Studies of aircraft noise perception

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Additional Information:

- A Doctoral Thesis. Submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy of Loughborough University.

Metadata Record: https://dspace.lboro.ac.uk/2134/15352

Publisher: © John Bruce Ollerhead

Please cite the published version.
This item was submitted to Loughborough University as a PhD thesis by the author and is made available in the Institutional Repository (https://dspace.lboro.ac.uk/) under the following Creative Commons Licence conditions.

For the full text of this licence, please go to: http://creativecommons.org/licenses/by-nc-nd/2.5/
STUDIES OF AIRCRAFT NOISE PERCEPTION

Submission for the Degree of Doctor of Philosophy

of the

Loughborough University of Technology

by

John Bruce Ollerhead B.Sc.(Eng)., M.Sc.(Eng).,
C.Eng., M.R.Ae.S., F.I.O.A.

October
1982
FOREWORD

The author's long-standing interest in aircraft noise is reflected in a number of papers and reports on several aspects of the problem. These concern the basic noise generation mechanisms of aircraft propulsion systems (jets, rotors, propellers and fans), the subjective aspects of aircraft noise perception and the long-term community impact of noise near airports.

This compilation brings together the results of eight particular studies which have been central to the author's work on aircraft noise perception and impact. Much of the content has been extracted directly from original technical reports in a sequence which reflects the logical development of the subject matter. Reference is made to related publications where appropriate although these are essentially abbreviated versions of the original reports.
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>SUBJECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
</tr>
<tr>
<td>2</td>
<td>PROBLEMS OF HELICOPTER NOISE ESTIMATION AND REDUCTION</td>
</tr>
<tr>
<td>3</td>
<td>HELICOPTER AURAL DETECTABILITY</td>
</tr>
<tr>
<td>4</td>
<td>A REDETERMINATION OF THE NOY CONTOURS</td>
</tr>
<tr>
<td>5</td>
<td>SCALING AIRCRAFT NOISE PERCEPTION</td>
</tr>
<tr>
<td>6</td>
<td>LABORATORY STUDIES OF SCALES FOR MEASURING HELICOPTER NOISE</td>
</tr>
<tr>
<td>7</td>
<td>A FURTHER SURVEY OF SOME EFFECTS OF AIRCRAFT NOISE IN RESIDENTIAL COMMUNITIES NEAR LONDON (HEATHROW) AIRPORT</td>
</tr>
<tr>
<td>8</td>
<td>VARIATIONS OF COMMUNITY RESPONSE TO AIRCRAFT NOISE WITH TIME OF DAY</td>
</tr>
<tr>
<td>9</td>
<td>PREDICTING PUBLIC REACTION TO NOISE FROM MIXED SOURCES</td>
</tr>
</tbody>
</table>
INTRODUCTION

The author's interest in the subjective aspects of aircraft noise stemmed initially from, and has subsequently paralleled, studies of physical mechanisms of aircraft noise generation (2, 3, 7, 12, 14).* These in turn followed earlier work on the unsteady aerodynamics of helicopter rotors (eg ref 1) which led directly to studies of helicopter rotor noise (6, 8, 9).

Following a preliminary review of the potential effects of rocket launcher noise on NASA ground crews (4), physical and subjective aspects of noise were brought together in an analysis of the effects of rotor design parameters on the aural detection range of military helicopters (10) which is described in Chapter 2. This work identified the need for more reliable aural detection criteria and an experimental study was subsequently performed (16) which is the subject of Chapter 3.

Attention was also turned to the problems of noise certification of civil aircraft which was being introduced by American and international civil aviation authorities. The rules had been formulated for subsonic jet transports but not for propeller driven aircraft or supersonic transports. For these types there were unanswered questions about the applicability of the current noise scaling methodology. Two studies were therefore addressed at these questions. The first (5) concerned general aviation aircraft noise and the second, which is the subject of Chapter 4, was concerned with noisiness perception at high levels and low frequencies (11).

* The author's publications on aircraft noise and related topics are listed chronologically at the end of this chapter. Reference numbers here, as elsewhere, refer to literature cited at the end of the current chapter.
Both involved laboratory experiments designed to test the accuracy of the current measurement techniques for estimating the subjectively perceived magnitude of aircraft flyover noise.

This work led to more extensive studies of the problem (13, 15, 25, 40) and two of these are reported in Chapters 5 and 6. The first (25) covers a large scale laboratory experiment to evaluate current noise scaling methodology for various classes of aircraft. Since the publication of that work international noise certification regulations have been extended to cover most types of fixed wing aircraft but helicopters are still excluded. The second study described in Chapter 6 therefore involved a further extensive experiment aimed at resolving some of the difficulties which have so far hindered the introduction of noise certification rules for helicopters (40).

A further area of interest has been the associated subject of airport noise impact which concerns the serious adverse effects of aircraft noise upon people who live near airports (17, 18, 22, 24, 27-32, 35, 39). By comparison with laboratory experimentation, many more factors have to be taken into account when considering the long-term effects of continual noise intrusion upon people in their own homes and quantitative study involves social survey research (29, 30, 31). Chapter 7 describes one of the author's surveys carried out in the vicinity of Heathrow airport. One of the main objectives was to evaluate the Noise and Number Index methodology used in the UK for airport planning purposes but numerous alternative approaches were examined. Of particular interest was the relative nuisance of the noise during day, evening and night and Chapter 8 investigates this question further (35).
Aircraft noise is of course one particular component of the general acoustic environment in which people live and although around airports it is a very dominant one it is not always appropriate to treat the problem in isolation. Therefore throughout the period of the research described herein attention has been given to the more general problems of environmental noise evaluation (19, 20, 21, 23, 26, 33, 36, 37). This is complicated by the fact that people appear to react differently to noise from different sources, e.g., aircraft and trains, but Chapter 9 describes a procedure proposed for the purpose of predicting public reaction to noise from mixed sources.
Reports and Papers on Aircraft Noise and Related Topics
by
J B OLLERHEAD


* Those marked with an asterisk are included, wholly or partly, in this compilation


18 "Noise and Airport Planning", Proc. of Conference on Airports of the Future, Tel Aviv University, November 1972.


26 "Environmental Noise Nuisance", Sozial-und-Präventivmedizin, Switzerland, 19, pp. 169-175, 1974.


33* "Predicting Public Reaction to Noise from Mixed Sources", Inter-noise 78, San Francisco, May 1978.


35* "Variation of Community Response to Aircraft Noise with Time of Day", Noise Control Engineering, September-October 1978, 68-78 (Also presented at Inter-noise 77, Zurich, March 1977).


The author's collaboration with M V Lowson led to the development of techniques for the prediction of helicopter rotor noise which are widely used by the helicopter manufacturing industry for design purposes (references 6, 8 and 9 in Chapter 1). This paper, which was prepared by the author, outlines the Lowson-Ollerhead theory (in Section 3) and its application to investigate the effects of rotor design and operating parameters on aural detection distance. The rotor noise theory, which uses an empirical blade airload model to apply Lowson's theory of rotating source noise to the special geometry of helicopter rotors, agrees well with measured data for both main and tail rotors and shows the importance of the higher harmonic airloads to the noise generation process.

The criterion of aural detection in this analysis was a received noise level of 40 PNdB. This was little more than an indication of the probability of detection but more realistic criteria were not then available. In view of significant role of aural detectability in tactical situations, the need for research in this area was pointed out.
PROBLEMS OF HELICOPTER NOISE ESTIMATION AND REDUCTION

J. B. Oillerhead and M. V. Lawson

Wyle Laboratories Research Staff

Huntsville, Alabama

Abstract

The general problem of helicopter rotor noise generation, propagation and reception is reviewed in the light of recent theoretical work. Prediction methods are described and in several comparisons of theoretical and experimental results good agreement is found. These methods have been programmed for computer solution and a parameter study is presented which demonstrates the effects of the significant variables on the aural detection of rotor noise. A set of design charts is given in the paper which can be used with reasonable accuracy to estimate rotationnal noise spectra for helicopter rotors as a function of design and flight variables. The paper is concluded with recommendations for future work and measures to reduce the noise radiated by a helicopter.

Nomenclature

\( A_{n\lambda} \) \( B_{n\lambda} \) In phase and quadrature pressure amplitude of \( n^{th} \) sound harmonic due to \( \lambda^{th} \) loading harmonic.

\( \sigma_0 \) Atmospheric speed of sound.

\( a_0 \) \( \alpha \) \( \lambda \) \( c \) \( D_0 \) \( f \) \( J_n \) \( M \) \( M_E \) \( M_F \) \( M_r \) \( m \) \( n \) \( P_0 \) \( R \) \( R_0 \) \( r \) \( S \) \( T_0 \) \( x \) \( y \) \( \alpha \) \( \theta \) \( \lambda \) \( \phi \) \( \psi \) \( \Omega \) Rotational Mach number of point airloads \( (M = \Omega R/a_0) \).

Effective rotational Mach number \( \lambda = M V/(1 - M) \).

Translational Mach number of rotor hub.

Component of \( M_r \) in direction of observer.

Sound harmonic number based on blade passage frequency.

Sound harmonic number based on rotational frequency.

Sound pressure amplitude (rms).

Radius of action of point airloads.

Total distance between observer and rotor hub.

Total blade surface area.

Rotor thrust.

Blade thickness.

Blade tip speed.

Axial distance between observer and rotor hub measured positive in direction of thrust action.

Component of \( r \) in rotor disc plane.

Blade pitch angle.

Angle measured from disc plane, positive toward rotor shaft.

Loading harmonic number based on rotational frequency.

Angle measured from rotor shaft, positive toward rotor disc.

Rotor blade azimuth position.

Rotational speed \( (\text{rads/sec}) \)

Nomenclature

\( a_0 \) \= \text{Atmospheric speed of sound.}

\( a_0 \) \= \text{Atmospheric sound speed.}

\( \alpha \) \= \text{Axial distance between observer and rotor hub measured positive in direction of thrust action.}

\( \beta \) \= \text{Angle measured from blade passage frequency.}

\( \beta \) \= \text{Angle measured from disc plane, positive toward rotor shaft.}

\( \lambda \) \= \text{Angle measured from rotor shaft, positive toward rotor disc.}

\( \psi \) \= \text{Angle measured from rotor shaft, positive toward rotor disc.}

\( \Omega \) \= \text{Rotation speed (rads/sec).}

\( M \) \= \text{Rotational Mach number of point airloads (M = \( \Omega R/a_0 \)).}

\( M_E \) \= \text{Effective rotational Mach number} \( \lambda = M V/(1 - M) \).

\( M_F \) \= \text{Translational Mach number of rotor hub.}

\( M_r \) \= \text{Component of \( M_r \) in direction of observer.}

\( m \) \= \text{Sound harmonic number based on blade passage frequency.}

\( n \) \= \text{Sound harmonic number based on rotational frequency.}

\( P_0 \) \= \text{Sound pressure amplitude (rms).}

\( R \) \= \text{Radius of action of point airloads.}

\( R_0 \) \= \text{Total distance between observer and rotor hub.}

\( r \) \= \text{Total blade surface area.}

\( T_0 \) \= \text{Rear thrust.}

\( V_T \) \= \text{Blade thickness.}

\( x \) \= \text{Blade tip speed.}

\( y \) \= \text{Axial distance between observer and rotor hub measured positive in direction of thrust action.}

\( \alpha \) \= \text{Component of \( r \) in rotor disc plane.}

\( \alpha \) \= \text{Blade pitch angle.}

\( \theta \) \= \text{Angle measured from disc plane, positive toward rotor shaft.}

\( \lambda \) \= \text{Angle measured from rotor shaft, positive toward rotor disc.}

\( \psi \) \= \text{Rotor blade azimuth position.}

\( \Omega \) \= \text{Rotation speed (rads/sec).}

1
1.0 Introduction

As military and commercial utilization of the helicopter becomes more and more widespread, so too does the concern about helicopter noise. In its present form the helicopter represents a complex but nevertheless reliable and extremely valuable form of transport vehicle. However, in the quest for improvement in performance, the acoustic problem has to large extent been ignored and we are now faced with the consequence that one of the major hurdles obstructing further advancement is that of noise. An increasing amount of effort has been devoted to this subject in recent years but there still exists a considerable amount of confusion. The purpose of this paper is to review the problem of helicopter noise in the light of recent theoretical work, to indicate potential methods for its control and to recommend possible avenues for further research.

The helicopter noise problem can be divided into three distinct areas as illustrated in Figure 1. The noise is generated by a large number of sources, predominant amongst which are the rotors, the engine compressor and exhaust, and the gearboxes. The occupants of the machine are protected to some extent by the fuselage structure and whatever soundproofing materials have been installed. However, it is very difficult to attenuate the low frequency sound which is a major proportion of rotor noise and internal sound levels are generally high. This is a problem to all helicopter operators from the standpoints of comfort, safety and communications. The external noise causes two problems. At short distances, of less than about one thousand feet, it is annoying to the people exposed to it. This is a problem to commercial operators, to civic authorities, and to the communities they serve. At much greater distances, of the order of several thousand feet, we come to the military problem of aural detection which is a significant factor in tactical operations. It is important to recognize that although these three aspects of helicopter noise are obviously strongly related, they are different problems and should be treated independently. For example, the helicopter which is judged to be the least noisy by people immediately below its flight path is not necessarily the least detectable at greater distances and vice versa. Furthermore, the potential benefits of noise reduction may be significantly different. Whereas community annoyance can probably be eased by lowering noise levels by a few decibels, the detectability could conceivably be increased if such a reduction were accompanied by a decrease in performance to give a net increase in detection time. The delicate balance between noise and performance requires particularly careful evaluation in the problem of aural detection.

2.0 Noise Propagation and Reception

Before proceeding to the main problem of helicopter noise generation and reduction at the source, it is as well to examine the basic effects of sound propagation and the reception of sound by people. These two factors are of vital importance to the over-all problem and yet, unfortunately, they are very poorly defined at the present time. This is particularly true in the case of aural detection where each is equally important as the noise source characteristics.

2.1 Propagation Effects

Each of the five major factors which influence the sound actually observed by a listener is discussed below in turn.

2.1.1 Spherical Spreading. In an ideal medium, the total sound power radiated through an expanding spherical wavefront remains constant so that sound pressure levels are reduced by 6 dB each time the distance from the source doubles. Two effects cause deviations from this rule at small source to receiver distances which are within the "near field." Firstly, the near field is a region in which the physical dimensions of the source region are important so that sound reaches the observer from various directions. Secondly, at distances which are small compared with typical sound wavelengths, non-propagating "hydrodynamic" pressure fluctuations amplify the apparent sound levels.

If the source is within a few wavelengths of the ground, sound reflection effects will affect propagation characteristics. These include amplifications due to an effective increase in power when the height is small compared with a wavelength and due to interference between the direct and reflected signals. Variations in the far field sound levels of up to 6 dB are possible.

2.1.2 Atmospheric Absorption. This is due to atmospheric energy transfer processes of which two types predominate: Molecular absorption losses associated with resonance effects in polyatomic gases, and classical absorption losses which are inherent in all gases due to basic gas transport phenomena. Classical absorption causes excess atmospheric attenuation (over and above divergence losses) which, when measured as an attenuation constant (in dB per 1,000 ft) is proportional to frequency squared. There is also a component of molecular absorption which has the same dependence on frequency and taken in combination, these components yield an observed attenuation constant of $5.3 \times 10^{-8} f^2$ dB/1,000 ft, where $f$ is the frequency in Hz. Molecular absorption absorbs acoustic energy through a relaxation phenomenon of air molecules which are excited into resonance by the sound wave. In air the principal effect involves an interaction of water vapor...
molecules with oxygen molecule resonances so that molecular absorption is highly dependent upon atmospheric humidity. New empirical equations for molecular losses derived in Reference 1 are plotted as a function of frequency and humidity at a temperature of 59° F in Figure 2. Also shown in the same figure is the curve represented by the classical loss equation given above. This illustrates the importance of molecular absorption at low frequencies, typical of helicopter rotational noise, where it is in fact the only known absorption mechanism. Even so, at a frequency of 10 Hz the attenuation constant is a mere 0.01 dB/1,000 ft. At 5,000 Hz the total access attenuation constant is of the order of 20 dB/1,000 ft, so that the low atmospheric attenuation of the low frequencies is clearly a significant factor in helicopter detectability.

2.1.3 Ground Attenuation. Terrain effects range from the dissipation of acoustic energy at the edge of a sound wave traveling parallel to the ground to the direct impedance offered to the path of a sound wave penetrating a very leafy jungle. Very little data on these effects are available and those experiments which have been performed (e.g., References 2, 3, 4 and 5) have provided very limited results, largely because of the large number of variables which should be taken into consideration. Even the various results which are available conflict with each other.

Loewy6 condensed the results of References 2 and 7 into a single table which is reproduced here as Table 1 in order to show the magnitude of terrain absorption. These results represent measured attenuation coefficients for sound waves traveling parallel to the ground over or through the described vegetation. They are expressed, as before, in dB/1,000 ft units but Dobbins and Kindick4, in a more recent study, found that attenuation constants are extremely sensitive to absolute distance from the sound source. Their results, measured for pure tones in a dense jungle are summarized in Table II. The variations of attenuation constants with distance are seen to be large and the difficulty of making use of these data for prediction purposes is apparent.

Probably the most valuable set of experimental data available, although still somewhat limited for practical use, was provided by Wiener and Keast5. They measured ground absorption coefficients for octave bands of random noise propagation over a two-mile stretch of scrubby grassland about one foot in depth. The terrain was extremely flat over the whole test range.

Table 1

<table>
<thead>
<tr>
<th>Octave Band Number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency (Hz)</td>
<td>20</td>
<td>1.75</td>
<td>150</td>
<td>300</td>
<td>600</td>
<td>1200</td>
<td>2400</td>
<td>4800</td>
</tr>
<tr>
<td>Conditions</td>
<td>To</td>
<td>To</td>
<td>To</td>
<td>To</td>
<td>To</td>
<td>To</td>
<td>To</td>
<td>To</td>
</tr>
<tr>
<td>Dense Jungle</td>
<td>0.4</td>
<td>30</td>
<td>46</td>
<td>60</td>
<td>75</td>
<td>110</td>
<td>160</td>
<td>210</td>
</tr>
<tr>
<td>Sparse Jungle</td>
<td>0.4</td>
<td>6.8</td>
<td>11</td>
<td>13.5</td>
<td>18</td>
<td>27</td>
<td>45</td>
<td>65</td>
</tr>
<tr>
<td>Grass 18 In. High</td>
<td>---</td>
<td>2.0</td>
<td>11</td>
<td>27</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>32</td>
</tr>
</tbody>
</table>
TABLE II

**TERRAIN LOSS COEFFICIENTS COMPUTED FOR VARIOUS DISTANCES THROUGH JUNGLE (From Reference 4)**

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>Distance Between Reference and Exploring Microphones (ft)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25 (dB/ft.)</td>
</tr>
<tr>
<td>63</td>
<td>-0.096</td>
</tr>
<tr>
<td>250</td>
<td>0.348</td>
</tr>
<tr>
<td>1000</td>
<td>0.004</td>
</tr>
<tr>
<td>4000</td>
<td>0.032</td>
</tr>
<tr>
<td>8000</td>
<td>2.68</td>
</tr>
</tbody>
</table>

* Least Squares

Their results are shown in Figure 3 where the attenuation values have been collapsed on the basis of the parameter \( f_r \), the product of band center frequency and source-receiver distance. The data is seen to collapse in the form of two straight lines with zero attenuation below \( f_r < 4 \times 10^5 \) and increasing at 3 dB per doubling of \( f_r \) for \( f_r \geq 4 \times 10^5 \). However, significant differences are found in some measurements in the 300 to 600 Hz band, where the attenuation is substantially increased at the lower values of \( f_r \). Although there is theoretical support for the trends observed in the other frequency bands, the reason for this latter effect, which is inevitably found in sound propagation experiments, is not clear at the present time. Two possible explanations are that it is a function of the heights of the source and microphone above the surface or, that it results from variations in ground impedance with frequency. Whatever the reason this effect has an important influence on helicopter detection.

Wiener and Keast's data were measured for sound propagation parallel to the ground. In general, however, aircraft sound rays are not parallel to the ground but inclined at an angle to it. As might be expected, it is found that the importance of ground cover effects increase as this angle decreases. In fact, Reference 8 shows that when the elevation angle is less than 7 degrees, ground absorption loss exceeds that due to divergence and all atmospheric effects. This is illustrated in Figure 4 which shows the attenuation constant for ground cover absorption of sound in the 150-300 Hz octave band as a function of elevation angle. It is important to note that the audible ranges of helicopters in tactical situations are generally such that the elevation angle is small so that terrain effects represent a major factor in aural detectability.

2.1.4 Refraction Effects. Atmospheric wind velocity and temperature gradients cause bending of sound rays by refraction, resulting in propagation phenomena known as "shadow zones" and "focusing." A shadow zone within which the sound cannot be heard, will form whenever the...
speed of sound decreases with altitude, due to changes in either temperature or wind speed, and causes an upward refraction of the sound rays. For a negative temperature gradient the shadow zone boundary is a circle with the sound source as its center. This phenomenon is undoubtedly important from an operational standpoint because atmospheric temperature gradients vary significantly during the day, especially in summer. The consequent variation in sound propagation characteristics is easily observed, for example, in the vicinity of airfields. Early in the morning, aircraft ground movements can be heard several miles from the airfield, whereas late in the afternoon, when ground temperature has risen appreciably, this sound rarely travels more than a mile or so close to the ground. This "temperature refraction" alone could alter aural detection range by an order of magnitude. In a wind gradient the shadow zone boundary begins upwind and takes a shape similar to that sketched in Figure 5. In fact, some sound does penetrate a wind generated shadow zone but the attenuation is high, its exact value depending on frequency and direction.

In contrast, sound focusing decreases nominal attenuation losses. This reinforcement is the reverse of the shadow zone phenomenon and occurs when the speed of sound increased with altitude, as it does for example, in a temperature inversion. Focusing can also occur downwind of a sound source although the effect is generally the result of wind speed variations at high altitude. However, it can cause increases in sound pressure levels of some 10 to 20 dB at distances of several miles from the source. Both effects, due to upward and downward refraction are illustrated in Figure 6.

2.1.5 Scattering by Turbulence. This is a further important source of atmospheric attenuation for low frequency sound. It involves, not a dissipation of sound energy but a redirection, and is akin to the refraction effects described above. Its principal effect, in an irregular sound field such as that produced by a helicopter, is to equalize acoustic energy propagating in all directions at large distances from the source. This is a direct result of scattering of the sound field by the non-uniform sound velocity distribution in atmospheric turbulence. Thus, a highly directional sound profile can be gradually rounded out, tending to a nondirectional pattern at great distances from the source. Another important but non-quantifiable effect of scattering is to cause random fluctuations in the sound intensity at long distances from the source. The magnitude of such fluctuations can be as high as 20 dB.
2.2 Human Receiver Characteristics

The threshold of hearing for an average young man is shown in Figure 7. This represents the level at which a pure tone just becomes audible as a function of its frequency. In fact, there is considerable variation of this threshold from person to person but this figure, from Reference 9, is widely used and represents a good average. The same curve also holds good for the audible levels of bands of noise, whose widths are critical bands, which will be discussed shortly. This curve shows that the hearing mechanism is most sensitive to sound in the 3,000 to 4,000 Hz frequency range and that sensitivity is much reduced at low frequencies. The ability of a listener to hear a particular acoustic signal is a function of its spectrum in relationship to the threshold level and also the extent to which it is masked by background noise.

It is relevant to consider the masking of a pure tone by noise, defined as sound with its energy spread over a band of frequencies. If the level of the pure tone is decreased until it is just inaudible in a wide band of noise whose spectrum is flat, it is found that the tone remains inaudible as the bandwidth of the noise is decreased to a value known as the critical bandwidth (with a center frequency equal to that of the tone). As the bandwidth is further diminished the tone is heard and becomes increasingly "loud" relative to the background noise. Furthermore, it is found that at the critical bandwidth the tone is just audible when its level is equal to that of the noise. Figure 8 shows the variation of the critical bandwidth with frequency.

