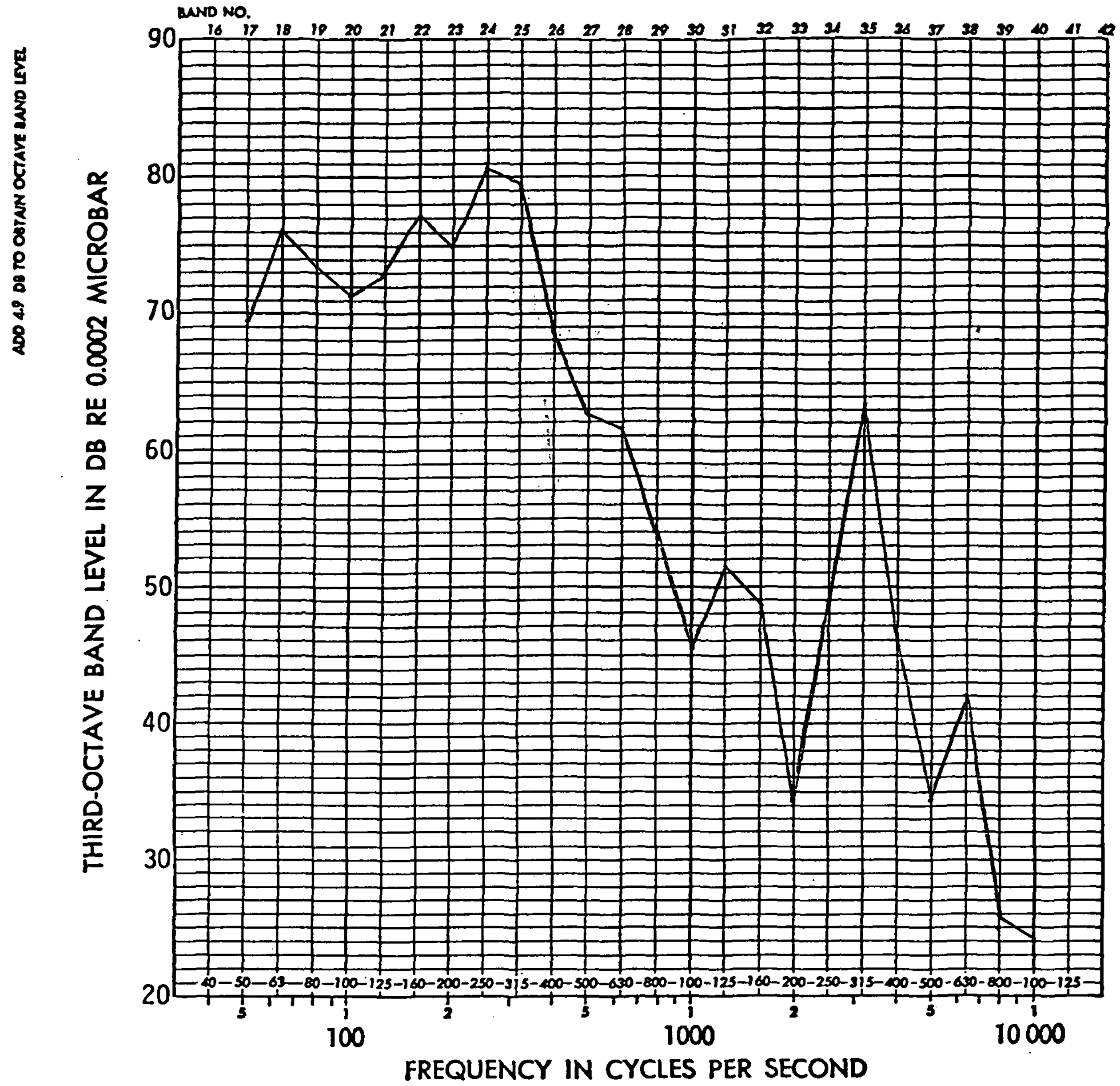

Figure 14: 707 Take-off - C Synthesis
1/3 octave x 1/2 second spectrum giving peak dBA and dBAT
(tone-corrected according to FAR 36)

Table I: Tone corrections calculated from mean presentation levels averaged over six seats

	(av) real	(av) v_1t_1	c	v_1t_2	v_2t_1
727L					
Peak (PNdBT-PNdB)	1.09	2.74	2.52	3.33	2.47
EPNdB-PNdBD	1.38	2.17	2.61	2.64	1.86
av	1.24	2.46	2.57	2.99	2.17
707T					
Peak (PNdBT-PNdB)	4.00	4.35	5.41	5.16	4.60
EPNdB-PNdBD	2.82	3.94	4.99	4.71	4.26
av	3.41	4.15	5.20	4.94	4.43
727T					
Peak (PNdBT-PNdB)	1.73	1.23	1.63	1.26	1.80
EPNdB-PNdBD	1.40	1.31	1.32	1.76	1.44
av	1.57	1.27	1.48	1.51	1.62

Table II: Duration corrections calculated from mean presentation levels averaged over six seats

	real	v_1t_1	c	v_1t_2	v_2t_1
727L *					
PNdBD-PNdB+10	5.50	3.93	4.69	4.33	8.24
EPNdB-PNdBT+10	5.79	3.35	4.78	3.64	7.63
av	5.65	3.62	4.74	3.99	7.94
727T					
PNdBD-PNdB+10	6.74	8.21	8.20	7.52	11.20
EPNdB-PNdBT+10	6.41	8.29	7.89	8.02	10.84
av	6.58	8.25	8.05	7.77	11.02
707T					
PNdBD-PNdB+10	10.32	5.41	5.84	5.39	8.07
EPNdB-PNdBT+10	9.14	5.00	5.42	4.94	7.73
av	9.73	5.21	5.63	5.17	7.90

* 10 dB is added to difference so that larger differences correspond to longer flyover durations

It is remarkable that raising the tonal components by 10dB (producing the v_1t_2 synthesis) has very little effect on this measure of the tone correction: 0.8dB at most for the 707T/0, and 0.2 dB for the 727 landing.

The duration corrections (Table II) also point out differences between the actual events and the syntheses. The simulated 707 T/0 is very plainly too short compared with the actual recording over the 10 dB--down duration. The 727 landing is also short, for the v_1t_1 simulation, though the combination simulation is too long and the 727 take-off is slightly long. The effect of increasing the simulation duration by a factor of 2.5 has increased this measure of the duration correction by between 2 1/2 and 4 1/2 dB.

Theory would give a 4 dB correction for a duration increased by x2.5, so this would agree well. However, theory would suggest raising the "tones" by 10 dB should increase the tone correction by between 1 and 3 dB, which does not seem to be so accurately reflected in these simulations.

It is evident that the simulations do not correspond exactly to the real recordings.

EXPERIMENT DESCRIPTION

SIGNALS

As it was realised that the simulations did not sound exactly like the originals, a simple discrimination test was not in order. Instead, a two-fold test was designed. If the syntheses were very close to the originals, it was postulated that there would be no more difference in their subjective effect compared to that of the original than there would be between two presentations of the original. To test this, both the v_1t_1 synthesis and the original were included twice in the design.

On a lower order, it was suggested that the synthesis might differ no more from the actual recording than another event of the same type (e.g. a 707 take-off). Therefore, two more recordings of each of the three events (707 take-off, 727 take-off, 727 landing) recorded at the same site were included in the design.

Additionally, the other three syntheses for each original (v_1t_2, v_2t_1 and C) were included to see firstly if the combination synthesis was more realistic than the single-pass one, and secondly if the usual tone and duration corrections would adequately account for the parameter changes.

The twenty-seven experimental signals are listed in Table III. Each signal was presented at four different levels, 6 dB apart. A standard signal was used, which was a sample of USASI noise, faded up and down in level to resemble one of the synthesized flyovers.

To produce sounds similiar to those that would be heard indoors, the signals were played through a 1/3-octave shaping filter set to simulate the attenuation of an average house, as given in AIR 1081. All the signals passed through this filter, including the standard sounds.

Table III: 27 flyover sounds used as experimental signals

Signal #	
1	Real 727 landing used as original
2	Repeat of #1
3	Second example of 727 landing
4	Third example of 727 landing
5	Real 727 take-off used as original
6	Repeat of #5
7	Second example of 727 take-off
8	Third example of 727 take-off
9	Real 707 take-off used as original
10	Repeat of #9
11	Second example of 707 take-off
12	Third example of 707 take-off
13	Synthesized 727 landing - v_1t_1
14	Synthesized 727 landing - repeat of #13
15	Synthesized 727 landing - v_1t_2
16	Synthesized 727 landing - v_2t_1
17	Synthesized 727 landing - C^2
18	Synthesized 707 take-off - v_1t_1
19	Synthesized 707 take-off - repeat of #18
20	Synthesized 707 take-off - v_1t_2
21	Synthesized 707 take-off - v_2t_1
22	Synthesized 707 take-off - C^2
23	Synthesized 727 take-off - v_1t_1
24	Synthesized 727 take-off - repeat of #23
25	Synthesized 727 take-off - v_1t_2
26	Synthesized 727 take-off - v_2t_1
27	Synthesized 727 take-off - C^2

METHOD

The experimental method used was the Magnitude Estimation method. This psychophysical method was introduced by S.S. Stevens (Refs. 1 and 2) and has been used widely as a method of relating human response evaluations to physical stimuli. Results from a number of studies indicate that the relationship between sensation and the physical stimulus is a power function (Ref. 1, p. 166). The relationship is:

$$\psi = kI^n$$

where ψ = subjective response
I = stimulus intensity
k = constant of proportionality
n = constant exponent

If the intensity is expressed in decibels, then the equation after rearranging becomes:

$$\log_{10} \psi = \frac{n}{10} \times \text{dB} + \text{constant}$$

Consequently, a log-log plot of subjective response versus stimulus power gives a linear relation with a slope of n/10. The quantity n has been determined experimentally for many stimuli. For noise in particular it has the approximate value of 0.3.

