NASA Technical Paper 1386

Effects of Duration and Other Noise Characteristics on the Annoyance Caused by Aircraft-Flyover Noise

David A. McCurdy and Clemans A. Powell

MARCH 1979

NASA

NASA Technical Paper 1386

Effects of Duration and Other Noise Characteristics on the Annoyance Caused by Aircraft-Flyover Noise

David A. McCurdy and Clemans A. Powell
Langley Research Center
Hampton, Virginia

National Aeronautics
and Space Administration

**Scientific and Technical
Information Office**

1979

SUMMARY

A laboratory experiment was conducted to determine the effects of duration and other noise characteristics on the annoyance caused by aircraft-flyover noise. A newly developed aircraft-noise synthesis system was used to synthesize 54 aircraft-flyover noise stimuli representing the factorial combinations of 3 durations, 3 aircraft velocities, 3 sound pressure levels, and 2 tone conditions. Forty-eight test subjects made annoyance judgments on the test stimuli in a subjective test facility simulating an outdoor acoustic environment. The judgments were made by using a graphical scale procedure similar to numerical category scaling.

Statistical analyses comparing the subjective judgments with the acoustical characteristics of the stimuli in terms of several rating scales were conducted to determine the effects of duration on annoyance and the appropriate duration correction. The effects of tonal content and Doppler shift were also studied.

A duration correction with a magnitude of 3 dB per doubling of effective duration, as used in the effective-perceived-noise-level procedure, was found to account most accurately for the effect of duration and resulted in a significant improvement in the annoyance-prediction ability of the rating scales. Current tone-correction procedures did not adequately account for the effects of the interaction of tonal content with sound pressure level. No significant effect of Doppler shift was found.

INTRODUCTION

The question of how the duration of an aircraft-flyover noise affects the annoyance of an observer on the ground has been examined in a number of previous studies. These studies, however, have yielded widely varying conclusions concerning the existence and magnitude of such an effect. For example, reference 1 indicates that annoyance increases as duration increases and requires a correction to the noise rating scales. The magnitude of the correction was found to vary as a function of duration from 2 dB to 6 dB per doubling of duration. Reference 2 indicates that annoyance increases as duration increases and recommends a constant duration correction of 3 dB per doubling of effective duration. (The effective duration is defined as the duration of a continuous-level signal with energy equal to the energy contained in the flyover-noise signal. The energy contained in the flyover signal is based on the numerical integration of energy between the first and last points at which the flyover signal is 10 dB down from the maximum sound level.) Reference 3 concludes that duration does not affect annoyance and no correction is needed. Although regulatory agencies such as the Federal Aviation Administration have adopted a duration correction of 3 dB per doubling of effective duration, recent studies, such as that reported in reference 4, raise doubt concerning both the magnitude and methodology of the appropriate correction.

A review of previous duration studies identifies three factors that appear to be the sources of many of the differences in results. As a first factor, many of the studies (refs. 1 and 5 to 8) used simulated aircraft-flyover noises as test stimuli. The simulated noises ranged from octave bands of broadband noise to jet-engine test-stand noise. These stimuli often had time histories and spectra that were unrealistic and not representative of real aircraft.

The second factor involves studies in which recordings of real aircraft were used as test stimuli (refs. 2, 4, 7, 8, and 9). In order to obtain recordings of real aircraft noises with a wide range of durations it is necessary to vary the distance from the take-off/landing threshold at which the recordings are made. The effects of this change in distance on duration, spectral content, and Doppler shift are illustrated in figure 1. For a given aircraft-noise source, figure 1 shows the relative values of duration, spectra, and Doppler shift under the flight path at three different distances from threshold. When using recordings of real aircraft in duration studies, long-duration stimuli will have less high-frequency noise and a lower rate of change in frequency than stimuli of shorter duration. As a consequence of this inability to vary duration and other noise parameters independently, any effects on annoyance attributed to duration may, in fact, be caused by changes in spectral content or Doppler shift.

The third factor causing differences in the results of duration studies is variations in experiment design and methodology. The foremost example of this factor is found in reference 5. Studies in which subjects received instructions associating the word "duration" with the annoyance of a noise produced results indicating a duration effect of greater magnitude than the duration effect indicated by the results of studies in which the instructions contained no mention of duration. Two opposing explanations for this variation are suggested. First, in experiments where a duration cue is given, subjects may tend to rank noise stimuli by duration and thereby give too much emphasis to duration in making their annoyance judgments. The opposite view is that, in experiments where no duration cue is given, subjects tend to de-emphasize duration in their annoyance judgments by ranking noise stimuli according to peak intensity levels and by making their judgments prior to the end of the noise stimuli.

The purpose of the study reported herein was to determine the effects of duration on the annoyance caused by aircraft-flyover noise in such a manner as to avoid the factors that appear to have affected the results of previous studies. To avoid the problems associated with recordings of simulated and real aircraft noise, a newly developed aircraft-noise synthesis system was used to generate a set of simulated aircraft-flyover noises in which duration, tonal content, aircraft velocity, and sound pressure level were independently controlled and individually varied. These synthesized noise stimuli had time histories and spectra representative of real aircraft and sounded very similar to real aircraft noise. To prevent instruction bias, a short tone or beep audio cue was placed at the end of each test stimulus and the subjects were instructed to wait until they heard the audio cue before making their annoyance judgments. No mention of duration was made in the instructions given the subjects. The purpose of this method was to insure that the subjects' judgments were based on the entire stimulus noise but were not biased by the specific mention of duration.

In this study, 48 test subjects made judgments of the annoyance of a set of 54 synthesized aircraft-noise stimuli. The stimuli set included combinations of duration, tonal content, aircraft velocity, and sound pressure level that encompassed the range of noise parameters associated with V/STOL, SST, and CTOL (conventional take-off and landing) aircraft. Statistical analyses of the subjective judgments address the effects of duration, tonal content, aircraft velocity, sound pressure level, and the interactions of these parameters. The effects of Doppler shift are also addressed. The ability of several noise rating scales to predict annoyance and the effect of duration corrections on that predictive ability are examined.

SYMBOLS AND ABBREVIATIONS

The following rating scales have been used in the acoustical analysis of the aircraft noises used in this study. Additional descriptive information concerning frequency weightings and computational procedures can be found in references 10 and 11.

EPNL effective perceived noise level (equivalent to IPNLT), EPNdB

L_A A-weighted sound pressure level, based on the 1/3-octave bands from 50 Hz to 10 kHz, dB

L_D D-weighted sound pressure level, based on the 1/3-octave bands from 50 Hz to 10 kHz, dB

OASPL overall sound pressure level, dB

PL_1 perceived level, according to the Stevens Mark VII procedure, PLdB

PL_2 perceived level, calculated as suggested by Higgins in reference 11, PLdB

PNL perceived noise level, PNdB

PNLT tone-corrected perceived noise level (FAR 36 procedure), PNdB

The use of the capital letter "I" preceding the abbreviations of the rating scales other than EPNL (e.g., IL_A, IOASPL, IPL_2, and IPNLT) denotes the addition of a duration correction to the calculation procedure. The correction procedure used is the same as that incorporated in the EPNL calculations and has a magnitude of 3 dB per doubling of effective duration. Effective duration is defined as the duration of a continuous-level signal with energy equal to the energy contained in the flyover-noise signal. The energy contained in the flyover signal is based on the numerical integration of energy between the first and last points at which the flyover signal is 10 dB down from the maximum sound level.

Other abbreviations and symbols used herein are as follows:

ANSI American National Standards Institute

a_0, a_1, a_2 constant coefficients

$\left.\begin{array}{l} b_0, b_1, b_2 \\ c_0, c_1, c_2, c_3 \\ d_0, d_1, d_2, d_3 \\ e_0, e_1, e_2, e_3 \end{array}\right\}$ regression coefficients

CTOL conventional take-off and landing

c speed of sound, 340 m/sec

D overall duration correction based on a duration-correction magnitude of 3 dB per doubling of effective duration, dB

D' overall duration correction based on the optimum duration-correction magnitude expressed in terms of decibels per doubling of effective duration, dB

FAR Federal Aviation Regulation

f frequency, Hz

f_s source frequency, Hz

h distance of closest approach of aircraft to ground observer, m

J subjective annoyance response

L rating-scale level, dB

r Pearson product-moment correlation coefficient

SPL sound pressure level, dB

SST supersonic transport

t duration, sec

V/STOL vertical or short take-off and landing

v aircraft velocity, m/sec

β Doppler shift parameter, v^2/ch, sec^{-1}

ϕ angle of departure

angle formed at the observer location between the aircraft-noise source and the ground track of the flight path such that $\theta < 90^{\circ}$ during the approach of the source

Subscript:

max   maximum

EXPERIMENTAL METHOD

Noise Stimuli

Aircraft-noise synthesizer.- A newly developed aircraft-noise synthesis system was used to generate the noise stimuli used in this study. Program input consists of aircraft flight parameters and acoustical reference parameters. The flight parameters include aircraft velocity, angle of approach or departure, and the listener's ground location in terms of sideline and downrange distances from the point of lift-off or touchdown. The acoustical reference parameters are comprised of descriptive reference spectra, consisting of broadband and narrowband noise and harmonic tones, and of associated directivity patterns. The input parameters are used for computations of time-varying aircraft position and narrowband random noise, which include the appropriate broadband and narrowband components, Doppler shift, directivity, and atmospheric effects.

A block diagram of the system is shown in figure 2. The system uses a general-purpose computer for computations used in the generation of digital representations of the predicted aircraft-noise waveform. The keyboard and printer and the paper tape reader and punch are used for operator control and data input. The system computes an updated spectrum for each 0.08-sec increment of the flyover noise. These spectra are converted from the frequency domain into digital waveforms in the time domain by inverse Fourier transformation. The digital waveforms are then stored sequentially on magnetic discs. Upon completing the generation and storage of the digital waveforms for the entire noise stimulus, the waveforms are read into the buffer memory of the computer and through a 12-bit digital-to-analog converter at a constant rate. The overall time-dependent amplitude of the analog signal is controlled by a programmable attenuator in order to maintain the full dynamic range throughout the simulated flyover noise. The audio system consists of an analog tape recorder and an amplifier and loudspeaker for monitoring the generated flyover noise.

Test stimuli.- The stimuli used in this experiment consisted of loudspeaker-reproduced tape recordings of 54 synthesized aircraft-flyover noises in which duration, Doppler shift, and tonal content were individually controlled by specifying aircraft velocity, altitude, and reference spectra. As shown in figure 3, the noises represented the 54 factorial combinations of 3 durations, 3 velocities, 3 sound pressure levels, and 2 tone conditions. Based on the A-weighted sound pressure level, the 10-dB down durations of the stimuli were 10, 20, and 40 sec. Typical time histories for stimuli of each duration are presented in figure 4. The aircraft velocities were 40, 80, and

160 m/sec. This range of velocities brackets the typical velocities of STOL, CTOL, and SST aircraft. The combinations of aircraft altitude and velocity used in the synthesizer program to obtain the desired nine factorial combinations of duration and velocity yielded five different patterns of Doppler shift as shown in table I. During the preparation of the presentation tapes, the three sound pressure levels of each combination of duration, velocity, and spectrum were manually set at intervals of approximately 9 dB. The broadband content of both spectra was based on that of a B-727 departure. One spectrum contained no tonal components and the other spectrum contained strong tonal components centered at 1100 Hz and 2200 Hz. The acoustical reference input parameters were adjusted so that each synthesized stimulus had the same broadband spectrum at the time of maximum A-weighted sound pressure level. Figures 5(a) and 5(b) show the range and mean of each 1/3-octave-band sound pressure level occurring at and normalized by the maximum A-weighted level for the group of stimuli without tones and the group of stimuli with tones, respectively. Comparison of the two figures shows that the broadband portion of the spectra compares very well across groups and within groups of stimuli. From these figures the effectiveness of the synthesis system in holding the spectra constant across 4-to-1 ratios of both duration and velocity and five patterns of Doppler shift is apparent.