Figure 8 shows some typical octave band background noise levels measured in a jungle environment from Reference 4. Combining the lower boundary from this figure with the octave band threshold level from Figure 7 yields a tentative detection threshold for aircraft noise which is plotted in Figure 10. Superimposed on this graph are a number of curves which show the variation of a typical helicopter noise spectrum with range. This figure illustrates the suppression of high frequency sound by the attenuation effects previously described and indicates the increasing importance of the low frequency sound at increasing distances. Furthermore, it is important to recognize that the data represented in Figure 7 were obtained included in Figure 7. In the presence of background noise, which is always the case in reality, the problem becomes more complicated, especially when the signal consists of a complex combination of harmonic, pulsatile and modulated random noise as does helicopter noise.
using earphones. There is a distinct possibility that other non-auditory physiological effects (such as "feel") could be important in defining the low frequency threshold of subjects who are completely exposed to the sound field. In addition, sound waves can excite ground vibrations which again might be detected by an observer before the sound is heard.

However, with the exception of reciprocating engine exhaust-noise the "mechanical" sound sources are important to the internal and near-external sound fields only. In the far external field the aerodynamically generated sound is dominant. At medium distances from a single lifting rotot helicopter, Cox and Lynn found that the various sources, listed in the order of their importance to the subjectively judged magnitude of the sound, or loudness, are:

- Blade slap (when it occurs)
- Piston engine exhaust noise
- Tail rotor "rotational" noise
- Main rotor "vortex" noise
- Main rotor "rotational" noise
- Gearbox noise
- Turbine engine noise
- Other sources.

Some of these sources are identified in the spectrum shown in Figure 11 which is adapted from Reference 10.

However, at very great distances from the helicopter, where its sound is barely audible, we may expect sound propagation effects to modify the above rank listing to:

- "Blade slap"
- Main rotor "rotational" noise.

Depending on the specific helicopter configuration, piston engine exhaust noise and tail rotor "rotational" noise should possibly be included with these two principal sources although it is likely that their levels will be appreciably lower. This example illustrates the importance of observer position in the helicopter noise problem. In the absence of "blade slap," main rotor "rotational" noise is predominant at great distances although it is nearer to the bottom of the list at closer positions.

The quotation marks have been used with the terms "blade slap", "rotational" and "vortex" because the need, or justification, for distinguishing between these various descriptions of rotor noise is unclear. Furthermore, the use of these expressions has led to considerable confusion in the understanding of rotor noise. The main distinction between them is in fact mostly subjective and it is as well to describe them in terms of their subjective characteristics. "Rotational" noise is the component of rotor noise which is positively identified as harmonic in nature and dominates the low frequency end of the spectrum. For a main rotor, it is the steady thumping sound with a frequency equal to the blade passage frequency. The characteristic sound of a tail rotor or a conventional propeller is almost entirely rotational noise. "Vortex" noise is heard as the higher frequency swishing sound at points close to the main rotor and is essentially random noise modulated in amplitude and frequency by the motion of the blades. "Blade slap" is familiar as a sharp cracking or banging sound which occurs during certain maneuvers at high speed and conditions of low inflow. The sources of each type of sound will now be discussed in turn.

![Figure 10. Variation of Helicopter Sound Spectrum with Range](image)
3.1 Rotational Noise

The fundamental mechanism of rotational noise has been understood since the work of Gutin11 gave a theoretical solution for the sound radiated by a propeller under the action of steady thrust and torque forces. By considering the rotary motion of a single aerodynamic force having both lift and drag components Gutin showed that noise was generated at all harmonics of the fundamental blade passage frequency. The noise arises merely because the point of action of the force undergoes an oscillatory motion with respect to a listener. In the notation of this paper, the sound pressure amplitude of the mth sound harmonic is given as

\[ C_{mb} = \frac{mB\Omega}{2\pi a_r} \left( \frac{T_0}{r} - \frac{D}{rM} \right) \]

which has a finite value for all values of m. This formula is well known and has proved useful for estimating propeller noise for at least the first few harmonics. However, when it was used in attempts to calculate the noise of helicopter rotor it was found that it grossly underestimated the sound pressure levels of all harmonics other than the first. An obvious possible source of this discrepancy lay in the fact that helicopter blade airloads are far from steady. Loewy and Sutton12 and Schlegel et al.13 extended Gutin's work to investigate this possibility and, using digital computers, were able to obtain numerical solutions for rotor noise due to both steady and harmonically varying lift and drag forces. Their results showed that the fluctuating airload components made a very substantial contribution to rotor noise and they were able to improve the correlation between theory and experiment by an order of magnitude. Using available experimental airload data which gave up to 10 harmonics of the differential blade pressure fluctuations, reasonable agreement was found with the first three or four sound harmonics. However, beyond these the correlation again deteriorated rapidly.

In recent theoretical studies of the rotor noise problem14, 15, 16 the authors took a closer look at the effects of the fluctuating airloads. Starting from Logan's earlier result for the sound field of forces in motion17 two approaches were followed. In the first, the general acoustic equation was integrated numerically, by computer, for the particular case of the helicopter rotor. This analysis admitted higher harmonic blade motions as well as airloads and included all near field effects. In the second, by making a number of assumptions which are valid for the far field case (that is for points which are sufficiently far from the rotor for it to be considered a point source) a closed form solution for the noise field was developed. This offered the significant advantage that computer time was reduced by a factor of around 100 which in turn enabled a fairly thorough parameter study to be performed.
where the argument of all Bessel Functions \( j \) is \( NMy/r \).

In this equation the three force components are defined by the Fourier summations:

\[
C_n = \sum_{\lambda=0}^{\infty} \frac{i^{(n-\lambda)}}{4\pi} \left[ \frac{\alpha\lambda T}{\alpha} \left( \frac{1}{n-\lambda} \right)^{\lambda} \right] \left( \frac{1}{n-\lambda} \right)^{\lambda} J_n(\lambda n+\lambda) \\
-\frac{b\lambda T}{\alpha} \left( \frac{1}{n-\lambda} \right)^{\lambda} J_n(\lambda n+\lambda) \\
+ \frac{b\lambda T}{\alpha} \left( \frac{1}{n-\lambda} \right)^{\lambda} J_n(\lambda n+\lambda)
\]

\[
T(\psi) = a_T + \sum_{\lambda=1}^{\infty} \alpha_{\lambda} T \cos \lambda\psi + b_{\lambda} T \sin \lambda\psi
\]

\[
D(\psi) = a_D + \sum_{\lambda=1}^{\infty} \alpha_{\lambda} D \cos \lambda\psi + b_{\lambda} D \sin \lambda\psi
\]

\[
C(\psi) = a_C + \sum_{\lambda=1}^{\infty} \alpha_{\lambda} C \cos \lambda\psi + b_{\lambda} C \sin \lambda\psi
\]

The radial component arises due to out of plane deflections of the blade surface, for example due to coning.

If the radial forces and the harmonic components of thrust and drag are put equal to zero this equation reduces to Equation (1) which is Gustin's result for the steady loaded propeller.

For a rotor with \( B \) blades, Equation (2) gives an identically zero result for all values of \( m \) which are not integral multiples of \( B \), so that \( m \) may be replaced by \( mB \) where \( m \) is thus the sound harmonic number. Also, the harmonics which are multiples of \( B \) are additive so that the force terms must be multiplied by \( B \) to become the total rotor thrust, drag and radial forces based on the blade passage frequency. It may be seen from Equation (2) that any sound harmonic receives contributions from all loading harmonics \( \lambda \) (where \( \lambda \) is the loading harmonic number based on the rotor rotational frequency \( B \)). Physically, this is a manifestation of the Doppler effect. If a single frequency sound source moves in a circular path, then as it approaches a listener an increase of frequency is observed whereas a lower frequency is heard as the source recedes. The net result is a continuous modulation of the observed frequency which, when reduced to Fourier terms, consists of an infinite number of harmonic components, with a peak in the region of the basic source frequency. If now the source has a large number of harmonic components, each generates sound at each harmonic of the blade passage frequency as shown by Equation (2). The effect is illustrated in Figure 12 which shows the calculated contributions of the first 60 loading harmonics on a four blade rotor to a number of sound harmonics. It will be seen that each

\[
mB(1 - M) < \lambda < mB(1 + M)
\]

should be included to calculate the level of the \( mB^{th} \) sound harmonic. More importantly, this figure indicates the necessity to include an adequate number of loading harmonics in any calculation of the noise. For example significant contributions to the eighth sound harmonic (\( m = 8 \)) are made by all loading harmonics between the fourteenth and the fiftieth. Omission of any of these loads may be expected to result in a significant error. It was shown in fact in Reference 14 that all loading harmonics in the approximate range

\[
\text{Figure 12. Acoustic Contribution of Loading Harmonics Calculated for a Field Point 10 Degrees Below Rotor Disc}
\]
A study of the available full scale blade loading data of References 18 and 19 revealed that the amplitudes of the airload harmonics, as obtained from integrated differential pressures measured around the 80 percent radius station, decayed approximately as some inverse power of harmonic number, at least within the range of the data which covered the first ten harmonics. Some typical plots are shown in Figure 14. Surprisingly little variation of the exponent of this power law was found between all the steady forward-flight cases examined, even though the rotor advance ratio varied between zero and 0.3. Furthermore, the data were obtained for two very different helicopter types, namely the two-blade UH-1 and the four-blade CH-34. For steady flight out of ground effect the optimum value for the exponent was found to be -2.0 so that the amplitude of the $\lambda$th loading harmonic is proportional to $\lambda^{-2.0}$.

![Figure 13. Range of Effective Contribution of Loading Harmonics to Sound Radiation](image.png)

Consider first of all the chordwise pressure variations. Previous investigators have considered a variety of pressure profiles ranging from rectangular to experimentally measured variations. However, chordwise variations will only be of importance for acoustic frequencies which have wavelengths of the order of the chord length and less. Typcally this corresponds to frequencies greater than about 500 Hz, i.e., harmonics greater than about the fifteenth. Even assuming a uniform chordwise pressure distribution, several hundred loading harmonics would be required to accurately define the sound field. Only at higher frequencies could improved definition of the chordwise pressure pattern be expected to further increase accuracy. Inclusion of these variations at lower frequencies therefore only results in spurious changes to the results, effectively introducing just a part of the high frequency inputs. For this reason the authors consider that a point chordwise load is perfectly adequate for present purposes.

Before discussing spanwise loading variations it is well to consider the source of the harmonic blade airloads and what we know about them. They result from basic rotor flow asymmetries, wake effects and from non-linear effects at high angles of attack and high Mach numbers. Rotor aerodynamics is an exceedingly complex three-dimensional problem and at the present time even the accurate prediction of low frequency fluctuations, for the purposes of calculating blade vibration response, is a formidable task. Since the frequencies of importance to noise are generally much higher, the acoustic problem is even more difficult and available experimental and theoretical airload data are of limited value.

![Figure 14. Rotor Loading Harmonic Laws at Various Advance Ratios. Data from Scheiman (18)](image.png)

In the studies reported in References 14 and 16 it was assumed that this law could be extrapolated indefinitely to higher frequencies in order to provide some estimate of the higher airload harmonic levels. However, before this result could be used as a basis for noise computations account had to be taken of the loading phase variations around the rotor azimuth and along the blade span. Phase variations over a distributed acoustic source are of first order importance and lack of phase information caused some early difficulties. No obvious phase trends could be isolated from the measured data and as expected, the computed acoustic levels were very sensitive to the phase details assumed. This was particularly true at the higher frequencies. Consequently it was assumed that the phases could be randomized. In the case of the spanwise loading variation this was accomplished by the introduction of a "correlation length" concept such as is commonly used in turbulence theory. This allows for the decreasing correlation between the fluctuating pressures at two points on the blade surface as the distance between them increases. Very simply the correlation length can be thought of as the distance over which pressure fluctuations act in phase with
each other. Pressures outside this region have no coherent phase relationship with those inside it. In this case it was assumed that the correlation length is inversely proportional to frequency (or proportional to wavelength) and this had the approximate net effect of adding a further 0.5 to the exponent of the loading power law. In the azimuth sense the phase problem was eliminated by decorrelating the contributions of successive loading harmonics to any particular sound harmonic. If the phase relationships are known then the amplitude of the $n^{th}$ sound harmonic is

$$C_n = \sqrt{\sum_{\lambda} A_{n\lambda}^2 + \sum_{\lambda} B_{n\lambda}^2} \tag{5}$$

where $A_{n\lambda}$ and $B_{n\lambda}$ are the in-phase and quadrature sound pressure amplitudes due to the $\lambda^{th}$ loading harmonics. If the phases between the various $\lambda$ contributions are known to be completely random, then the pressure amplitude can be written

$$C_n = \sqrt{\sum_{\lambda}(A_{n\lambda}^2 + B_{n\lambda}^2)} \tag{6}$$

The accuracy of this assumption is unknown although it is likely to improve with increasing frequency as the source of the airload fluctuations becomes more random.

The only remaining problem is that of dividing the radial airload distributions into a set of point forces for the purposes of computation. To investigate this, cases were computed after segmenting the loading into 1, 5, and 10 radially spaced components. If the loading phase was retained, very large differences were found between the three cases and the discrepancies increased with frequency. However, under the assumption of random phase the error was practically eliminated and Figure 15 shows a comparison of the computed acoustic spectrum for the three loading breakdowns. The largest difference between them was approximately 2 dB and at harmonics above the 4th the difference was negligible. Consequently, the use of a single loading point at the 80 percent radial station was adopted since this appears to give adequate results for a helicopter rotor and minimizes the computational effort.

Figure 16 shows a comparison between theory and experiment for the H-34 main rotor noise. The measured data are from Reference 13 and the theoretical results were computed using Equation (5) and the various assumptions outlined above, including a loading law based on the $-2.5^\text{th}$ power of harmonic number. For comparison the theoretical results of Schlegel et al.13 and those using Gutin's theory11 are also included. The large errors of the simple theory are apparent. Further comparisons between the present theory and experimental results are shown in Figures 17, 18, and 19, all of which correspond to hovering rotors. Figure 17 includes measured data for the UH-1 two blade main rotor from three sources. One set was obtained from the 6 percent bandwidth spectrum analysis shown in Figure 11 and the other two were obtained by Wyle Laboratories through a 2 Hz bandwidth analysis of two separate noise
recordings. All three sets of data have been normalized on the basis of the third harmonic sound pressure levels because uncertainties regarding the low frequency response of the instrumentation and microphone location caused doubts about the over-all level. The curve is only useful therefore to compare the spectrum shapes and the agreement is seen to be good. Figures 18 and 19 compare results for the main and tail rotors respectively of a light observation helicopter and in this case, the correlation of both over-all level and spectrum shape is satisfactory.

3.1.1 Effect of Forward Speed. It was shown in Reference 17 how the effects of steady translation of the rotor hub could be accounted for by replacing the term \( r \) in the stationary case (Equation (3)) by \( r (1 - M) \) where \( M \) is the component of the hub translation Mach number in the direction of the observer. However, in applying this transformation it is important to remember that both \( r \) and \( M \) relate to the "retarded" position of the hub, that is its location when the rotor emitted the sound under consideration. Its position at the instant of observation is different and to relate the two requires another transformation discussed in References 14 and 15.

![Figure 18. Correlation of Measured and Predicted Main Rotor Rotational Noise Harmonics for Two-Bladed Rotor](image1)

![Figure 19. Correlation of Measured and Predicted Tail Rotor Rotational Noise Harmonics for Two-Bladed Tail Rotor](image2)

![Figure 20. Variation of Sound Directivity Patterns With Forward Speed for Various Harmonics](image3)
Figure 20 shows the computed effects of rotor translational velocity. Figure 20(a) illustrates the variation of harmonic sound pressure levels, in a plane containing the rotor shaft, at a distance of 1,000 ft around a typical hovering rotor. The main features of this plot are a directional peak at approximately 30 degrees below the disc plane and a distinct minimum some 10 degrees above it. As far as is known these features have not been experimentally confirmed for a rotor although the existence of a similar effect for a propeller is known. Parts (b) and (c) of the same figure correspond to translational Mach numbers of 0.125 and 0.25. The symmetry of the directivity patterns about the shaft axis exhibited by the hover case has been replaced by a forward movement of the lobes so that more sound is radiated forward than aft. These figures take into account the retarded position effects discussed above so that the results correspond to sound pressure levels observed 1,000 feet from the instantaneous position of the rotor. The directivity lobes determined with respect to the helicopter would show even more exaggerated profiles.

3.1.2 Design Curves. In Reference 14, it was shown how the closed form solution for rotor noise, Equation (2), could be revised and, by means of a further approximation, simplified to enable a set of design charts to be developed for the purposes of estimating rotational noise levels. The charts themselves and instructions for their use, are presented in the Appendix. In this section the method and its limitations are briefly described.

Equation (2) gives the result for the magnitude of the ninth sound harmonic of a B-bladed hovering rotor (when \( n \) is replaced by \( m \)). For the purposes of this discussion this equation can be written, in a simplified form:

\[
C_n = \sum_{\lambda=0}^{\infty} K \frac{1}{r} \left\{ \frac{nM}{R} \sin \theta C_{\lambda T} J_1' - \frac{C_{\lambda D}}{R} J_2' \right\}
\]

\[+ \frac{nM}{R} \cos \theta C_{\lambda C} J_3' \]  

where \( K \) is a constant and \( \theta \) is the angle between the observer and the disc plane. The \( J \) terms are complex collections of Bessel functions of argument \( nM \cos \theta \). Now if the helicopter is translating at Mach number \( M_r \) the sound that reaches the observer at \( r \), (relative to the helicopter) was generated at some previous instant, when the relative position of the observer was \( r', \theta' \). When the necessary transformations are applied to account for the forward speed, Equation (7) becomes:

\[
C_n = \sum_{\lambda=0}^{\infty} K \frac{1}{r'} \left\{ \frac{nM}{r'(1-M_r^2)} \sin \theta' C_{\lambda T} J_1' - \frac{C_{\lambda D}}{R} J_2' \right\}
\]

\[+ \frac{nM}{R} \cos \theta' C_{\lambda C} J_3' \]  

where the Bessel functions involved now have the argument \( nM \cos \theta'/r'(1-M_r^2) \), and \( M_r' \) is the component of \( M_r \) along the line \( r' \). But \( r'(1-M_r^2) = r \) so that

\[
\frac{\sin \theta'}{1-M_r'^2} = \sin \theta \quad \text{and} \quad \cos \theta' = \frac{\cos \theta}{1-M_r^2}
\]

and Equation (8) simplifies to

\[
C_n = \sum_{\lambda=0}^{\infty} \frac{1}{r} \left\{ \frac{nM}{R} \sin \theta C_{\lambda T} J_1' - \frac{C_{\lambda D}}{R} J_2' \right\}
\]

\[+ \frac{nM}{R} \cos \theta C_{\lambda C} J_3' \]  

Note that Equation (10) is identical to Equation (9) with the exception of the additional \( (1-M_r^2) \) in the denominator of the thrust term. However, in the case of the helicopter the direction of translation lies close to the rotor disc plane where \( \sin \theta \), and consequently the importance of the thrust term is small. At larger values of \( \theta \) where the thrust term becomes significant, \( M_r' \) becomes small and so also does the error due to the inclusion of the \( (1-M_r^2) \) term in the denominator. To evaluate this possibility, two cases were computed. The first was obtained using the correct forward speed transformation and the second was computed for hovering conditions but using the appropriate value of \( M_r' \) for each observer position. A comparison between the two cases is shown in Figure 21 for \( M = 0.5, \) \( M_r = 0.125. \) The sound pressure level variations for the first, second, fourth and eighth harmonics of the four blade rotor are included. Figure 21 shows that the maximum error in the "Effective Mach Number" case is less than 2 dB, and the error is generally considerably less than 1 dB.

The advantage of this approach is that it enables forward speed transformations to be made in order to interpret hover results as equivalent forward speed results. If a constant relationship is assumed between the thrust, drag and radial components of the blade airloads, and adopting the inverse power law form for the harmonic airloads, the sound intensity for the ninth harmonic can be represented by a relationship of the form

\[
C_n n \propto \frac{T^2}{R^4} f(mB_r, M, \theta).
\]
where A is the rotor disc area and the function, f, can be inferred from the equations above.

Thus the harmonic sound pressure level is reduced to a function of the three variables \( m_B \), \( M \), and \( \theta \) and the design charts given in the Appendix are simply computed sound intensities for a range of these variables (specifically \( m_B \) = 2 through 60, \( M \) = 0.1 through 1.0, and \( \theta \) varies between \(+\)15 degrees and -90 degrees above the rotor disc). Values of \( C_L^2 \) are read from these charts and corrected for total rotor thrust, observer distance and disc area. The ratio between thrust, drag and radiol component of force was chosen as 10:1:1 which is thought to be reasonably typical. The loading power low exponent was, as before, -2.5.

Quite clearly, the charts are based on many assumptions and for this reason their validity may be questioned for general purpose use. They are presented however, as an example of how the results of the theory can be condensed into a useful form. They have certainly proved valuable in performing basic design calculations and comparisons of predicted and experimental results have shown remarkably good agreement.

3.2 Broadband Noise

Broadband rotor noise, commonly known as "vortex" noise has a very distinctive character, subjectively, but probably due to the difficulties of both theoretical and experimental analysis a detailed physical description has yet to be put forward. It has generally been thought that the broadband component dominates main rotor noise at frequencies above about 150 Hz or so, but ultra-narrow band analysis of helicopter noise reported in Reference 14 indicates that the harmonic noise is significant at least up to frequencies of about 400 Hz. Figure 22 shows this 2 Hz bandwidth analysis and the existence of harmonic peaks up to fairly high frequencies can be seen. Nevertheless the peaks do become less clearly defined with increasing frequency, implying a gradual increase of "randomness".

This result is not surprising. Those airflow fluctuations which make significant contributions to the very low frequency rotational noise harmonics are highly repetitive in successive blade revolutions and it is to be expected that the resulting sound has a well defined harmonic form. At higher frequencies the airflow fluctuations result from smaller scale aerodynamic disturbances which are more irregular and azimuthal variations can be expected to differ in different blade revolutions. The net effect of such differences on the sound spectrum will be to increase the width of the "discrete frequency" peaks. When the source pressure fluctuations become sufficiently random the peaks merge into each other and become indistinguishable as peaks. Thus it seems logical to consider rotor noise as being a gradual transgression from harmonic to broadband noise as frequency increases.
It is worthwhile to note at this point that one of the problems that has possibly hindered an understanding of rotor noise is the difficulty of performing adequate narrow band analysis of tape recordings. Main rotor noise harmonics are typically spaced at frequency intervals of 10 - 15 Hz requiring a very narrow filter bandwidth, preferably of the order of 2 Hz or less. This in turn requires a very long recording sample which is almost impossible to obtain accurately. Further, the signal-to-noise ratios of present day tape recorders are not adequate to give any resolution at the higher harmonics. For example, in Figure 22, the harmonics around 400 Hz are as much as 45 dB below the overall level. Good tape recorders have signal-to-noise ratios of around 50 dB and it is clear that special recording techniques are necessary to analyze data at frequencies above 400 Hz or so.

Another problem is that conventional random data analysis techniques do not give information on the modulation frequencies or amplitudes. The broadband noise is modulated by the same mechanisms discussed in Section 3.1, but spectrum analyzers give time averaged levels. Very wide band noise should not fluctuate in amplitude to any significant degree due to frequency modulation because the loss of power from any given band should be balanced by a gain in power from an adjacent band. Even with a very peaked source spectrum and a frequency modulation of the order of an octave, no more than 2 or 3 dB amplitude modulation is likely to result in any given frequency band. Direct amplitude modulations due to varying source distance as the blades rotate will also be very small at moderate distances from the rotor, although they should become significant at points very close to the rotor.