The magnitude estimation method is then utilized to obtain a "Subjective db" for each noise (Ref. 4). Subjective dB is the mechanism for evaluating various engineering calculation procedures. Subjective dB answers the following question: "For a particular engineering calculation procedure as applied to a noise event, do the judges place the noise at the same level as does the engineering procedure and if there is a difference between the judged and calculated level, how great is that difference?" The Subjective dB method for investigating various engineering calculation procedures can best be understood by reference to Figure 15.

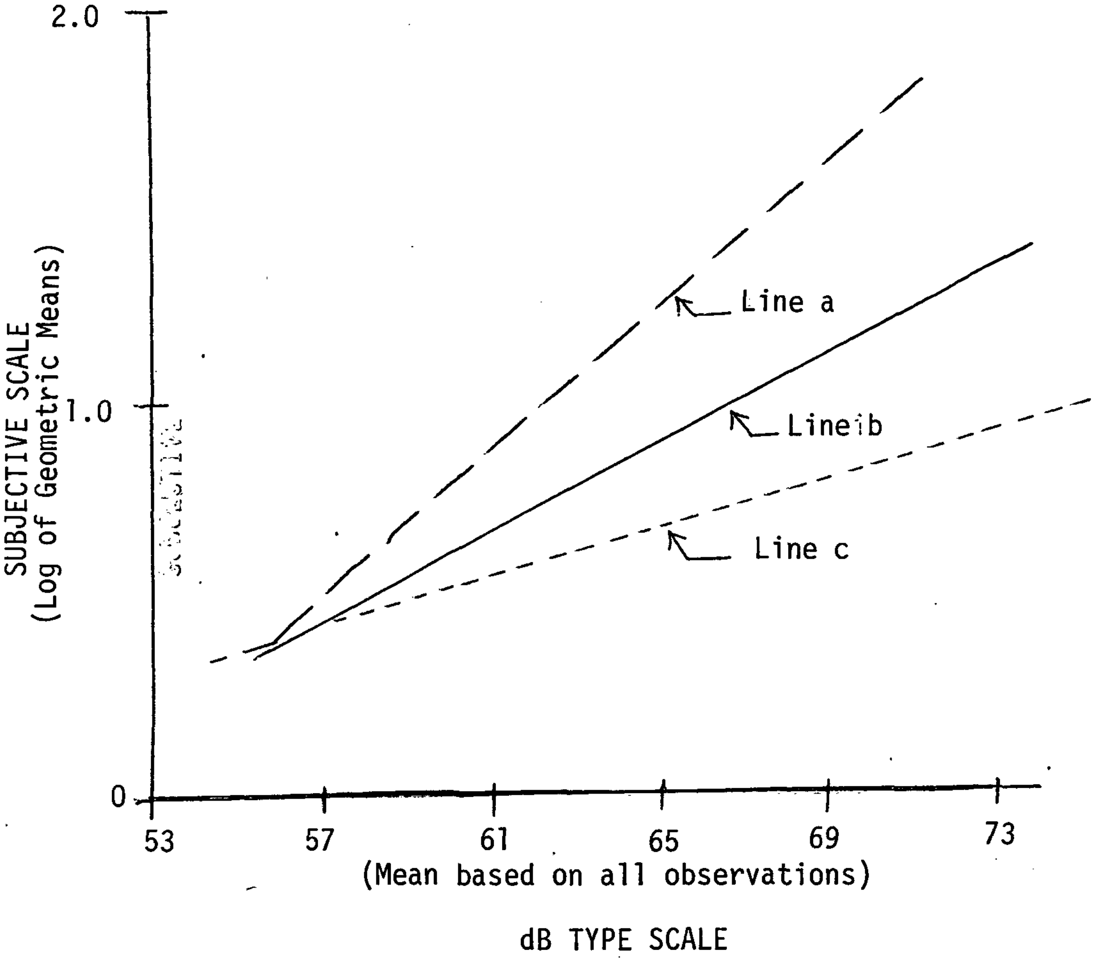

Figure 15: Derivation of Subjective dB

Two assumptions form the basis for acquiring a Subjective dB for any one noise. These assumptions are:

- That the group of subjects is matching numbers in a manner that reflects the amount of annoyance.
- That rate of change of annoyance is different across noises and is a function of a particular noise under investigation.

The abscissa in Figure 15 gives values for a particular calculation procedure under investigation while the ordinate represents the mean evaluations by the judges. Line b is the least squares, best-fitting straight line based on judgments to all noises at all levels. Line b would be based on 108 points, 27 noises at 4 levels. Lines a and c are best-fitting lines for two hypothetical, individual noises (both lines a and c would be based on the four levels for a particular noise or on four points).

The operations in calculating a Subjective dB are:

(1) Obtain equation for best-fitting line using all levels of all noises investiated. This gives an estimate of how well an engineering calculation procedure performs for a wide variety of noises.

(2) Obtain equation for best-fitting line for each individual noise (Lines a, c, ...)

(3) Using the mean of a particular engineering calculation procedure, find, for each individual noise (Lines a and c), the subjective response score predicted by this grand mean.

(4) Using the subjective response score obtained in (3), calculate the engineering calculation procedure value using best-fitting line based on all observations (Line b). This value is the Subjective dB for ME.

Using results from Figure 15 as an example: For the noise on which Line a is based, when the noise is calculated to be at 65 on a dB-type scale, the judges place it at approximately 71, Sub. dB is 71. For the noise on which Line c is based, when the noise is calculated to be at 65 on a dB-type scale, the judges place it at approximately 61, Sub. dB is 61. Each of the 27 noises investigated will be assigned a Sub. dB as described. The predicted results, for each engineering calculation system investigated, will be similar to results presented in Figure 15.

DATA COLLECTION DETAILS

The twenty-seven signals, each presented at four levels, resulted in 108 experimental sounds which were randomly ordered and divided into eight groups. Each group was recorded on a different tape. Each tape started with the standard sound, and this was repeated between the seventh and eighth presentations on each tape.

All the signals were recorded through a DBX compander to improve the signal to noise ratio. They were played back through a DBX compander, an ALTEC 1/3 octave band spectrum shaper to simulate the filter effects of a house wall, and into the Exterior Effects Room (EER) at the NASA-Langley noise research facility. Six overhead speakers were used to give a spatially distributed sound.

A monitor microphone, situated approximately in the center of the EER, was used to ensure signal levels were consistent throughout the experiment.

Sixty subjects were used, all of whom were audiologically normal. Nine were males.

Six subjects participated simultaneously, seated near the center of the room at positions designated Seats #1 to #6.

The six subjects taking part in one experimental session were first given the instruction sheet to read. This was as follows:

INSTRUCTIONS

We are asking you to help answer this question. How annoying are various kinds of sounds? We will ask you to listen to some sounds and rate them in terms of annoyance. The sounds you are to rate will be presented to you one at a time. Listen to all of each sound before making your judgment. In a moment, we will have you listen to a sound with an annoyance score of 10. Use that sound as a standard, and judge each succeeding sound in relation to that standard. For example, if a sound seems <u>twice</u> as annoying as the standard, you will write "20" in the space for that sound on the answer sheet. If it seems only <u>one-quarter</u> as annoying, write "2 1/2." If it seems <u>three</u> times as annoying, write "30." If <u>slightly more than twice</u> as annoying, you may choose to write "21" or "22" or "23" whatever is appropriate. If <u>slightly less annoying</u> than the standard, use the number that best expresses the difference, such as "7" or "8" and so on.

We will also ask you to judge if each sound you hear would be acceptable to you if you experienced it in your home four or five times an hour during your waking hours. This required a simple "yes" or "no" answer in the space provided on the answer sheet.

Your ratings should reflect only your own opinion of the sounds; that is what we want. Each sound is numbered to correspond to the numbers on your answer sheet.

You will now hear the standard sound with an annoyance rating of 10, followed by 13 or 14 more sounds. Rate each of the sounds following the standard as previously instructed; a score of "20" if twice as annoying, "5" if half as annoying, and so on. Also indicate your judgment of the acceptability of each sound.

After a chance to ask any questions the subjects might have had, they were taken into the EER and seated in the designated positions. There was a short break between each tape, to allow collection of the answer sheets and, after the fourth tape, there was a longer break. After the eighth and final tape, the subjects were asked to fill in a questionnaire.

The order of presentation of the eight tapes was varied in a balanced Latin Square design (see Table IV). The final two sessions repeated the order of the first two.

SIGNAL ANALYSIS

Initial investigation of the six listening positions showed that signal levels varied between positions by up to 4 dBA because of variations in the frequency characteristics of the room. It was therefore decided that the presentation levels should be measured at each position separately.

However, it was decided that the change in level for different presentations of the same sound was equivalent for the six positions, from considerations of dBA and dB linear measurements at different positions.