Test Subjects

The 48 subjects used in the experiment were randomly selected from a pool of local residents with a wide range of socioeconomic backgrounds and were paid to participate in the experiments. All the subjects had previously participated in aircraft-noise-related experiments. All test subjects were given audiograms prior to the experiment to verify normal hearing within 20 dB (ANSI 1969). Table II gives the sex and age data for the subjects.

Reproduction System and Test Facility

Audio reproduction system.- A diagram of the basic noise reproduction system is shown in figure 6. The monophonic recordings of the synthesized aircraft-noise stimuli were played back on a studio-quality tape recorder. A commercially available noise-reduction system which provided a nominal 30-dB increase in signal-to-noise ratio was used to reduce tape hiss to inaudible levels.

Test facility.- The exterior effects room (EER) of the Langley aircraft noise reduction laboratory was used as the test facility in the experiment. This room has seating for 39 subjects and a volume of approximately 340 m^3. The reverberation time for the room is approximately 0.5 sec at 1000 Hz. The stimuli were presented by means of six overhead loudspeakers. The 4 seating locations used by the subjects during each of the 12 test sessions are indicated in figure 7.

Experiment Design

A procedure similar to that of numerical category scaling which incorporates a graphical scale was chosen as the psychophysical method for the experiment described in this report. The choice of a method similar to numerical category scaling was made primarily to conserve test time and allow test subjects to make as many judgments as possible during a fixed length of time. The use of a graphical scale on which subjects could indicate judgments between numbered points was made in order to provide the subjects with more flexibility in making their judgments and possibly decrease the variability in judgments of an individual stimulus. The scale selected was a unipolar, 10-point scale from 0 to 9. The end points of the scale were labeled "Not at all Annoying" and "Extremely Annoying." The term "ANNOYING" was defined in the subject instructions as "UNWANTED, OBJECTIONABLE, DISTURBING, or UNPLEASANT."

Two sets of three tape recordings of the various stimuli were prepared for presentation to the subjects. The first set consisted of tapes I, II, and III and the second set consisted of tapes IV, V, and VI. Tapes IV, V, and VI contained the same stimuli as tapes I, II, and III, but in reverse order. The order of the stimuli on each tape is given in table III. The particular orders were based on random selection with two constraints to provide some measure of balance. The first constraint was that each of the three durations, three aircraft velocities, three sound pressure levels, and two tone conditions should occur an equal number of times in each tape. The second constraint was that each of the three levels should occur once in succeeding groups of three stimuli, starting at the beginning of a tape. A period of 6 sec was provided between stimuli for the subjects to make and record their judgments. Each tape recording required approximately 25 min for playback and served as a test session for the subjects.

The subjects were divided into 12 groups of 4 subjects. Each group was presented one of the two sets of three tapes followed by a fourth tape from the other set. The fourth tape contained the same stimuli as the first tape, but in reverse order. The tape recordings were presented to each group in a different order, as shown in table IV. The purpose of changing the tape order was to provide a balanced presentation to prevent subject fatigue or other temporal effects from unduly influencing the results.

Procedure

Upon arrival at the laboratory, the subject groups were seated in a conference room and given a set of instruction sheets, a consent form, a practice scoring sheet, and a set of scoring sheets. Copies of these items are given in the appendix. After reading the instruction sheets, the subjects completed the consent form which is required of all subjects who participate in subjective experiments in the laboratory. The subjects were given a brief verbal explanation of the scoring sheets and then asked by the test conductor if they had any questions about the test. Throughout the experiment, the same person served as the test conductor.

The subjects were then ushered by the test conductor into the test facility and seated according to subject numbers, which were randomly assigned in the conference room. A demonstration of three practice stimuli, listed in table III, was given while the test conductor remained in the test facility. In order for subjects to gain experience in scoring the sounds, they were instructed to make and record judgments of the practice stimuli on the practice scoring sheet. Afterwards, the test conductor again asked if there were any questions concerning the test. The test conductor left the facility and the first of four test sessions began. After the conclusion of each 25-min session the test conductor reentered the test facility, collected the scoring sheets, and issued new sheets for the next session. Between the second and third sessions, the subjects were given a 15-min rest period outside the test facility.

RESULTS AND DISCUSSION

Acoustic Data Reduction

The stimuli were measured, with no subjects present, at the head position of the subject pictured in figure 7 in the first row to the reader's right. A 1/3-octave-band analysis of the measurements (analog filtering with digital sampling, root-mean-square detection, and integration) was used to provide time histories for computations required by the rating scales. The frequency range for analysis was band limited from 22.5 Hz to 22.5 kHz. For all scales other than OASPL, the frequency range was 50 Hz to 10 kHz and the values were calculated from the measured 1/3-octave-band levels.

Maximum levels and duration-corrected levels were obtained for each stimulus for each rating scale. The duration-corrected levels were calculated by using a correction of 3 dB per doubling of effective duration, as described in the section entitled "Symbols and Abbreviations." The maximum and duration-corrected levels for the highest level stimulus of each combination of duration, velocity, and tonal content are given in table V. The values of the first six rating scales were calculated as specified in reference 10. The PL_2 calculations were based on a new procedure recommended by Higgins in reference 11.

Subjective Data Reduction

The mean values of the judgments were calculated for each of the 54 stimuli. These mean values were used as the subjective scores for the stimuli in the various regression analyses involving the rating scales.

Reliability of Subjective Judgments

Each subject judged 72 noise stimuli. The first 54 stimuli judged consisted of the complete set of 54 test stimuli. The last 18 stimuli judged were repeats of the first 18 stimuli in reverse order. Regression analyses, the results of which are given in table VI, were performed on these repeated judgments in two ways. The first was a regression of each individual subject's second judgment (dependent variable) on his first judgment (independent variable) for each stimulus. The second was a regression of the mean (over subjects) of the second judgments on the first judgments for each of the 54 stimuli. The Pearson product-moment correlation coefficients for the two regression analyses were 0.722 and 0.961, respectively.

Analysis of Variance

In order to determine if the main parameters of duration, velocity, sound pressure level, and tonal content and/or their interactions affected annoyance, an analysis of variance was performed on the subjective judgments. No attempt was made to distinguish between first judgments and repeated judgments. For each of the 54 stimuli there were 64 judgments, 16 of which were repeats. The results of the analysis of variance are given in table VII. Of the four main parameters, only duration, sound pressure level, and tonal content were found to be significant at the 0.01 level. The velocity parameter was not significant at the 0.05 level. Of the 11 possible interactions, only the following 2 were found to be significant at a level of 0.01: duration with level and tonal content with level. No other interactions were significant at the 0.05 level.

Figure 8 illustrates the effects of duration, sound pressure level, and the interaction of duration with level. The figure shows the relationships between the mean annoyance rating and L_A for each of the three stimuli durations. "Mean annoyance rating" is the average of the subjective annoyance judgments of all the stimuli having the combination of parameters specified. From figure 8 it is seen that increased duration causes increased annoyance. For a given level, the magnitude of the change in annoyance resulting from a doubling of duration remains relatively constant. This result is in agreement with the general practice of equating the increase in annoyance due to duration with an increase in maximum level and expressing a duration correction in terms of a constant number of decibels per doubling of duration. Although the analysis of variance indicated a significant interaction of duration with level, no consistent effect is apparent in figure 8. The analysis of variance method assumes that for a given level condition, the level is constant across duration. In actuality, as can be seen from figure 8, there were small variations in level across duration. These variations, and not a real interaction, are believed to be responsible for the analysis of variance result. This conclusion is supported by the application of an analysis of covariance technique to the data. The results of this analysis method are discussed in a subsequent section.

The effects of tonal content and the interaction of tonal content with level are illustrated in figure 9. The mean annoyance rating is plotted against L_A for the stimuli without tonal components and the stimuli with tonal components. As is generally accepted, the stimuli with tonal components are more annoying than stimuli with no tones. Of more interest, perhaps, is the interaction of tonal content with level. As level increases, the difference in annoyance between the stimuli with and without tones decreases. The effects illustrated in figures 8 and 9 will be discussed further in the following sections.

Effects of Duration on Annoyance

Having confirmed the existence of an effect on annoyance resulting from a parameter or interactions of parameters, the question arises as to how to include that effect in the methods used to predict annoyance to noise. Presently, the most common method of correcting for duration is the one incorporated in the EPNL calculation procedure used by the Federal Aviation Administration in aircraft certification (ref. 10). The magnitude of the correction is equivalent to 3 dB per doubling of effective duration. To determine the optimum magnitude of the duration correction in terms of decibels per doubling of effective duration the following analyses were performed.

If the duration corrections based on the optimum magnitude are applied to the maximum levels of the stimuli, the subjective annoyance judgments can be represented by the linear equation

$$J = a_0 + a_1(L_{max} + D') \tag{1}$$

where J is the subjective annoyance judgment, L_{max} is the maximum level, and D' is the optimum duration correction. This equation can be expanded to the form

$$J = a_0 + a_1 L_{max} + a_1 D' \tag{2}$$

However, if the duration corrections are calculated by using a nonoptimum magnitude and if the maximum levels and durations are not correlated, the equation best fitting the data would be of the form

$$J = a_0 + a_1 L_{max} + a_2 D \tag{3}$$

where a_1 is not equal to a_2 and D is the nonoptimum correction. Combining equations (2) and (3) yields

$$a_1 D' = a_2 D \tag{4}$$

which gives

$$D' = \frac{a_2}{a_1} D \tag{5}$$

Duration corrections based on 3 dB per doubling of effective duration (i.e., the difference between the duration-corrected level and the respective maximum level for each rating scale) were used in multiple regression analyses of the form of equation (3). The optimum duration-correction magnitudes were then calculated from equation (5) with D set equal to 3 dB per doubling of effective duration. These calculations were made for each of the rating scales for stimuli without tones, stimuli with tones, and all stimuli combined. The resulting optimum magnitudes, in terms of equivalent decibels per doubling of effective duration, are given in table VIII. In general, the optimum magnitudes are very near 3 dB, with over half being within ±0.25 dB and over 85 percent within ±0.50 dB. The agreement is even better when only the subgroups of stimuli without tones and stimuli with tones are considered.

The correlation coefficients of the regressions of the mean subjective judgments on the rating scales with optimum duration corrections were compared with the coefficients of the regressions of the mean subjective judgments on the scales with corrections based on 3 dB per doubling of effective duration. These comparisons indicated that the correlations were fairly insensitive to the differences between the 3-dB values and the optimum values. The largest change in coefficient was 0.007. From these results it appears that a duration correction of 3 dB per doubling of effective duration provides highly accurate and satisfactory results.

Predictive Ability of Measurement Scales

Table IX presents the results of linear regressions of the mean subjective judgments on the duration-corrected scales (3 dB per doubling of effective duration) for all stimuli, stimuli without tones, and stimuli with tones. The scales are ranked in terms of decreasing correlation coefficients. The correlation coefficients were compared by using a two-tailed t-test for the significance of difference between correlation coefficients when samples are not independent (ref. 12). Results of these tests are given in table X. The performance of three of the rating scales (IPNLT, IPL_1, and IPL_2) is of particular interest.