By far the most significant modulation effects are likely to result from source directivity patterns. The basic "figure-of-eight" dipole directivity pattern of the sound radiated by blade surface pressure fluctuations becomes asymmetric due to the blade motion, with an increase in intensity of the sound radiated at acute angles to the direction of motion. The effect on the observed sound is shown in Figure 23. In this case the difference in sound level between approaching and receding blades could be as much as 10 dB and this would be the level of an amplitude modulation occurring at the blade passage frequency.

3.2.1 Sources of Broadband Noise. At high frequencies there are several mechanisms which could contribute to broadband rotor noise. These include

- Blade boundary layer turbulence.
- Random trailing edge vorticity.
- Turbulence in the wake and the oncoming airstream.

As discussed above, at lower frequencies, broadband noise sources can be regarded as essentially the same as those which cause rotational noise and in the transition region can probably be mostly attributed to the vortex wake.

The turbulent boundary layer on the blade causes fluctuating pressure to act on the blade surfaces and these in turn radiate sound. However, the relatively small turbulent intensities in the boundary layer are unlikely to represent a very significant source of broadband rotor noise.

It is generally assumed that random vortex shedding is the major source and this has led to the somewhat misleading title "vortex" noise. The reasons for the importance attached to this phenomenon can be traced to Yudin who performed experiments using rotating cylindrical rods which generate well defined Karman vortex sheet wakes. Yudin derived the basic power laws for the magnitude of this "vortex" noise and these were later refined and applied to broadband propeller noise by Hubbard.
Turbulence in the oncoming stream causes a randomly fluctuating downwash at the blade, thus generating a fluctuating lift and an associated sound radiation. The presence of fairly large scale, high intensity turbulence in the rotor wake suggests that this source is a major contributor to broadband rotor noise. The importance of this source was pointed out by Kramer who performed experiments to show that many of the effects observed in earlier rotating blade (or rod) experiments could be attributed to the passage of the blades through the wakes of preceding blades. Sharland reached similar conclusions from measurements of the noise radiated by blades immersed in laminar and turbulent flow.

The implications of these results are that for blades which operate in clean air, the random vortex shedding is likely to be of major importance whereas blades operating in the turbulent wake shed by other blades are likely to generate a significant proportion of broadband noise due to substantial oncoming turbulence effects. Since it is known that large rotor lift fluctuations are caused by wake effects it is reasonable to assume that the latter source of random rotor noise is very significant and possibly dominant in certain flight modes.

3.2.2 Prediction of Broadband Noise Characteristics. To define the broad features of the noise at any point around the rotor the source characteristics must be defined in terms of overall level, spectrum shape and directonality.

Overall Level. Two recent studies have been addressed at broadband helicopter noise. The first was by Davidson and Hargreaves who gave the following expression for the overall sound pressure level at 500 ft:

$$d_B = 20 \log \frac{V_T}{10^5} + 20 \log T - 10 \log S - 25.5$$  \hspace{1cm} (12)

Schlegel et al. suggested:

$$d_B = 20 \log \frac{V_T}{10^5} + 20 \log T - 10 \log S - 43$$  \hspace{1cm} (13)

Both expressions show the same functional dependency on tip speed, thrust and blade area and Lowe25 derived the same relationship through dimensional arguments. The constant terms in the two equations are different although most of this difference can be explained by the use of different field points (i.e., different elevations from the rotor plane) and different definitions of the sound pressure levels. In fact, Equation (13) which corresponds to a point 20 degrees below the disc plane gives a slightly higher result than Equation (12), when appropriate directivity corrections are made and is thus recommended for prediction purposes.

Frequency Spectrum. Davidson and Hargreaves gave no predictions of typical frequencies and Yudin's results for rotating rods are too far removed from the present problem. Fortunately, Schlegel et al. obtained some broadband noise data from a test stand rotor. They modified Hubbard's formula for the Strouhal frequency (at which the broadband noise spectrum peaks) and constrained it to fit their experimental data by appropriate choice of Strouhal Number, $S_T$. The formula they suggest is

$$f = \frac{S_T \cdot 0.7 \cdot V_T}{T \cdot \cos \theta + S \cdot \sin \alpha} \quad \text{with} \quad S_T = 0.28$$  \hspace{1cm} (14)

A comparison of this Strouhal Number with other results for vortex shedding phenomena23-24 indicated that its value was consistently high. On the other hand, it is the correct result for helicopter rotor noise. This suggests that the vortex shedding phenomenon is not wholly responsible for the broadband noise components. As discussed above, another likely cause is the wake shed by previous blades.

Very little data on the turbulent properties of vortex wakes are available and probably the most useful paper is that of Spreiter and Sacks. They give an empirical formula for the distance behind a wing at which the trailing tip vortex can be regarded as fully rolled up. This distance is proportional to $s/A/C_L$ where $s$ is the semispan, $A$ the aspect ratio and $C_L$ the lift coefficient. They also give the core radius as being proportional to the semispan $s$.

The most significant feature of these results is the inverse dependence of rollup distance on lift coefficient. This suggests that as lift increases, the vortex cores become more tightly rolled up and consequently the turbulence scale decreases. This corresponds to an increase in frequency. This is directly opposed to the trends of formula (14) above which shows a frequency decrease as the lift coefficient (angle of attack) increases. These opposite trends make it difficult to predict the effect of thrust on frequency.

On the other hand, rotor scale does seem to be an important parameter and it is suggested that in the light of present knowledge an appropriate frequency formula is simply

$$f = \frac{V_T}{K R_0}$$  \hspace{1cm} (15)

where the value of $K$ can be obtained from Schlegel et al.'s data as 0.035.

A major influence on the acoustic frequencies radiated by a particular pressure source spectrum on the blade is that of frequency modulation by blade rotation, as discussed in Section 3.1. A source frequency $f$ on the blade will be observed to vary approximately between $(1 - M)f$ and $(1 + M)f$ at points away from the rotor. Thus the observed spectrum will broaden as tip speed increases. This effect was demonstrated experimentally by Von Witten27.

Lowe25 accommodated this effect into a spectrum prediction formula by assuming a source spectrum of the form

$$p_0^2(f) = \frac{f}{(f_0^2 + f^2)^2}$$  \hspace{1cm} (16)

and integrating this over the range $(1 - M)f$ to $(1 + M)f$.

The result, for the observed spectrum level is

$$S(M,f') = \frac{8 M}{f_0} \left( \frac{1 + M}{1 - M} \right)^3 \left\{ 1 + \frac{f'}{2 + (1 + M)^2 f^2 + (1 - M)^2 f^2} \right\}$$  \hspace{1cm} (17)
A comparison of this theoretical result with the experimental data from References 13 and 24 is shown in Figure 24, and shows very favorable agreement.

![Figure 24. Comparison Between Spectral Predictions and Experimental Data for Broadband Noise](image)

Directionality. This presents the most difficulty in the prediction of broadband rotor noise since very little data are available. The best approach appears to be the adoption of a dipole radiation pattern. However, this gives zero sound radiation in the plane of the disc which is unrealistic for several reasons, two of which are that blade angle of attack and blade flapping incline the dipole axis away from the axis of rotation. An appropriate expression for use with Equation (13) for the overall level, is

$$D = 10 \log \left( \frac{\cos^2 \phi + 0.1}{\cos^2 70^\circ + 0.1} \right)$$  \hspace{1cm} (18)

where $\phi$ is measured from the shaft axis and the constant terms are included to prevent a zero in the disc plane ($\phi = 90^\circ$).

### 3.3 Blade Slap

The harsh cracking or banging noise generated by rotors in certain flight conditions is the most serious and undesirable of the rotor noise types. It is now generally accepted that there are two types of blade slap although the mechanisms behind them are by no means resolved.

#### 3.3.1 Wake Interaction Slap

This is a true "slapping" phenomenon and is due to the interaction of a blade with a vortex trailing from a preceding blade. It occurs on a single rotor helicopter during maneuvers which cause low inflow conditions, for example during a tight turn or landing flare. On a tandem rotor helicopter it can also arise when the wake from the forward rotor interferes with the aft rotor. The blade/vortex interaction causes a sharp change of blade angle of attack with the consequent generation of an impulsive type of sound. Flow separation and stall may exaggerate this condition. Figure 25 shows an analysis of the noise of a hovering CH-47 tandem rotor helicopter both without blade slap and in the presence of a moderate amount of slap. This figure is taken from Reference 14. The slap was somewhat sporadic over the duration of the recording sample and substantially increased levels may be anticipated for more continuous and heavy slap conditions. However, the plot does reveal the increase of harmonic levels during blade slap which is the significant feature of impulsive noise.

![Figure 25. Measured Harmonic Sound Levels for CH-47B Chinook Helicopter (Unpublished Wyle data)](image)

Analytically, the impulsive loading caused by wake interactions causes an increase in the higher harmonic airloads and this in turn leads to higher calculated acoustic levels at the higher frequencies. The theory described in Section 3.1 is quite consistent with the experimental data. The main practical problem lies in the definition of the harmonic airloads. Adequate experimental measurement of these loads would require particularly wide frequency range instrumentation. The best approach at the present time appears to be to adopt a somewhat lower value for the loading law exponent discussed in Section 3.1. Such a step would be consistent with the observation$^{14}$ that under "rough running" conditions, e.g., during a landing flare, the airload of Reference 18 shows precisely this effect.
3.3.2 Advancing Blade Slap. Rotors operating at very high advance ratios generate impulsive noise for a different reason. Although this noise has in the past been attributed to the formation of shock waves near the tip of the advancing blade, these shock waves do not propagate to an observer unless the blade actually exceeds Mach 1. In fact, the effect is basically acoustic and theory shows that higher harmonic levels rapidly increase with tip speed. Figure 26 shows this effect. These results are calculated for a point 10 degrees below the plane of a four blade main rotor. A loading power law exponent of -2.0 was assumed and noise spectrum levels for both this and steady loading only (i.e., steady lift, drag and radial components) are included. Considering the harmonically loaded case first of all it can be seen that the high frequency energy increased very rapidly with an increase of M. For example, the 10th harmonic level increases by almost 40 dB when the rotational Mach number is raised from 0.75 to 1.0. It is even more important to note that although, at low tip speeds (low M), there is a very large difference between the steady loading case and that which includes the fluctuating airloads, when M = 1.0 the difference is negligible. This indicates that at high tip speeds the fluctuating load levels are of little concern in the slap problem. The slap would be practically as serious even if all fluctuating loads were removed. As shown in Section 3.1.2, a high tipspeed is acoustically equivalent to the combination of a moderate tip speed and a high forward speed although in the latter case there is a very substantial increase in the forward radiated noise.

Again, the rotational noise theory predicts the significant trends of advancing blade slap and in this case it is of less importance to define the fluctuating airloads levels with accuracy if the advancing blade Mach number approaches unity.

4.0 Design Considerations In Rotor Noise Control

In this section a simplified parameter study is described which illustrates the practical problems of helicopter noise control, at the design stage, and the usefulness of the analytical tools described in the previous sections.

The study was accomplished using a computer program which calculates rotational noise, broadband noise, atmospheric and ground attenuation and reduces the computed acoustic spectrum to a Perceived Noise Level. The latter is computed from the details of the noise spectrum and is a single number which gives a measure of the subjective noisiness of the sound. In the present context it is being used as an aural detection parameter. The justification for doing so is not particularly strong but the predicted trends should be realistic and it is a convenient technique in the absence of valid detection criteria. The methods used in the computer program are those described in this paper.

The noise of a single lifting rotor was computed to examine the influence of each of the five variables: disc loading, number of blades, tip speed, forward speed and altitude. Each was varied independently of the rest as shown in Table III. The central column gives the standard case and comparison cases were chosen by varying each parameter in turn, to the values shown in the other columns. The total rotor thrust and blade loading were held constant by chord variation. In the case of velocity variations, appropriate disc angle of attack corrections were made. The results which are shown in Figure 27 include full account of altitude and forward velocity effects.

### TABLE III

**DESIGN VALUES FOR PARAMETER STUDY**

<table>
<thead>
<tr>
<th>Constants:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Thrust = 10,000 lbs</td>
<td></td>
</tr>
<tr>
<td>Blade Loading = 80 lb/ft²</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables</th>
<th>Std. Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disc Loading (lb/ft²)</td>
<td>1 2 4 8 16</td>
</tr>
<tr>
<td>Number of Blades</td>
<td>2 3 4 5 6</td>
</tr>
<tr>
<td>Tip Speed (ft/sec)</td>
<td>400 550 700 850 1000</td>
</tr>
<tr>
<td>Forward Speed (Kts)</td>
<td>50 100 150 200 250</td>
</tr>
<tr>
<td>Altitude (ft)</td>
<td>50 100 200 500 1000</td>
</tr>
</tbody>
</table>
It is assumed for present purposes that the rotor is audibly detectable when the observed perceived noise level reaches 40 PNdB. Figure 27 gives the calculated ground distance between the observer and the approaching rotor at the instant when a 40 PNdB level is observed. In all cases the rotor is moving toward the observer, and the “detection distance” is plotted against the variable in each case.

Possibly the most striking result is the large effect of blade tip speed, particularly for values in excess of about 600 ft per second. This clearly illustrates the very substantial benefits of low tip speed for helicopter noise control. In fact the curve plotted in Figure 27 gives a slightly exaggerated picture since the advancing blade tip reached Mach 1.0 at a rotational tip speed of about 860 ft per second, so that higher tip speeds are somewhat excessive for conventional type helicopters.

The results for variations of disc loading and rotor altitude show no unexpected effects with detection range increasing steadily with increases of either parameter. At low altitudes the increased ground attenuation is mostly responsible for the reduction in detection range.

Surprisingly little effect of blade number is to be found, with little more than a 20 percent change in detection distance for blade numbers between 2 and 6. This may be attributed to two opposing effects. It was shown in Reference 14 that sound power output is approximately proportional to \((mB)^{2-k}\) where \(k\) is the loading law exponent which in this case was taken as 2.5. Thus power output varies as the inverse cube of the number of blades. This beneficial reduction in noise output however, is offset by a shift in the spectrum to higher frequencies. This in turn leads to an increase in perceived noise level since the ear is more sensitive at the higher frequencies. Also, it should be noted that as \(k\) decreases, which is effectively the case in rough rotor inflow conditions, larger blade numbers will not be so effective and blade slap levels could be significantly increased. However, additional wake interaction effects which may have a critical effect on fluctuating airflow levels could result from blade number changes and these have not been considered.

The most interesting finding is the effect of forward velocity whose curve shows a distinct minimum at a velocity of approximately 150 Kts. This is basically the effect of the minimum in the rotor noise radiation pattern immediately above the rotor. A change in forward velocity causes a change in rotor attitude, or angle of attack, such that under certain velocity/height combinations this minimum radiation is directed towards the observer.

Even more significant is the effect this has on warning time, i.e., the time interval between the instants of detection and passage overhead. This is plotted in Figure 28 which shows that a dramatic reduction in warning time is achieved at velocities around 150 Kts.

This exercise shows the importance of considering each aspect of the detection problem, namely source characteristics, propagation effects and observer characteristics in defining helicopter noise control requirements.
5.0 Potential Methods for Helicopter Rotor Noise Reduction

The main steps which should be taken to reduce rotor noise (assuming that other sources are inherently of low intensity or effectively silenced) are (1) reduce rotor tip speed, (2) reduce disc loading and (3) reduce the magnitude of the fluctuating airloads on the blades.

The first two of these are a matter of fundamental design whereas the third is not easy to control. The main sources of fluctuating airloads can be summarized as

- Asymmetrical flow over the blades.
- Vortex wake interactions.
- Rotor/fuselage interference.

Of these three problems the last is obviously the most straightforward to alleviate. The clearance between the main rotor and the fuselage can be increased and the tail rotor can be isolated from the influence of the main rotor and tail fin. Such actions are likely to introduce control problems but should result in substantial noise reductions, particularly for the tail rotor.

One obviously undesirable feature from the wake interaction point of view is the use of multiple rotor helicopters where one rotor can interact with the wake of the other. From the general viewpoint it might be possible to design a tandem rotor helicopter to minimize vortex interactions by careful choice of rotor phasing while at the same time allowing sufficient rotor separation to allow the wake to diffuse. However, this seems to pose a difficult design problem.

Of the other two sources the first can essentially be ignored since little can be done about it except by limiting forward speeds. Furthermore the airflow harmonics which result from flow asymmetries are basically of low frequency and of little importance to the noise field. The remaining problem, that of the vortex wake interactions, is difficult to analyze but several possible techniques offer some hope for its alleviation. The predominant cause of large lift fluctuations is that the vorticity shed by a blade rolls up into an intense, concentrated vortex tube in which air velocities are very high. If some means could be found to diffuse this energy or delay roll-up, the induced velocities and associated blade lift fluctuations could be diminished.

This is a matter of detailed blade design, in particular of the outboard blade sections. Conventional blades are designed to generate as much lift as possible and in practice the section lift increases right up to the last two percent or so of the blade radius. Beyond that point it falls off rapidly so that a very large change of bound circulation occurs causing intense shedding of vorticity. By changing the blade geometry near the tip this lift decay can be made more gradual, spreading the vorticity shedding over a greater length of blade. In Reference 13, Schlegel et al report that a trapezoidal shaped blade tip, which has just such an effect, reduced the vortex noise by 7 dB. The particular rotor studied was not run under conditions of "blade slap" but it may be expected that in such an extreme case the modified tip would have been more effective. Spencer et al performed some detailed experiments on the use of various tip forms for "vortex thickening" and found that very substantial reductions of induced velocities could be obtained at the price of some increased drag. One particular tip, a 60 degree sweep delta shape, effectively diffused the tip vortex with no performance penalty. It may well be that they accidentally selected a shape which caused the phenomenon known as vortex breakdown. This effect has been observed in the flow over slender delta wings and is best described as a catastrophic change in the structure of a vortex which results in a substantial diffusion of the concentrated energy. The promising results noted by Spencer et al suggest that the use of swept-back tips is worthy of further study and optimization of tip configuration should significantly improve their noise reduction potential.

An alternative approach, to effectively obtain the same result is to increase the number of blades in the rotor. The angular separation between adjacent pairs is reduced so that the vorticity shed by the preceding blade has less time to roll up before the following blade reaches it. That this fact may help to reduce noise was noted in Reference 14 where comparisons of "vortex" noise levels for 2, 3, 5 and 6 bladed rotors, as reported in References 13 and 24, indicate that the radiated sound power decreases with the number of blades. However, adding blades increases frequency toward regions where the hearing is more sensitive. On the other hand, higher frequencies are more rapidly attenuated by atmospheric absorption to that the frequency shift could contribute to reduced detectability.

It is important to note that the reductions in tip speed and disc loading may not always prove as effective as the basic theory indicates since there may be secondary effects on the harmonic airloads. Reference 14 showed that, experimentally, there is an optimum collective pitch setting for a given rotor. This is really a disc loading effect for, if the change in noise level due to the direct effect of collective pitch variation on thrust is accounted for, it is found that at low disc loading the noise level is high due to the close proximity of the wake. As the disc loading is increased the wake moves away more rapidly resulting in noise reductions. As the disc loading is further increased however, certain portions of the blades begin to stall and the noise level rises again. Of course if the rotor solidity were increased this latter rise could be prevented to some extent but the example does serve to show that all aspects of any design change must be carefully considered. This is not easy because the high frequency fluctuating blade loads will always be difficult to estimate.

A further possible method for reducing the noise is to increase the blade chord. Results derived in Reference 29 showed that, for a fixed frequency input, the integrated aerodynamic load was proportional to the inverse square root of the chord length. Thus sound output would be predicted as inversely proportional to chord. Also an increase in chord at fixed thrust and collective pitch causes a decrease in tip speed with its associated acoustic benefits.
Perhaps any increase in rotor blade area would be more efficiently employed as an extra blade. Nevertheless, it is thought that a more detailed study of blade chord effects would be valuable.

One last feature of potential use is that of the directionality characteristics. A quite definite minimum just above the plane of the disc is predicted. It may be possible to design or fly a helicopter so that this minimum occurs at the position where minimum noise is desired. It should be particularly noted however, that this minimum has not been confirmed experimentally.

It would seem well worthwhile to perform detailed (possibly scale model) experiments to study some of these noise control methods. The experiments should be deliberately designed to cover cases outside normal operating ranges so that the trends can be well defined. Such experiments could be of considerable value in reducing helicopter rotor noise radiation.

The major design requirements for minimum noise can be summarized as follows:

- Low tip speed
- Large number of blades
- Low disc loading
- Large blade chord
- Minimum interference with rotor flow
- Any features which will reduce the high frequency airflow fluctuations.

The parameter study described in Section 4 highlighted the importance of propagation and receiver characteristics to the overall problem of helicopter noise, at least from the standpoint of aural detection. At the present time very little is known about sound attenuation by ground cover and valid aural detection criteria are urgently needed. Experimental research in both areas is required to advance the state-of-the-art in helicopter noise.

Appendix: Instructions for Use of Design Charts - Figure 29

Parameters Required (The notation in this appendix differs in some cases from that in the main text).

The following parameters are required for use in the noise calculations using the design charts.

- $x$, $y$, $z$: Field point coordinates relative to helicopter measured in feet with $x$ measured positive in the direction of motion (parallel to ground), $y$ measured sideways in the plane of the disc, $z$ measured downwards from helicopter. (Results for $y$ equal results for $-y$.)
- $A$: Disc area, ft$^2$ (or $T/A =$ disc loading, lb/ft$^2$)
- $\Omega$: Rotor angular velocity, rad/sec ($\Omega =$ rpm x $2\pi/60$)
- $V$: Flight velocity, ft/sec
- $a_0$: Speed of sound in free air, ft/sec
- $i_d$: Disc incidence (angle between disc and $x$-axis), degrees
- $m$: Sound harmonic (equals 1 for fundamental, 2 for second harmonic, etc.)
- $B$: Number of blades
- $T$: Thrust, lb
- $R$: Rotor Radius, ft

To calculate the rotational noise spectrum occurring instantaneously at any point, $r$, $\theta$ relative to the rotor center and its direction of motion:

1. Calculate range $r = \sqrt{x^2 + y^2 + z^2}$
2. Calculate the rotational Mach Number $M$
   $$M = 0.8 \frac{\Omega R}{a_0}$$
3. Calculate the flight Mach number $M_F = \frac{V}{a_0}$
4. Calculate the angle $\theta'$ between the flight direction and the line joining the rotor and the field point
   $$\theta' = \cos^{-1} \left( \frac{x}{r} \right)$$
5. Calculate the Effective Rotational Mach Number
   $$M_E = \frac{M}{1 - M_F \cos \theta'}$$
6. Calculate the angle $\theta$ between the rotor plane and the line $r$. If the disc incidence is $i_d$, this is given by
   $$\theta = \tan^{-1} \left( \frac{z}{\sqrt{x^2 + y^2}} \right) - \tan^{-1} \left( \frac{x}{\sqrt{x^2 + y^2}} \right)$$
7. Using the values of $M_E$ and $\theta$ look up each chart to obtain values of the harmonic sound pressure level for $n = 2$, 3, 4, 6, 8, 10, 12, 16, 20, 30, 40 and 60.
8. Correct the values obtained for thrust, disc loading and distance according to
   $$SPL_n = 1 + 11 + 10 \log \frac{T}{(\frac{1}{A})} \text{ db re: } 0.002 \mu \text{bar}$$
9. Plot $SPL_n$ against $n$ and fit smooth curve.
Figure 29: Rotor Noise Harmonic Sound Pressure Levels as Functions of Harmonic Number, Rotational Mach Number, and Angle from Disc Plane.
The sound pressure levels from this curve for \( n = 8, 2B, 3B, \ldots \) give the required harmonic level at the point \( x, y, z \).