Therefore, all 108 sound presentations were recorded at one of the listening positions (seat #3) and only the top level presentation of each signal (a total of 21 signals, ignoring the repeat presentations) at

Table IV: Tape presentation order

Experimental Session #	Ten Orders of Tape Presentation							
1	1	2	8	3	7	4	6	5
2	2	3	1	4	8	5	7	6
3	3	4	2	5	1	6	8	7
4	4	5	3	6	2	7	1	8
5	5	6	4	7	3	8	2	1
6	6	7	5	8	4	1	3	2
7	7	8	6	1	5	2	4	3
8	8	1	7	2	6	3	5	4
9	1	2	8	3	7	4	6	5
10	2	3	1	4	8	5	7	6

each of the other five positions. All these recorded signals were analyzed at MAN - Acoustics and Noise, Inc. laboratory in Seattle, using a GR1921 real-time 1/3-octave band analyzer and a PDP-11/10 computer.

Using the 108 sounds recorded at seat #3, the differences in level for each signal between the top-level presentation and the three lower level sounds, as well as the repeat presentations, were calculated and used to adjust the twenty-one signal levels for each of the other five positions. Thus the complete set of 108 levels was calculated for each position.

CALCULATION PROCEDURES

The physical calculation procedures used in the analysis were:

01 peak dB(linear) OASPL
02 peak PNdB
03 duration corrected PNdB (PNdBD)
04 tone-corrected PNdB (PNdBT)
05 duration corrected PNdBT (EPNdB)
06 Peak dBA
07 duration corrected dBA (dBAD)
08 tone-corrected dBA (dBAT)
09 duration corrected dBAT (EdBA)
10 peak dBH
11 duration corrected dBH
12 tone-corrected dBH (dBHT)
13 duration corrected dBHT (EdBH)

Tone and duration corrections were calculated using the procedures in FAR-36. The dBH is a unit based on a weighting curve proposed by T. H. Higgins. The curve consists of two straight-line segments, the first rising by 6dB/octave from 50 Hz to 4 KHz, passing through zero at 1 KHz. From 4 KHz, the curve falls at a slope of 6 dB/octave to 10 KHz (Ref. 3).

RESULTS

As described under the EXPERIMENT DESCRIPTION section, magnitude estimation ratings in conjunction with the subjective dB method (Ref. 5) are to be utilized to obtain the sixty subjects' evaluations of the twenty-seven flyover signals. The basic aim of this approach is to evaluate engineering calculation procedures. However, at present there is no engineering calculation procedure which applies in a perfect manner to a diverse collection of noise events. It is this fact, no engineering calculation procedure is generally adequate for all noise signals, that provides the basis for comparing response to the actual flyovers to response to the simulated flyovers. Basically, results will be examined to determine if human response results to the actual and simulated recordings conform to usual or expected findings when applying a particular engineering calculation procedure such as PNdB or dBA.

Differences in signal presentation levels at the six listening positions were large enough to lead to a decision to relate magnitude estimation judgments to only those presentation levels which subjects experienced. An alternative approach would have involved investigation of subjective response to an average of presentation levels over all six listening positions. Table V provides mean presentation levels for each of the thirteen calculation procedures at each listening position; the means are based on the twenty-seven signals at four levels. The mean over all listening positions and the range of means across positions are also given. The range of mean levels across the six positions varies with the calculation procedure. The greatest difference is based on OASPL (dB linear) where difference between position "2" and position "6" is approximately

3.9dB. The least mean difference between positions involves EdBH with a mean difference of approximately 1.7 EdBH between position "1" and position "6". In any event, relating the judgment data to presentation levels at the individual positions increases precision and does not bias the outcome in the direction of a particular calculation procedure.

For each of the six listening positions, subjective dB was calculated for the thirteen engineering calculation procedures by relating log magnitude estimation results to the four levels of a particular flyover signal. The subjective dB determinations were computed for each of the sixty subjects on an individual basis as opposed to an approach which relates the mean response for the group to the four levels of a flyover. Thusly, each person's set of subjective dB's, for the twenty-seven noise signals but using a particular calculation procedure, are independent in the sense that they are not based on responses from other persons taking part in the experiment. Table VI provides some summary information which is based on these individual subjective dB determinations. The basic datum of Table VI is the mean subjective dB (based on 60 independent subjective dB's) for each of the twenty-seven flyover signals. Mean, range, and standard deviation (S.D.) for the twenty-seven flyovers are provided for the thirteen engineering calculation procedures. Primary interest is in measures of scatter such as the range of mean subjective dB and standard deviation. The smaller the range and standard deviation, the greater the applicability of a particular engineering calculation procedure to a set of noise signals. As expected, OASPL (peak dB linear) has the largest standard deviation indicating that it is not as effective as calculation procedures which are designed to reflect annoyance and/or

Table V: Mean presentation levels at six listening positions for thirteen engineering calculation procedures

Unit		Listening Position						Mean	Range
		1	2	3	4	5	6		
1	OASPL	77.55	79.27	76.77	78.38	77.53	75.39	77.48	3.88
2	PNdB	80.23	80.60	79.06	80.39	79.49	78.26	79.67	2.34
3	PNdB$_D$	77.01	77.57	75.83	77.03	76.33	75.12	76.48	2.45
4	PNdB$_T$	83.00	83.07	81.87	83.05	82.54	80.78	82.39	2.29
5	EPNdB	79.43	79.67	78.26	79.30	78.79	77.40	78.81	2.27
6	dBA	67.84	67.75	65.85	67.50	66.65	65.88	66.91	1.99
7	dBA$_D$	64.90	65.04	63.13	64.56	63.84	63.03	64.08	2.01
8	dBA$_T$	70.58	70.15	68.61	70.01	69.66	68.29	69.55	2.29
9	EdBA	67.35	67.20	65.55	66.84	66.33	65.29	66.43	2.06
10	dBH	68.07	68.11	67.11	67.74	67.42	66.10	67.43	2.01
11	dBH$_D$	64.76	65.01	63.88	64.74	64.02	63.22	64.27	1.79
12	dBH$_T$	70.84	70.65	69.94	70.52	70.52	68.72	70.20	2.12
13	EdBH	67.23	67.11	66.33	67.03	66.57	65.51	66.63	1.72

Table VI: Summary information for mean subjective dB: mean, range for 27 noises, standard deviation (S.D.), rank for range and S.D.

Unit	Mean	Range	S.D.	Rank Range	Rank S.D.
OASPL	77.5	12.9	4.1	11	13
Peak PNdB	79.6	7.5	2.0	4.5	3.5
PNdB$_D$	76.4	9.2	2.2	7	6
PNdB$_T$	82.4	6.4	1.9	2	2
EPNdB	78.8	8.6	2.1	4.5	5
Peak dBA	66.9	9.4	2.5	8	7
dBA$_D$	64.1	11.7	2.9	9.5	9.5
dBA$_T$	69.6	9.1	2.7	6	8
EdBA	66.4	11.7	2.9	9.5	9.5
Peak dBH	67.5	13.7	3.3	12	11
dBH$_D$	64.3	6.6	2.0	3	3.5
dBH$_T$	70.3	14.3	3.6	13	12
EdBH	66.4	5.2	1.6	1	1

loudness effects. The smallest of the thirteen standard deviations (S.D.) is for EdBH which also shows the smallest range; the S.D. is 1.6 EdBH with a range of 5.2 EdBH. The group of calculation prodedures based on a dBA weighting perform less adequately than those based on PNdB. The dBH calculation prodedure is unusually effective when corrected for duration in that the standard deviation for mean subjective dB using EdBH is 1.6 dB and is 2.0 dB using dBH_D. Since there is evidence that annoyance response is related to the duration of a signal (total weighted energy), it can be concluded that the most valid calculation procedure is one that is most improved by including duration effects; i.e., all effective aspects of a signal are weighted by the calculation procedure and not just the peak level. Correcting variants of PNdB and dBA for duration produces larger standard deviations and ranges for the mean subjective dB's while smaller standard deviations and ranges result from duration correcting dBH (Table VI). Thusly, it can be concluded that of the procedures evaluated, dBH is the most accurate or valid. Not only does it provide the smallest standard deviation and range (S.D. is 1.6 dB for EdBH) but for both Peak dBH and Peak dBH_T, accuracy is markedly improved when corrected for duration effects.

COMPARISONS BASED ON PAIRS OF IDENTICAL SIGNALS

As described above, one strategy for determining if response to the simulated flyovers was similar to that for actual flyover recordings was to obtain judgment data on two presentations of the same four actual flyover levels. Mean subjective dB for OASPL comparisons are given for 727-L, 727-T, and 707-T in Table VII.

Table VII: Comparison of mean subjective dB for pairs of actual recordings and pairs of simulated recordings (v_1t_1) based on OASPL

a.

	Presentation		
Recording	1	2	Diff.
727-L (AR)*	84.6	84.5	0.1
727-L v_1t_1	74.8	76.3	-1.5
Diff.	9.8	8.2	

b.

	Presentation		
Recording	1	2	Diff.
727-T (AR)	73.5	73.7	-0.2
727-T v_1t_1	72.2	72.8	-0.6
Diff.	1.3	0.9	

c.