When the stimuli without tones and the stimuli with tones are combined into one group, the IPNLT is ranked highest and its coefficient is significantly higher (0.01 level) than that of any of the other scales. This improvement can be attributed to the fact that PNLT is the only rating scale incorporating a tone correction. However, when the stimuli are separated by tonal content the ranking of IPNLT drops to the midrange of scales where its correlation coefficients are significantly lower (0.01 level) than the IPL_1 correlation coefficients for the stimuli without tones and are significantly lower than the IPL_1 (0.05 level) and IPNL (0.01 level) correlation coefficients for the stimuli with tones. From these results it appears that the tone correction used in the IPNLT aids in comparing the annoyance of stimuli with distinctly different tonal content but may slightly degrade the prediction ability when used in comparing stimuli of similar tonal content. The IPL_1, on the other hand, consistently ranks at the top for all three groupings of stimuli. For the combined stimuli, IPL_1 ranks second behind IPNLT and, with the exception of IPNL, its correlation coefficient is significantly higher (0.01 level) than those of the scales ranked below it. The IPL_1 ranks first for both of the subgroups of stimuli, and its correlation coefficients are significantly higher than those of any of the other scales in either subgroup with the exceptions of IPL_2 in the subgroup of stimuli without tones and IPNL and IL_D in the subgroup of stimuli with tones.

The PL_2 scale is a new calculation procedure recently suggested by Higgins as a method for predicting human response to noise (ref. 11). The procedure is based on a weighting curve consisting of two straight-line segments. The first segment rises by 6 dB per octave from 50 Hz to 4 kHz and passes through zero at 1 kHz. The second segment falls at a slope of 6 dB per octave from 4 kHz to 10 kHz. From tables IX and X it can be seen that IPL_2 did not show any improvement over the other scales in predicting annoyance and, in particular, ranked low in comparing stimuli with different tonal content. Although it did rank second in the subgroup of stimuli without tones, its correlation coefficient was not significantly better than those of the scales ranked below it. Similar comparisons of the rating scales without duration corrections also gave no indication that the PL_2 procedure predicted annoyance better than any of the other calculation procedures.

For each of the three stimuli groupings the correlation coefficient of the lowest ranking duration-corrected scale was greater than the correlation coefficient of the highest ranking scale without a duration correction. Comparisons of the correlations of the duration-corrected and uncorrected scales for a given calculation procedure showed improvements of 0.037 to 0.079 in the correlation coefficients. From these results it is clear that the addition of an accurate duration correction significantly improves the annoyance-prediction ability of the rating scales.

Interaction of Duration With Level and of Tonal Content With Level

The analysis of variance of the subjective data indicated that the interaction of duration with level and of tonal content with level significantly affected annoyance. Since variations in L_A appeared to be responsible for the indication of a significant interaction of duration with level, an analysis of covariance technique was applied to the data for each interaction. This analysis was performed in order to confirm the existence of these interactions and to determine if the rating scales account for the effects of the interactions.

Linear least-squares regression analyses were performed with the mean subjective judgments for each group of stimuli separated by duration or tonal content as the dependent variable and with the corresponding rating-scale values as the independent variable by using three different models. The first model assumed a common slope and a common mean for all stimuli. The second model assumed a common slope but separate means for the stimuli with different values of duration or tonal content. The third model assumed separate slopes and separate means. Details for this analysis can be found in reference 13 and basically consisted of comparing the residual mean squares between the three models with appropriate F-tests. First, a null hypothesis of common slopes was tested by comparing the second and third models. Rejection of the null hypothesis of common slopes indicates the existence of an interaction and indicates that the rating scale does not adequately account for the effect of the interaction. If the null hypothesis of common slopes was not rejected, then a test of the null hypothesis of common means, assuming common slopes, was tested by comparing the first and second models. This type of analysis was performed for both interactions with each rating scale. The results in terms of calculated and tabulated F-values are presented in table XI.

For the duration and level interaction the analysis was performed by using the rating scales without duration corrections. The null hypothesis of common slopes was not rejected, indicating that no interaction effect exists or that the rating scales adequately account for the effect. This tends to support the conclusion stated in a previous section that the indication of a significant duration and level interaction by the analysis of variance procedure resulted from slight variations in levels presented to the subjects across durations. Since the hypothesis of common slopes was not rejected, the null hypothesis of common means was tested and was rejected for all the rating scales. This confirms the existence of a duration effect and demonstrates the need for a duration correction to the rating scales.

For the interaction of tonal content and level the analysis was performed by using the duration-corrected rating scales. For this interaction, the null hypothesis of common slopes was rejected for all of the rating scales. Rejection of this hypothesis indicates the existence of an interaction of tonal content and level which is not adequately accounted for by the rating scales. Since the hypothesis of common slopes was rejected, the test for common means was not applicable.

Effects of Tonal Content on Annoyance

As indicated by the analysis of variance, annoyance is significantly affected by the tonal content of a noise. The addition of tones to a noise increases annoyance. The improvement in annoyance prediction resulting from the addition of a tone correction is evident from the significantly better correlation for all stimuli obtained from the IPNLT when compared with the other rating scales. However, when the stimuli without tones and the stimuli with tones are considered separately, the correlations with IPNL are greater than the correlations with tone-corrected measure IPNLT. As shown in table X(c), the difference between the correlations with IPNL and IPNLT is significant for stimuli with tones.

The results of the study indicate two areas of possible improvement in the PNLT tone-correction method. First, a change in the procedure to account for the apparent interaction of tonal content and sound pressure level previously discussed may improve the method. Second, a modification to the procedure to prevent the application of a tone correction to stimuli which contain no tones may improve the consistency and accuracy of the procedure. In this study, the PNLT tone-correction procedure applied overall tone corrections (PNLT - PNL) ranging from 0.6 to 2.7 dB and averaging 1.8 dB to the stimuli without tones. Based on the results of this study, the prediction of the effects of tonal content appears to be the largest remaining source of variation in the prediction of overall annoyance response.

Effects of Doppler Shift on Annoyance

As shown in table I, the stimuli used in this study represented five different patterns of Doppler shift. Since Doppler shift did not vary independently of duration and velocity, the effects of Doppler shift on annoyance could not be determined from the analysis of variance performed on the subjective judgments. In order to study the effects of Doppler shift, multiple regressions of the mean subjective judgments on the rating scales and duration corrections were performed with and without a Doppler shift term for the group of stimuli without tones and the group of stimuli with tones. The regressions without a Doppler shift term were of the form

$$J = b_0 + b_1 L_{max} + b_2 D \tag{6}$$

Three different Doppler shift models were assumed in the regressions with a Doppler shift term. The resulting regressions were

$$J = c_0 + c_1 L_{max} + c_2 D + c_3 \sqrt{\beta} \tag{7}$$

$$J = d_0 + d_1 L_{max} + d_2 D + d_3 \beta \tag{8}$$

$$J = e_0 + e_1 L_{max} + e_2 D + e_3 \beta^2 \tag{9}$$

where the Doppler shift parameter β is the aircraft velocity squared divided by the speed of sound and the distance of closest approach (that is, v^2/ch). The Doppler shift parameter is an approximation of the maximum rate of change of frequency.

The resulting coefficients of the Doppler shift terms, c_3, d_3, and e_3, varied in both magnitude and sign between rating scales and between models. Overall, the effect of the Doppler shift terms on the predicted subjective response was small relative to the effects of level, duration, and tonal content. Comparisons of the correlations for the regressions with a Doppler shift term with the correlations of the regressions without a Doppler shift term showed no significant improvement resulting from the addition of the Doppler shift term. Based on this analysis, Doppler shift does not appear to have a significant effect on the annoyance of either stimuli without tones or stimuli with tones.

CONCLUSIONS

A laboratory subjective listening test was performed to investigate the effects of duration and other noise characteristics on the annoyance caused by aircraft-flyover noise. The stimuli represented the 54 factorial combinations of 3 durations, 3 sound pressure levels, 3 aircraft velocities, and 2 tone conditions. The stimuli were synthesized by using a newly developed aircraft-noise synthesis system. Forty-eight test subjects made annoyance judgments on a total of 72 stimuli in a subjective test facility simulating the outdoor acoustic environment. The judgments were made by using a graphical scale procedure similar to numerical category scaling. The following conclusions were noted:

1. The duration of an aircraft-flyover noise significantly affects annoyance and should be taken into account in the quantification of aircraft-noise annoyance. Rating scales incorporating an accurate duration correction predict annoyance significantly better than scales without duration corrections.

2. A duration correction of 3 dB per doubling of effective duration was found to result in the greatest increase in the accuracy of the rating scales considered. This duration correction is identical to the one incorporated in the effective-perceived-noise-level (EPNL) procedure used by the Federal Aviation Administration in its FAR 36 aircraft certification rules.

3. The rating scales found to be most accurate in predicting the annoyance for all the stimuli were the effective perceived noise level (EPNL, IPNLT), the Stevens Mark VII perceived level with duration corrections (IPL_1), and the perceived noise level with duration corrections (IPNL). The IPL_1 was the most consistent in predicting annoyance of stimuli with similar tonal content.

4. Tonal characteristics appear to be the largest remaining source of variation in the prediction of overall annoyance response.

5. No significant effect of Doppler shift on annoyance was found for either noise with strong tonal components or noise without tonal components.

Langley Research Center
National Aeronautics and Space Administration
Hampton, VA 23665
January 26, 1979

APPENDIX

INSTRUCTIONS, CONSENT FORM, AND SCORING SHEETS

Copies of the instructions, consent form, and scoring sheets used in the experiment are presented in the following pages.

APPENDIX

Instructions

The experiment in which you are participating will help us understand the characteristics of aircraft sounds which can cause annoyance in airport communities. We would like you to judge how ANNOYING some of these aircraft sounds are. By ANNOYING we mean - UNWANTED, OBJECTIONABLE, DISTURBING, or UNPLEASANT.

The experiment consists of four 25 minute sessions. During each session 18 aircraft sounds will be presented for you to judge. Before each session you will be given a rating sheet with 18 scales like the one below.

Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying

After listening to each sound, please indicate how annoying you judge the sound to be by placing a mark across the scale. If you judge a sound to be only slightly annoying, then place your mark closer to the NOT ANNOYING AT ALL end of the scale. Similarly, if you judge a sound to be very annoying then place your mark closer to the EXTREMELY ANNOYING end of the scale. A moderately annoying judgment should be marked in the middle portion of the scale. A mark may be placed anywhere along the scale, not just at the numbered locations. Each aircraft sound will be followed by a beep or short tone. Please do not make your judgments until after the beep. You will have about five seconds after the beep to make and record your judgment. There are no right or wrong answers; we are only interested in your judgment of each sound.

Before the first session begins you will be given a practice rating sheet and three sounds will be presented to familiarize you with making and recording judgments. I will remain in the testing room with you during the practice time to answer any questions you may have.

Thank you for your help in conducting the experiment.

APPENDIX

Voluntary Consent Form for Subjects for Human Response to

Aircraft Noise and Vibration

I understand the purpose of the research and the technique to be used, including my participation in the research, as explained to me by the Principal Investigator (or qualified designee).

I do voluntarily consent to participate as a subject in the human response to aircraft noise experiment to be conducted at NASA Langley Research Center on ______________________________.

Date

I understand that I may at any time withdraw from the experiment and that I am under no obligation to give reasons for withdrawal or to attend again for experimentation.