The fundamental frequency is

\[
\Omega B / \left( 2 \pi (1 - M_F \cos \theta) \right) \text{ Hz}
\]

Example - Calculate the rotational noise spectrum 1000 ft from a 3-blade rotor at an angle of 20 degrees below the flight path for the following parameters:

\( T = 10,000 \text{ lb}, T/A = 7 \text{ lb/ft}^2, V = 200 \text{ ft/sec}, i_d = 5 \text{ degrees}, \Omega R = 600 \text{ ft/sec}, \theta_0 = 1117 \text{ ft/sec}. \)

(1) \( r = 1000 \text{ ft} \)
(2) \( M = 0.8 \times 600 / 1117 = 0.429 \)
(3) \( M_F = 200 / 1117 = 0.179 \)
(4) \( \theta = 20 \text{ degrees} \)
(5) \( M_E = \frac{0.429}{1 - 0.179 \times 0.938} = 0.516 \)
(6) \( \theta = 20^\circ - 5^\circ = 15^\circ \)
(7) From Charts

\[
\begin{array}{ccccccccccc}
 n & 2 & 3 & 4 & 6 & 8 & 10 & 12 & 16 & 20 & 30 & 40 & 60 \\
\hline
\text{SPL} & 84.5 & 82.5 & 81.5 & 77.5 & 72.67 & 63 & 54 & 48 & 44.5 & 38.5 & 35 & 30 \\
\end{array}
\]

(8) Correction = \(-10 \log_{10} \left( \frac{10,000}{1000^2} \cdot 7 \right) + 11 = -0.5 \text{ dB} \)

\[
\begin{array}{ccccccccccc}
 n & 2 & 3 & 4 & 6 & 8 & 10 & 12 & 16 & 20 & 30 & 40 & 60 \\
\hline
\text{SPL} & 84 & 82 & 81 & 77.15 & 66.563.5 & 56.5 & 53.5 & 47.5 & 44.38 & 38 & 35 & 30 \\
\end{array}
\]

The results of Steps 9 and 10 can be seen in Figure 30 where the harmonic levels corresponding to \( n = 1, 2, 3, \ldots \) are drawn as vertical lines.

![Figure 30. Hand Calculated Rotational Noise Spectrum](image)

The fundamental frequency in this case is

\[
\frac{\Omega B}{2\pi (1 - M_F \cos \theta)} = \frac{600 \times 3}{2 \pi \sqrt{\frac{A}{\pi} (1 - M)}} = 16.1 \text{ Hz}
\]

References


The military value of helicopters for tactical and surveillance missions is reduced by their characteristic, high level noise signatures. In many situations a helicopter can be heard approaching from distances between 5 and 10 miles which, at helicopter speeds, gives several minutes warning to the enemy.

The principal objective of this study was to develop a methodology for predicting helicopter aural detection thresholds. Various aspects of the helicopter noise problem were considered including measurement and analysis, and propagation over ground cover. The extract presented here describes a laboratory experiment designed to develop and test an empirical procedure for estimating the probability of detection of a helicopter noise signal in the presence of a masking noise.

A procedure was devised for calculating detection threshold levels from either octave or \( \frac{1}{3} \)-octave band level estimates of the helicopter signal and the ambient noise. In tests involving combinations of 21 different helicopter sounds and 8 ambient noise spectra the model was found to agree with measured thresholds to within an accuracy of \( \pm 4 \) dB.

The procedure was subsequently tested in field experiments using real helicopter approach tests. These confirmed the accuracy of the method for predicting median detection thresholds.

The method is believed to be used in classified military studies of helicopter operations.

4.0 EXPERIMENTAL PROGRAM

The main purpose of this study is to develop methods for calculating the aural detection ranges of helicopters. To do this it is necessary to: (a) specify an analytical/empirical model of the aural detection process and (b) define its applicability and accuracy. This section describes the experimental study which was conducted for these purposes.

4.1 EXPERIMENTAL REQUIREMENTS

An accurate, practical and useful method for estimating the aural detection thresholds of helicopter sounds should take account of all variables which are known to be of first-order importance to the problem. These include the acoustic characteristics of the helicopter, the effects of propagation over long distances on the observed sound, the ambient noise environment in the vicinity of the observer, and finally, the hearing acuity of the observer himself.

It is obviously desirable for each variable to be specified in terms of quantities which can be conveniently measured, or more importantly from a design standpoint, estimated. Of equal importance is the need to recognize the degree of accuracy with which each can be specified. Although it is evident that the psychoacoustic variables themselves have wide confidence intervals, there is little point in demanding greater resolution than can be expected of the physical inputs. In this section each of the main factors are examined in the light of these requirements and for their applicability to a potential model.

Source Radiation

At the present time, the state of the art in helicopter noise estimation for design purposes is such that the first few (12) harmonics of rotor noise can be estimated with reasonable confidence (+ 2 - 3 dB) and the remainder of the spectrum with somewhat lesser accuracy (+ 5 dB). The spectral details, associated with these estimates, in terms of energy distribution, can be predicted fairly well but phase information, which has an important bearing on the pulsatile nature of the total sound, is beyond the present state of the art. It may be confidently expected that as knowledge advances, improvements in all areas will be forthcoming, but the very nature of the problem suggests that definition of high-frequency spectral details will always be difficult. This is particularly true of such transient phenomena as blade slap.

Helicopter noise can be measured with as much accuracy as the instrumentation will allow. Modern techniques can provide very high quality data provided very rigorous experimental procedures are followed. In practice it is difficult to maintain
ideal conditions and instrumentation limitations make themselves felt. Measurements at long distances from the helicopter are extremely sensitive to environmental conditions while short range measurements present problems with nonstationarity.

Taking all these factors into account it would seem that source noise should be measured or estimated in terms of 1/3 octave band spectrum levels. This bandwidth is sufficiently narrow to allow fairly detailed spectral resolution, particularly being close to the critical bandwidth over a wide frequency range (Figure 28), and yet wide enough to avoid serious errors due to nonstationarity in the analysis of flyover data. Also, the reduction of design predictions to this format is fairly convenient. Furthermore, commercial analysis equipment for this purpose is readily available in a wide variety of forms.

In addition to frequency selectivity, there is the question of time averaging. A judgement on an appropriate analysis time constant must be made on the basis of both psychoacoustic considerations and the significance of short-time scale signal fluctuations such as blade passage modulation.

**Propagation**

The effects of the atmosphere and the terrain are of profound importance to the aural detectability problem, particularly the latter in the case of low-flying aircraft. Unfortunately, although atmospheric absorption can be estimated with some reliability, very little is presently known about terrain effects. Also of probable significance are the effects of random signal level fluctuations due to atmospheric inhomogeneities and other causes. Although unpredictable, these are always present and, like other propagation effects, will eventually become better documented. Some account of their influence is thus considered desirable.

**Masking Noise**

Masking noise may of course be specified in practically any terms, depending upon what is known about the ambient noise in a particular environment. In general, it seems unlikely that there would be any necessity to be more specific than an octave band level spectrum; but again, for flexibility, the model should accommodate a 1/3-octave band level definition. The effects of temporal variations of level could be considered, but lack of detailed knowledge would probably render this superfluous in the majority of applications.

**Human Observer Characteristics**

Hearing acuity varies significantly from person to person and also from community to community and must be included in the model as a variable. For convenience it should be appropriate to include this variable as a pure tone absolute threshold function.
Specific Objectives

In the light of the foregoing considerations, the experimental program was divided into two phases with the following objectives:

Phase I: To provide the necessary supporting data to establish an adequate analytical model of the aural detection process. Subsidiary goals of the Phase I tests were specifically:

1. To develop a reliable experimental technique
2. To measure absolute and masked thresholds for tones, tonal complexes and bands of noise, both stationary and modulated
3. To investigate the critical band concept as applied to detection of helicopter noise.

Phase II: To test and or refine the model through application to actual helicopter sounds.

4.2 EXPERIMENTAL METHOD

Psychophysical Test Procedure

Although an audibility threshold is defined as that specific level at which finite neural activity is stimulated, it is not possible to measure the threshold level with the precision that this description implies because of a difference between the levels at which the stimulus is definitely audible and definitely inaudible. The magnitude of this difference is a function of many factors, including whether or not the signal is increasing or decreasing in level, its duration, the degree of concentration of the subject, whether or not he is warned of the signal's existence, what to listen for, and so on. Many techniques have been established for the measurement of audibility thresholds and a choice between them inevitably rests upon the desired compromise between accuracy and speed; as in most measurements, higher precision generally requires more time.

In the psychophysical method which was originally proposed for the present study, subjects were to listen to a helicopter sound that was gradually increasing in intensity and were to respond when they first detected the sound. Unfortunately, this method suffers from two well-known types of error often observed in psychophysical experiments; errors of anticipation and errors of habituation. The former refers to the tendency of subjects to consistently respond too early, i.e., below their actual detection threshold, and the latter refers to the tendency of subjects to wait too long before reporting their detection of the stimulus, i.e., they respond well.
above their detection threshold. These errors are typically cancelled out by presenting both increasing and decreasing stimulus intensity sequences: a procedure which requires a relatively large number of trials with each stimulus to be effective and which, conceptually, at least, does not fit the field detection situation. The original technique, therefore, could not be counted on to yield reliable results. Furthermore, it was very inefficient; each threshold determination would probably require at least 10 separate threshold determinations.

The standard psychophysical Method of Adjustment appeared to provide a satisfactory balance between the requirements of reliability and efficiency. A pilot test of the Method of Adjustment was performed with eight subjects. Each subject adjusted a logarithmic potentiometer to control the headphone (TDH-39) level of a computer-simulated helicopter sound spectrum which was repeatedly "turned on" for approximately 1.75 seconds and "off" for approximately 0.5 seconds. Subjects manipulated the potentiometer until they were satisfied that the sound was just at their detection threshold. They then brought the signal to suprathreshold levels and repeated the adjustment procedure for a total of 20 threshold estimates for each subject. Thresholds estimated varied over a 22 dB range for the eight subjects. An individual subject was, however, quite reliable at picking and remaining with a particular threshold value from trial to trial. The average standard deviation of the subjects around their own mean thresholds was 1.7 dB. This small amount of variability indicated that the Method of Adjustment could provide a reliable indication of detection thresholds for complex acoustic stimuli.

Although the method appeared to be very reliable, it soon became clear that it would be too time consuming to investigate the large number of different helicopter sound characteristics that contribute to detection. Each adjustment required between 30 and 60 seconds. Thus, if thresholds were to be found for only 100 different stimuli, and each threshold estimate were composed of only 10 different adjustments/subject, each subject would have to make 1000 adjustments requiring a total of 500 to 1000 minutes (8 to 16 hours for this limited number of stimulus values).

An alternative approach was required because considerably more than 100 data points/subject were desired. Von Bekésy described an audiometric technique for determining pure-tone thresholds as a continuous function of frequency, which used a modified Method of Adjustment. The technique has been favorably evaluated by Hirsch, who found that it was a quick and reliable means of obtaining auditory thresholds across an entire audible frequency spectrum. Bekésy's audiometer consists of a variable-frequency oscillator that is coupled mechanically to a rotary drum on which is mounted an audiogram blank. The listener controls the direction of an attenuator motor, continuously adjusting the signal level between the points where it becomes audible and then inaudible. A writing device inscribes the amount of attenuation on the vertical axis of the audiogram blank so that the result is a continuous line that moves up and down between points of audibility and inaudibility as a
function of frequency. The Békésy method may be classified as a modified Method of Adjustment, which is reliable and yields a large number of threshold determinations in a short time.

The methodology used for the Békésy audiometer was, therefore, suited to the present task and the entire experiment was designed around its use. The test signals were recorded on analog magnetic tape or generated in such a way that the stimulus parameter under study, normally frequency, was varied slowly with time. A pure tone audiogram, for example, was obtained during a 5 minute frequency sweep from 12 Hz to 12,000 Hz. Whatever the signal, the subject, who was able to control the signal level, was asked to continually adjust it to the just-audible point for the entire test duration. The control and data acquisition system developed for this purpose is described below.

Control and Data Acquisition System (DAS)

In order to obtain a statistically adequate number of measurements for the range of variables envisaged, it was clear from the outset that some form of automated test procedure would be required to obtain them reliably and accurately. Accordingly, a significant proportion of the effort during this project was directed toward the development of an automatic test control and data handling system for large-scale Békésy audiometry.

The system was centered around a 120-dB, continuously variable attenuator, which controlled the level of the stimulus sound being presented to the subject. The setting of the attenuator was controlled by a bi-directional electric motor, which was in turn controlled by the subject. The tracking rate of the attenuator was 2 dB per second in either direction.

The subject listened to the stimulus sounds inside an acoustic test chamber. He was furnished with a simple hand-held pushbutton cord switch and instructed to push down the button as long as he was able to hear the stimulus, releasing it when the signal became inaudible. Pushing the button drove the attenuator in the direction of increasing attenuation and releasing it caused the sound level to increase again.

During the course of the test the automatic data system sampled the position of the attenuator setting at intervals of approximately one second, recording these data on punched paper tape. These were subsequently subjected to computer analysis by programs which converted the punched numbers to sound pressure levels and related these levels to the temporally varying characteristics of the stimulus sounds. The results were made available as listings or plots of the means and deviations of the threshold levels, either for individual subjects or as average results for an entire test jury. An example plot is shown in Figure 30.
The data system and associated software have been described in detail by Adcock. The main design features of the system together with operation and calibration procedures and data analysis methods are described in Appendix II to this report.

Test Environment

It was originally intended to present all stimuli to the subjects in a totally progressive wave environment in Wyle Laboratories' 1500-cu ft progressive wave chamber, which has been described in References 48 through 51. However, recent experience in another project revealed a serious difficulty in maintaining a known stimulus level at the subjects' ears at frequencies above 1 kHz. This problem is caused by head diffraction patterns which vary between subjects and with head orientation. Accordingly, it was decided to avoid this problem through the use of wide-frequency-range headphones. At the same time, the possible importance of total body exposure at low frequencies was recognized and to retain the effect of nonauditory stimulation, the tests were performed inside a newly developed low-frequency progressive wave chamber. A crossover network was used so that the test subject seated in the working section of this facility was totally exposed to frequencies below 65 Hz generated by loudspeakers while listening to higher frequencies through high-quality binaural headphones.

A cutaway view of the acoustic chamber in Figure 31 shows its three sections. The first is a 1300-cu ft loudspeaker enclosure containing four 30-inch-diameter Electrovoice W30 speakers. These generate a test sound pressure level in excess of 120 dB at frequencies down to less than 10 Hz. To damp out resonances, two wedges containing 170 lb of low-density glass fibers are installed in this enclosure. The speakers are mounted in a reinforced wooden baffle and are driven by the parallel 190 watt channels of a Crown DC-300 solid-state amplifier. The total system has a very low harmonic distortion of less than 0.3% at levels less than 100 dB. The middle test section, which is 10 ft long x 8 ft wide x 7 ft high, can accommodate four seated subjects, although the present adjustment tests involved only one subject at a time. Behind the working section and designed to absorb the total speaker power output of more than 50 acoustic watts, are four 20 ft long fiberglass wedges, each spanning the full height of the chamber and expanding to a maximum width of 2 ft at the rear wall. In all the facility is 50 ft long and is constructed of 12-in.-thick concrete to provide high attenuation of external noise.

The headphones used were the newly available Koss ESP-9 electrostatic units which have a nominal frequency response (+ 5 dB) of 10 - 18,000 Hz. To minimize self-generated noise, the AC-powered voltage source was replaced for the tests by a dry cell to maintain the polarization voltage of 500 volts. The E-9 energizer was driven by a single 20-watt channel of a Crown D-40 solid-state amplifier, like the
DC-300 an extremely stable, wide-dynamic range-amplifier with excellent frequency response characteristics.

Figure 32 is a schematic diagram of the entire sound generation and instrumentation system. The frequency characteristics of the system, as measured in a Koss B&K 6-cc coupler using a ½-in.B&K 4133 microphone, are presented in Figure 33. These calibrations were performed at a sound pressure level of approximately 100 dB. The overall response curves, one for each earphone, are presented in Figure 33 (a), and the lower diagram (b) shows the separate free field (loudspeaker) and pressure field contributions (headphones). It should be noted that in the loudspeaker frequency range, the headphone coupler arrangement is totally transmissive since precisely the same function is measured both with the headphones removed and with the microphone removed from the coupler. Details of the calibration procedures used are described in Appendix II.

4.3 PHASE I TESTS

The main purpose of the Phase I tests was to validate the experimental procedures and equipment and to investigate certain aspects of aural detection of relevance to the helicopter problem which do not appear to have been covered in previous research. Specifically, these included measurements of absolute and masked thresholds for tones, bands of noise, both stationary and harmonically modulated, and finite bands of multiple harmonic noise. The precise combinations of signals and noise included are presented in Section 5.0. Altogether, more than two hundred and fifty individual tests were run over a period of 2 months for a total test duration of approximately 65 hours (including Phase II).

Subjects

Initial Phase I tests were repeated with five subjects. However, experience showed that equally consistent data could be obtained with three trained subjects selected from Wyle Laboratories engineering staff, so to cover the maximum ground in the time available, the bulk of the experiments were performed with three subjects. The absolute audibility function for pure tones averaged over these three subjects showed good agreement with a variety of previous determinations taken from the literature. This comparison is discussed in Section 5.0; it suffices here to state that the agreement was sufficiently close for the subjects to be regarded as having normal hearing acuity.

Signal Generation

Early experiments caused abandonment of the original plan to perform all experiments using a remote controlled tape reproducer as a signal source. This was due to problems associated with the extraordinary dynamic range of the ear as
witnessed by the threshold contours of Figure 27. Very simply, it was discovered that "tape hiss" was considerably more audible than the signal for tones and other sounds with low frequencies and this problem required major alterations to both the equipment and the test plan.

All pure-tone thresholds were obtained by what was effectively a standard Bekesy audiometric procedure using the BFO as a signal source with direct input to the DAS amplifier. The automatic sweep facility of this oscillator was used to vary the frequency at a rate of 0.83 octave per minute in the range of 10 Hz to 12.5 kHz.

Stationary random noise was investigated in octave or 1/3-octave bandwidths by on-line filtering of pink noise reproduced from an FM tape recording (see Figure 32). During the test, the tape was played continuously into the stepping filter of the B&K 2112 Audio Frequency Spectrometer which was automatically switched from band to band at 0.83 octave per minute between 12.4 Hz and 10 kHz (1/3-octaves) or between 16 Hz and 8 kHz (octaves).

The same technique was also used to generate modulated random noise and multiple harmonic noise. The modulated noise was initially recorded on FM tape by modulating the same source of pink noise with an electro-optical amplitude modulator. This unit was driven by a modulating signal from an HP 650 A oscillator. Modulation depths* of up to 12 dB and frequencies of up to 40 Hz were applied to pink noise signals with energy between 10 and 12,500 Hz. Some oscilloscope records of these signals were shown in Figure 13. The harmonic sounds containing one hundred harmonics and fundamental frequencies of 10, 20 and 40 Hz were generated digitally by an XDS Sigma V computer, converted to analog form by a high-speed digital to analog recorder and recorded on an Ampex AG 500 1/4-inch direct record tape machine.

Test Procedures

Test participation demands considerable concentration on the part of the test subject, so to avoid fatigue, test runs were limited to the shortest possible duration. For this reason all preliminary checks and setup procedures were completed before the subject entered the test chamber. Upon entering, the subject was seated and warned of an imminent start by the test controller. A two-way intercom was installed, and the controller was able to hear the subject at all times when a test was in progress. The subject himself could monitor the progress of the test by watching a slave console which relayed the status of the DAS. Illumination of an amber "STANDBY" lamp

* Defined in terms of the peak to trough rms levels.
indicated that the DAS was readied and the test could commence at any time. When the display switched to "READY" (also amber), the system was energized and the attenuator motor was running. At this time, the subject should have been wearing his headphones and begin to perform his test function. When a green "RUN" lamp lit, data was being acquired. The simultaneous illumination of a red "LIMIT" lamp warned that the attenuator had reached the end of its travel and had thus automatically switched off the motor. This situation could only be remedied by the controller as described in Appendix II, and to be sure that the abort had come to his attention the subject was asked to advise the controller whenever the red lamp was lit. At the end of the run the display switched from "RUN" to "STANDBY" at which point the subject was usually asked to relax and leave the chamber.

Should an abnormal situation arise during a run, the subject could operate a guarded "ALARM" switch on the panel which lit a warning lamp on the main control console and also automatically shut down the test. For cases of extreme emergency, a switch was also installed within reach of the subject which cut off the electrical supply to the audio power amplifiers. This precaution was taken to protect the subject in the event of a signal runaway. However, no abnormal situations were experienced at any time during the program.

Typical Phase I tests lasted between 5 and 10 minutes. In every case, the first 60 seconds of the stimulus signal was maintained constant to allow the subject to acclimatize himself and to give the motor time to move the attenuator to the vicinity of the appropriate working range. Similarly, the final stimulus signal was maintained for an additional 30 seconds as a check that the subject had indeed tracked his threshold accurately and was still concentrating at the end of the test. For wide frequency range sweeps, the 100 dB dynamic range of the system was insufficient to give a reasonable chance of avoiding running out of attenuator range. This problem was remedied by inserting an optional 30-dB attenuator into the system at the DAS output. The procedure for using this was to hold the frequency sweep in the vicinity of 400 Hz for a period of 20 seconds. During this time, the attenuator was switched into the circuit, leaving the subject sufficient time to adjust to a new attenuator setting before continuing the sweep.

4.4 PHASE II TESTS

The purpose of this second series of tests was to provide comprehensive experimental confirmation of the validity of the aural detectability criteria for practical application. The experiment was designed around the Bekésy audiometric procedure but involved the use of recorded helicopter sounds in place of the "artificial" stimuli used previously and a wide range of ambient noise spectra.
The helicopter noise recordings were obtained from the U. S. Air Force Flight Dynamics Laboratory, the Acoustics Branch of NASA Langley Research Center Dynamics Loads Division, and from Wyle Laboratories' magnetic tape library. Many signals were examined for their ability to meet the following criteria listed: in order of importance:

1. Good signal-to-noise ratio and signal quality;
2. Freedom from wind noise, insect sounds, bird calls, vehicle movements, voices, and other spurious sounds;
3. Long duration and steadiness;
4. A diversity of source characteristics;
5. Large distance between source and microphone.

In fact, these requirements were difficult if not impossible to meet collectively, and in almost all cases a compromise of some kind was necessary. Probably of most significance in this regard is that it was generally criterion number 5 that suffered, and most of the sounds selected were recorded at distances substantially less than detection range. The 21 signals selected for study are listed in Table I.

Most of the original recordings were made on wide-frequency-range FM equipment, and the initial intention was to use an FM reproducer to generate the test stimuli in order to include the frequencies below 20 - 25 Hz. Unfortunately, severe problems were encountered with the PS-207 remote operation facility which could not be overcome during the available test period. Accordingly, it was necessary to copy data to a direct record system for reproduction according to the arrangement shown in Figure 32.

Because of the finite travel rate of the DAS attenuator, any sudden, large changes of level essentially cause the loss of threshold data while the potentiometer travels to its new equilibrium problem. To minimize the occurrence of such discontinuities, the 21 signals were copied in sequence onto a master test tape through an amplifier whose gain was continually adjusted to maintain an approximately constant overall level.

The 1/3-octave band levels were read at 15-second intervals from time history analyses of this tape made with a 300-msec averaging time (see Section 5.4). These histories were further averaged by eye to smooth out low-period random fluctuations for a total effective averaging time of the order of 10 seconds.