	Presentation		
Recording	1	2	Diff.
707-T (AR)	81.4	81.8	-0.4
707-T v_1t_1	74.1	73.9	0.2
Diff.	7.3	7.9	

The results of Table VII a. show that the mean subjective dB's for the two presentations of the actual recording of 727-L (AR) are almost identical (0.1 dB difference) while the difference is 1.5 dB for the two presentations of the simulated flyover (727-L v_1t_1). Neither difference is considered unusually large and the conclusion is that repeatability is adequate. However, the difference between mean subjective dB pairs of actual and simulated recordings averages some 9 dB. Clearly the response to the simulations utilizing OASPL as a calculation procedure is not similar to response to recordings of the actual flyover. Table VII b. provides the same comparison data for the 727 take-off. For both the actual and simulated flyover, repeatability is almost perfect. Also, the difference based on pairs of actual versus simulated recordings averages approximately one dB. For the 727 take-off, it is concluded that there is similarity of response to the actual and simulated recordings but not to the extent that there is similarity of response to identical presentations of the same flyover. The results for the 707 take-off are almost identical to those for the 727 landing. As Table VII c. shows, repeatability is unusually high (differences of -0.4 and 0.2 dB), but the simulations are, on average, judged at some 7.6 dB lower than the actual recordings.

---

*(AR) means actual recording

Comparisons based on identical presentations of pairs of actual and pairs of simulated recordings for the three basic calculation procedures (PNdB, dBA and dBH) are given in Tables VIII through XVI. Tables VIII, IX and X provide results for the three aircraft operations utilizing the PNdB group of procedures. As for the peak OASPL comparisons, repeatability is unusually high for results based on response to the two presentations of the actual flyover recordings. While repeatability of results for the pairs of simulations is not as high as repeatability for pairs of actual recordings, it is considered satisfactory in that differences range from 0.0 to 1.6 dB. Subjective dB differences between pairs of actual and simulated recordings are, in general, much less pronounced than comparisons based on peak OASPL. However there are inconsistencies across the three flyovers investigated. For example, $PNdB_T$ shows similarity of response to the actual and simulated 727 takeoff with differences of 1.1 and -0.2 $PNdB_T$ (Table IX, c.) while differences for the 707 takeoff are 6.1 and 6.3 $PNdB_T$ (Table X, c.). Examining results for all three flyovers leads to the conclusion that $PNdB_D$ shows the greatest amount of similarity of response to the actual and simulated recordings with differences ranging from 0.2 to 2.4 $PNdB_D$.

For the four calculation procedures based on the dBA weighting network (Tables XI, XII and XIII), repeatability remains high for both the actual and simulated pairs but with greater consistency of response to the actual recordings. The range of differences for pairs of subjective dB based on the actual recordings is 0.0 to 0.4 dB while it is 0.0 to 1.7 dB for the simulated recordings. However, when comparing response differences based on the actual vs. the simulated recordings, inconsistencies across the three flyovers are readily apparent. For the 727 landing comparisons, the simulations are rated from 3.1 to 7.3 dB less annoying than are the actual recordings. The range is less for the 727 takeoff in that it is from 0.0

Table VIII: Comparison of mean subjective dB for pairs of actual recordings and pairs of simulated recordings (v_1t_1) for 727 landing (727-L): PNdB set of calculation procedures

a. Peak PNdB

Recording	Presentation		Diff.
	1	2	
727-L (AR)	79.6	79.1	0.5
727-L v_1t_1	76.9	78.0	-1.1
Diff.	2.7	1.1	

b. $PNdB_D$

Recording	Presentation		Diff.
	1	2	
727-L (AR)	78.0	77.7	0.3
727-L v_1t_1	76.3	77.5	-1.2
Diff.	1.7	0.2	

c. $PNdB_T$

Recording	Presentation		Diff.
	1	2	
727-L (AR)	84.2	83.8	0.4
727-L v_1t_1	79.3	80.9	-1.6
Diff.	4.9	2.9	

d. EPNdB

Recording	Presentation		Diff.
	1	2	
727-L (AR)	81.4	81.0	0.4
727-L v_1t_1	78.6	80.1	-1.5
Diff.	2.8	0.9	

Table IX: Comparison of mean subjective dB for pairs of actual recordings and pairs of simulated recordings (v_1t_1) for 727 take-off (727-T): PNdB set of calculation procedures

a. Peak PNdB

Recording	Presentation		Diff.
	1	2	
727-T (AR)	78.1	78.1	0.0
727-T v_1t_1	76.8	77.7	0.9
Diff.	1.3	0.4	

b. $PNdB_D$

Recording	Presentation		Diff.
	1	2	
727-T (AR)	74.9	75.0	-0.1
727-T v_1t_1	72.5	73.0	-0.5
Diff.	2.4	2.0	

c. PNdBT

Recording	Presentation		Diff.
	1	2	
727-T (AR)	81.8	81.7	0.1
727-T v_1t_1	80.7	81.9	-1.2
Diff	1.1	-0.2	

d. EPNdB

Recording	Presentation		Diff.
	1	2	
727-T (AR)	78.0	78.0	0.0
727-T v_1t_1	75.6	76.3	-0.7
Diff.	2.4	1.7	

Table X: Comparison of mean subjective dB for pairs of actual recordings and pairs of simulated recordings (v_1t_1) for 707 take-off (707-T): PNdB set of calculation procedures

a. Peak PNdB

	Presentation		
Recording	1	2	Diff.
707-T (AR)	83.8	84.2	-0.4
707-T v_1t_1	78.3	78.4	-0.1
Diff.	5.3	5.8	

b. PNdB_D

	Presentation		
Recording	1	2	Diff.
707-T (AR)	77.6	78.1	-0.5
707-T v_1t_1	76.6	76.6	0.0
Diff	1.0	1.5	

c. PNdBT

	Presentation		
Recording	1	2	Diff.
707-T (AR)	85.4	85.7	-0.3
707-T v_1t_1	79.3	79.4	-0.1
Diff.	6.1	6.3	

d. EPNdB

	Presentation		
Recording	1	2	Diff.
707-T (AR)	79.5	80.0	-0.5
707-T v_1t_1	77.2	77.1	0.1
Diff.	2.3	3.1	

Table XI: Comparison of mean subjective dB for pairs of actual recordings and pairs of simulated recordings (v_1t_1) for 727 landing (727-L): dBA set of calculation procedures

a. Peak dBA

Recording	Presentation 1	Presentation 2	Diff.
727-L (AR)	68.1	67.8	0.3
727-L v_1t_1	63.2	64.5	-1.2
Diff.	4.9	3.3	

b. dBA_D

Recording	Presentation 1	Presentation 2	Diff.
727-L (AR)	67.8	67.6	0.2
727-L v_1t_1	63.1	64.5	- 1.4
Diff.	4.7	3.1	

c. dBA_T

Recording	Presentation 1	Presentation 2	Diff.
727-L (AR)	72.6	72.4	0.2
727-L v_1t_1	65.3	67.0	- 1.7
Diff.	7.3	5.4	

d. EdBA

Recording	Presentation 1	Presentation 2	Diff.
727-L (AR)	71.2	70.9	0.3
727-L v_1t_1	65.5	66.7	- 1.2
Diff.	5.7	4.2	

Table XII: Comparison of mean subjective dB for pairs of actual recordings and pairs of simulated recordings (v_1t_1) for 727 take-off (727-T): dBA set of calculation procedures

a. Peak dBA

Recording	Presentation 1	Presentation 2	Diff.
727-T (AR)	65.0	65.2	-0.2
727-T v_1t_1	63.6	64.4	-0.8
Diff.	1.4	0.8	

b. dBA_D

Recording	Presentation 1	Presentation 2	Diff.
727-T (AR)	62.6	62.6	0.0
727-T v_1t_1	59.5	60.2	- 0.7
Diff.	3.1	2.4	

c. dBA_T

Recording	Presentation 1	Presentation 2	Diff.
727-T (AR)	68.6	68.6	0.0
727-T v_1t_1	67.6	68.6	-1.0
Diff.	1.0	0.0	

d. EdBA

Recording	Presentation 1	Presentation 2	Diff.
727-T (AR)	65.7	65.7	0.0
727-T v_1t_1	62.8	63.5	-0.7
Diff.	2.9	2.2	

Table XIII: Comparison of mean subjective dB for pairs of actual recordings and pairs of simulated recordings (v_1t_1) for 707 take-off (707-T): dBA set of calculation procedures

a. Peak dBA

Recording	Presentation 1	Presentation 2	Diff.
707-T (AR)	72.2	72.5	-0.3
707-T v_1t_1	64.6	64.6	0.0
Diff.	7.6	7.9	

b. dBA_D

Recording	Presentation 1	Presentation 2	Diff.
707-T (AR)	65.8	66.2	-0.4
707-T v_1t_1	63.5	63.4	0.1
Diff.	2.3	2.8	

c. dBA_T

Recording	Presentation 1	Presentation 2	Diff.
707-T (AR)	74.3	74.4	-0.1
707-T v_1t_1	65.4	65.5	-0.1
Diff.	8.9	8.9	

d. EdBA

Recording	Presentation 1	Presentation 2	Diff.
707-T (AR)	67.8	68.2	-0.4
707-T v_1t_1	64.1	63.9	0.2
Diff.	3.7	4.3	