I undertake to obey the regulations of the laboratory and instructions of the Principal Investigator regarding safety, subject only to my right to withdraw declared above.

I affirm that, to my knowledge, my state of health has not changed since the time at which I completed and signed the medical report form required for my participation as a test subject.

____

Signature of Subject

APPENDIX

Practice Rating Sheet

Subject No.__________ Group__________

Practice Sound		Judgment										
I	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
II	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
III	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying

APPENDIX

Rating Sheet

Page 1

Subject No.________ Group________ Session________ Tape________

<u>Sound</u>

Sound												
1	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
2	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
3	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
4	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
5	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
6	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
7	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
8	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
9	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying

APPENDIX

Rating Sheet

Page 2

Subject No._______ Group_______ Session_______ Tape_______

<u>Sound</u>

Sound												
10	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
11	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
12	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
13	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
14	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
15	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
16	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
17	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying
18	Not at all Annoying	0	1	2	3	4	5	6	7	8	9	Extremely Annoying

REFERENCES

1. Pearsons, Karl S.: The Effects of Duration and Background Noise Level on Perceived Noisiness. ADS-78, FAA, Apr. 1966.

2. Ollerhead, J. B.: An Evaluation of Methods for Scaling Aircraft Noise Perception. NASA CR-1883, 1971.

3. Parry, H. J.; and Parry, J. K.: The Interpretation and Meaning of Laboratory Determinations of the Effect of Duration on the Judged Acceptability of Noise. J. Sound & Vib., vol. 20, no. 1, Jan. 8, 1972, pp. 51-57.

4. Shepherd, Kevin P.: The Subjective Evaluation of Noise From Light Aircraft. NASA CR-2773, 1976.

5. Little, J. W.; and Mabry, J. E.: Sound Duration and Its Effect on Judged Annoyance. J. Sound & Vib., vol. 9, no. 2, Mar. 1969, pp. 247-262.

6. Pearsons, Karl S.: Combination Effects of Tone and Duration Parameters on Perceived Noisiness. NASA CR-1283, 1969.

7. Pearsons, Karl S.; and Bennett, Ricarda L.: The Effects of Temporal and Spectral Combinations on the Judged Noisiness of Aircraft Sounds. FAA-NO-69-3, June 1969.

8. Kryter, K. D.; Peeler, D. J.; Dobbs, M. E.; and Lukas, J. S.: Reliability of Laboratory Tests of VSTOL and Other Long-Duration Noises. NASA CR-2471, 1974.

9. Kryter, K. D.; Johnson, P. J.; and Young, J. R.: Judgment Tests of Flyover Noise From Various Aircraft. NASA CR-1635, 1970.

10. Pearsons, Karl S.; and Bennett, Ricarda L.: Handbook of Noise Ratings. NASA CR-2376, 1974.

11. Higgins, Thomas H.: Human Response to Sound: The Calculation of Perceived Level, PLdB (Noisiness or Loudness) Directly From Physical Measures. Rep. No. FAA-RD-76-1, Nov. 17, 1976.

12. Klugh, Henry E.: Statistics - The Essentials for Research. John Wiley & Sons, Inc., c.1970.

13. Edwards, Allen L.: Experimental Design in Psychological Research. Third ed. Holt, Rinehart and Winston, Inc., c.1968.

TABLE I.- DOPPLER SHIFT PATTERNS CORRESPONDING TO FACTORIAL COMBINATIONS OF DURATION AND VELOCITY

		Aircraft velocity		
		v/2	v	2v
Noise duration	t	[a]β/2	β	2β
	2t	β/4	β/2	β
	4t	β/8	β/4	β/2

[a] $\beta = \frac{v^2}{ch}$, where $\left[\frac{d(f/f_s)}{dt}\right]_{max} = \beta \times [\mathrm{Function}\ (\theta)]$.

TABLE II.- TEST SUBJECTS

Sex	Number of participants	Mean age	Median age	Age range
Male	7	26	25	21 to 34
Female	41	37	36	18 to 65
All subjects	48	35	34	18 to 65

TABLE III.- PRESENTATION ORDER OF STIMULI ON TAPES

Practice tape	Tape I	Tape II	Tape III	Tape IV	Tape V	Tape VI
2221	2221	1332	3112	1121	3232	2311
2132	1312	2321	3231	3312	1321	1222
2311	3132	1111	2322	2132	2312	3131
	2331	3122	1221	3121	3111	2112
	3221	3211	2211	1231	1132	1322
	1112	1232	3332	2212	2121	3331
	1212	2232	1122	1311	3322	2222
	2131	3311	2111	3222	1211	1131
	3321	2122	1331	2332	2231	3212
	2332	2231	3212	3321	2122	1331
	3222	1211	1131	2131	3311	2111
	1311	3322	2222	1212	2232	1122
	2212	2121	3331	1112	1232	3332
	1231	1132	1322	3221	3211	2211
	3121	3111	2112	2331	3122	1221
	2132	2312	3131	3132	1111	2322
	3312	1321	1222	1312	2321	3231
	1121	3232	2311	2221	1332	3112

Stimuli key			
A	B	C	D
Duration, sec (a)	Aircraft velocity, m/sec	Nominal L_A, dB	Total content
1 = 10 2 = 20 3 = 40	1 = 40 2 = 80 3 = 160	1 = 70 2 = 79 3 = 88	1 = No tones 2 = Strong tones

[a]Time between the first and last points at which the noise signal is 10 dB down from the maximum sound level.

TABLE IV.- ORDER OF TAPES PRESENTED TO TEST-SUBJECT GROUPS

Test-subject group	Tapes presented during sessions -			
	1	2	3	4
1	I	II	III	IV
2	II	III	I	V
3	III	I	II	VI
4	III	II	I	VI
5	II	I	III	V
6	I	III	II	IV
7	IV	V	VI	I
8	V	VI	IV	II
9	VI	IV	V	III
10	VI	V	IV	III
11	V	IV	VI	II
12	IV	VI	V	I

TABLE V.- HIGHEST MEASURED LEVELS OF STIMULI FOR EACH COMBINATION OF DURATION, VELOCITY, AND TONAL CONTENT

STIMULI	OASPL	IOASPL	L_A	IL_A	L_D	IL_D	PNL	IPNL	PNLT	IPNLT	PL_1	IPL_1	PL_2	IPL_2
1131	93.8	91.4	89.0	86.3	93.1	90.7	98.3	96.2	99.6	97.6	89.7	87.6	94.1	91.8
1132	91.0	88.3	86.5	83.7	91.0	88.4	96.8	94.5	101.8	98.0	88.0	84.1	91.4	88.9
1231	92.8	89.5	88.3	84.4	92.5	88.8	98.3	94.5	100.1	95.9	89.5	85.8	93.5	89.9
1232	90.5	89.7	86.1	84.4	90.9	89.4	97.0	95.3	103.0	99.0	88.2	86.6	90.9	90.2
1331	94.0	92.9	89.5	86.6	93.6	91.6	98.9	97.1	100.6	98.7	90.2	88.4	94.5	93.2
1332	91.0	90.6	85.8	84.7	91.9	90.2	97.9	96.0	104.0	98.8	88.0	87.1	91.7	91.5
2131	93.0	92.5	87.8	87.5	92.0	91.9	97.3	97.6	99.1	99.1	88.7	88.9	93.4	93.0
2132	89.8	91.1	85.2	86.0	89.6	90.7	95.1	96.9	99.9	100.2	86.2	88.1	90.2	91.5
2231	96.8	96.3	91.8	91.1	96.1	95.5	101.8	101.0	103.7	102.4	93.2	92.4	97.3	97.4
2232	91.5	92.9	86.8	87.7	91.9	92.6	97.6	98.6	103.7	102.5	88.8	89.7	91.7	93.3
2331	95.3	96.4	89.9	90.0	94.4	95.2	99.6	100.6	101.6	102.1	91.0	92.0	95.7	96.8
2332	92.8	94.1	86.3	87.8	91.8	93.5	97.7	99.2	103.7	102.1	88.9	90.3	93.1	94.7
3131	95.0	98.0	89.9	92.9	94.2	97.3	99.7	103.1	101.2	104.6	90.9	94.4	95.3	98.4
3132	93.8	97.5	88.7	92.4	93.6	96.9	98.7	102.9	104.4	106.5	89.8	94.2	94.6	97.7
3231	97.0	99.4	91.6	94.2	96.4	98.6	101.9	104.0	103.5	105.6	93.2	95.3	97.5	99.8
3232	93.0	97.5	87.8	92.3	92.8	97.1	99.1	102.8	104.0	106.2	90.1	94.0	93.5	98.0
3331	95.3	100.8	90.6	93.4	94.6	99.1	100.4	104.3	102.8	105.9	91.8	95.5	96.1	101.3
3332	95.6	101.4	89.2	93.9	95.2	99.9	100.2	105.2	105.9	108.5	91.4	96.2	97.2	101.9

Stimuli key			
A	B	C	D
Duration, sec (a)	Aircraft velocity, m/sec	Nominal L_A, dB	Total content
1 = 10 2 = 20 3 = 40	1 = 40 2 = 80 3 = 160	1 = 70 2 = 79 3 = 88	1 = No tones 2 = Strong tones

[a]Time between the first and last points at which the noise signal is 10 dB down from the maximum sound level.

TABLE VI.- REGRESSION ANALYSES FOR REPEATED JUDGMENTS

Regression	Intercept	Slope	Correlation coefficient
Individual judgments	1.540	0.705	0.722
Means over subjects	0.614	0.935	0.961

TABLE VII.- ANALYSIS OF VARIANCE

Source of variation	Degrees of freedom	Sum of squares	Mean square	F-ratio (a)
Duration	2	1483.57233	741.78617	185.30082*
Velocity	2	15.48889	7.74445	1.93459ns
Level	2	9206.72516	4603.36258	1149.93632*
Tones	1	486.90042	486.90042	121.62945*
Interaction of:				
Duration and velocity . . .	4	20.84255	5.21064	1.30164ns
Duration and level	4	64.45180	16.11295	4.02507*
Duration and tones	2	6.86960	3.43480	0.85803ns
Velocity and level	4	15.75420	3.93855	0.98386ns
Velocity and tones	2	22.35262	11.17631	2.79188ns
Level and tones	2	104.39521	52.19760	13.03915*
Duration, velocity, and level	8	48.55150	6.06894	1.51604ns
Duration, velocity, and tones	4	29.43382	7.35845	1.83817ns
Duration, level, and tones	4	29.53707	7.38427	1.84462ns
Velocity, level, and tones	4	17.55384	4.38846	1.09625ns
Duration, velocity, level, and tones	8	36.75415	4.59427	1.14766ns
Residual	3402	13 618.70156	4.00315	
Total	3455	25 207.88471	7.29606	

[a]ns indicates not significant at 0.05 level; * indicates significant at 0.01 level.