The tape, which was initiated and terminated by 60 seconds and 30 seconds respectively of a 100-Hz control tone for setup and calibration purposes, lasted about
<table>
<thead>
<tr>
<th>Ident. No.</th>
<th>Helicopter Type</th>
<th>Flight Configuration</th>
<th>Estimated Ground Distance ft.</th>
<th>Signal Duration sec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CH-47B</td>
<td>Hover in ground-effect</td>
<td>200</td>
<td>57</td>
</tr>
<tr>
<td>2</td>
<td>UH-1B</td>
<td>Hover and approach</td>
<td>5,000</td>
<td>61</td>
</tr>
<tr>
<td>3</td>
<td>UH-1B</td>
<td>Hover in ground-effect</td>
<td>600</td>
<td>70</td>
</tr>
<tr>
<td>4</td>
<td>CH-47B</td>
<td>Hover in ground-effect</td>
<td>300</td>
<td>94</td>
</tr>
<tr>
<td>5</td>
<td>UH-1B</td>
<td>Hover in ground-effect</td>
<td>2,500</td>
<td>53</td>
</tr>
<tr>
<td>6</td>
<td>HH-43B</td>
<td>Hover in ground-effect</td>
<td>200</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>HH-43B</td>
<td>Hover at 50 ft altitude</td>
<td>200</td>
<td>23</td>
</tr>
<tr>
<td>8</td>
<td>HH-43B</td>
<td>Hover at 200 ft altitude</td>
<td>200</td>
<td>20</td>
</tr>
<tr>
<td>9</td>
<td>HH-43B</td>
<td>Hover at 500 ft altitude</td>
<td>200</td>
<td>29</td>
</tr>
<tr>
<td>10</td>
<td>CH-47A</td>
<td>Flyover at 1100 ft, 100 kt</td>
<td>from 10,000</td>
<td>62</td>
</tr>
<tr>
<td>11</td>
<td>CH-47A</td>
<td>Flyover at 750 ft, 100 kt</td>
<td>from 6,000</td>
<td>37</td>
</tr>
<tr>
<td>12</td>
<td>CH-47A</td>
<td>Flyover at 450 ft, 100 kt</td>
<td>from 6,000</td>
<td>37</td>
</tr>
<tr>
<td>13</td>
<td>CH-47A</td>
<td>Flyover at 250 ft, 100 kt</td>
<td>from 6,000</td>
<td>36</td>
</tr>
<tr>
<td>14</td>
<td>QH-50</td>
<td>Flyover at 125 ft, 30 kt</td>
<td>from 4,000</td>
<td>80</td>
</tr>
<tr>
<td>15</td>
<td>QH-50</td>
<td>Flyover at 1000 ft, 40 kt</td>
<td>from 3,500</td>
<td>51</td>
</tr>
<tr>
<td>16</td>
<td>YOH-6</td>
<td>Flyover at 500 ft, 100 kt</td>
<td>from 10,000</td>
<td>57</td>
</tr>
<tr>
<td>17</td>
<td>YOH-6</td>
<td>Hover at 500 ft</td>
<td>200</td>
<td>36</td>
</tr>
<tr>
<td>18</td>
<td>CH-3E</td>
<td>Flyover at 1000 ft, 60 kt</td>
<td>from 8,000</td>
<td>85</td>
</tr>
<tr>
<td>19</td>
<td>CH-3E</td>
<td>Hover at 500 ft</td>
<td>200</td>
<td>39</td>
</tr>
<tr>
<td>20</td>
<td>UH-1F</td>
<td>Flyover at 1000 ft, 60 kt</td>
<td>from 9,000</td>
<td>91</td>
</tr>
<tr>
<td>21</td>
<td>UH-1B</td>
<td>Ground run</td>
<td>50</td>
<td>81</td>
</tr>
</tbody>
</table>
23 minutes. The entire tape was played for each test in combination with one of eight different ambient masking sounds (including "zero" ambient for the determination of absolute thresholds). These ambient sounds all comprised gaussian random noise with spectrum levels designed to provide the severest possible test of the analytical threshold model. Only two spectrum shapes were involved: flat "pink" noise and noise whose 1/3-octave band level decayed at the rate of 6 dB per octave. These sounds were recorded on two separate tapes which were reproduced at different levels to obtain the specified conditions.

To provide improved statistical reliability in these main tests, ten subjects were used. These were all men in their twenties and early thirties who were selected from approximately thirty applicants on the basis of acceptable hearing ability. No related experience was required and, in an attempt to derive realistic results typical of "untrained" listeners in the tactical situation, no extensive training was given. Each subject was paid for his services and participated in each of the eight test runs described above. In addition, sine sweep audiograms were measured on a number of occasions. Because of the long duration of these tests, each subject was allowed at least 30 minutes' rest period between successive tests.

The written instructions given to the subjects are presented in Appendix III. The participants were given ample time to study these and to ask any questions to satisfy themselves and the Test Director that they fully understood what was required. In addition they were allowed short practice runs. It should be noted that the instructions made specific reference to aircraft sounds. This was felt to be important after preliminary tests revealed that a difference normally existed between the level at which an unspecified stimulus difference was detected and the level at which the signal was recognized as the sound of a helicopter.
5.0 RESULTS AND DISCUSSION

The experimental data presented in Sections 5.1 and 5.2 were read from the computer plots such as shown in Figure 30. Curves were fitted through the average threshold points by eye, and from it values were read and tabulated at 1/3-octave intervals. The data points shown in the various figures are these 1/3-octave values. All results are averaged for the same three experienced subjects and it should thus be noted that each 1/3-octave data point effectively represents the contributions of approximately 45 individual measurements. The variability of these individual measurements, due to both the differences between subjects and the adjustment oscillation, had an average standard deviation of approximately 4 dB. However, deviations tended to increase to around twice this value at frequencies above 4000 Hz.

5.1 ABSOLUTE THRESHOLD DETERMINATIONS

The absolute threshold of audibility for a pure tone is shown in Figure 34 as a function of frequency. Two sets of data are plotted, measured on two occasions separated by several weeks, which indicate the degree of repeatability obtained. Very small differences are observed at frequencies below 2000 Hz, but discrepancies are apparent at higher frequencies. This reflects the increased data scatter at high frequencies referred to above and is probably attributable to the difficulties of accurately establishing the true sound pressure level in this region (see the headphone response diagram, Figure 33).

A best-fit curve has been fairied through the data points for use as a basic pure-tone reference in subsequent discussions. The slight hump in the curve around 80 - 100 Hz is purely a function of the stimulus presentation system as Figure 35 reveals. This compares the thresholds measured in three ways: with headphones only, with loudspeakers only, and with both headphones and loudspeakers connected through a 65-Hz crossover network. The loudspeaker-alone data is practically undistinguishable from the "combined" curve, but a marked increase in the threshold level may be seen for the headphone presented sound. This difference, with a maximum of 12 dB at 25 - 31.5 Hz, is probably due to the different methods by which the sound pressure levels were measured and the fact that the headphone levels differ between the coupler used for calibration purposes and the normal head fitting position due to leakage through the seal. The rapid convergence of the two curves above 63 Hz suggests that a higher frequency crossover would have been more appropriate, but the choice is not considered detrimental since the findings of the study are based upon relative threshold measurements.

The pure-tone function is compared with previous threshold determinations in Figure 36. Fletcher and Munson measured the average threshold of eleven subjects in 1933 using headphones, and this curve has been and continues to be widely used.
and quoted. Their curve agrees closely with present data in the mid-frequency range (63 – 1000 Hz), although differences at higher frequencies become quite large. The second curve, which demonstrates the increase of threshold level with increasing age, is due to Robinson and Dadson, who performed a painstaking experiment with 90 subjects to determine threshold levels for totally free field exposure (Minimum Audible Field). The derived levels are those measured at the center of the head position in the absence of the subject, and this will account to some extent for the lower threshold evident in the 20-year-old curve around 4000 Hz. As Wiener has shown, sound diffraction patterns around the head cause an increase in sound pressure level at the entrance to the ear, relative to that in the undisturbed field, by 10 or more dB at frequencies above 1000 Hz. At frequencies below 1000 Hz, the Robinson and Dadson data is considerably lower than the other curves and is probably due to differences in subject performance and extra-auditory effects not present with headphone stimulation. At very low frequencies the curves tend to converge, although the previous curves terminate at 25 Hz. One of the few previous studies of very low frequency noise was performed by Von Bekesy in 1936, and a report in Reference 53 explains that this curve corresponds to the Minimum Audible Pressure measured at the eardrum. In any event, this data is rather higher in level than that from the other sources. As noted previously, the subtle differences between the various results are of no concern here since the only requirement was to determine all thresholds in the same way to provide comparative results for difference sound sources.

The absolute threshold for 1/3-octave bands of stationary random noise is compared with the pure-tone curve in Figure 37. Again two separate sets of data are shown, and the same comments regarding agreement apply. The differences between the two curves are small but consistent. At low frequencies the noise threshold is lower than the tone threshold, whereas at frequencies above 1000 Hz the converse is true. It is likely that the low-frequency difference is related to the fact that low-frequency narrow bands of random noise differ from pure tones mainly in that their rms levels vary with time. In fact, subjectively, a low-frequency band of noise sounds precisely like a tone with the same center frequency whose intensity fluctuates in a random manner. Based on the analysis presented in Section 2.1, Figure 38 has been prepared to show the level in dB relative to the true, long time averaged level, which is exceeded by a narrow band of noise for 10% of the time. An averaging time of 200 msec was assumed to be typical of the hearing system as discussed in Section 3.4. If it is appropriate to suppose that a listener can detect the most intense 10% of the signal, then the fact that the curve of Figure 38 agrees closely with the difference between the tone and 1/3-octave noise thresholds in Figure 40 supports the value of 200 msec for the averaging time for the hearing system. However, the choice is arbitrary, and a higher percentage would imply a smaller averaging time.

The difference between the two curves at high frequencies is largely a critical bandwidth effect. Accepting that the pure-tone curve is also correct for critical
bands of noise (since both sounds stimulate the same region of the basilar membrane),
the fact that the 1/3-octave bandwidths are greater than critical bandwidths at high
frequencies explains the increase in threshold level. This is discussed further in
Section 5.3.

Also shown in Figure 37 is the 1/3-octave band spectrum of the ambient noise
in the test chamber. Measurements above 160 Hz are uncertain due to inherent micro-
phone noise, but levels are generally more than 10 dB below the threshold level.

The thresholds measured for octave bands of noise are presented in Figure 39 in
comparison with both the tone and the 1/3-octave curves. Similar comments are
applicable to the tone versus noise comparison, although the differences are smaller
at low frequencies and greater at high frequencies. This is precisely as might be
expected. Because of the increased bandwidth, the rms level fluctuations are decreased
at low frequencies, as shown by the octave band curve in Figure 38; whereas, the
increased difference at high frequencies reflects the higher octave/critical bandwidth
ratio.

The results of various degrees of amplitude modulation of the random noise are
shown in the 1/3-octave band thresholds plotted in Figure 40, where the curve for
stationary noise is also included. The modulations were all impressed at a frequency of
10 Hz, typical of helicopter main rotor blade passage frequencies, and at levels of 3, 6
and 9 dB (corresponding to peak-to-trough pressure ratios of 1.4, 2 and 2.8 respec-
tively). Although some slight differences may be observed, there are no obvious trends,
and it is felt that these cannot be regarded as significant, particularly at the high fre-
quencies. It should be noted that these modulation depths are equivalent, in the
case of rotor broadband noise, to tip Mach numbers of 0.17, 0.33, and 0.47, which
are perhaps rather low, but were restricted by the capacity of the modulator available.
The corresponding peak-to-mean sound pressure level ratios (crest factors) are computed
to be 1.2, 2.2 and 2.9 dB respectively.

Absolute thresholds were also measured for 1/3-octave bands of harmonic noise
with fundamental frequencies of 10, 20 and 40 Hz respectively. The signals were
generated with zero interharmonic phase, but instrumentation response may be expected
to have a significant effect upon the observed phase differences. In each case the
test included the 1/3-octave bands covering the range between the 3rd and 100th
harmonics of the fundamental. The three sets of results are compared with the pure-
tone threshold in Figure 41. Only slight differences between the various data may
be noted, and again these are somewhat random. This is almost certainly true
between 50 and 160 Hz where the low-frequency bands pass signals which are
essentially sinusoidal with only very slight amplitude variations due to interharmonic
beating (see Figure 20). The differences above 1000 Hz, where the threshold
should be influenced by the critical band effect, are too small to warrant much
discussion. Of particular significance here is that the outputs of the higher
frequency bands are essentially highly modulated (Figure 20) with crest factors of as much as 9 dB. The amount of modulation perceived will be highly dependent upon aural averaging time as indicated by Figure 22, which shows the peak-to-mean SPL ratio as a function of this factor. If, as has been suggested by studies of amplitude modulated random noise\(^4\), the averaging time is much less than 100 msec, then for the 10-Hz signal, a listener would hear intermittent levels considerably in excess of the rms levels indicated in Figure 41. It is quite likely, therefore, that these observed level fluctuations depress the threshold (just as do the random fluctuations in low-frequency bands of noise) below the value which might be expected on a critical band basis. However, even though the modulations are very apparent in all bands, the auditory averaging time and consequently the depth of the perceived modulations are unknown. It can be stated that the apparent perceived modulation diminishes as the modulation (fundamental) frequency increases, a fact which again points to the role of auditory temporal averaging.

5.2 MASKED THRESHOLD DETERMINATIONS

In order to investigate the masking effects of ambient noise, similar experiments were repeated in which the signals were mixed with wideband noise which was essentially flat as measured by a constant-percentage bandwidth analyzer. The masking of pure tones by this noise at two levels is illustrated in Figure 42. The broken lines show both the absolute threshold for pure tones and the 1/3-octave band levels of the masking noise. It is apparent that the low-frequency thresholds are controlled by the absolute hearing ability, whereas the high frequency thresholds are controlled by the presence of the masking noise. Although there are some differences between the threshold signal-to-noise ratios at the two levels, an attempt has been made to minimize these differences in the fixed curves. It may be seen that the level of the just-audible tone decreases, relative to the 1/3-octave band level of the ambient, as frequency increases, being typically 10 dB below it at the high frequencies. It is interesting to compare Figure 42 with the results of Hawkins and Stevens\(^3\), which are reproduced in Figure 43. Figure 43 shows masked tone thresholds for four different masking levels of "white" noise. Although their data extends down only to a frequency of 100 Hz, they bear a good resemblance to the present ones, although detailed inspection reveals some notable differences in the mid-frequency range.

Results for 1/3-octave bands of noise, both stationary and modulated, are presented in Figure 44. It may be seen that in the region where the masking noise is well above absolute threshold, the differential threshold is roughly constant at about -5 dB. In other words, the band of noise is just detectable when the existing level of noise in that band is raised by 1 dB (since the addition of two uncorrelated signals which differ in level by 5 dB gives a combined level 1 dB greater than that of the highest level). In terms of the auditory mechanism it can be stated that a noise signal
is audible in an ambient noise when the combined critical band level is increased by 1 dB. This is greater than the generally accepted just noticeable difference (JND) of 0.5 dB (which is produced by a noise 9 dB less than ambient).

Use of the critical band concept leads to an explanation of the difference between the masked threshold for tones and bands of noise which is evident in Figure 44. Since we may assume that the tone is also audible when it raises the combined critical band level by 1 dB, the above ratio gives a direct measure of the critical bandwidth. Use will be made of this in Section 5.3.

In Figure 45 the results for the masking of octave bands of noise are presented. These are entirely consistent with the 1/3-octave band data since the 1/3-octave band components of the just-masked octave bands lie approximately 5 dB below ambient 1/3-octave band levels.

Figure 46 compares the masked thresholds for tones and 1/3-octave bands of harmonic noise; again, as in the case of absolute thresholds, the two curves are essentially coincident. In this case, however, it is clear that the threshold for filtered harmonic noise is decidedly lower than that for 1/3-octave bands of random noise. It can only be concluded that this difference is attributable to the high modulation level in the case of the harmonic complexes. Depending upon averaging times, the peak-to-mean SPLs for these signals can be as high as 9 dB (Figure 22) and substantially greater than those associated with the modulated noise signals studied (up to 3 dB).

5.3 AUDITORY FREQUENCY AND TEMPORAL RESOLUTION

Critical Bands

It is clear from the results plotted in Figure 44 that the "critical ratio" measured by Fletcher and Munson and later by Hawkins and Stevens (see Figure 28) is a function of two parameters. It first depends upon the width of the critical band, the discriminatory filter of the hearing mechanism, and secondly, the minimum perceptible differences in the critical band level caused by the addition of the tone. The factor of 2.5 noted by Zwicker et al. is between their critical band function and the "critical ratio" is in fact the just-noticeable signal increment which, expressed in logarithmic units, is 4 dB. This corresponds closely to the value of 5 dB observed directly for bands of noise in Figures 44 and 45*. Figure 44 can also be used to obtain a direct measure of the critical bandwidth by equating the energy in the just-audible tone to that in the critical bandwidth of the just-audible noise; i.e., since

*But see paragraph 5.5
\[ L = N_s + 10 \log_{10} \Delta f' \]  

(24)

where \( L \) is the SPL of the just-audible tone, \( N_s \) is the PSD of the just-audible noise signal, and \( \Delta f' \) is the critical bandwidth. Since \( N_s = N_3 - \Delta f_3 \), where the subscript 3 denotes 1/3-octave quantities, then

\[ 10 \log_{10} \Delta f' = L - N_3 + 10 \log_{10} \Delta f_3 \]  

(25)

The critical band function derived in this way is compared with those of Zwicker et al. and Greenwood in Figure 47.

The same equation (25) should be true whether the threshold is a masked threshold or an absolute threshold (in quiet). It has therefore been applied to the data from Figures 40 and 42 to obtain further estimates of the critical bandwidth function, which are presented in Figure 50.

The three curves are perhaps more notable for their differences than their similarities, a fact which corroborates the conclusions of Swets et al., De Boer, and others, that critical bandwidths are very difficult to measure, being a function of the measurement method, the assumed filter function, and many other psychosensory variables. On the other hand, the results do at least straddle the previously obtained values, tending to favor that due to Greenwood. Because of this, the fact that Greenwood did use measurements made at low frequencies and because his function has a convenient and simple mathematical description, it seems most appropriate to rely upon it for an aural detection model.

Multiband Detection

An important question which arises in the practical application of threshold data for tones and narrow bands of noise is whether the simultaneous detection of more than one band or frequency component influences the combined threshold level. To investigate this problem, a test was performed to measure the masked threshold level of a noise signal which varied in bandwidth steps between one single 1/3 octave (centered at 500 Hz) and 13 bands covering the range 125 to 2000 Hz. The spectrum of the signal was adjusted so that each band was equally detectable according to the finding that the differential threshold for bands of noise is -5 dB. The results, illustrated diagrammatically in Figure 48(a), demonstrated that the masked threshold decreased at a slow rate as the number of just-detectable bands increased. The rate from Figure 48(b) is approximately \(-N/4 \) dB where \( N \) is the number of bands. Although this data is very limited, it does suggest that the depression of the threshold by multiple band detection is a small effect since in general it is likely that detection will be confined to a relatively small region of the frequency spectrum.
Averaging Time

Although the literature cites auditory averaging times between 10 msec and 200 msec, several considerations suggest that a value nearer to the latter is probably more accurate. The first evidence is described in Section 5.1 in connection with the different threshold levels for low-frequency bands of noise. The second is related to the fact that level fluctuations in bands of noise become less perceptible as frequency increases. For 1/3-octave bands of Gaussian random noise, the transition from unsteady to steady sound occurs around 4000 Hz. For octave bands it occurs at a little lower frequency. Naturally, this is a highly subjective phenomenon, and the above statement is based upon very few observations, but it does agree with the peak-to-steady rms data presented in Figure 38 for a 200 msec averaging time. The curves for octave and 1/3-octave bands of noise cross the 0.5-dB just-noticeable difference line around the above-mentioned frequencies. In a similar way, the perception of amplitude modulations in filtered harmonic noise decreases as modulation frequency increases. The 40-Hz modulation frequency used in the experiments appeared to approach the limit of perception. This would seem unlikely if the averaging time were 10 msec (i.e., $0.25 \times$ modulation period), giving a peak-to-true level ratio of around 5 dB. For similar reasons, the agreement between both absolute and masked thresholds for filtered harmonic noise would probably diverge widely if the averaging time were very small.

5.4 MODEL FOR HELICOPTER AURAL DETECTABILITY

The experimental results presented in the previous sections provide the basis for calculating the aural detectability of helicopter noise in the light of the following conclusions:

1) It is reasonable to assume that a unique absolute audibility threshold function exists which is the same for constant-amplitude tones and for critical bands of random-noise where the latter should ideally be measured as the 90th percentile level obtained from the output of a sound pressure level detector with an averaging time of around 200 msec.

2) High levels of amplitude modulation do appear to cause an increase in signal detectability. Although random signals with amplitude modulations of up to 3 dB peak-to-average SPL (corresponding to 9 dB peak-to-trough SPL) did not reveal this increase, modulations as high as 9 dB in the case of filtered harmonic noise did indicate a noticeable lowering of the threshold. Again, it seems that peak sound pressure levels recorded by a system with a 200-msec averaging time are appropriate for the specification of detection level.
3) The critical band function established by Greenwood provides a convenient explanation for the observed differences between both masked and absolute thresholds of the various sounds studied.

4) At levels well above the absolute threshold, a narrow-band signal is audible in a noise background when the combined level of any critical band is increased by 1 dB, i.e., when the critical band level of the signal is increased to within 5 dB of that of the masking noise.

5) Simultaneous detection of many adjacent frequency bands causes a small depression of the threshold. However, the effect is sufficiently small to be ignored for practical purposes.

6) In the presence of masking noise, the combined threshold level may be calculated by (decibel) addition of the absolute threshold for tones and the critical band masking level.

Since measuring instruments incorporating critical band filters cannot be obtained commercially, it is necessary in practice to use filters that can yield an adequate approximation to the critical band spectrum. One-third octave band filters are most convenient for this purpose, although other bandwidths can be used with greater or lesser accuracy.

In any event, to compute the audibility threshold level of a helicopter noise spectrum in a particular ambient noise environment, it is necessary (a) to convert the ambient noise data to the form of a critical band spectrum, (b) to define a critical band masking level which is 5 dB less than the critical band ambient spectrum, (c) to combine an appropriate absolute threshold of hearing with the masking level (by decibel addition) to establish a combined threshold function, (d) to convert the helicopter noise data to a critical band spectrum, and finally, (e) to adjust the overall level of this spectrum to the highest value at which no critical band level exceeds the combined threshold level.

When computing a detection distance, step (e) is a lengthy process because the observed helicopter spectrum changes its frequency dependence with distance due to sound absorption. It thus becomes necessary either to examine the variation of each individual critical band level with distance to determine which one is critical, or to estimate a detection distance, compute the difference between the signal and the combined threshold and iterate toward an exact solution based on the magnitude of the error.

Detailed procedures for the calculation of both detection thresholds and detection distances, together with methods for converting both octave and 1/3-octave band data to critical band spectra are presented in Appendix IV. The next section includes
the results of the Phase II experiments and demonstrates the level of accuracy which may be expected of these procedures.

5.5 APPLICABILITY OF THRESHOLD PREDICTION PROCEDURES

Figure 49 shows the eight ambient noise conditions established for the eight tests which comprised the Phase II experiments. The lowest ambient (Test 1) is the noise floor of the test chamber and corresponds to "quiet" conditions. The remaining masking noises were mixed with the signal and generated by the loudspeakers and headphones. The broken line superimposed on these profiles is the average pure-tone since inspection of individual results indicated that two of the original ten subjects performed very poorly (unacceptably high standard deviations of their threshold levels). Even for the remaining eight, the standard deviation of the threshold (as previously defined) was approximately 10 dB.

The eight combined critical band thresholds computed from the data shown in Figure 49 by the method described in Appendix IV, Method A, are presented in Figure 50. Note that the two lowest ambient levels cause only a very small deviation from the absolute tone threshold (lowest curve).

The 1/3-octave band spectrum of the helicopter recording was measured (for all frequencies between 12.5 Hz and 10,000 Hz) at 15-second intervals throughout its length and digitized for computer analysis. The absolute threshold level of the signal was computed in each case by applying the measured attenuation level averaged for all eight subjects. Each spectrum level corrected in this way for overall level and for the system frequency response was converted to an equivalent critical band level by the method of Appendix IV, Method A. The differences between the estimated threshold level $L_{ne}$ and the actual threshold level $L_n$ in each band at all instants of time were analyzed to derive the results presented in Figures 51 through 58 and Tables II and III. Figures 51 through 58 show (in 1/3-octave rather than critical band levels) the distributions of the measured threshold band levels about the theoretical values. Three curves are shown in each figure which correspond to the 75th, 92nd and 97.5th percentiles of the measured level distributions. These are levels exceeded by the measured threshold band levels 2.5%, 8% and 25% of the time respectively. These diagrams show very clearly that the theoretical threshold levels are exceeded, at some part of the audible frequency range, for around 25% of the time. Of direct interest is the average differences between the measured and theoretical threshold levels.