Table XIV: Comparison of mean subjective dB for pairs of actual recordings and pairs of simulated recordings (v_1t_1) for 727 landing (727-L): dBH set of calculation procedures

a. Peak dBH

Recording	Presentation		Diff.
	1	2	
727-L (AR)	61.1	60.8	0.3
727-L v_1t_1	66.0	67.4	-1.4
Diff.	-4.9	-6.6	

b. dBH_D

Recording	Presentation		Diff.
	1	2	
727-L (AR)	60.5	60.2	0.3
727-L v_1t_1	65.6	66.6	-1.0
Diff.	-5.1	-6.4	

c. dBH_T

Recording	Presentation		Diff.
	1	2	
727-L (AR)	65.6	65.6	0.0
727-L v_1t_1	68.9	70.6	-1.7
Diff.	-3.3	-5.0	

d. EdBH

Recording	Presentation		Diff.
	1	2	
727-L (AR)	64.3	63.9	0.4
727-L v_1t_1	68.0	69.0	-1.0
Diff.	-3.7	-5.1	

Table XV: Comparison of mean subjective dB for pairs of actual recordings and pairs of simulated recordings (v_1t_1) for 727 take-off (727-T): dBH set of calculation procedures

a. Peak dBH

Recording	Presentation		Diff.
	1	2	
727-T (AR)	68.8	68.6	0.2
727-T v_1t_1	67.0	68.5	-1.5
Diff.	1.8	0.1	

b. dBH_D

Recording	Presentation		Diff.
	1	2	
727-T (AR)	64.7	64.8	-0.1
727-T v_1t_1	62.2	63.1	-1.1
Diff.	2.5	1.7	

c. dBH_T

Recording	Presentation		Diff.
	1	2	
727-T (AR)	72.8	72.5	0.3
727-T v_1t_1	71.6	73.4	-1.8
Diff.	1.2	-0.9	

d. EdBH

Recording	Presentation		Diff.
	1	2	
727-T (AR)	68.1	68.0	0.1
727-T v_1t_1	65.5	66.6	-1.1
Diff.	2.6	1.4	

Table XVI: Comparison of mean subjective dB for pairs of actual recordings and pairs of simulated recordings (v_1t_1) for 707 take-off (707-T): dBH set of calculation procedures

a. Peak dBH

Recording	Presentation		Diff.
	1	2	
707-T (AR)	69.1	69.7	-0.6
707-T v_1t_1	68.4	68.3	0.1
Diff.	0.7	1.4	

b. dBH_D

Recording	Presentation		Diff.
	1	2	
707-T (AR)	63.9	64.6	-0.7
707-T v_1t_1	66.4	66.5	-0.1
Diff.	-2.5	-1.5	

c. dBH_T

Recording	Presentation		Diff.
	1	2	
707-T (AR)	70.4	70.8	-0.4
707-T v_1t_1	69.0	69.0	0.0
Diff.	0.5	1.8	

d. EdBH

Recording	Presentation		Diff.
	1	2	
707-T (AR)	65.8	66.3	-0.5
707-T v_1t_1	66.6	66.6	0.0
Diff.	-0.8	-0.3	

to 3.1 dB while it ranges from 2.3 to 8.9 dB for the 707 takeoff. Clearly, a particular variant of an engineering calculation procedure shows comparability of response to the actual and simulated recordings for one flyover but not for another. Utilizing dBA_T, the two differences between pairs are 1.0 and 0.0 dB for the 727 takeoff (Table XII, c.) while for the other two flyovers the differences range from 5.4 to 8.9 dBA_T. Due to the fact that for the 707 takeoff, peak dBA and dBA_T results show that the simulations are rated some 7 to 9 dB less annoying than the actual flyovers, it is concluded that the dBA weighting approach has more generalized application when utilized in conjunction with a correction for duration.

Turning to the group of calculation procedures based on dBH (Tables XIV, XV and XVI), as expected the pattern of comparisons is, for the most part, similar to that for the previous two basic calculation procedure groups. There is almost identical response to pairs of actual recordings (differences ranging from 0.0 to 0.7 dB) while response to the pairs of simulated recordings shows but adequate comparability (differences ranging from 0.0 to 1.8 dB). Again, there are inconsistencies for comparisons involving actual and simulated recordings. For example, the comparison of the simulated vs. the actual recording for the 707 takeoff shows that the simulation is judged, at most, 0.8 EdBH more annoying than the actual recording (Table XVI, d.) while for the 727 landing, the simulated recording is judged some 6.6 dBH more annoying than the actual recording (Table XIV, a.). The variant of dBH which provides the highest comparability between the actual and simulated recordings across all three flyovers is EdBH (differences range from 0.3 to 5.1 dBH) although dBH_T is almost equally effective. The dBH group of engineering calculation procedures does provide one difference of interest over the PNdB and dBA groups. Some of the comparisons between actual and simulated pairs show that the simulations are judged more annoying than the actual recordings; this result was not obtained for any of the comparisons based on PNdB or dBA where the simulations were, in all comparisons, judged less annoying than the actual recordings. For the 24 comparisons involving dBH, 13 of the simulations are judged more annoying than the actual recordings and 11 are judged less annoying than the corresponding actual recordings. This is an expected result (approximately equal dispersion of more or less annoying) if a particular calculation procedure is genuinely applicable to a diverse group of noise signatures.

COMPARISONS BASED ON ALL SIGNALS UTILIZING FOUR CALCULATION PROCEDURES

The just preceding analyses were based on comparisons of subjective dB results for pairs of identical signals. A general conclusion from these analyses for pairs of identical signals is that similarity of response to the actual and simulated recordings is very much a function of both the flyover signal and a particular engineering calculation procedure. Another approach for investigating comparability of response to the two kinds of flyover involves utilization of subjective response to all twenty-seven flyovers. With the exception of the dBH calculation procedures, there is a marked tendency from the pair analyses for the actual recordings to be rated more annoying than the simulated recordings. Does this same situation hold if all 12 actual recordings and 15 simulated recordings are considered as a group? For this analysis, the variant of a particular basic group with the smallest standard deviation (S. D.) is selected for consideration. Thusly results based on PNdB$_T$, Peak dBA and EdBH are selected (see Table VI). Since EPNdB is also widely used to evaluate response to flyover noise, it was also included.

If there is a tendency for the 12 actual recordings vs. the 15 simulated recordings to demonstrate comparable annoyance response, approximately 6 or 50% of the actual recordings should rank in a high "twelve" annoyance set while some 7 to 8 of the simulated recordings would rank in a high "fifteen" annoyance set. Table XVII shows the results for this comparison.

Table XVII: Percent of Actual and Simulated Recordings Falling in High Annoyance Sets Based on Twenty-Seven Flyovers

Calculation Procedure	Actual Recordings	Simulated Recordings
PNL_T	67%	33%
EPNdB	67%	40%
Peak dBA	67%	33%
EdBH	58%	53%

Consideration of results to all 27 flyovers shows that, on the whole, there is a tendency for the actual recordings to be judged more annoying than the simulated recordings utilizing PNL_T, EPNdB and Peak dBA. However, there is much more annoyance comparability between the actual and simulated recordings using the EdBH calculation procedure. As for the pair analyses, this result supports a conclusion that the EdBH procedure has broader application to a diverse group of noise signatures than the conventional calculation procedures.

STATISTICAL INFERENCE COMPARISONS AMONG ACTUAL AND SIMULATED RECORDINGS

For the four calculation procedures of the previous section ($PNdB_T$, EPNdB, Peak dBA, and EdBH), an analysis of variance based on repeated measurements but "nested" on the six listening-position groups was completed. A summary of the results is given in Table XVIII. As expected, subjective dB's among the twenty-seven noises are significantly different for all four calculation procedures. That there are significant differences among the groups for each of the six listening positions is quite possibly an artifact. The calculated value for obtaining the subjective dB's for each procedure (calculated $PNdB_T$ as opposed to judged $PNdB_T$ as an example) is the mean of the presentation level at each of the six listening positions.

Table XVIII: Summary of Analyses of Variance for Four Engineering Calculation Procedures

Engineer. Calc. Proc.	Source of Variation	Sum of Squares	df	Mean Square	F-ratio	Signif. Point
$PNdB_T$	Noises(N)	5709.8	26	219.6	11.93	P<.005
	Groups(G)	1199.4	5	239.9	13.03	P<.005
	N x G	10557.9	130	81.2	4.41	P<.005
	Error	25854.8	1404	18.4		
EPNdB	Noises(N)	6835.8	26	262.9	14.44	P<.005
	Groups(G)	1046.1	5	209.2	11.49	P<.005
	N x G	10452.7	130	80.4	4.42	P<.005
	Error	25560.6	1404	18.2		
Peak dBA	Noises(N)	10119.6	26	389.2	24.8	P<.005
	Groups(G)	1195.3	5	239.1	15.2	P<.005
	N x G	13701.4	130	105.4	6.7	P<.005
	Error	22054.6	1404	15.7		
EdBH	Noises(N)	3801.1	26	146.2	6.39	P<.005
	Groups(G)	745.1	5	149.0	6.52	P<.005
	N x G	7985.8	130	61.4	2.68	P<.005
	Error	32111.2	1404	22.9		

Using such an approach results in an almost identical mean subjective dB and mean presentation level at each listening position. Table XIX shows this correspondence between the mean presentation levels and mean subjective db's for the six groups using $PNdB_T$. The differences are trivial in that they range from 0.0 to 0.3 $PNdB_T$. Using the mean presentation level at each listening position as the calculated value for obtaining the subjective db's is considered an oversight. A more appropriate approach uses the mean of the presentation levels over all listening positions as the calculated value for obtaining subjective dB's. To make certain that these statistically

Table XIX: Mean $PNdB_T$ Presentation Levels and Mean Subjective dB's for the Six Listening Position Groups

Group	Mean Presentation Level	Mean Subjective dB
1	83.0	83.0
2	83.1	83.2
3	81.9	81.9
4	83.0	83.1
5	82.5	82.2
6	80.8	80.7

significant group differences were due to the analysis method and not to actual differences among the groups, individual analyses of variance were completed for each of the six groups using $PNdB_T$, EPNdB, Peak dBA, and EdBH. For these 24 independent-across-groups analyses, if there are no reliable differences among the ten subjects at each listening position, this finding would support a conclusion that differences between groups is indeed an artifact (subjects not different on the dependent measure). For these 24 individual analyses, variance attributable to individual differences among subjects at each listening position was less than the error variance in 22 of the 24 analyses. The Group 2 subject F-ratio for EdBH was significant at the $P < .05$ point but this can be considered a chance finding since 1 of 20 analyses would, by chance alone, produce a difference at the $P < .05$ point. Subjects do not differ from Group to Group on the dependent measure.