TABLE VIII.- OPTIMUM DURATION-CORRECTION MAGNITUDES

[Optimum duration-correction magnitudes are in decibels per doubling of effective duration]

Scale	All stimuli	Stimuli without tones	Stimuli with tones
OASPL	3.35	2.62	2.70
L_A	3.87	3.05	3.19
L_D	3.44	2.94	3.06
PNL	3.57	3.12	3.25
PNLT	2.80	3.16	3.08
PL_1	3.61	3.09	3.06
PL_2	3.41	2.63	2.72
Mean across scales . . .	3.44	2.94	3.01

TABLE IX.- RESULTS OF LINEAR REGRESSIONS OF MEAN SUBJECTIVE

JUDGMENTS ON DURATION-CORRECTED[a] SCALES

Rank	All stimuli				Stimuli without tones				Stimuli with tones			
	Scale	r	Slope	Intercept	Scale	r	Slope	Intercept	Scale	r	Slope	Intercept
1	IPNLT	0.973	0.2146	-15.628	IPL_1	0.982	0.2484	-16.594	IPL_1	0.990	0.2166	-13.141
2	IPL_1	.956	.2317	-14.799	IPL_2	.977	.2373	-16.698	IPNL	.989	.1955	-12.954
3	IPNL	.953	.2101	-14.684	IPNL	.977	.2262	-16.553	IL_D	.988	.1975	-12.016
4	IL_D	.946	.2109	-13.602	IPNLT	.977	.2252	-16.802	IL_A	.987	.1962	-10.896
5	IL_A	.941	.2082	-12.338	IOASPL	.976	.2402	-16.855	IPNLT	.985	.1978	-13.851
6	IOASPL	.939	.2194	-14.567	IL_A	.976	.2273	-14.352	IOASPL	.982	.2061	-12.944
7	IPL_2	.937	.2142	-14.198	IL_D	.974	.2283	-15.526	IPL_2	.981	.1997	-12.454

[a]Correction based on 3 dB per doubling of effective duration.

TABLE X.- COMPARISON OF CORRELATION COEFFICIENTS OF MEAN SUBJECTIVE JUDGMENTS WITH DURATION-CORRECTED RATING SCALES

(a) All stimuli[a]

Rank	Rating scale	t-statistic						
		IPNLT	IPL_1	IPNL	IL_D	IL_A	IOASPL	IPL_2
1	IPNLT	-------						
2	IPL_1	3.862**	-------					
3	IPNL	5.182**	1.694[ns]	-------				
4	IL_D	5.842**	4.928**	4.500**	-------			
5	IL_A	6.399**	6.950**	6.103**	2.054[ns]	-------		
6	IOASPL	5.737**	5.376**	3.584**	2.485*	0.464[ns]	-------	
7	IPL_2	6.010**	5.422**	4.055**	3.046**	.755[ns]	0.533[ns]	----

[a]ns indicates not significant; * indicates significant at 0.05 level (t_{51} = 2.007); and ** indicates significant at 0.01 level (t_{51} = 2.676).

TABLE X.- Continued

(b) Stimuli without tones[a]

Rank	Rating scale	t-statistic						
		IPL_1	IPL_2	IPNL	IPNLT	IOASPL	IL_A	IL_D
1	IPL_1	-------						
2	IPL_2	1.761^{ns}	-------					
3	IPNL	3.669**	0.115^{ns}	-------				
4	IPNLT	4.380**	.152^{ns}	0.133^{ns}	-------			
5	IOASPL	2.255*	.556^{ns}	.258^{ns}	0.240^{ns}	-------		
6	IL_A	2.976**	.410^{ns}	.733^{ns}	.682^{ns}	0.125^{ns}	-------	
7	IL_D	6.340**	1.070^{ns}	2.508*	2.587*	.686^{ns}	0.522^{ns}	----

[a]ns indicates not significant; * indicates significant at 0.05 level (t_{24} = 2.064); and ** indicates significant at 0.01 level (t_{24} = 2.797).

TABLE X.- Concluded

(c) Stimuli with tones[a]

Rank	Rating scale	t-statistic						
		IPL_1	IPNL	IL_D	I_{LA}	IPNLT	IOASPL	IPL_2
1	IPL_1	-------						
2	IPNL	0.963^{ns}	-------					
3	IL_D	1.407^{ns}	0.810^{ns}	-------				
4	IL_A	2.076^{*}	1.555^{ns}	0.364^{ns}	-------			
5	IPNLT	2.759^{*}	2.999^{**}	1.255^{ns}	1.587^{ns}	-------		
6	IOASPL	3.193^{**}	2.178^{*}	2.460^{*}	1.332^{ns}	0.671^{ns}	-------	
7	IPL_2	3.135^{**}	2.606^{*}	3.441^{**}	1.565^{ns}	$.857^{ns}$	0.235^{ns}	----

[a]ns indicates not significant; * indicates significant at 0.05 level (t_{24} = 2.064); and ** indicates significant at 0.01 level (t_{24} = 2.797).

TABLE XI.- RESULTS FROM ANALYSIS OF COVARIANCE TECHNIQUE FOR INTERACTION OF DURATION WITH LEVEL AND OF TONAL CONTENT WITH LEVEL

Rating scale	Duration with level, F-statistic[a], Slope	Duration with level, F-statistic[a], Mean	Rating scale	Tonal content with level, F-statistic[a], Slope
OASPL	1.43^{ns}	$11.45^{\dagger}$	IOASPL	6.88^{*}
L_A	1.10^{ns}	$14.86^{\dagger}$	IL_A	6.97^{*}
L_D	1.16^{ns}	$17.77^{\dagger}$	IL_D	6.37^{*}
PNL	1.04^{ns}	$20.38^{\dagger}$	IPNL	7.00^{*}
PNLT	2.05^{ns}	$41.71^{\dagger}$	IPNLT	5.20^{*}
PL_1	$.95^{ns}$	$19.10^{\dagger}$	IPL_1	7.62^{**}
PL_2	1.39^{ns}	$11.15^{\dagger}$	IPL_2	8.93^{**}

[a]ns indicates not significant at 0.05 level ($F_{2,48} = 3.19$); † indicates significant at 0.01 level ($F_{2,50} = 5.06$); * indicates significant at 0.05 level ($F_{1,50} = 4.03$); and ** indicates significant at 0.01 level ($F_{1,50} = 7.17$).

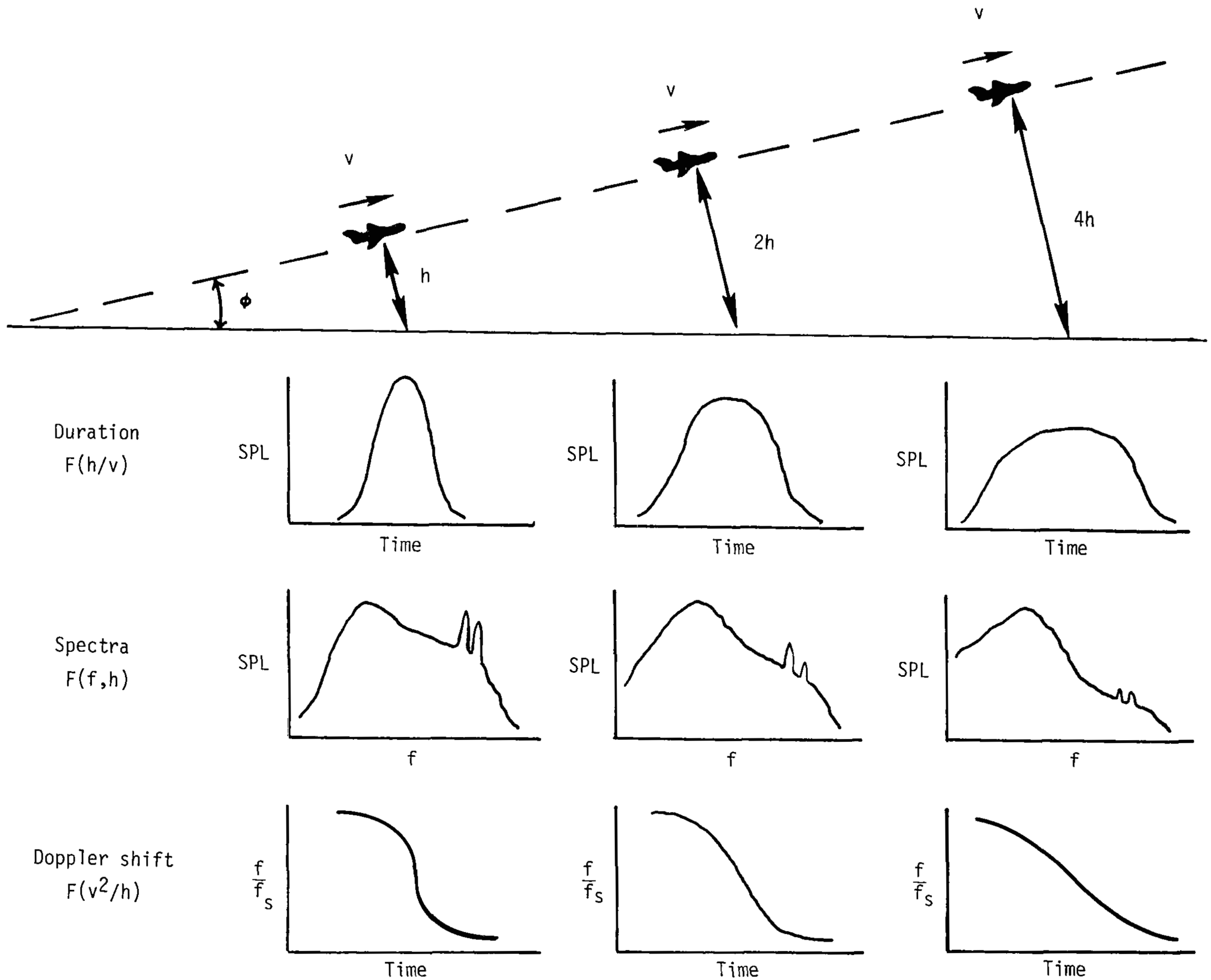


Figure 1.- Confounding of parameters associated with recordings of real aircraft noise.