This information is listed in Table II in terms of the average for each individual test and the grand average for all tests. Two errors (differences) are analyzed: the
TABLE II. ANALYSIS OF THE THRESHOLD PREDICTION ERRORS

<table>
<thead>
<tr>
<th>TEST NO.</th>
<th>Error $L = \sum_{dB} (L_n - L_{ne})^*$</th>
<th>Error $L_{min} = (L_n - L_{ne})_{min}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>std. devn.</td>
</tr>
<tr>
<td>1</td>
<td>+4.8</td>
<td>2.4</td>
</tr>
<tr>
<td>2</td>
<td>+5.3</td>
<td>2.5</td>
</tr>
<tr>
<td>3</td>
<td>+4.4</td>
<td>2.9</td>
</tr>
<tr>
<td>4</td>
<td>+6.0</td>
<td>3.5</td>
</tr>
<tr>
<td>5</td>
<td>+3.3</td>
<td>2.8</td>
</tr>
<tr>
<td>6</td>
<td>+2.2</td>
<td>2.9</td>
</tr>
<tr>
<td>7</td>
<td>+3.6</td>
<td>3.4</td>
</tr>
<tr>
<td>8</td>
<td>+2.0</td>
<td>4.0</td>
</tr>
<tr>
<td>ALL</td>
<td>+4.0</td>
<td>3.6</td>
</tr>
</tbody>
</table>

$L_n =$ measured threshold level; $L_{ne} =$ theoretical threshold level.
minimum error \( L_{\text{min}} = (L_n - L_{\text{ne}})_{\text{min}} \), which is the minimum difference between the measured and theoretical levels found in any of the 30 frequency bands at one instant of time, and composite error \( \sum L = \sum_{\text{dB}} (L_n - L_{\text{ne}}) \), the decibel sum of all 30 differences at one instant of time. The latter is effectively the total intensity level of the signal relative to the threshold level.

The grand average errors are +1.0 dB and +4 dB respectively. In other words, the helicopter sounds were just detectable, on the average, when an individual critical band level increased to 1 dB above the theoretical combined threshold*, or when the decibel sum of such differences reached a value of 4 dB. The associated standard deviations of 3.9 dB and 3.6 dB suggest, in fact, that the second criterion is a little more consistent. However, the increase in practical complexity does not seem justified by the small reduction in variability. Furthermore, the average minimum difference, \( L_{\text{min}} \), seems to be more consistent from test to test, i.e., as the ambient level varies, than does \( \sum L \) (inter-test standard deviations of the mean are 1.3 dB and 1.5 dB respectively). The standard deviations of between 3 and 4 dB seem satisfactory in light of the subjective variability of approximately 7 dB, the average standard deviation for this experiment. Also, it is unlikely that the acoustic stimuli, both signal and noise, could ever be specified with greater accuracy; indeed for most applications it is probable that significantly larger errors might be expected.

It is of interest to examine the frequency distribution of the minimum error \( L_{\text{min}} \) presented for each test in Table III. For the lowest threshold levels (1), (2), and (7), which would be encountered in practice only in the quietest forest or jungle environments, detections are confined to a limited band of mid-frequencies. This is also clearly illustrated in Figures 51 and 52. As the ambient level of the "flat" ambient noise is increased (Tests 3 through 6), there is a noticeable shift of the most frequent detection bands to lower frequencies, as might be expected. When the "sloping" ambient noise is used in Test 8, the combined threshold curve slopes at practically the same rate as the typical helicopter noise spectrum and indicates that detections occur over a wide range of frequencies.

In the interpretations of these results, it is most important to recognize that the sounds studied were selected to include as wide a range of objective and subjective characteristics as possible. For the most part they are not typical of the sound observed at distances of tens of thousands of feet typical of helicopter detection ranges. In

* A possible explanation for this 1 dB increment is that it is the margin required for the listener to identify the sound as that of a helicopter (see Section 4.5).
<table>
<thead>
<tr>
<th>TEST NO.</th>
<th>CRITICAL BAND CENTER FREQUENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50</td>
</tr>
<tr>
<td>1</td>
<td>2 3 13 13 14 30 14 8 8</td>
</tr>
<tr>
<td>2</td>
<td>2 17 13 14 30 13 5 8</td>
</tr>
<tr>
<td>3</td>
<td>2 2 2 6 23 9 5 20 5 2 3 2</td>
</tr>
<tr>
<td>4</td>
<td>2 2 6 9 8 17 14 23 5 5 2 2 2 2 2</td>
</tr>
<tr>
<td>5</td>
<td>5 3 27 16 19 8 3 13 2 2 2 2</td>
</tr>
<tr>
<td>6</td>
<td>6 5 9 13 47 5 6 2 2 2 2 2</td>
</tr>
<tr>
<td>7</td>
<td>3 3 13 16 2 2 25 13 8 17</td>
</tr>
<tr>
<td>8</td>
<td>3 6 3 5 9 3 13 5 2 13 6 2 6 2 2 5 16</td>
</tr>
</tbody>
</table>
retrospect, it is perhaps unfortunate that the choice of signals and the unavoidable change from FM to direct record/reproduction equipment were jointly responsible for a shift of emphasis to higher frequencies. The Phase I results certainly permit confidence that the low frequency threshold functions are accurate, but further research is required to specify the magnitude and applicability of critical band functions in that region with more precision.

A comparison of Figures 49 and 36 reveals that the average pure-tone threshold for the subjects who participated in the tests is rather higher than the free field threshold for the 20-year old men studied by Robinson and Dodson. For general application, it is recommended that the free-field curve measured by Robinson and Dodson be used because (a) it is directly applicable to the case of helicopter detection conditions (at least in open country), and (b) it was obtained in experiments involving a large number of subjects. For convenience, the data from Reference 34 has been extrapolated down to 12.5 Hz on the basis of the present results and tabulated in Appendix IV.
LITERATURE CITED


64


47. Adcock, B. D., AN AUTOMATIC DATA SYSTEM FOR LARGE SCALE AUDIOMETRIC TESTING, Wyle Laboratories Technical Memorandum, Wyle Laboratories Research Staff, Hampton, Virginia (to be published).


APPENDIX IV

METHODS FOR CALCULATING HELICOPTER AURAL DETECTION THRESHOLDS

Three methods for the calculation of helicopter detection thresholds are presented in this appendix. A choice among them may be made on the basis of the degree of resolution of the available input data, the computational equipment available, the accuracy required in the specific application, and the time available to perform the computations. The first method, which in general is practical only for machine computation, requires specification of the 1/3-octave band level spectra for both the helicopter noise signal and the ambient noise at the observer location. These data are converted, as accurately as possible, to critical bandwidth spectra, and the transformed signal spectrum is compared to a combined threshold function. This is the method which was tested as described in Section 5.0 of the main report. The second method is still based upon the use of 1/3-octave band level data but is simplified by the adoption of an approximate method for the inclusion of the critical bandwidth effect. It is thought that this approach will be only slightly less accurate than the first method, and yet it offers considerable simplification of the computational steps. The final version requires only octave band spectral resolution and is otherwise identical to the second method. It is probably less accurate than either of the alternatives, but it is amenable to hand calculation.

Whatever the choice, the basic calculation indicates whether the particular helicopter noise spectrum is audible in the particular masking noise and by what margin. This margin is obtained as the greatest (or least) difference between the signal and a combined threshold. The true threshold level for this particular signal spectrum can be obtained by applying the appropriate dB adjustment to reduce the above difference to zero. In general, however, a change of signal level requires an adjustment of the helicopter position, and this in turn, due to frequency dependent attenuation effects, requires a modification to the spectrum shape. In this case the threshold is best established by iteration, basing successive estimates of the correct result upon the previous error. Again, this procedure is best performed by machine calculation.

In the following instructions, frequent use is made of the summation notations $\sum dB$ and $(L_1 + L_2 + \ldots)_dB$. These are used to denote the decibel summation of sound pressure levels according the formula

$$\sum_{n} dB (L_n) = (L_1 + L_2 + L_3 + \ldots)_dB$$

$$= 10 \log_{10} (10^{L_1/10} + 10^{L_2/10} + 10^{L_3/10} + \ldots)_dB$$

102
As an aid to manual calculation, a tabular method for sound pressure level summation is presented in Table VII at the end of this Appendix.

**METHOD A. "EXACT" CALCULATION USING 1/3-OCTAVE BAND LEVELS**

**Data Required**

$L_n$ for $n = 1$ to 30: the 1/3-octave band levels, in dB, of the helicopter noise spectrum at the observer location for the frequencies $f_n = 12.5, 16, 20, \ldots, 10,000$ Hz (or for whatever frequency range the data are available).

$M_n$ for $n = 1$ to 30: the 1/3-octave band levels of the ambient masking noise at the observer location for the same frequency range.

**Method**

**Step 1:** Convert the 1/3-octave band levels $L_n$ and $M_n$ to critical band levels $L'_n$ and $M'_n$ using the "exact" relationships described in Table VIII.

**Step 2:** Calculate the combined critical band threshold level $T'_n$ at each frequency as the decibel sum of the absolute threshold $A'_n$ (from Table IX) and the masking threshold $(M'_n - 5)$ dB; i.e.,

$$T'_n = A'_n + (M'_n - 5) \text{ dB}$$

Note that if the difference between $A'_n$ and $(M'_n - 5)$ exceeds 13 dB, it is sufficiently accurate to put $T'_n$ equal to the greater of $A'_n$ and $(M'_n - 5)$.

**Step 3:** Subtract the combined thresholds $T'_n$ from the critical band signal levels $L'_n$. If the greatest value is greater than +1 dB, it may be assumed that the signal is audible.

**Example**

A hypothetical example is worked in Table X. The 1/3-octave band levels of the helicopter signal and the ambient noise are listed in columns 4 and 5 and the absolute tone threshold is copied from Table IX into column 3.
Step 1 is executed in Tables XI and XII according to the instructions provided in Table 11.2, and the results are transferred to columns 6 and 9 of Table X. The critical band masking level obtained by subtracting 5 dB from the ambient levels is given in column 7. The combined threshold which is the decibel summation of the absolute tone threshold (3) and the masking level (7) is tabulated in column 8. Column 10 is the detection level which is the difference between the signal level and the threshold level. Column 11 shows the audible level in each critical band (all positive values of the column 10 entry minus 1 dB).

Thus, in this example the signal is audible in the four critical bands at 50, 63, 80 and 100 Hz. The most audible band is at 63 Hz, and the audibility threshold for the signal is 3 dB below that specified in column 4. The results of the example calculations are shown graphically in Figure 66.

**METHOD B. APPROXIMATE CALCULATION BASED ON 1/3-OCTAVE BAND LEVELS**

*Data Required*

As in Method A, L and M for n = 1 to 30, the 1/3-octave band levels of the helicopter signal and the ambient noise.

*Method*

For this approach, the absolute audibility threshold has been converted to an effective 1/3-octave band level threshold $A_n$ so that no critical band conversions are required. The computational steps required are as follows:

**Step 1:** Calculate the 1/3-octave band masking level by subtracting 5 dB from each of the 1/3-octave ambient noise levels.

**Step 2:** Establish the 1/3-octave band combined threshold by the decimal summation of absolute threshold $A_n$ (from Table IX) and the masking level; i.e.,

$$T_n = \left[ A_n + (M_n - 5) \right] \text{ dB}$$

*This method is only approximately correct for helicopter type spectra which decay fairly uniformly in the lowest bands. Errors will be greater for different type spectra.*
Step 3: Compute the detection level in each band by subtracting the threshold level from the signal level.

Step 4: The audible level is the amount by which each detection level exceeds 1 dB.

Example

The previous hypothetical data is reused for this example which is worked in Table XIII. Comparing the final results with those obtained using the exact method (Table X) it may be seen that the detection levels agree to within 1 dB. This is typical of the relative accuracy which may be expected.

METHOD C. APPROXIMATE CALCULATION BASED ON OCTAVE BAND LEVELS*

Data Required

$L_k^*$ and $M_k^*$ for $k = 1$ to $8$, the octave band sound pressure levels of the helicopter signal and the ambient noise in the frequency range 16 Hz to 8 kHz.

Method

The procedure is precisely the same as that for Method B above except that the equivalent octave band thresholds $L_k^*$ from Table XIV are used. Again, the procedure should be restricted to helicopter type spectra.

Step 1: Calculate the octave band masking levels by subtracting 5 dB from each of the octave band ambient noise levels.

Step 2: Establish the octave band combined threshold by the decibel addition of the absolute threshold $L_k^*$ and the masking level; i.e.,

$$T_k = \left[ L_k^* + (M_k^* - 5) \right] \text{ dB}$$

*This method is subject to the same restrictions regarding spectrum shape as Method B (q.v.).
Step 3: Compute the detection level in each band by subtracting the threshold level from the signal level.

Step 4: The audible level is the amount by which each detection level exceeds 1 dB.

Example

A complete example is worked in Table XV. The octave band levels listed in columns 3 and 5 correspond to the $1/3$-octave levels presented in the previous two examples. The final result is again very similar, i.e., an audible level of 3 dB in the octave band centered on 63 Hz. This demonstrates the usefulness of this procedure.
### TABLE VII. TABLE FOR THE ADDITION OF SOUND PRESSURE LEVELS
(TO THE NEAREST 0.5 dB)

<table>
<thead>
<tr>
<th>dB</th>
<th>-20</th>
<th>-10</th>
<th>-0</th>
<th>+10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.010</td>
<td>0.100</td>
<td>1.00</td>
<td>10.00</td>
</tr>
<tr>
<td>0.5</td>
<td>0.011</td>
<td>0.112</td>
<td>1.122</td>
<td>11.22</td>
</tr>
<tr>
<td>1.0</td>
<td>0.013</td>
<td>0.126</td>
<td>1.258</td>
<td>12.58</td>
</tr>
<tr>
<td>1.5</td>
<td>0.014</td>
<td>0.141</td>
<td>1.412</td>
<td>14.12</td>
</tr>
<tr>
<td>2.0</td>
<td>0.016</td>
<td>0.159</td>
<td>1.585</td>
<td>15.85</td>
</tr>
<tr>
<td>2.5</td>
<td>0.018</td>
<td>0.178</td>
<td>1.780</td>
<td>17.8</td>
</tr>
<tr>
<td>3.0</td>
<td>0.020</td>
<td>0.200</td>
<td>2.00</td>
<td>20.0</td>
</tr>
<tr>
<td>3.5</td>
<td>0.022</td>
<td>0.224</td>
<td>2.24</td>
<td>22.4</td>
</tr>
<tr>
<td>4.0</td>
<td>0.025</td>
<td>0.251</td>
<td>2.51</td>
<td>25.1</td>
</tr>
<tr>
<td>4.5</td>
<td>0.028</td>
<td>0.282</td>
<td>2.82</td>
<td>28.2</td>
</tr>
<tr>
<td>5.0</td>
<td>0.032</td>
<td>0.316</td>
<td>3.16</td>
<td>31.6</td>
</tr>
<tr>
<td>5.5</td>
<td>0.036</td>
<td>0.355</td>
<td>3.55</td>
<td>35.5</td>
</tr>
<tr>
<td>6.0</td>
<td>0.040</td>
<td>0.398</td>
<td>3.98</td>
<td>39.8</td>
</tr>
<tr>
<td>6.5</td>
<td>0.045</td>
<td>0.447</td>
<td>4.47</td>
<td>44.7</td>
</tr>
<tr>
<td>7.0</td>
<td>0.050</td>
<td>0.501</td>
<td>5.01</td>
<td>50.1</td>
</tr>
<tr>
<td>7.5</td>
<td>0.056</td>
<td>0.562</td>
<td>5.62</td>
<td>56.2</td>
</tr>
<tr>
<td>8.0</td>
<td>0.063</td>
<td>0.631</td>
<td>6.31</td>
<td>63.1</td>
</tr>
<tr>
<td>8.5</td>
<td>0.071</td>
<td>0.708</td>
<td>7.08</td>
<td>70.8</td>
</tr>
<tr>
<td>9.0</td>
<td>0.080</td>
<td>0.794</td>
<td>7.94</td>
<td>79.4</td>
</tr>
<tr>
<td>9.5</td>
<td>0.089</td>
<td>0.891</td>
<td>8.91</td>
<td>89.1</td>
</tr>
</tbody>
</table>

**Method:** Subtract the decade of the highest level from all values and convert each to an energy value using the table. Convert the sum of the energies to the nearest 1/2 dB level, remembering to replace the decade.

**Example:** To calculate $\sum dB (58 + 64.5 + 73.5 + 71.5) dB$, subtract 70 from each:

8.0 (-20), 4.5 (-10), 3.5 (-0), 1.5 (-0)

Energy values: 0.063 + 0.282 + 2.24 + 1.412 = 4.01

Nearest dB level: 6.0

Add back original 70: $\sum dB = 76.0 dB$
<table>
<thead>
<tr>
<th>n</th>
<th>f</th>
<th>Bn, (dB)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.5</td>
<td>9.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>8.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-8.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>7.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>6.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>31.5</td>
<td>6.0</td>
<td>-</td>
<td>-9</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>5.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-3.5</td>
<td>0</td>
<td>-1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>4.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-9</td>
<td>0</td>
<td>0</td>
<td>-2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>63</td>
<td>3.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-1</td>
<td>0</td>
<td>-3.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>80</td>
<td>3.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-2.5</td>
<td>0</td>
<td>-4.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-3.5</td>
<td>0</td>
<td>-6.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>125</td>
<td>1.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-5</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>160</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-10.5</td>
<td>0</td>
<td>-12.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>200</td>
<td>0.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>250</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>315</td>
<td>0.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>400</td>
<td>-0.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>500</td>
<td>-1.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>630</td>
<td>-1.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>800</td>
<td>-1.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>1,000</td>
<td>-1.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>1,250</td>
<td>-1.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>1,600</td>
<td>-2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>23</td>
<td>2,000</td>
<td>-2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>2,500</td>
<td>-2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>25</td>
<td>3,150</td>
<td>-2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>26</td>
<td>4,000</td>
<td>-2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>27</td>
<td>5,000</td>
<td>-2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>28</td>
<td>6,300</td>
<td>-2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>29</td>
<td>8,000</td>
<td>-2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>30</td>
<td>10,000</td>
<td>-2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**1) Exact Method:**
For frequencies below 250 Hz (i.e., n = 13) the critical bandwidth is greater than the 1/3-octave bandwidth at the same center frequency so that the critical band level must be obtained by summing the total or partial energies from a number of adjacent 1/3-octave bands. The formula for this addition is:

\[ L' = \sum_{i=1}^{14} \left[ L_i + B_{n_i} \right] \]

where \( L_i = L_n \) for \( i = n \)

(Note that the \( B_n \) are added algebraically to the \( L_n \) before the decibel summation across \( i \).)

**Example:**
Compute the sound pressure level in the critical band centered at 40 Hz when the 1/3-octave band levels at 31.5, 40, 50 and 63 Hz are 58, 54, 53 and 50.5 dB respectively (remaining levels may be ignored):

\[ L' = \sum_{i=1}^{14} \left[ 49 + 54 + 53 + 48.5 \right] = 58 \text{ dB} \]

For frequencies of 250 Hz and above, the exact method is the same as the approximate method below.

**2) Approximate Method:**
Add algebraically the increment \( R_{n_i} \) to the 1/3-octave level \( L_n' \); i.e., \( L_n' = L_n + R_n' \).
### TABLE IX. ABSOLUTE THRESHOLDS OF AUDIBILITY (IN QUIET) FOR PURE TONES (CRITICAL BANDS) AND 1/3-OCTAVE BAND NOISE LEVELS FOR FREE-FIELD LISTENING CONDITIONS.

(Adapted from the data of Robinson & Daidson)

<table>
<thead>
<tr>
<th>n</th>
<th>f_n (Hz)</th>
<th>A'_n</th>
<th>A_n</th>
<th>n</th>
<th>f_n (Hz)</th>
<th>A'_n</th>
<th>A_n</th>
<th>n</th>
<th>f_n (Hz)</th>
<th>A'_n</th>
<th>A_n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.5</td>
<td>83.0</td>
<td>79.5</td>
<td>11</td>
<td>125</td>
<td>21.5</td>
<td>20.0</td>
<td>21</td>
<td>1,250</td>
<td>3.0</td>
<td>4.5</td>
</tr>
<tr>
<td>2</td>
<td>16.0</td>
<td>77.0</td>
<td>71.5</td>
<td>12</td>
<td>160</td>
<td>18.0</td>
<td>17.0</td>
<td>22</td>
<td>1,600</td>
<td>2.5</td>
<td>4.5</td>
</tr>
<tr>
<td>3</td>
<td>20.0</td>
<td>70.0</td>
<td>62.0</td>
<td>13</td>
<td>200</td>
<td>14.5</td>
<td>14.0</td>
<td>23</td>
<td>2,000</td>
<td>1.5</td>
<td>3.5</td>
</tr>
<tr>
<td>4</td>
<td>25.0</td>
<td>63.0</td>
<td>53.0</td>
<td>14</td>
<td>250</td>
<td>12.0</td>
<td>12.0</td>
<td>24</td>
<td>2,500</td>
<td>0.0</td>
<td>2.0</td>
</tr>
<tr>
<td>5</td>
<td>31.5</td>
<td>56.5</td>
<td>48.5</td>
<td>15</td>
<td>315</td>
<td>9.0</td>
<td>9.5</td>
<td>25</td>
<td>3,150</td>
<td>-2.5</td>
<td>-0.5</td>
</tr>
<tr>
<td>6</td>
<td>40.0</td>
<td>50.0</td>
<td>44.0</td>
<td>16</td>
<td>400</td>
<td>7.0</td>
<td>7.5</td>
<td>26</td>
<td>4,000</td>
<td>-4.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>7</td>
<td>50.0</td>
<td>43.5</td>
<td>38.5</td>
<td>17</td>
<td>500</td>
<td>6.0</td>
<td>7.0</td>
<td>27</td>
<td>5,000</td>
<td>-1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>8</td>
<td>63.0</td>
<td>36.5</td>
<td>32.0</td>
<td>18</td>
<td>630</td>
<td>5.0</td>
<td>6.0</td>
<td>28</td>
<td>6,300</td>
<td>5.0</td>
<td>7.0</td>
</tr>
<tr>
<td>9</td>
<td>80.0</td>
<td>31.0</td>
<td>28.0</td>
<td>19</td>
<td>800</td>
<td>4.0</td>
<td>5.5</td>
<td>29</td>
<td>8,000</td>
<td>14.0</td>
<td>16.0</td>
</tr>
<tr>
<td>10</td>
<td>100.0</td>
<td>26.0</td>
<td>24.0</td>
<td>20</td>
<td>1,000</td>
<td>4.0</td>
<td>5.5</td>
<td>30</td>
<td>10,000</td>
<td>16.0</td>
<td>18.0</td>
</tr>
</tbody>
</table>

\[ n = \text{Band number} \]
\[ A'_n = \text{Tone (critical band) threshold} \]
\[ f_n = \text{Center frequency, Hz} \]
\[ A_n = \text{1/3-octave band threshold} \]

\[ \text{to nearest 0.5 dB} \]