A product of the analyses of variance for a repeated measurement design is an error term which permits an evaluation of subjective differences between individual flyovers. Such questions can be posed as, "How large must a difference between ratings of two signals be for the difference to be accepted as a reliable (non-chance) difference?" For the present study, the extent that responses to actual and simulated recordings are in the same statistical inference set can be examined. Duncan's Multiple Range Test (Ref. 5) is used for this analysis and the results are given in Tables XX through XXIII. The results for this multiple range analysis provide a finite number of overlapping sets where all mean values within a particular set are considered as, "not being reliably different" or that any differences can be considered as due to measurement or experimental error. For the mean subjective dB's using PNdBT there are 14 overlapping sets, for EPNdB there are 10 sets, for PeakdBA there are eleven, and the dBH calculation procedure provided 6 overlapping sets. Using the EPNdB based analysis (Table XXI) as an example, Set 1 covers the lowest rated five flyovers with subjective dB's ranging from 74.5 to 76.4. It is concluded that there is no reliable difference among these five noise signals with a range of 1.9 EPNdB. The rule is that any two means that are not overlapped by two contiguous sets are reliably different. Thusly for the EPNdB analysis, signal 26 with a rank of 1 is outside of set 2 and signal 19 with a rank of 6 is not included in set 1 (both are not overlapped by two contiguous sets), so these two means are reliably different; a difference of 2.6 EPNdB results in a reliable difference.

Utilizing this multiple range data relative to the aims of the present study, the question can be raised concerning tendencies to be members

Table XX: Significant Differences Among 27 Noises at $P<.01$ Level Using Duncan's Multiple Range Test - $PNdB_T$

RANK	$PNdB_T$	A/C	DESIG.	Orig. Code No.	Actual or Sim.
1	79.3	707-T	v_1t_1	18	S
2	79.3	727-L	v_1t_1	13	S
3	79.4	707-T	v_1t_1-*rp	19	S
4	80.2	707-T	v_1t_2	20	S
5	80.7	707-T	c	22	S
6	80.7	727-T	v_1t_1	23	S
7	80.9	727-L	v_1t_1-rp	14	S
8	80.9	727-L	v_1t_2	15	S
9	81.1	727-T	c	27	S
10	81.3	727-L	3'd-**eg	4	A
11	81.4	707-T	v_2t_1	21	S
12	81.7	727-T	orig.-rp	6	A
13	81.8	727-T	orig.	5	A
14	81.9	727-T	v_1t_1-rp	24	S
15	82.9	727-T	2'd-eg	7	A
16	83.2	707-T	2'd-eg	11	A
17	83.3	727-T	3'd-eg	8	A
18	83.6	727-T	v_2t_1	26	S
19	83.6	727-T	v_1t_2	25	S
20	83.8	727-L	orig.-rp	2	A
21	84.0	727-L	c	17	S
22	84.1	707-T	3'd-eg	12	A
23	84.2	727-L	orig.	1	A
24	84.3	727-L	v_2t_1	16	S
25	85.1	727-L	2'd-eg	3	A
26	85.4	707-T	orig.	9	A
27	85.7	707-T	orig.-rp	10	A

2 3 4 5 6 7 8 9 10 11 12 13 14

*rp means repeated

**eg means example of an actual recording

Table XXI: Significant Differences Among 27 Noises at P<.01 Level Using Duncan's Multiple Range Test - EPNdB

RANK	EPNdB	A/C	DESIG.	ORIG. CODE NO.	ACTUAL OR STM.	1	2	3	4	5	6	7	8	9	10
1	74.5	727-T	v_2t_1	26	S	│									
2	75.6	727-T	v_1t_1	23	S	│	│								
3	76.2	727-T	c	27	S	│	│	│							
4	76.3	727-T	v_1t_1-*rp	24	S	│	│	│	│						
5	76.4	707-T	v_2t_1	21	S	│	│	│	│						
6	77.1	707-T	v_1t_1-rp	19	S		│	│	│	│					
7	77.2	707-T	v_1t_1	18	S		│	│	│	│	│				
8	77.9	707-T	v_1t_2	20	S			│	│	│	│	│			
9.5	78.0	727-T	orig.-rp	5	A			│	│	│	│	│			
9.5	78.0	727-T	orig.	6	A			│	│	│	│	│			
11	78.1	727-T	2'd-**eg	7	A			│	│	│	│	│			
12	78.2	727-T	v_1t_2	25	S			│	│	│	│	│			
13	78.2	707-T	c	22	S			│	│	│	│	│			
14	78.3	727-T	3'd-eg	8	A			│	│	│	│	│			
15	78.6	727-L	v_1t_1	13	S				│	│	│	│			
16	79.2	727-L	v_2t_1	16	S					│	│	│	│		
17	79.5	707-T	orig.	9	A						│	│	│		
18	80.0	707-T	orig.-rp	10	A							│	│		
19	80.1	727-L	v_1t_1-rp	14	S							│	│		
20	80.1	727-L	v_1t_2	15	S							│	│	│	
21	80.2	707-T	2'd-eg	11	A							│	│	│	
22	80.2	707-T	3'd-eg	12	A							│	│	│	
23	81.0	727-L	orig.-rp	2	A								│	│	│
24	81.1	727-L	3'd-eg	4	A								│	│	│
25	81.4	727-L	orig.	1	A								│	│	│
26	82.3	727-L	c	17	S									│	│
27	83.1	727-L	2'd-eg	3	A										│

*rp means repeated
**eg means example of an actual recording

Table XXII: Significant Differences Among 27 Noises at P<.01 Level Using Duncan's Multiple Range Test - Peak dBA

RANK	PK dBA	A/C	DESIG.	ORIG. CODE NO.	ACTUAL OR SIM.
1	63.2	727-L	v_1t_1	13	S
2	63.6	727-T	v_1t_1	23	S
3	64.5	727-T	v_1t_1-*rp	24	S
4	64.5	727-L	v_1t_1-rp	14	S
5	64.6	727-T	c	27	S
6	64.6	707-T	v_1t_1-rp	18	S
7	64.6	707-T	v_1t_1	19	S
8	65.0	727-T	orig.	5	A
9	65.2	727-T	orig.-rp	6	A
10	65.6	727-L	v_1t_2	15	S
11.5	65.9	727-T	v_2t_1	25	S
11.5	65.9	727-T	v_1t_2	26	S
13	66.4	727-T	3'd-**eg	8	A
14	66.5	707-T	v_1t_2	20	S
15	66.6	727-T	2'd-eg	7	A
16	67.1	707-T	c	21	S
17	67.1	707-T	v_2t_1	22	S
18	67.3	727-L	3'd-eg	4	A
19	67.4	727-L	v_2t_1	16	S
20	67.8	727-L	orig.-rp	2	A
21	68.1	727-L	orig.	1	A
22	68.4	727-L	c	17	S
23	69.8	727-L	2'd-eg	3	A
24	70.7	707-T	3'd-eg	12	A
25	71.1	707-T	2'd-eg	11	A
26	72.2	707-T	orig.	9	A
27	72.5	707-T	orig.-rp	10	A

*rp means repeated

**eg means example of an actual recording

TABLE XXIII: Significant Differences Among 27 Noises at P<.01 Level Using Duncan's Multiple Range Test - EdBH

RANK	EdBH	A/C	DESIG.	ORIG. CODE NO.	ACTUAL OR SIM.	1	2	3	4	5	6
1	63.8	727-L	3'd-*eg	4	A	│					
2	63.9	727-L	orig.-**rp	2	A	│					
3	64.3	727-L	orig.	1	A	│	│				
4	64.3	727-T	v_2t_1	26	S	│	│				
5	65.0	707-T	c	22	S	│	│	│			
6	65.2	707-T	v_1t_2	20	S	│	│	│			
7	65.6	727-T	v_1t_1	23	S	│	│	│	│		
8	65.6	707-T	v_2t_1	21	S	│	│	│	│		
9	65.8	707-T	orig.	9	A	│	│	│	│	│	
10	66.3	707-T	orig.-rp	10	A	│	│	│	│	│	│
11	66.4	727-T	c	27	S	│	│	│	│	│	│
12	66.6	707-T	v_1t_1-rp	19	S		│	│	│	│	│
13	66.6	707-T	v_1t_1	18	S		│	│	│	│	│
14	66.6	727-T	v_1t_1-rp	24	S		│	│	│	│	│
15	66.8	727-L	$v1t_2$	15	S		│	│	│	│	│
16	66.9	707-T	3'd-eg	12	A		│	│	│	│	│
17	66.9	707-T	2'd-eg	11	A		│	│	│	│	│
18	67.0	727-L	2'd-eg	3	A			│	│	│	│
19	67.9	727-T	3'd-eg	8	A				│	│	│
20	68.0	727-L	v_1t_1	13	S				│	│	│
21	68.0	727-T	orig.-rp	6	A				│	│	│
22	68.0	727-T	2'd-eg	7	A				│	│	│
23	68.1	727-T	orig.	5	A				│	│	│
24	68.5	726-T	v_1t_2	25	S					│	│
25	68.8	727-L	c	17	S						│
26	69.0	727-L	v_2t_1	16	S						│
27	69.0	727-L	v_1t_1-rp	14	S						│

*eg means example of an actual flyover
** rp means repeated flyover

of the same set for the subjective dB's based on actual recordings of a particular aircraft operation. These results are given in Table XXIV for the four engineering calculation procedures. Only results for 727-T (727 Takeoffs) show all examples of the four actual recordings with no reliable response differences across all four calculation procedures.