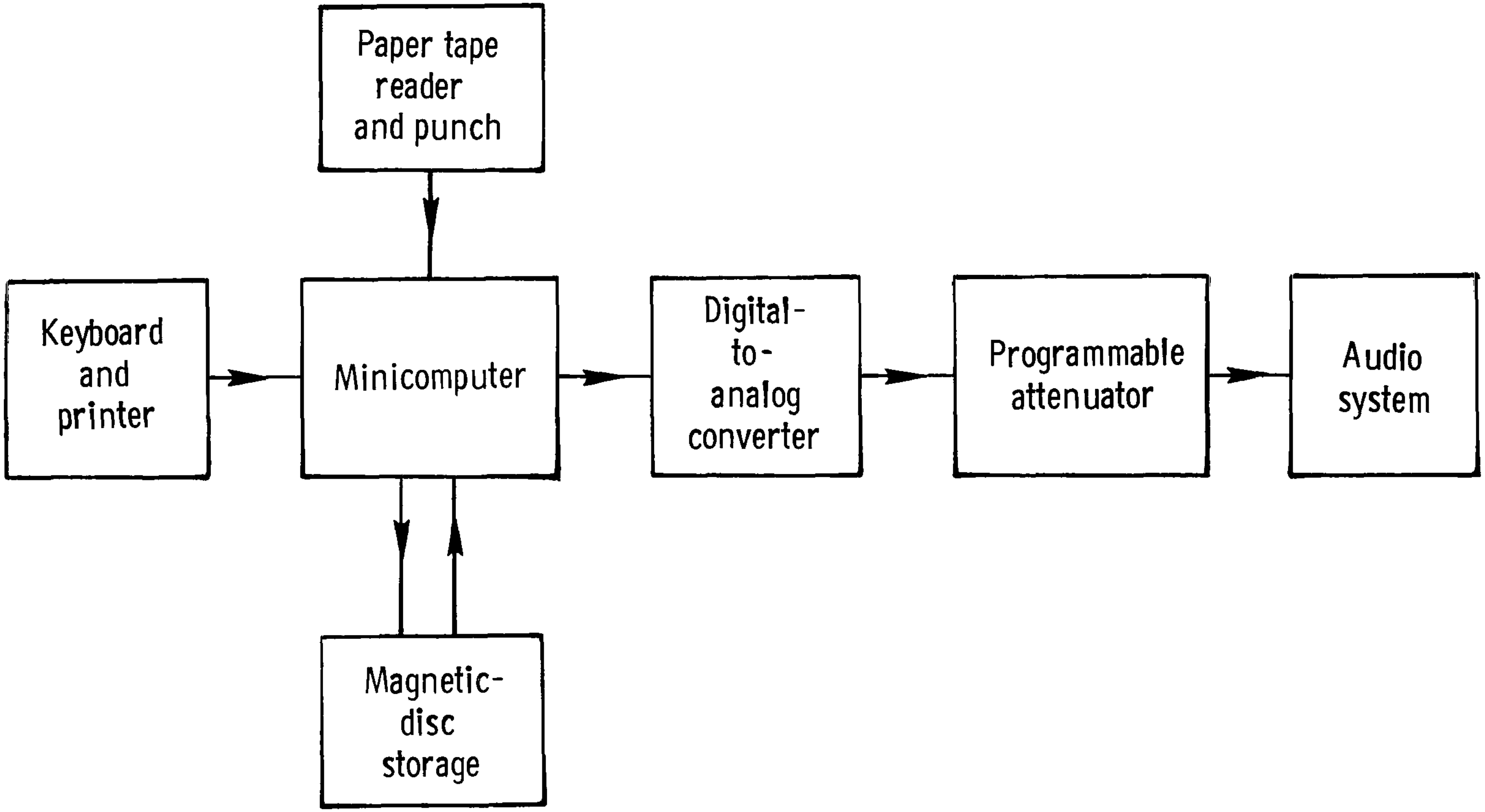


Figure 2.- Block diagram of aircraft-noise synthesis system.

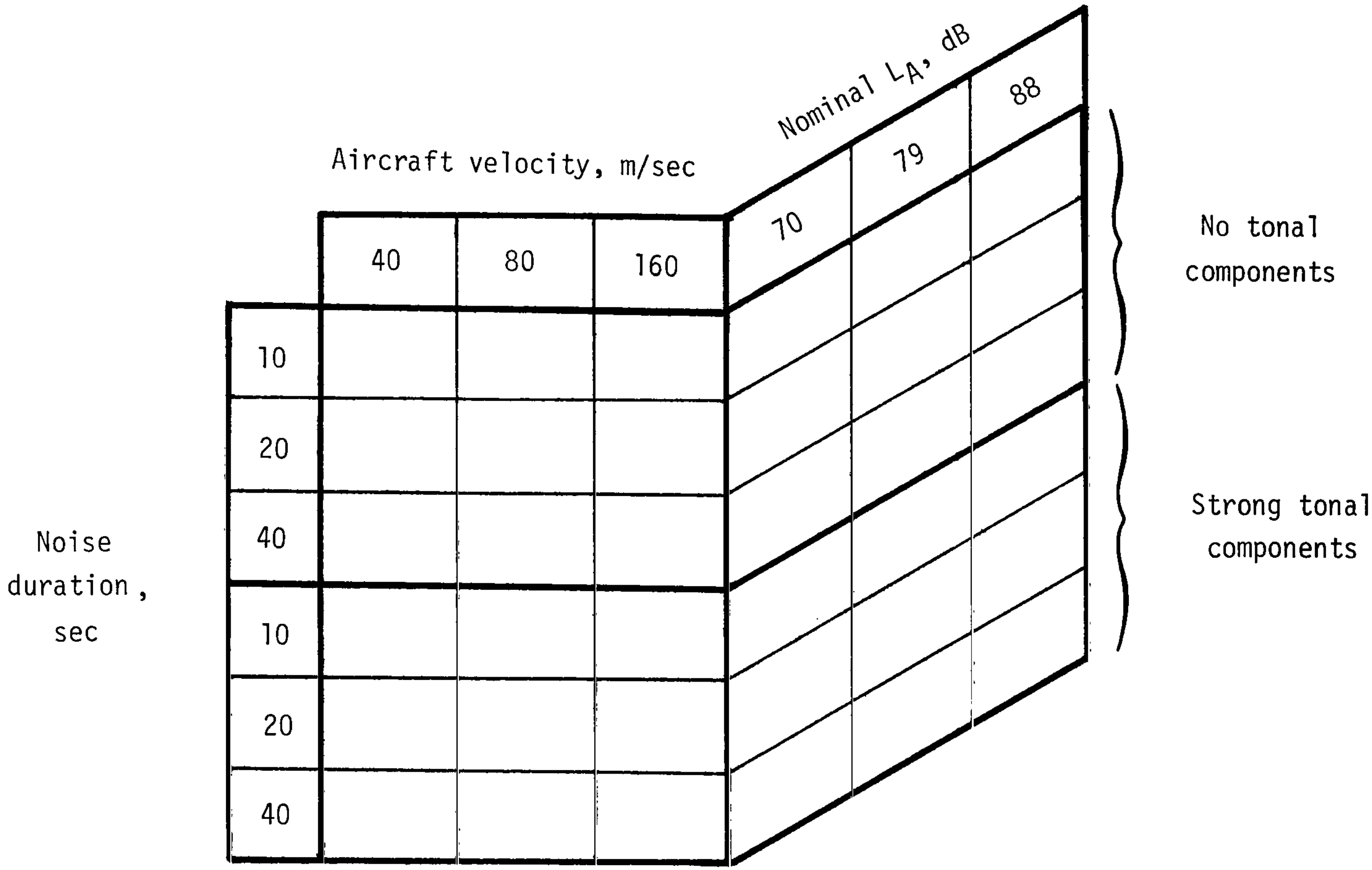
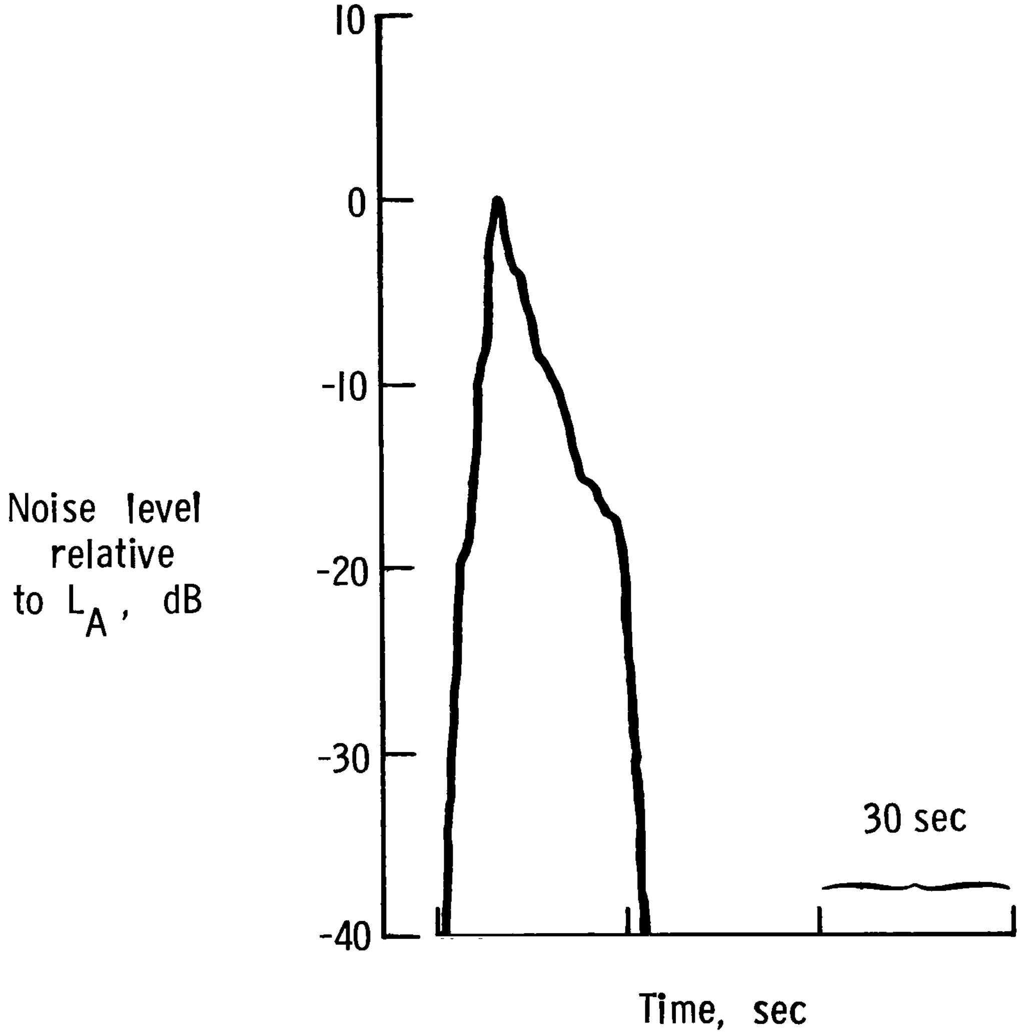
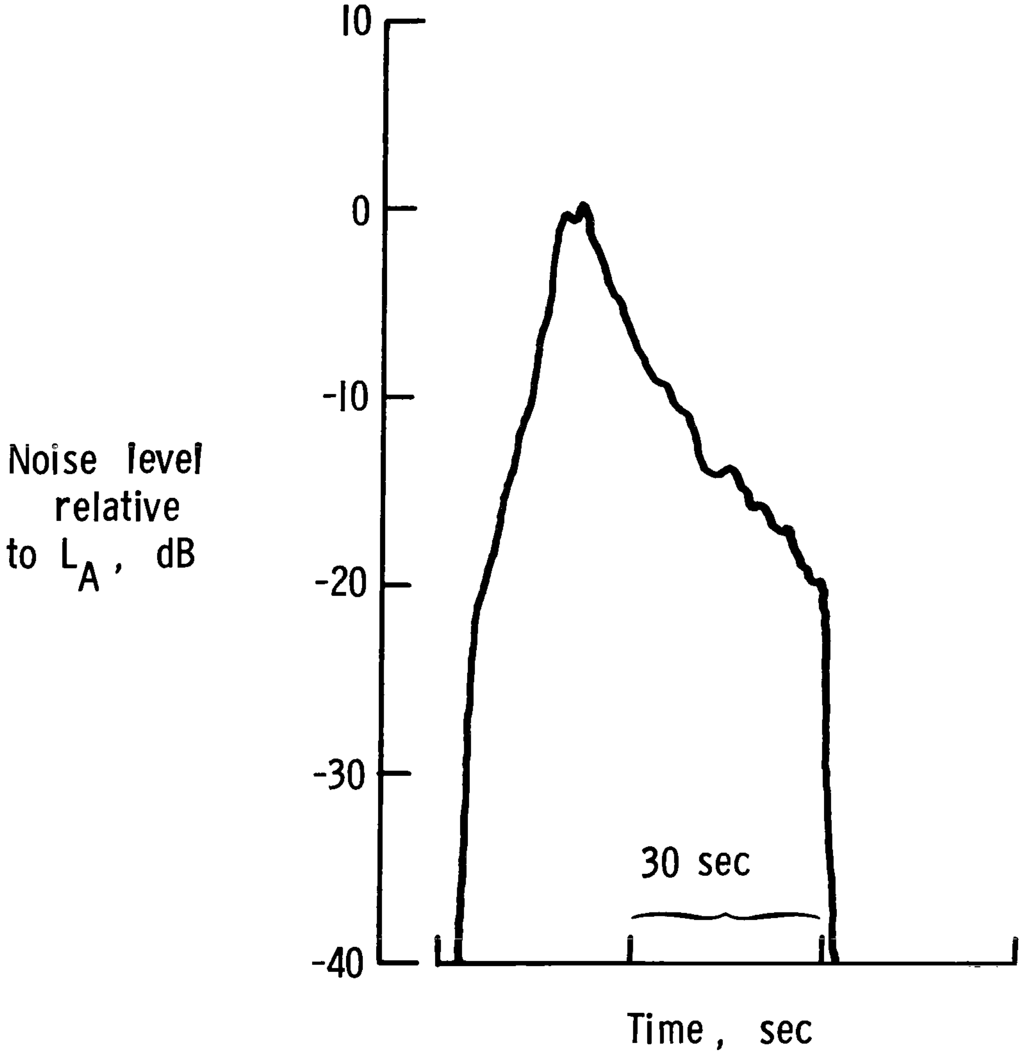


Figure 3.- Experiment design. Noise duration denotes the time between the first and last points at which the noise signal is 10 dB down from the maximum sound level.