**Note:** The 1/3-octave band thresholds \( A_n \) below 250 Hz are approximately correct for 1/3-octave band spectra which decay at 6 dB per octave, which are typical of helicopter noise. Errors will increase for different spectra. For narrow band sounds, e.g., for single 1/3-octave bands of noise below 200 Hz, the tone threshold \( A' \) gives the correct result.
<table>
<thead>
<tr>
<th>n</th>
<th>f&lt;sub&gt;n&lt;/sub&gt;</th>
<th>A&lt;sub&gt;n&lt;/sub&gt;</th>
<th>L&lt;sub&gt;n&lt;/sub&gt;</th>
<th>M&lt;sub&gt;n&lt;/sub&gt;</th>
<th>M&lt;sup&gt;1&lt;/sup&gt;-5&lt;sup&gt;th&lt;/sup&gt;</th>
<th>T&lt;sub&gt;n&lt;/sub&gt;</th>
<th>L&lt;sup&gt;1&lt;/sup&gt;&lt;sub&gt;n&lt;/sub&gt;</th>
<th>n</th>
<th>L&lt;sup&gt;1&lt;/sup&gt;&lt;sub&gt;T&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.5</td>
<td>83</td>
<td>54</td>
<td>30</td>
<td>35.5</td>
<td>30.5</td>
<td>83</td>
<td>57</td>
<td>-26</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>77</td>
<td>52</td>
<td>30</td>
<td>36</td>
<td>31</td>
<td>77</td>
<td>57</td>
<td>-20</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>70</td>
<td>48</td>
<td>30</td>
<td>37</td>
<td>32</td>
<td>70</td>
<td>57.5</td>
<td>-12.5</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>63</td>
<td>46.5</td>
<td>30</td>
<td>36.5</td>
<td>31.5</td>
<td>63</td>
<td>57.5</td>
<td>-5.5</td>
</tr>
<tr>
<td>5</td>
<td>31.5</td>
<td>56.5</td>
<td>44</td>
<td>30</td>
<td>36</td>
<td>31</td>
<td>56.5</td>
<td>52.0</td>
<td>-4.5</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>50</td>
<td>42</td>
<td>30</td>
<td>35</td>
<td>30</td>
<td>50</td>
<td>48.5</td>
<td>-1.5</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>43.5</td>
<td>40</td>
<td>30</td>
<td>34</td>
<td>29</td>
<td>43.5</td>
<td>45</td>
<td>1.5</td>
</tr>
<tr>
<td>8</td>
<td>63</td>
<td>36.5</td>
<td>36.5</td>
<td>30</td>
<td>33.5</td>
<td>28.5</td>
<td>37</td>
<td>41</td>
<td>4.0</td>
</tr>
<tr>
<td>9</td>
<td>80</td>
<td>31</td>
<td>32</td>
<td>30</td>
<td>33</td>
<td>28</td>
<td>33</td>
<td>36.5</td>
<td>3.5</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>26</td>
<td>28</td>
<td>30</td>
<td>32</td>
<td>27</td>
<td>29.5</td>
<td>31.5</td>
<td>2.5</td>
</tr>
<tr>
<td>11</td>
<td>125</td>
<td>21.5</td>
<td>23.5</td>
<td>30</td>
<td>31.5</td>
<td>26.5</td>
<td>27.5</td>
<td>26.5</td>
<td>1.0</td>
</tr>
<tr>
<td>12</td>
<td>160</td>
<td>18</td>
<td>18</td>
<td>30</td>
<td>31</td>
<td>26</td>
<td>26.5</td>
<td>19</td>
<td>7.5</td>
</tr>
<tr>
<td>13</td>
<td>200</td>
<td>14.5</td>
<td>12.5</td>
<td>30</td>
<td>30.5</td>
<td>25.5</td>
<td>26</td>
<td>12.5</td>
<td>13.5</td>
</tr>
<tr>
<td>14</td>
<td>250</td>
<td>12</td>
<td>6</td>
<td>30</td>
<td>30</td>
<td>25</td>
<td>25</td>
<td>6</td>
<td>-19</td>
</tr>
<tr>
<td>15</td>
<td>315</td>
<td>9</td>
<td>0.5</td>
<td>30</td>
<td>29.5</td>
<td>24.5</td>
<td>24.5</td>
<td>0</td>
<td>-24.5</td>
</tr>
<tr>
<td>16</td>
<td>400</td>
<td>7</td>
<td>-9</td>
<td>30</td>
<td>29.5</td>
<td>24.5</td>
<td>24.5</td>
<td>-9.5</td>
<td>-34</td>
</tr>
<tr>
<td>17</td>
<td>500</td>
<td>6</td>
<td>-</td>
<td>30</td>
<td>29</td>
<td>24</td>
<td>24</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>630</td>
<td>5</td>
<td>-</td>
<td>30</td>
<td>29</td>
<td>24</td>
<td>24</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>800</td>
<td>4</td>
<td>-</td>
<td>30</td>
<td>28.5</td>
<td>23.5</td>
<td>23.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>1,000</td>
<td>4</td>
<td>-</td>
<td>30</td>
<td>28.5</td>
<td>23.5</td>
<td>23.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>1,250</td>
<td>3</td>
<td>-</td>
<td>30</td>
<td>28.5</td>
<td>23.5</td>
<td>23.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>1,600</td>
<td>2.5</td>
<td>-</td>
<td>30</td>
<td>28</td>
<td>23</td>
<td>23</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>23</td>
<td>2,000</td>
<td>1.5</td>
<td>-</td>
<td>30</td>
<td>28</td>
<td>23</td>
<td>23</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>2,500</td>
<td>0</td>
<td>-</td>
<td>30</td>
<td>28</td>
<td>23</td>
<td>23</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>25</td>
<td>3,150</td>
<td>-2.5</td>
<td>-</td>
<td>30</td>
<td>28</td>
<td>23</td>
<td>23</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>26</td>
<td>4,000</td>
<td>-4</td>
<td>-</td>
<td>30</td>
<td>28</td>
<td>23</td>
<td>23</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>27</td>
<td>5,000</td>
<td>-1</td>
<td>-</td>
<td>30</td>
<td>28</td>
<td>23</td>
<td>23</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>28</td>
<td>6,000</td>
<td>5</td>
<td>-</td>
<td>30</td>
<td>28</td>
<td>23</td>
<td>23</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>29</td>
<td>8,000</td>
<td>14</td>
<td>-</td>
<td>30</td>
<td>28</td>
<td>23</td>
<td>23.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>30</td>
<td>10,000</td>
<td>16</td>
<td>-</td>
<td>30</td>
<td>28</td>
<td>23</td>
<td>24</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
### TABLE XI. METHOD A EXAMPLE: "EXACT" COMPUTATION OF AMBIENT NOISE CRITICAL BAND LEVELS FROM 1/3-OCTAVE BAND LEVELS.

<table>
<thead>
<tr>
<th>i</th>
<th>( M_i + B ) ( B ) from Table VIII</th>
<th>( n_i )</th>
<th>( n_i )</th>
<th>( \Sigma_{dB} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.5</td>
<td>30</td>
<td>30</td>
<td>35.5</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>30</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>30</td>
<td>30</td>
<td>37</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>30</td>
<td>30</td>
<td>36.5</td>
</tr>
<tr>
<td>5</td>
<td>31.5</td>
<td>21</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>26.5</td>
<td>30</td>
<td>35</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>21</td>
<td>30</td>
<td>34</td>
</tr>
<tr>
<td>8</td>
<td>63</td>
<td>29</td>
<td>30</td>
<td>33.5</td>
</tr>
<tr>
<td>9</td>
<td>80</td>
<td>27.5</td>
<td>30</td>
<td>33</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>26.5</td>
<td>30</td>
<td>32</td>
</tr>
<tr>
<td>11</td>
<td>125</td>
<td>25</td>
<td>30</td>
<td>31.5</td>
</tr>
<tr>
<td>12</td>
<td>160</td>
<td>19.5</td>
<td>30</td>
<td>31</td>
</tr>
<tr>
<td>13</td>
<td>200</td>
<td></td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>n</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td>250</td>
<td>315</td>
<td>400</td>
<td>500</td>
<td>630</td>
<td>800</td>
<td>1000</td>
<td>1250</td>
<td>1600</td>
<td>2000</td>
<td>2500</td>
<td>3150</td>
<td>4000</td>
<td>5000</td>
<td>6300</td>
<td>8000</td>
<td>10K</td>
</tr>
<tr>
<td>( R_{n} )</td>
<td>0</td>
<td>-0.5</td>
<td>0.5</td>
<td>-1.0</td>
<td>-1.0</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>( M_{n} )</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>( M'_{n} )</td>
<td>30</td>
<td>29.5</td>
<td>29.5</td>
<td>29.5</td>
<td>29.5</td>
<td>28.5</td>
<td>28.5</td>
<td>28.5</td>
<td>28.5</td>
<td>28.5</td>
<td>28.5</td>
<td>28.5</td>
<td>28.5</td>
<td>28.5</td>
<td>28.5</td>
<td>28.5</td>
<td>28.5</td>
</tr>
</tbody>
</table>
### TABLE XII. METHOD A EXAMPLE: "EXACT" COMPUTATION OF HELICOPTER SIGNAL CRITICAL BAND LEVEL FROM 1/3-OCTAVE BAND LEVELS

<table>
<thead>
<tr>
<th>( i )</th>
<th>( L_i )</th>
<th>( n_f )</th>
<th>( L_i + B ) (( B ) from Table VIII)</th>
<th>( \sum_{dB} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.5</td>
<td>16</td>
<td>54</td>
<td>57</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>54</td>
<td>52</td>
<td>57</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>54</td>
<td>52</td>
<td>57.5</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>54</td>
<td>52</td>
<td>57.5</td>
</tr>
<tr>
<td>5</td>
<td>31.5</td>
<td>43</td>
<td>48</td>
<td>52.0</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>43</td>
<td>44</td>
<td>48.5</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>35</td>
<td>42</td>
<td>45</td>
</tr>
<tr>
<td>8</td>
<td>63</td>
<td>39</td>
<td>36.5</td>
<td>41</td>
</tr>
<tr>
<td>9</td>
<td>80</td>
<td>34</td>
<td>32</td>
<td>36.5</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>28.5</td>
<td>28</td>
<td>31.5</td>
</tr>
<tr>
<td>11</td>
<td>125</td>
<td>23</td>
<td>23.5</td>
<td>26.5</td>
</tr>
<tr>
<td>12</td>
<td>160</td>
<td>13</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>13</td>
<td>200</td>
<td>12.5</td>
<td>-8.5</td>
<td>12.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( n )</th>
<th>( f )</th>
<th>( R_n )</th>
<th>( L_n )</th>
<th>( L'_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>250</td>
<td>0</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>15</td>
<td>315</td>
<td>-0.5</td>
<td>-9</td>
<td>-9</td>
</tr>
<tr>
<td>16</td>
<td>400</td>
<td>-1.0</td>
<td>-1.5</td>
<td>-1.5</td>
</tr>
<tr>
<td>17</td>
<td>500</td>
<td>-1.5</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>18</td>
<td>630</td>
<td>-1.5</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>19</td>
<td>800</td>
<td>-1.0</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>20</td>
<td>1000</td>
<td>0</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>21</td>
<td>1250</td>
<td>-0.5</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>22</td>
<td>1600</td>
<td>-0.5</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>23</td>
<td>2000</td>
<td>-1.0</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>24</td>
<td>2500</td>
<td>-1.5</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>25</td>
<td>3150</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>26</td>
<td>4000</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>27</td>
<td>5000</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>28</td>
<td>6300</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>29</td>
<td>8000</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>30</td>
<td>10K</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-2.0</td>
</tr>
</tbody>
</table>

112
## TABLE XIII. WORKED EXAMPLE USING METHOD B.

<table>
<thead>
<tr>
<th>Band Number</th>
<th>Band Center Frequency (Hz)</th>
<th>1/3 Octave Threshold (From Table IX)</th>
<th>Helicopter 1/3 db Level (Input Data)</th>
<th>Masking Level (Input Data)</th>
<th>Combined 1/3 db Threshold ((= (3) + (5) \text{ db}))</th>
<th>Detection Level ((= (6) - 1 \text{ db}))</th>
<th>Audible Level ((= (7) - 1 \text{ db}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>( f_n )</td>
<td>( A_n )</td>
<td>( L_n )</td>
<td>( M_{-5} )</td>
<td>( T_n )</td>
<td>( D_n )</td>
<td>( S_n )</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>12.5</td>
<td>79.5</td>
<td>54</td>
<td>25</td>
<td>79.5</td>
<td>-25.5</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>71.5</td>
<td>52</td>
<td>25</td>
<td>71.5</td>
<td>-19.5</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>32</td>
<td>55</td>
<td>48</td>
<td>25</td>
<td>48.5</td>
<td>-4.5</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>44</td>
<td>42</td>
<td>25</td>
<td>42</td>
<td>-2</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>38.5</td>
<td>40</td>
<td>25</td>
<td>38.5</td>
<td>1.5</td>
<td>0.5</td>
</tr>
<tr>
<td>6</td>
<td>63</td>
<td>32.5</td>
<td>36.5</td>
<td>25</td>
<td>32.5</td>
<td>4.0</td>
<td>3.0</td>
</tr>
<tr>
<td>7</td>
<td>80</td>
<td>28</td>
<td>32</td>
<td>25</td>
<td>32</td>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>8</td>
<td>100</td>
<td>24</td>
<td>28</td>
<td>25</td>
<td>28</td>
<td>0.5</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>125</td>
<td>20</td>
<td>23.5</td>
<td>25</td>
<td>23.5</td>
<td>-2.5</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>160</td>
<td>17</td>
<td>18</td>
<td>25</td>
<td>18</td>
<td>-7.5</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>200</td>
<td>14</td>
<td>12.5</td>
<td>25</td>
<td>12.5</td>
<td>-13</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>250</td>
<td>12</td>
<td>6</td>
<td>25</td>
<td>6</td>
<td>-19</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>315</td>
<td>9.5</td>
<td>0.5</td>
<td>25</td>
<td>0.5</td>
<td>-24.5</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>400</td>
<td>7.5</td>
<td>-9</td>
<td>25</td>
<td>-9</td>
<td>-34</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>500</td>
<td>7</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>630</td>
<td>6</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>800</td>
<td>5.5</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>1,000</td>
<td>5.5</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>1,250</td>
<td>4.5</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>1,600</td>
<td>4.5</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>2,000</td>
<td>3.5</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>2,500</td>
<td>2</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>23</td>
<td>3,150</td>
<td>-0.5</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>4,000</td>
<td>-2</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>25</td>
<td>5,000</td>
<td>1</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>26</td>
<td>6,300</td>
<td>7</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>27</td>
<td>8,000</td>
<td>16</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>28</td>
<td>10,000</td>
<td>18</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
### TABLE XIV.
**Absolute Thresholds (in Quiet) for Octave Band Noise Levels -- Free Field Listening**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>k</td>
<td>f_k</td>
<td>A_k</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>16</td>
<td>76.5</td>
</tr>
<tr>
<td>2</td>
<td>31.5</td>
<td>53.5</td>
</tr>
<tr>
<td>3</td>
<td>63</td>
<td>37.5</td>
</tr>
<tr>
<td>4</td>
<td>125</td>
<td>24.5</td>
</tr>
<tr>
<td>5</td>
<td>250</td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>500</td>
<td>11</td>
</tr>
<tr>
<td>7</td>
<td>1,000</td>
<td>9.5</td>
</tr>
<tr>
<td>8</td>
<td>2,000</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>4,000</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>8,000</td>
<td>12</td>
</tr>
</tbody>
</table>

k = Band Number  
f_k = Center Frequency (Hz)  
A_k = Octave Band Threshold Level (dB)

N.B. Values only valid for helicopter type spectra.

### TABLE XV.
**Worked Example of Aural Detectability Calculation Using Method C -- Octave Band Data**

<table>
<thead>
<tr>
<th>Band</th>
<th>f_k</th>
<th>Threshold (from Table IV)</th>
<th>Helicopter Level (Input Data)</th>
<th>Masking Level (Input Data)</th>
<th>Combined Threshold (3)+(5) dB</th>
<th>Detection Level</th>
<th>Audible Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>76.5</td>
<td>57</td>
<td>30</td>
<td>76.5</td>
<td>-19.5</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>31.5</td>
<td>53.5</td>
<td>49.5</td>
<td>30</td>
<td>53.5</td>
<td>-4</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>63</td>
<td>37.5</td>
<td>42</td>
<td>30</td>
<td>38</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>125</td>
<td>24.5</td>
<td>29.5</td>
<td>30</td>
<td>31</td>
<td>-1.5</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>250</td>
<td>17</td>
<td>13.5</td>
<td>30</td>
<td>30</td>
<td>-16.5</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>500</td>
<td>11</td>
<td>-9</td>
<td>30</td>
<td>30</td>
<td>-39</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>1,000</td>
<td>9.5</td>
<td>-</td>
<td>30</td>
<td>30</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>2,000</td>
<td>7</td>
<td>-</td>
<td>30</td>
<td>30</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>4,000</td>
<td>3</td>
<td>-</td>
<td>30</td>
<td>30</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>8,000</td>
<td>12</td>
<td>-</td>
<td>30</td>
<td>30</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 26. Deformation Patterns of the Basilar Membrane for One Cycle of a 1000-Hz Tone (at 45° Intervals).

Figure 27. Various Determinations of the Threshold of Audibility and the Threshold of Feeling.
Figure 28. Comparison of Various Critical Bandwidth Measurements.
Figure 29. Typical Experimental Data on the Width of the Critical Band (from Reference 40).
Figure 31. Progressive Wave Acoustic Chamber.
Figure 32. Detection Experiments - Test Instrumentation.
Figure 33. System Frequency Response to Sinusoidal Input.
Figure 34. Absolute Audibility Threshold for Pure Tones.
Figure 35. Comparison of Apparent Thresholds for Headphone, Loudspeaker, or Combined Presentation of Stimulus.
Figure 36. Comparison of Present Results with Previous Pure Tone Threshold Data.
Figure 37. Audibility Threshold for 1/3-Octave Bands of Stationary Random Noise - In Quiet.
Figure 38. 90-th Percentile Levels for Bands of Random Noise with 200 msec Averaging Time.
Figure 39. Audibility Thresholds in Quiet - Comparison for Tones, 1/3-Octave and Octave Bands of Stationary Noise.
Figure 40. Audibility Thresholds for 1/3-Octave Bands of Noise - in Quiet.
Figure 41. Absolute Audibility Thresholds for 1/3-Octave Bands of Harmonic Noise.
Figure 42. Masked Threshold Levels for Pure Tones.
Figure 43. The Masking of Pure Tones by Wideband Noise (Data from Hawkins and Stevens).
Figure 44. Masked Thresholds for 1/3-Octave Bands of Random Noise.
Figure 45. Masked Thresholds for Octave Bands of Noise.
Figure 46. Masked Thresholds for 1/3-Octave Bands of Harmonic Noise.
From comparisons of absolute thresholds for tones and 1/3-octave bands of noise (Data from Figure 37).

From comparisons of absolute thresholds for tones and octave bands of noise (Figure 39).

From comparison of masked thresholds for tones and bands of noise (Figure 44).

Figure 47. Comparisons of Critical Bandwidths Determined by Various Criteria.
Figure 48. Depression of Masked Threshold by Multiple Band Detection.
Figure 49. 1/3-Octave Band Level Spectra of Masking Noises Used in Validation Tests.
Figure 50. Combined Critical Band Audibility Thresholds Computed from Data of Figure 49.
Figure 51. Percentile Distributions of 1/3-Octave Band Signal Levels at Threshold - Test 1.
Figure 52. Percentile Distributions of 1/3-Octave Band Signal Levels at Threshold - Test 2.
Figure 53. Percentile Distributions of 1/3-Octave Band Signal Levels at Threshold - Test 3.
Figure 54. Percentile Distributions of 1/3-Octave Band Signal Levels at Threshold - Test 4.
Figure 55. Percentile Distributions of 1/3-Octave Band Signal Levels at Threshold - Test 5.
Figure 56. Percentile Distributions of 1/3-Octave Band Signal Levels at Threshold - Test 6.
Figure 57. Percentile Distributions of 1/3-Octave Band Signal Levels at Threshold - Test 7.
Figure 58. Percentile Distributions of 1/3-Octave Band Signal Levels at Threshold - Test 8.
A REDETERMINATION OF THE NOY CONTOURS

Condensed from FAA Report No-70-3 "The Noisiness of Diffuse Sound Fields at High Intensities" by J B Ollerhead.

This paper, which was extracted by the author from his detailed technical report with the assistance of K M Eldred summarises an investigation of "equal noisiness contours" under extremes of sound field diffusivity. Although the noy curves incorporated in the standard EPNL procedures remain unchanged, S S Stevens also concluded in contemporary studies that some kind of averaging was necessary and took account of the author's results in the revision of his Loudness Level procedures.

A REDETERMINATION OF THE NOY CONTOURS

J. B. Ollerhead and K. M. Eldred
Wyle Laboratories
Hampton, Virginia, and El Segundo, California

Presented at the 81st Meeting of the Acoustical Society of America
Washington, D. C.
April 1971

1.0 INTRODUCTION

This study was performed as part of a continuing effort to evaluate and refine the Effective Perceived Noise Level (EPNL) scale for the purposes of rating aircraft noise. This scale has been developed via a continuous process of test and refinement from a 1959 concept of Kryter, who noted that human judgment of aircraft sounds appeared to be more closely correlated with the attribute of noisiness than with loudness. Accordingly he modified Stevens' basic loudness level calculation procedure by substituting contours of equal noisiness, in noys, for the loudness contours in sones. The noy curves were based on earlier (1944) measurements of equal "annoyance" for narrow bands of noise. In 1963 Kryter and Pearsons revised the Perceived Noise Level (PNL) method to include new equal noisiness data. A new contour was obtained from an experiment performed in a large classroom, the data from which are shown in Figure 1. The lower continuous curve was tabulated for use in the PNL computation and, despite more recent measurements by Parnell, Nagel and Parry in 1967, has remained in use until the present time.

Practical application of the PNL method involves the assumption that the relative noisiness of different frequency bands depends only on their relative intensities, at least at the higher sound pressure levels. The objectives of the present study were to check this assumption at sound pressure levels up to 120 dB and down to frequencies of 31.5 Hz, and also to examine the effect of listening conditions on the shape of the noy contour.

2.0 METHOD

The experiments were performed in two different test chambers which provided two extreme acoustic field environments. An essentially free field condition was generated inside a progressive wave chamber illustrated in Figure 2. This facility seated four subjects who faced an array of five 15-inch low frequency loudspeakers and a high frequency multi-cellular horn. Behind the subjects, a set of 12 feet deep fiberglass wedges efficiently absorbed the sound throughout the frequency range of interest. With this configuration the sound pressure level could be maintained to within +2 dB at all subject positions and all frequencies.

* This work was supported by the Federal Aviation Administration, Office of Noise Abatement.
** All sound pressure levels are referred to 2 x 10^-5 N/M^2.
In complete contrast, a highly diffuse field was generated inside a 100,000 ft.\(^3\) reverberation room which is shown in plan view in Figure 3. Due to absorption losses, this room required high power input to maintain a sound pressure level of 120 dB at high frequencies. The difficulty of meeting this requirement with low distortion loudspeakers was avoided by constructing an inner enclosure from canvas. The material was selected to be transmissive at low frequencies and reflective at high frequencies allowing the high frequency level to be maintained with moderate power sources inside the inner room. The sound pressure levels attained with this dual system are shown in Figure 4.

Twenty-four subjects took part and the method of paired comparison was used to establish equally noisy levels for one-third octave bands of noise in the frequency range between 31.5 Hz and 10 KHz. Comparisons were made relative to a standard reference sound comprising an octave band of noise centered at 1000 Hz, and, in the diffuse field, noy curves were measured at levels of 78, 86, 98 and 104 PNdB. A single curve at a perceived noise level of 100 PNdB was measured in the free field. The study is reported in detail in Reference 4.

3.0 RESULTS

The diffuse field results are presented in Figure 5. It is apparent that although the shapes of the few contours are very similar to each other, they do differ substantially from Kryter and Pearson's result. This is shown more clearly in Figure 6 where the envelope of the four curves has been collapsed by subtracting the PNL in each case. This envelope is coincident with the Kryter-Pearson curve at 1000 Hz (by definition), but at other frequencies is somewhat higher. The differences would of course appear smaller if the relative vertical positions of the two curves were adjusted slightly although this step is not justified according to the strict definition of perceived noise level. The difference at the low frequency is thought to be related to the high reverberation time of the test chamber, measurements of which are plotted in Figure 7. It may be noted that the longest decay time (over 20 seconds) occurs at precisely the frequency where the diffuse field contours level out and it thus seems possible that the noisiness judgments were influenced by the variable rise and decay times of the test sounds.