Table XXIV: Number of Four Actual Recordings By Aircraft Operation Which Show Comparable Response Using Duncan's Multiple Range Test

Aircraft Operation	(14) PNdBT(sets)	(10) EPNdB (sets)	(11) dBA (sets)	(6) EdBH (sets)
727-L	3 in Set 14 1 in Set 7	4 in Set 10 - - -	3 in Set 9 1 in Set 8	3 in Set 1 1 in Set 3
727-T	4 in Set 6 - - -	4 in Set 3 - - -	4 in Set 3 - - -	4 in Set 4 - - -
707-T	3 in Set 14 1 in Set 12	4 in Set 7 - - -	3 in Set 11 1 in set 10	4 in Set 4 - - -

The EPNdB calculation procedure does place all 4 actual recordings of each aircraft operation in the same set; both peak dBA and EdBH are very close to placing response to similar actual recordings in identical sets, while PNdBT is the least effective of the four procedures investigated. This is exemplified by the fact that three of the actual recordings for both the 727-L and 707-T are in Set 14 while the remaining actual recording is in Set 7 and Set 12 respectively. These results do show that there is a tendency for response to be similar to actual recordings of different examples of the same aircraft operation.

A comparison that is consistent with the basic aim of this study involves placement in the response rankings of an actual recording, its repeated recording, the two simulations of the actual recording ($v_1 t_1$), and the simulation of the actual recording which was designated "C" meaning a combination syntheses approach. These results are given in Table XXV.

Table XXV: Comparability of Response to Actual and Simulated Recordings Using Duncan's Multiple Range Test

	PNdBT	EPNdB	dBA	EdBH
727-L	Set 14(1,2,&17)* Set 2 (13&14)	Set 10(1,2,&17) Set 7 (13 & 14)	Set 9(1,2,&17) Set 1 (13&14)	Set 1(1&2) Set 6(13,14 & 17)
727-T	Set 3(5,6,23, 24 & 27)	Set 3(5,6,24 & 27) Set 2(23)	Set 2(5,6,23, 24 & 27)	Set 4(5,6,23, 24 & 27)
707-T	Set 14(9 & 10) Set 1(18,19,22)	Set 7(9,10,22) Set 5(18 & 19)	Set 11 (9 & 10) Set 8(22) Set 3 (18 & 19)	Set 3 (9,10,18, 19 & 22)

* Numbers are original code numbers for the 27 flyover signals (See Table III).

The original flyover code numbers are shown along with the main statistical set of which the subjective dB's are members. For the 727-T flyover, all five subjective dB's with the minor exception of No. 23 (original v_1t_1) using EPNdB, are members of the same sets. Results based on the 727-L recording show that No. 17 (Combination or "C" simulation) is the same set as the two actual recordings for three of the four calculation procedures while for the 707-T flyover, No. 22 (Combination or "C" simulation) is in the same set for two calculation procedures (EPNdB and EdBH). On the whole, the combination synthesis is judged as being more like the actual recordings than the synthesized recordings designated as v_1t_1.

DURATION AND TONE CORRECTION EFFECTS ON SIMULATIONS

Since both duration and tone were separately increased for simulation examples, there is an opportunity to examine the effectiveness of the FAR-36 duration and tone correction procedures. Using the duration correction as an example, this involves comparison of subjective dB's between logical pairs of v_2t_1 (velocity decreased by a factor of 2.5) and v_1t_1 (simulation with no aim of altering duration). Table XXVI provides the results of these comparisons. For each of the flyovers investigated, the first column provides v_2t_1 subjective dB less v_1t_1 subjective dB with no correction for duration while the second column gives this difference with the duration correction. If the duration correction were to function perfectly, all values in the second or corrected column for each flyover would be zero and in any event would be appreciably less then the corresponding difference in the first column. Although the FAR-36 duration correction does overcorrect in every case but one (dBHT differences with duration correction for 727-L is +0.53), on the whole, there is pronounced improvement for the corrected over the uncorrected differences. With the exception of duration corrections for the 727-T flyover, using dBA and dBAT, correction for duration of the flyovers provides improvement in agreement between calculated and judged values for all the remaining comparisons. The greatest amount of improvement between calculated and judged levels is 5.36 dBH for the 727-L flyover while there is a decrease of 0.32 dBAT between calculated and judged levels for 727-T. The average improvement between calculated and judged levels is 3.8 dB for 727-L, 1.8 dB for 707-T and 0.8 dB for 727-T. The duration correction, clearly, has the greatest effectiveness when applied to the dBH procedures; improvement in agreement between calculated and judged levels ranges from 2.0 dBH to 5.4 dBH.

Table XXVI: Duration Correction Effects on Simulations -
Subjective dB for v_2t_1 minus average subjective dB for v_1t_1

	727-L		707-T		727-T	
UNIT	uncorrected	duration corrected	uncorrected	duration corrected	uncorrected	duration corrected
PNdB	3.71	-0.52	2.52	-0.35	2.45	-1.58
PNdB$_T$	4.16	-0.18	2.05	-0.75	2.28	-1.42
dBA	3.56	-1.29	2.45	-0.92	1.96	-2.43
dBA$_T$	3.75	-0.67	1.96	-1.36	1.98	-2.30
dBH	5.44	-0.08	3.72	-0.63	3.75	-1.77
dBH$_T$	5.68	0.53	3.18	-1.02	3.77	-1.74

Table XXVII: Tone Correction Effects on Simulations -
Subjective dB for v_1t_2 minus average subjective dB for v_1t_1

	727-L		707-T		727-T	
UNIT	uncorrected	tone corrected	uncorrected	tone corrected	uncorrected	tone corrected
PNdB	1.41	0.83	1.31	0.82	2.02	2.32
PNdB$_D$	1.19	0.74	1.17	0.74	2.49	2.20
dBA	1.78	1.21	1.88	1.43	1.96	2.19
dBA$_D$	1.68	1.49	1.78	1.42	2.16	1.88
dBH	-0.97	-1.83	-1.27	-1.97	3.34	3.71
dBH$_D$	-1.37	-1.72	-1.33	-1.46	2.84	2.41

Results for the FAR-36 tone correction, comparable to those for the duration correction, are given in Table XXVII. It is quite clear that the FAR-36 tone correction is not particularly effective in increasing the agreement between calculated and judged values. The average increase across all calculation procedures for 727-L is a slight 0.1 dB, 0.2 dB for 707-T, and for 727-T there is a small decrease in agreement between calculated and judged levels of 0.1 dB. The tone correction almost always degrades agreement for dBH based procedures which does contribute disproportionally to absence of any significant improvement. However, the greatest increase in agreement is 0.6 PNdB and 0.6 dBA for 727-L which is not perceptable or measurable in airport community noise exposure situations. Excluding the dBH calculation procedures, the tone correction ranges from a 0.30 PNdB decrease in agreement between calculated and judged levels to a 0.58 PNdB increase in agreement. One last comment concerning the FAR-36 tone correction involves the "slope" method for identifying a pure tone. For 727-T which showed a negative effect from using the tone correction, this flyover showed the least "tone" of the three flyovers (see Table I) and the 1/3-octave bands to which tones were added do not correspond to those which the method identified as requiring correction. There is certainly a need for improvement in the state-of-the-art relative to tone correction technology.

COMPARISONS CONCERNING ACCURACY OF EPNdB CALCULATION PROCEDURE

Whether-or-not two different noise signals which are calculated to be at the same level produce comparable annoyance or loudness response in receivers is a function of the accuracy of the engineering calculation procedure employed. Since no engineering calculation procedure is perfectly valid for a diverse group of noise signals, how great can the annoyance response difference be to two or more signals with identical calculated levels and continue to support a conclusion that response is comparable or similar? Data relevant to this particular question are not readily available since the usual emphasis in laboratory studies involving human response to noise studies has been on comparisons of the accuracy or validity of various calculation procedures. Somewhat comparable results relative to the accuracy of the widely used EPNdB calculation procedure are available from this study and two previous studies (Refs 6 & 7). A summary of the significant parameters for making these comparisons is given in Table XXVIII.