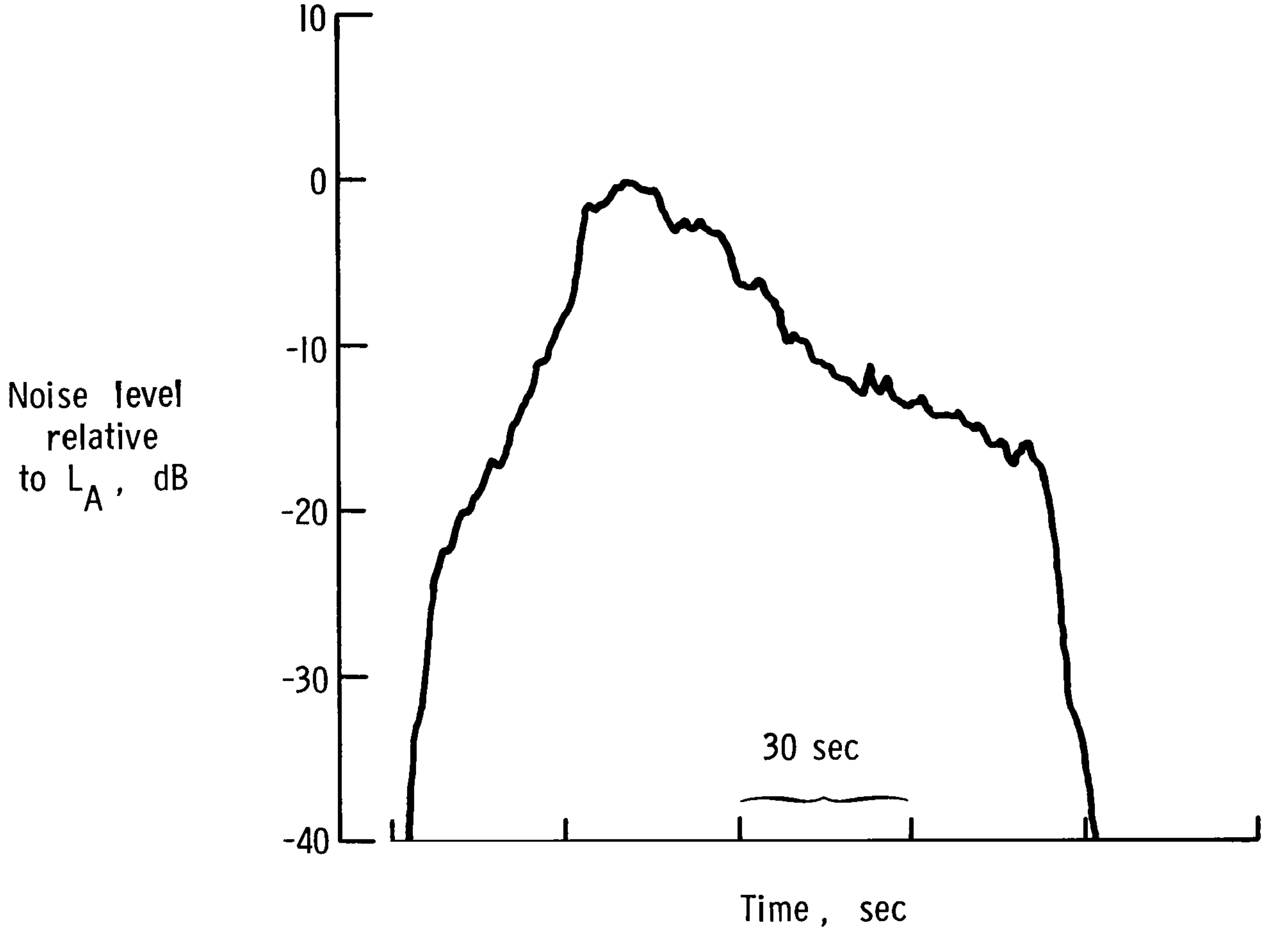
(a) Shortest duration.

Figure 4.- Typical time histories for stimuli of each duration.



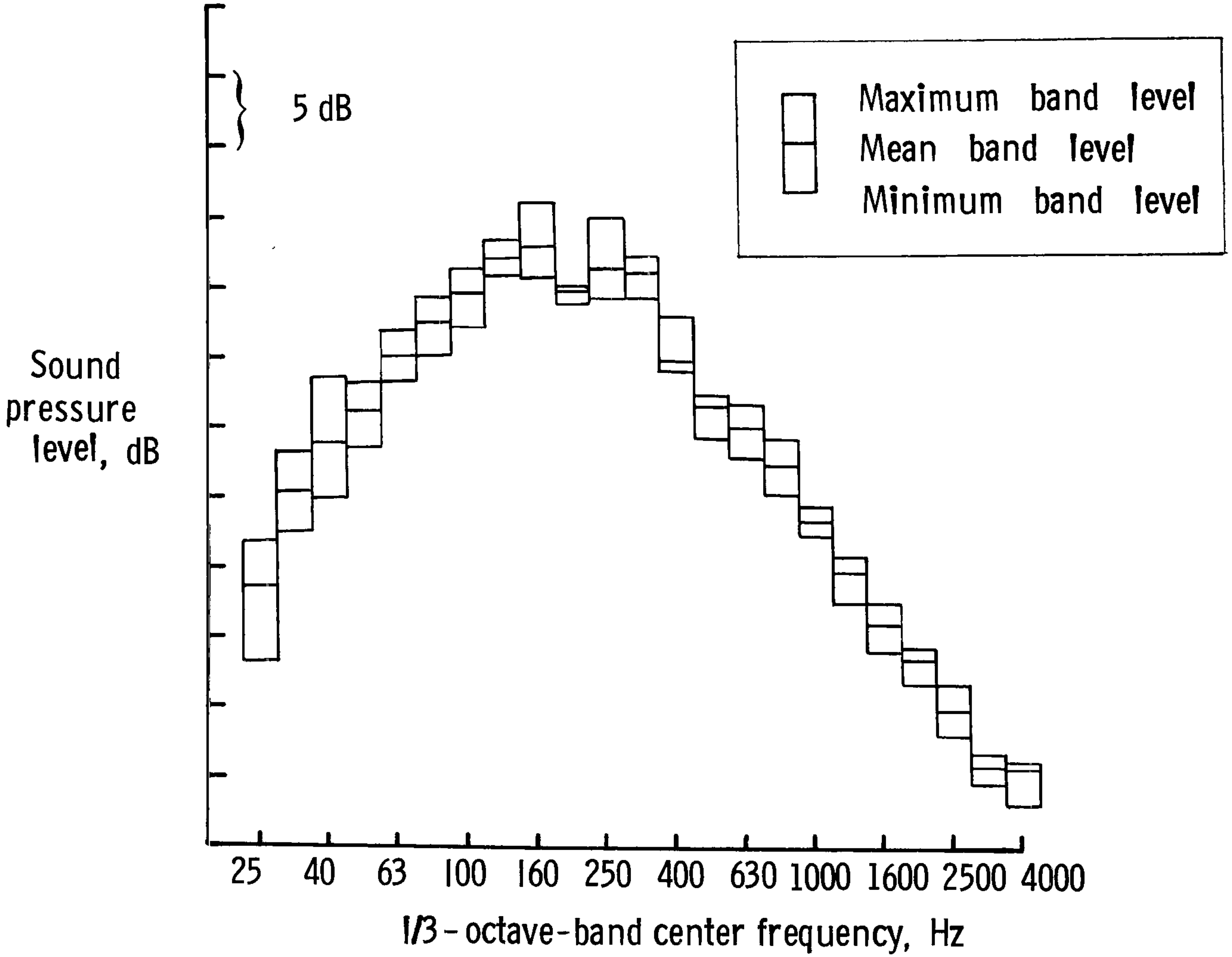
(b) Middle duration.

Figure 4.- Continued.



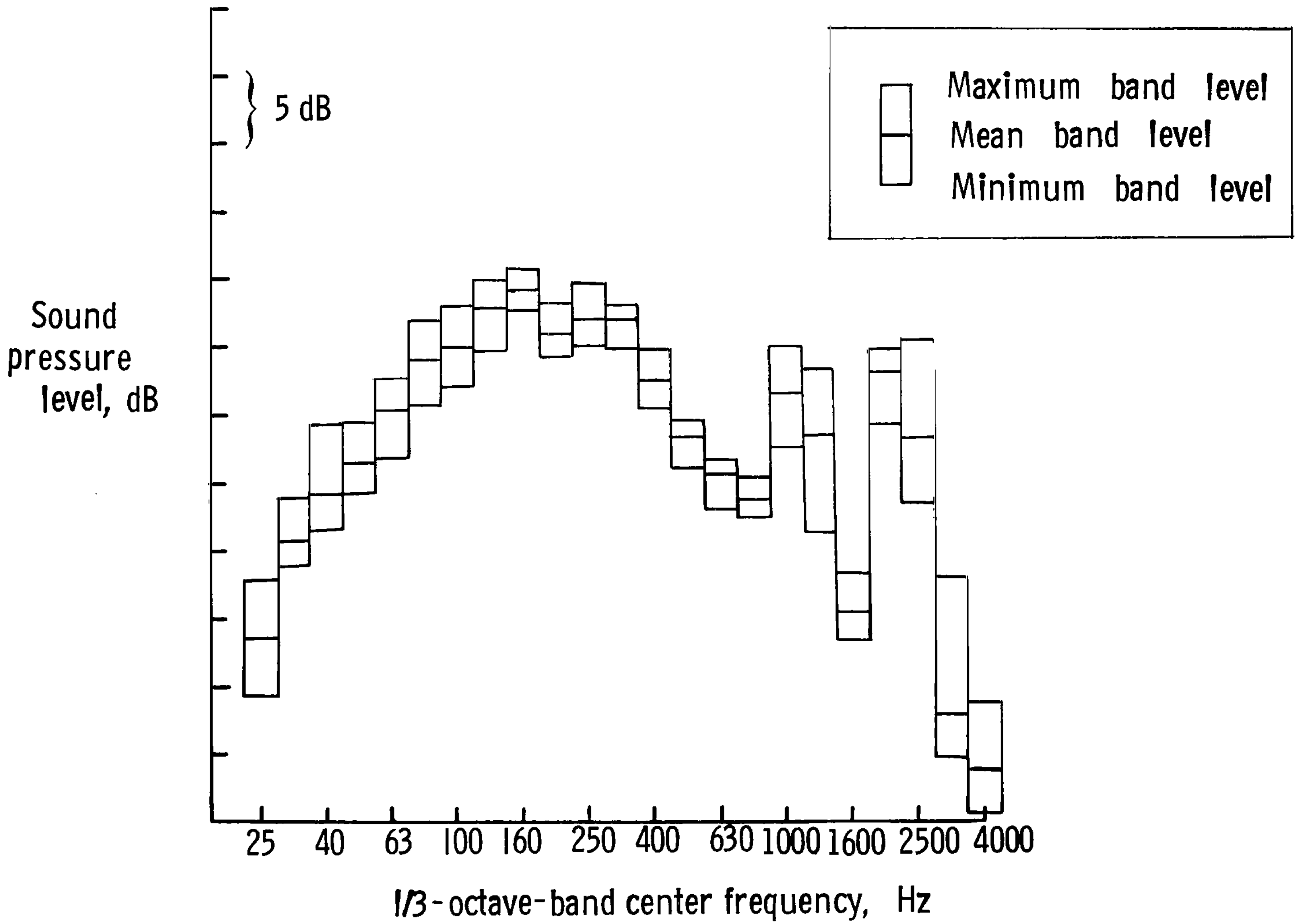
(c) Longest duration.