Table XXVIII: Comparison of Results of This Study to Comparable Studies Concerning Accuracy of the EPNdB Calculation Procedure

Study	Mean EPNdB Diff.	Range of Diff.	Range of Subj. dB	No. of Sets	No. of Noises	No. of Subjects	Error Mean Square	df for Error
Present	2.5	2.3-2.9	8.6	10	27	60	18.2	1404
STOL (Ref 6)	2.9	2.5-3.2	12.9	15	34	35	13.1	1122
Helicopter (Ref 7)	3.6	3.3-4.2	8.4	7	24	24	18.1	529

The basic data are provided in the first column under "Mean EPNdB Diff". This mean difference is based on the multiple range test (Table XXI for EPNdB of present study) with each difference being the least difference between mean subjective dB's which provides a reliable difference at the $P < .01$ point. The range of these statistically significant differences is given in the second column while the third column provides the range of mean subjective dB for each of the three studies. Since the mean EPNdB difference of column 1 is very much a function of the number of statistical sets obtained, number of noises, number of subjects, error term and degrees of freedom, this information is also provided. Refering to the results for the present study in columns 1 and 2, we would conclude that for the present study, EPNdB has an accuracy of 2.3 to 2.9 dB with an average accuracy of 2.5 EPNdB. This means that if two different signals are calculated at identical EPNdB levels, annoyance response could differ by approximately 2.5 EPNdB but the annoyance response to these two different signals would still be considered comparable or not different in magnitude. Somewhat different results were obtained for the STOL study of Reference 7 where the differences in annoyance response averages a 2.9 EPNdB difference for arriving at a conclusion that annoyance response is comparable to two signals which are calculated at identical levels. For the Helicopter study of Reference 8, the mean difference is increased to 3.6 EPNdB. This increase could be attributable to the diversity of noise signals investigated (actual recordings of helicopter noise, simulations of helicopter noise, propeller aircraft, commercial jet aircraft, and turboprop aircraft); however, increasing the number of subjects by a factor of 2.5 would also have provided a mean difference approximately equal to that of the present study. Utilizing the results of this study in conjunction with those from the two comparable studies, it is concluded that the accuracy of the EPNdB is in the neighborhood of 2.5 EPNdB. Thusly, a noise signature for a next generation aircraft could be predicted, synthesized and calculated; annoyance response could be 2 to 3 EPNdB greater than the calculated level based on actual flight of the aircraft; and it would be concluded that annoyance response is comparable to the synthesized and to the actual recording.

A final consideration involves the comparative accuracy of various engineering calculation procedures in conjunction with "annoyance" accuracy. The premise basic to evaluation and selection of engineering calculation procedures is that the most valid engineering calculation procedure is one which results in the greatest agreement between ratings of a diverse group of noise signals and a particular engineering calculation procedure. Also, this premise can be put in terms of the least amount of disagreement between noise ratings and the calculation procedure. Using the methods of the present study, what is an estimate of the "annoyance" accuracy of other engineering calculation procedures? Results are available for comparison using EdBH (S.D. for means of 1.6 dBH, Table VI) and Peak dBA with a S.D. of 2.5 dBA. Thusly EdBH would be considered the most valid procedure, EPNdB the next most valid (S.D. of 2.1 EPNdB, Table VI), and Peak dBA the least valid of those procedures. Using the multiple range results for EdBH and Peak dBA, along with that from Table XXVIII for EPNdB, provides the following mean least differences between the mean subjective dB's that are reliably different at the $P < .01$ point.

Calculation Procedure	Mean Least Differences	S.D. of Subjective dB Means
EdBH	2.8 EdBH	1.6 EdBH
EPNdB	2.5 EPNdB	2.1 EPNdB
Peak dBA	2.2 dBA	2.5 dBA

A very tentative conclusion that could be drawn from the above is, "annoyance" accuracy for a particular procedure can have a greater tolerance if validity is higher. Results for EdBH and Peak dBA do agree with the previous conclusion which is: annoyance response to two different signals which are calculated at the same level can differ by 2 to 3 dB and it is concluded that response to the two signals is comparable.

CONCLUSIONS

Although the main aim of this study involved comparisons between synthesized and actual recordings of aircraft flyovers, conclusions relative to a number of supporting or ancillary problems were also developed. Thusly, there are two parts to this section. The first part is concerned with the main aim involving response to synthesized and actual recordings while the second covers conclusions relative to ancillary problems involving human response to aircraft noise investigations.

MAIN AIM CONCLUSION

(1) Annoyance response is comparable to the two sets of recordings; i.e., the synthesized recordings and the actual recordings. This is a conclusion relative to the question, "Do persons perceive and respond to the synthesized recordings in the same manner as to actual recordings?"

(2) There is a requirement to obtain greater correspondence between 1/3-octave band levels of synthesized and actual recordings. The combination or "C" synthesis approach which has greater 1/3-octave band similarity to the actual recordings than the v_1t_1 synthesis has a much greater tendency to be judged similar to the actual recordings than the v_1t_1 synthesis (See Table XXV).

(3) There is also a requirement to obtain higher agreement between effective durations for the synthesized and actual recordings. Regardless of the engineering calculation procedure investigated, response was the most comparable for all synthesized and actual recordings for the 727 take-off (Table XXV). Effective duration was underdetermined for all simulations of the 727 landing and 707 take-off while it was somewhat too great for the

727 take-off which produced similarity of response to synthesized and actual recordings.

(4) An evaluation of a synthesized approach requires an engineering calculation procedure which is accurate or valid for the flyover signals under consideration. On the surface, this appears as an obvious but difficult to implement conclusion for any evaluation of diverse flyover signals. However, it is an essential conclusion. If only state-of-the-art engineering calculation procedures (PNdB, dBA, and their variants) were to have been used in this study, it would have been concluded that there was something about the synthesized recordings which resulted in their being less annoying than the actual recordings. The fact that the most valid engineering calculation procedure (EdBH) did distribute the 15 simulated and 12 actual recordings into "high annoyance" sets in a manner that is similar to previous investigations (Table XVII and XXIII), supports a conclusion that these state-of-the-art procedures are not sufficiently accurate or valid for evaluating the flyover signals under investigation. Without the utilization of the more valid engineering calculation procedure, it would have been concluded that there was some unknown and perhaps not detectable defect in the synthesizer.

(5) That the synthesizer cannot simulate the comb-filter effect discussed under "SYNTHESIS" or the less smooth time history of the actual recordings are in no manner serious short comings.

ANCILLARY CONCLUSIONS

The aim here is to emphasize results which contribute to or raise questions concerning human response to noise technology and/or application problems involving aircraft noise assessment.

(1) It is concluded that a correction for effective duration is an essential element for obtaining a valid measurement of annoyance response to aircraft noise. There was strong evidence that the duration of the signal very much influenced annoyance response. However, since FAR-36 duration correction effectiveness varied as a function of both calculation procedures and flyover signals, other methods of measuring effective duration could be considered.

(2) As shown by the results of Table XXVII, it can be concluded that the FAR-36 correction for "tone" is not effective. Since a number of previous studies have provided similar results, a serious review of this area merits consideration. It is quite likely that either a correction for "tone" is not required to provide a valid measure of aircraft noise annoyance or that the method for identifying and weighting "tone" effects requires improvement.

(3) Results show that the dBH (Ref. 3) engineering calculation procedure was unusually effective in conjunction with the FAR-36 duration correction. That this experimental procedure was more accurate than the two state-of-the-art procedures, suggests that it could be worthwhile to complete further research aimed at simplifying and improving the state-of-the-art in this area.

(4) It is concluded that the annoyance accuracy of current engineering calculation procedures is in the 2 to 3 dB range. As an example, two different noise signatures might be calculated at exactly the same level utilizing dBA but annoyance response (Subjective dB) to one signal is but 2 dBA greater than to the other signal. It would be concluded that annoyance response to these two signals is identical and that the dBA calculation procedure validly evaluates both signals. However, a difference of 3 dBA in annoyance response to two signals calculated at the same level, could lead to a conclusion that the dBA engineering calculation procedure does not accurately evaluate response

to these two aircraft flyover signals.

(5) It is concluded that the magnitude estimation method is a highly reliable method for evaluating engineering calculation procedures. One of the more remarkable results was the high agreement of response to repeated presentations of flyover signals. Agreement was almost perfect for repeated presentations of the actual recordings (differences ranging from 0.0 to 0.5 dB) and quite satisfactory for pairs of simulations (differences ranging from 0.0 to 1.7 dB).

REFERENCES

1. Stevens, S. S. "On the psychophysical law," Psycho. Rev., Vol. 64, No. 3, 1957.

2. Stevens, S. S. "Problems and method of psychophysics," Psycho. Bul., Vol. 55, No. 4, 1958.

3. *Higgins, T. H. "Human responses to sound: the calculation of perceived level, PLdB (noisiness or loudness) directly from physical measures," Report No. FAA-RD-76-1, Oct. 1976.

4. Mabry, J. E. (For Society of Automotive Engineers) "An evaluation of psychoacoustic procedures for determining human response to aircraft noise, Vol. I - Specifications for four experiments," Report No. FAA-RD-72-51, I, Oct. 1973.

5. Edwards, A. L. "Experimental design in psychological research," Holt, Rinehart and Winston, New York, 1960.

6. MAN-Acoustics & Noise, Inc. "Noise certification criteria and implementation considerations for V/STOL aircraft," FAA-RD-75-190, Nov. 1975.

7. MAN-Acoustics & Noise, Inc. "Noise certification considerations for helicopters based on laboratory investigations," FAA-RD-76-116, July 1976.

* The decision to utilize a second straight line segment at at 4 kHz with a negative slope of 6 dB resulted from a personal communication with T. H. Higgins.