Figure 4.- Concluded.



(a) Stimuli with no tonal components.

Figure 5.- Range and mean of 1/3-octave band levels of stimuli occurring at and normalized by maximum A-weighted sound pressure level.



(b) Stimuli with strong tonal components.

Figure 5.- Concluded.

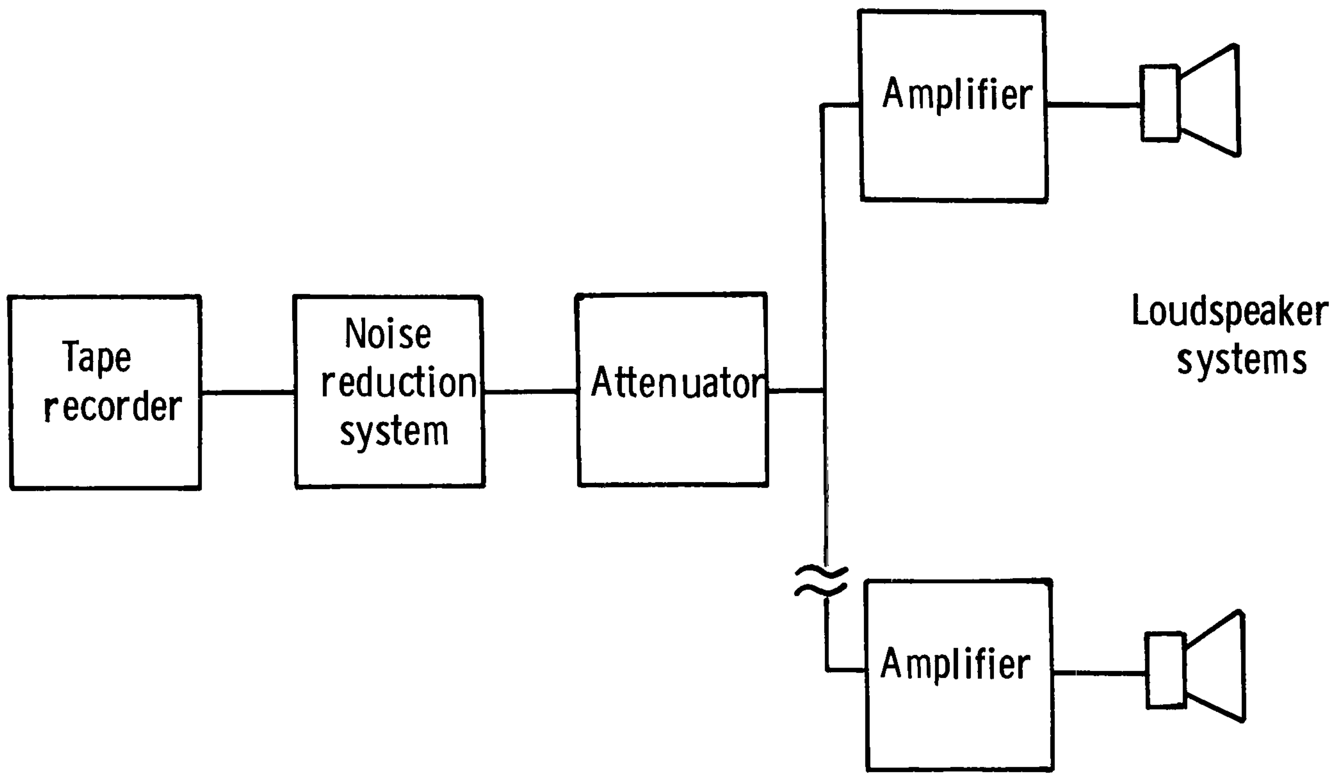


Figure 6.- Block diagram of audio reproduction system.

L-79-121

Figure 7.- Subjects in exterior effects room of the Langley aircraft noise reduction laboratory.

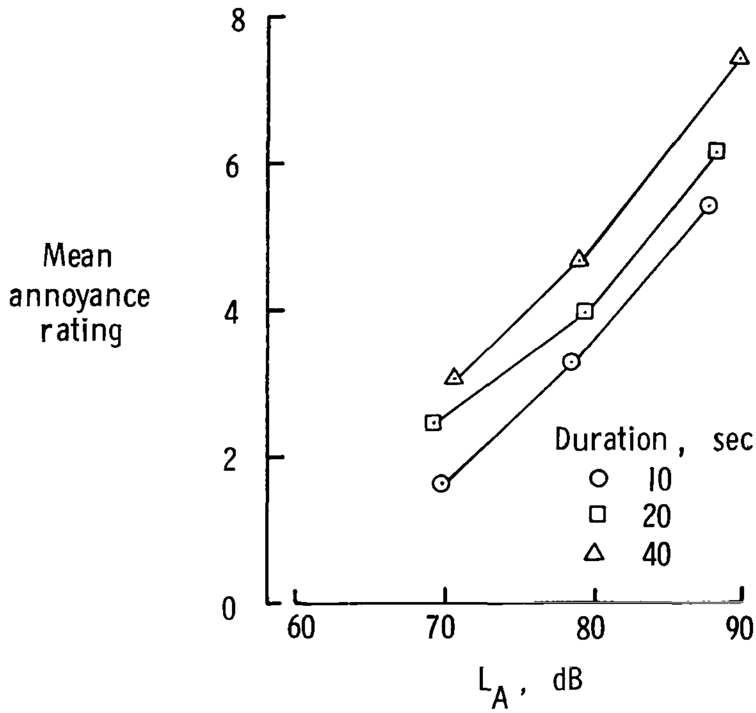


Figure 8.- Effects of duration and sound pressure level on annoyance.

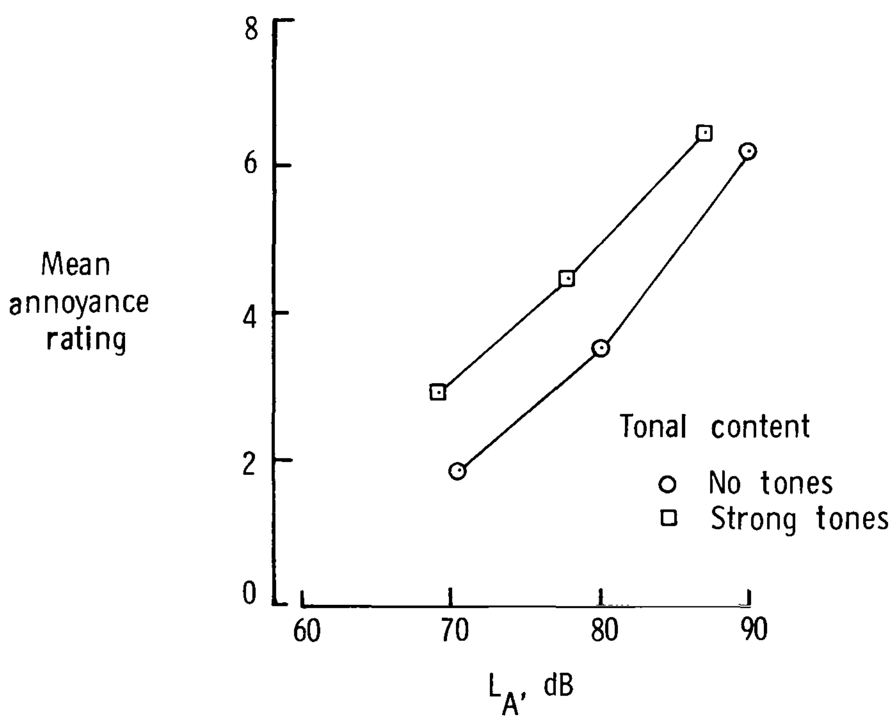


Figure 9.- Effects of tonal content and sound pressure level on annoyance.

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TP-1386		

4. Title and Subtitle	5. Report Date
EFFECTS OF DURATION AND OTHER NOISE CHARACTERISTICS ON THE ANNOYANCE CAUSED BY AIRCRAFT-FLYOVER NOISE	March 1979
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
David A. McCurdy and Clemans A. Powell	L-12579
	10. Work Unit No.
9. Performing Organization Name and Address	505-09-13-11
NASA Langley Research Center Hampton, VA 23665	11. Contract or Grant No.
	13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address	Technical Paper
National Aeronautics and Space Administration Washington, DC 20546	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

A laboratory experiment was conducted to determine the effects of duration and other noise characteristics on the annoyance caused by aircraft-flyover noise. The experiment was unique in that the duration, Doppler shift, and spectra were individually controlled by specifying aircraft operational factors, such as velocity, altitude, and spectrum, in a computer synthesis of the aircraft-noise stimuli. This control allowed the separation of the effects of duration from the other main factors in the experimental design: velocity, tonal content, and sound pressure level. In the experiment, 48 test subjects judged the annoyance of a set of noise stimuli which were comprised of factorial combinations of 3 durations, 3 velocities, 3 sound pressure levels, and 2 tone conditions. The judgments were made by using a graphical scale procedure similar to numerical category scaling. Each of the main factors except velocity was found to affect the judged annoyance significantly. The interaction of tonal content with sound pressure level was also found to be significant. The duration correction used in the effective-perceived-noise-level procedure, 3 dB per doubling of effective duration, was found to account most accurately for the effect of duration. No significant effect of Doppler shift was found.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Noise duration Noise annoyance Aircraft noise Subjective acoustics	Unclassified - Unlimited Subject Category 71

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	44	$4.50

* For sale by the National Technical Information Service, Springfield, Virginia 22161

NASA-Langley, 1979

National Aeronautics and
Space Administration

Washington, D.C.
20546

Official Business
Penalty for Private Use, $300

THIRD-CLASS BULK RATE

Postage and Fees Paid
National Aeronautics and
Space Administration
NASA-451

U.S.MAIL

3 1 1U,H, 022379 S00903DS
DEPT OF THE AIR FORCE
AF WEAPONS LABORATORY
ATTN: TECHNICAL LIBRARY (SUL)
KIRTLAND AFB NM 87117

S

NASA

POSTMASTER: If Undeliverable (Section 158
Postal Manual) Do Not Return