The free field result is shown as a broken line in Figure 8, and is remarkable for the substantial dip at 3150 Hz. To examine this phenomenon more closely a second test was performed using the method of adjustment at levels of 100 and 85 PNdB. The measurements are superimposed in Figure 8 where a marked difference may be seen, the latter contours bearing much more resemblance to the Kryter-Pearson curve.

However, the paired comparison result is very similar to that derived in a previous, independent experiment performed in the same facility (ref. 5). This may be seen in Figure 9 where several contours\(^1\) are compared, including Kryter's 1959 noy curve, the 1963 curve\(^2\), and a result from Parnell et al.\(^3\) All results are similar on the low frequencies but a large variation is apparent at frequencies above 2000 Hz.
4.0 CONCLUSIONS

It can only be concluded from a comparison of present and previous results that the apparent variation of noisiness with frequency is highly dependent upon the nature of the acoustic field and the test method employed. The difference between the equal noisiness curves measured by the methods of adjustment and paired comparison suggest that even the slightly different head orientations and motions adopted by the subjects cause large head diffraction variations in a free-field situation. Also, large differences were found between the diffuse field and free field results. With these observations in mind, the major conclusions of the study are as follows:

1) The equal noisiness curve currently used is substantiated for free-field listening conditions in the frequency range 80-2500 Hz.

2) The shape of the noy curve does not vary for levels up to 120 dB.

3) The existing curve can be confidently extrapolated down to 31.5 Hz.

4) The apparent judged noisiness of bands of noise at frequencies above 2500 Hz is very sensitive to the type of field and the test method.

To further improve the EPNL procedure for practical purposes, it is recommended that attempts be made to:

a) obtain an operational definition of an "average" real life environment, and

b) develop standardized noy curves for this standard field.

Implicit in recommendation (b) is that measurement techniques for noisiness contour evaluation be thoroughly investigated.
5.0 REFERENCES


FIGURE 1 - EQUAL NOISINESS CONTOURS FOR NARROW BANDS OF NOISE IN SEMI-REVERBERANT LISTENING CONDITIONS (KRYTER AND PEARSONS - REF 2)
FIGURE 2 - PROGRESSIVE WAVE TEST CHAMBER

- Anechoic Termination
- Screen
- Five 15-inch Low Frequency Loudspeakers
- One High Frequency Multicellular Horn Loudspeaker
FIGURE 3 - 100,000 FT³ REVERBERATION ROOM WITH CANVAS INNER ENCLOSURE
FIGURE 4 - SOUND PRESSURE LEVELS ACHIEVED IN REVERBERANT ROOM
FIGURE 5 - EQUAL NOISINESS CONTOURS - DIFFUSE FIELD - MEAN JUDGMENTS
FIGURE 6 - COMPARISON OF DIFFUSE FIELD CONTOURS WITH THAT OF KRYTER AND PEARSONS (REFERENCE 2)
FIGURE 7 - REVERBERATION TIME AND ACOUSTIC POWER REQUIRED TO MAINTAIN A SOUND PRESSURE LEVEL OF 120 dB
FIGURE 8 - COMPARISON OF FREE FIELD CONTOURS OBTAINED BY DIFFERENT METHODS
FIGURE 9 - COMPARISON OF FREE FIELD CONTOUR WITH PREVIOUS RESULTS
This is a condensed version of the full technical report which also included a comprehensive set of data. Both studies have been widely quoted and the raw data continues to be subjected to further analysis by other investigators. The scope of the experiment was unprecedented and the data set remains one of the largest in existence.

SCALING AIRCRAFT NOISE PERCEPTION†

J. B. OLLERHEAD

Department of Transport Technology,
University of Technology, Loughborough LE11 3TU, England

(Received 24 May 1972, and in revised form 14 November 1972)

Following a brief review of the background to the study an extensive experiment is described which was undertaken to assess the practical differences between numerous alternative methods for calculating the perceived levels of individual aircraft flyover sounds. One hundred and twenty recorded sounds, including jets, turboprops, piston aircraft and helicopters were rated by a panel of subjects in a pair comparison test. The results were analyzed to evaluate a number of noise rating procedures, in terms of their ability to accurately estimate both relative and "absolute" perceived noise levels over a wider dynamic range (84–115 dB SPL) than had generally been used in previous experiments. The performances of the different scales were examined in detail for different aircraft categories, and the merits of different band level summation procedures, frequency weighting functions, duration and tone corrections were investigated.

It was found that the complex procedures developed by Stevens, Zwicker and Kryter are generally more accurate than the weighted sound pressure level scales, particularly when integrated to include a signal duration allowance, although their main advantage lies with their ability to cope with signals over a wide range of bandwidth. However, Stevens' loudness level scale (Mk VI) and the Perceived Noise Level scale both overestimate the growth of perceived level with intensity because of a deficiency in the band level summation rule which is common to both. A simple correction is proposed which will enable both scales to properly account for the experimental observations. The differences between the various scales are small and it is shown that the D-weighted sound pressure level scale, perhaps somewhat fortuitously, provides a particularly reliable estimate of perceived level over a wide range of bandwidths. Indeed, when used with an integrated duration correction it is, on the whole, rather more accurate than the Effective Perceived Noise Level scale as presently formulated.

1. INTRODUCTION

The search for a suitable scale, upon which the subjective magnitude of aircraft sound can be accurately related to physical measurements, has been long and tortuous. After more than two decades of study, there remains considerable confusion about the relationships among a multitude of alternative approaches and the problem seems far from solved.

At the outset it must be stated that there are two distinct requirements. The first is for a procedure by which the long term effects of aircraft noise upon communities around airports can be estimated with a reasonable degree of confidence. In practice there are many factors in addition to the physical noise exposure which affect an individual's concern about aircraft noise, and attitude surveys have shown that the correlation between individual annoyance and noise level is relatively insensitive to the method used for measuring the latter. This problem will be the subject of a later paper; the subject of the present study is more closely related to the second requirement which is for a scale to measure the perceived level of an individual aircraft flyover sound. There is, of course, a certain compulsion to coordinate the two requirements on the intuitive grounds that there should only

† Work performed whilst the author was a member of Wyle Laboratories Research Staff, Hampton, Virginia, U.S.A.
be one solution to the problem which must by definition apply to both individual and multiple sounds, the one being merely some kind of summation of the other. This has led to the incorporation of the complex Effective Perceived Noise Level scale (EPNL) into the composite noise rating scale recently proposed [1] by the International Organisation for Standardization which in turn is similar to Noise Exposure Forecast (NEF), long used in the United States [2] for the purposes of airport planning. In the U.K., Robinson [3] has pursued his studies of the Noise Pollution Level concept to show that this rather elegant procedure, which accounts for fluctuations in level as well as the mean energy level, satisfies both requirements. In this case, however, the precise choice of a basic noise scale upon which to measure the levels is left to the user's discretion.

There is every reason to suppose that a major factor influencing community disturbance is the capacity of the aircraft noise to attract attention to itself. Humans can in fact adapt to surprisingly high levels of noise provided it has an inoffensive and unchanging character. Only when the sound changes in quality or interferes with communication (or rest) is it considered to be an intrusion. It was recognition of this effect that led to the inclusion of the "fluctuation term" in Noise Pollution Level [4]. However, other things being equal, it is assumed that once an aircraft sound is noticed, the listeners' objection to it will be a function of its judged loudness, noisiness or, in general, its "perceived level".

Factors which contribute to the perceived level of sound have been studied extensively for a period of more than forty years. The pioneering work of Steinberg [5], and Fletcher and Munson [6], revealed that the perceived level of sound is very sensitive to frequency and later Fletcher [7], Stevens [8] and Zwicker [9] hypothesized models of the process by which the ear integrates energy in different parts of the spectrum, defining rules for its mathematical simulation. In recent years the need for precision in the measurement of aircraft noise for flight certification purposes spurred considerable laboratory research, largely in the United States, into the possible roles of temporal variations in the sounds and the presence of intense concentrations of energy in very narrow segments of their spectra.

Unfortunately, perhaps because of inherent human variability and the difficulties of subjective measurement, these studies have led to little consensus of professional opinion and there has been much debate about the marginal differences between the accuracies of numerous procedures for scaling noise. Indeed this paper stems from the author's earlier participation in this contest [10] and reports an experiment [11] which was intended to seek a definitive solution to this problem. Before this is described, however, some of the earlier research will be briefly reviewed.

2. BACKGROUND

The earlier studies of perceived magnitude were concerned with the variations of loudness of pure tones with frequency. Fletcher and Munson [6] in 1933 produced the first set of loudness contours which formed the basis for the A, B and C weighting networks of modern sound level meters. These scales apply a linear transformation to the acoustic spectrum, weighting the sound energy by an approximation to the frequency response of the ear. In the A, B and C scales, however, the rather complicated energy summation process apparently applied by the ear is ignored.

These have been explained on the basis of the "critical band" hypothesis [7, 12], describing the role of the basilar membrane which apparently performs the aural frequency analysis. Sounds of different frequencies excite different regions of this long narrow organ allowing the ear to discriminate between them. However, since any single frequency component excites a finite length of the membrane, sounds of adjacent frequencies exert a
markedly reduced influence. In other words, the first component partially (or totally) masks the second. Zwicker [9] performed a very thorough analysis of aural masking and developed a graphical procedure for summing the “partial loudnesses” of adjacent bands of noise.

Figure 1. Loudness calculation chart for 1/3-octave band spectra with use demonstrated for 1/3-octave band of noise centred at 1000 Hz. (After Zwicker [24].)

An example of Zwicker’s procedure is shown in Figure 1 where the bold lines denote the “loudness energy” contained in the one-third-octave band of noise centered at 1 kHz when heard at each of the four sound pressure levels 60, 70, 80 and 90 dB SPL. The loudness in each case is proportional to the area under the curve and the right-hand “tails” account for the “upward spread of masking.” This illustrates that any sound is much more effective in masking sounds at higher frequencies than at lower ones. As the sound pressure level of this narrow band of noise is increased by 10 dB, the area under the curve, and thus the loudness, is doubled. It will be noticed that the widths of the 1/3-octave band “columns” varies with frequency. Zwicker’s original charts were in fact based on the equal width “critical bandwidths” but he revised the diagrams for use with the more convenient 1/3-octave bandwidths.

It should also be noted that if the energy is spread fairly uniformly over several adjacent bands the total loudness is still proportional to the total area under the curve so that “energy” in the tails is masked. Thus for wideband noise the loudness doubles for increases in sound pressure level which are rather less than 10 dB and can be as low as 8·3 dB. The significance of this will be discussed in more detail later.

In parallel studies to Zwicker’s, S. S. Stevens [8] developed an alternative procedure for computing loudness levels. Because of its more attractive simplicity this method tends to be much more widely used than the graphical technique.

In Stevens' method, the complex sound is again analyzed into narrow band levels and a “loudness index” is determined for each which includes the effect of the “sideband contribution.” The masking effect is accounted for in a rule for summing the individual loudness indices. The rule, very simply, follows from the assumption that the loudest band contributes fully to the total loudness. However, because of the masking effect, the con-
tributions of all other bands are inhibited to a fraction of their loudness when heard alone. The equation is

\[ S_i = S_m + F[\Sigma S - S_m], \]

where \( S_i \) is the total loudness, \( S_m \) is the loudness index of the loudest band and \( \Sigma S \) is the sum of all band loudness indices. The summation factor \( F \) was determined experimentally to be 0.3 for octave bands or 0.15 for 1/3-octave bands of noise.

In both Zwicker's and Stevens' loudness methods, the loudness level \((LL)\) is obtained from the total loudness by the formula

\[ LL = 33.3 \log_{10} S_i + 40. \]

This corresponds to the definition that unit loudness \((1 \text{ sone})\) is the loudness of a 1 kHz tone at a sound pressure level of 40 dB SPL and reflects the empirical power law relationship that

\[ \text{loudness} \sim (\text{intensity})^{0.3}. \]

In 1959 Kryter [13] performed a subjective experiment to determine how "noisy" the then new commercial jet aircraft were going to sound to people on the ground in comparison with existing propeller driven aircraft. He tested Stevens' loudness level scale \((LL_s)\) along with various other rating scales for their ability to accurately predict judged differences between a number of recorded aircraft sounds. He found that \( LL_s \) performed rather badly but when he substituted contours of "equal annoyance" for those of "equal loudness" in the loudness level computation, he discovered that the revised scale performed better than most of the alternatives. He concluded that the attribute of "noisiness" is different to that of loudness and is more relevant to the aircraft noise problem. He termed the revised scale "perceived noise level" \((PNL)\).

At a later date Little [14], followed by Kryter and Pearsons [15], observed that the \( PNL \) procedure did not adequately account for the presence of intense pure tones in an otherwise broadband sound, a combination typical of jet aircraft sound. Although the investigators disagreed on the magnitude of the tone effect its existence seemed proved and tentative procedures were proposed for a correction term based on conventional spectral analysis.

Kryter and Pearsons [15] also investigated the effects of signal duration and found that judged noisiness varied with signal duration in the range 1.5 to 12 seconds and that the increment was equivalent to 4.5 PNdB per doubling of duration. Pearsons [16] later extended these studies to cover durations up to 64 seconds and determined that the effect of duration on perceived noisiness is a continuously varying function of level, varying from 6 dB per doubling of duration at low durations to 2 dB per doubling at high durations. However an average value of 3 dB per doubling of duration, which is consistent with the notion that noisiness is proportional to the total incident energy, was adopted into a revised scale [17] known as Effective Perceived Noise Level \((EPNL)\). This included a duration allowance based upon the effective duration of the signal. Originally defined as the time interval for which the instantaneous perceived noise level was within 10 PNdB of the peak, the effective duration \( T_e \) is now established by integration according to the equation

\[ T_e = \frac{\int_{0}^{10 \text{ PNL/10}} df}{\int_{0}^{10 \text{ PNL/peak/10}} df}. \]

Methods for the calculation of \( EPNL \) (which includes both tone and duration corrections) have since been defined in precise detail and the scale is now used for the noise
certification of aircraft in both the U.S.A. [18] and the U.K. [19]. However its final formulation was one of innumerable versions which have been tested and retested in a large number of experiments performed by Kryter and others.

A review of this research into aircraft noise rating scales yields a somewhat confused picture but the main facts which appear to emerge may be summarized as follows.

(a) Since the earliest experiments involving aircraft noise, there has been a preponderance of emphasis upon the PNL scales. It has been extensively studied, revised, extended and varied in the minutest detail in attempts to improve correlation with experimental observation. Such attention has not been devoted to other basic scales. For example, the duration correction had not been applied by any investigator except the author to any of the sound pressure level scales until as late as 1969. Also, Zwicker's scale has hardly been used at all.

(b) Most scales and their refinements have been developed on the basis of studies involving steady state or highly controllable laboratory generated (synthetic) sounds, often with extreme spectral and temporal features. Few studies indicate any real, statistically significant differences between a multitude of variations upon the basic scales when they are used to rate real or recorded aircraft flyover sounds.

(c) Individual experiments have usually produced only a few data points which, because of the extremely variable nature of the problem, have represented small statistical samples. The potential sources of error, even under the most highly controlled test conditions, are numerous and many experiments, particularly those performed in "semi-diffuse" test rooms or out of doors, are prone to particularly large noise measurement errors.

It was this rather confused background which led to the inception of the present study. To fulfill what seemed to be a clear need for a thorough test of the basic noise rating scales, a large scale subjective test was designed to discriminate practical differences with a reasonable degree of confidence. For a full description of the study the reader is referred to the original technical reports [11, 20]. The experimental design and results are summarised in the following sections.

3. EXPERIMENTS

The basic objective of the experiment was to obtain as large a set of subjectively measured perceived levels of aircraft flyover sounds as possible and to correlate the judgments with the levels calculated by various rating procedures. In this way it would hopefully be possible to (a) make a realistic comparison of these various procedures, (b) find out how well they performed in an absolute sense, and (c) make recommendations for further improvements which might improve their validity for rating the sounds of aircraft.

A total of 120 aircraft flyover sound recordings, selected from various sources, were divided roughly equally into four major categories: jets (turbojet and turbofan), propeller turbine powered aircraft, piston engined aircraft, and helicopters.

The sounds included outdoor recordings of flyovers, take-offs and landings with the microphone located at various positions with respect to the flight path so that the sounds comprised a wide assortment of those which might be heard on or around a mixed traffic airport. Reference [11] contains a complete listing of the 120 sounds, which includes known or estimated data describing the aircraft type and classification, the flight mode, the slant distance between the aircraft and the microphone at its nearest point of approach, and the peak sound pressure level at the microphone location.

In order to rank the sounds upon an absolute scale of judged perceived level, each was
compared either directly or indirectly with a "standard reference" sound consisting of an octave band of "pink" noise (i.e., random noise with a uniform spectrum level as measured by a constant percentage bandwidth analysis) centered at a frequency of 1000 Hz. The measured perceived level of each could then be expressed in terms of the sound pressure level of the standard reference when it was judged to have an equal perceived magnitude. This is, of course, closely related to the basic definitions of many of the scales for calculating perceived level, particularly PN\*L. In fact, the judged level, obtained in this way, only differs from a measured PN\*L in the amount the bandwidth of the standard reference differs from an ideal octave due to the filter skirts and the finite signal-to-noise ratio of the sound generation system.

The basic test method employed in the subjective experiments was a pair comparison technique which had been developed and evaluated in earlier studies [10, 21]. In common with those and other studies, the subjects were asked to evaluate the sounds with respect to noisiness where the adjective "noisy" was alternatively described as "unwanted," "objectionable," or "disturbing."

In a single pair comparison, the aircraft sound in question ("comparison"), was compared to a reference sound by asking the subjects to rate one with respect to the other during ten repetitions of the pair. In five of these, the reference (variable level) sound appeared first; in the remainder, the comparison (fixed level) was first. In each of these two sets, the reference was played at five different levels, at increments of 5 dB over a range within which the "equally noisy" level was estimated to lie. The two orders of presentation were used so that a natural subjective bias towards the second sound of a pair could be eliminated by averaging. The pairs were randomly mixed with pairs associated with other comparisons so that the subjects could not recognize any regular presentation pattern.

The two sounds of each pair were separated by a one-second interval, and successive pairs by six seconds. The subjects were asked to rate the relative noisiness of the two sounds on a scale of ± 5 arbitrary units, using positive values if they considered the second sound to be more objectionable than the first, and negative numbers to indicate that they considered the second less objectionable than the first. Each subject recorded his scores on an IBM Portapunch card for subsequent computer analysis.

Although each sound could be compared directly with the standard octave band reference sound, during pilot experiments it was found undesirable to do so because (1) subjects reported difficulty in comparatively judging aircraft flyby sounds with respect to random noises, and (2) the constant repetition of a particular sound causes a rapid increase in the subjects' error rate. Specifically, subjects became confused as to whether they had heard the reference sound as the first or second of a pair. A high percentage of sign errors thus appeared on the score cards. Consequently, use was made of a number of intermediate reference sounds so that most aircraft noises were related to the standard reference by way of a two-stage comparison. Although this causes some accumulation of errors, they are not as serious as those caused by the sequence errors referred to above.

Thus, a system of pair comparisons was devised comprising three levels of reference, as shown in Figure 2. The simulated jet noise (resembling that of a jet engine ground runup), served as a "half-way point" between the Level 3 and 4 references and the standard. The Level 3 aircraft noise reference was selected to exhibit as short a duration as possible to minimize the difficulty of comparison with the four-second reference sounds.

The particular comparison arrangement was selected so that the perceived levels of the Level 2 and Level 3 references, with respect to that of the standard, could be carefully established by replicated measurements and by closing a number of 1-2, 2-3, 3-1 comparison triangles. The latter allowed a convenient check on the consistency of the subjective judgements.
Altogether, 150 sound pair comparisons were arranged. Each comparison involved 10 pairs of sounds so that the total experiment required 1500 sound pairs. These were divided into 60 tests, each test containing 25 pairs of sounds and lasting about 15 minutes, which has been found to be an acceptable duration for a single sitting.

The sounds were played to either 25 or 32 subjects in a progressive wave acoustic chamber in which four or five subjects at a time are seated facing an array of loudspeakers. Two separate systems were used, which included five 60-watt low-frequency units to generate frequencies below 500 Hz and a single multicellular horn for the higher frequency range. A set of 12 foot-long fibreglass wedges was installed behind the subjects to absorb the essentially unidirectional sound waves which are uniform across the facility test section to within ±3 dB at all frequencies. The sound system was capable of generating sound pressure levels in excess of 114 dB in the frequency range 20–5000 Hz and above 118 dB between 25 and 4000 Hz.

In order to maintain a uniform frequency response throughout the overall system, a spectrum shaper was included in the replay circuit. This was used to adjust the frequency response to give a flat spectrum at the center of the test section when pink noise was inserted at the signal source.

The 32 panel members were chosen to provide a reasonable distribution of age, occupation, and sex, the selection being made on the basis of audiometric measurements and an aptitude test. This was designed to test ability to make consistent judgments in a pair comparison test. The final panel selected comprised 20 females and 12 males with a median age of 26 years. The 25 members of the smaller panel were selected from the larger group. Individual tests were performed with four or five persons at a time, and took a total of eight weeks to complete.

The subjects recorded a total of 42,750 individual scores which were processed by a series of computer programs. For each sound pair comparison and each level of the reference sound, each subject’s two scores, $J_1$ and $J_2$, for “forward” and “reverse” orders
of presentation, were summed, it being remembered that the sequence reversal effectively changes the sign of the score. Thus,

\[ J = J_1 - J_2, \]

where

\[ J = \text{mean score in arbitrary units}, \]
\[ J_1 = \text{score (from } +5 \text{ to } -5\) assigned by the subject to that level with the reference presented first, \]
\[ J_2 = \text{score when the reference was presented second.} \]

Most subjects have a tendency to overemphasize the level of the second pair, and this summing procedure compensates for this “order effect.” The bias towards the second sound was, in fact to be, on average, equivalent to about 1 dB.

Curves were fitted by eye through the mean values at each relative level in order to obtain a “zero crossing,” at which relative level it is assumed the two sounds would be judged equally noisy. Curves were also fitted to the 20th and 80th percentile points to give some measure of the scatter of the data about this intercept. The zero-crossing intercepts were used, in conjunction with sound pressure level plots recorded during each of the sessions through a monitoring microphone located at the centre of the chamber test section, to derive an equivalent “equally noisy” level of the standard reference sound for each of the aircraft sounds. Each of the aircraft noise recordings was computer analyzed to yield various objective measures of perceived level. The 120 sounds were recorded in sequence onto master analysis tapes by the use of identical procedures to those used for making the test tapes themselves. These tapes were then played into an analogue-digital 1/3-octave band analysis system covering the 24 centre-frequencies between 50 Hz and 10 kHz. (Frequencies outside that range were in fact filtered from the signals played to the jury subjects.) The band levels were digitized at \( \frac{1}{4}\)-second intervals to yield, for each flyover sound, a matrix of band levels.

Various perceived level measures were computed from the \( \frac{1}{4}\)-octave band level matrices including several variations of the weighted sound pressure level, loudness level and perceived noise level scales. Each computation involved, for each time instant, some form of frequency band summation. Then, for each perceived level scale, and each sound under study, two values were retained for correlation with the subjective data. These are \( PL \), the largest value of the calculated level occurring during the flyover and referred to as simply the peak level, and \( EPL \), a time integrated or “duration corrected” value which is referred to as an “effective” level. This is calculated according to the summation equation

\[ EPL = 10 \log_{10} \frac{1}{T} \sum_{k=K_1}^{K_2} 10^{PL(k)/10} \Delta t, \]  

where \( PL(k) \) is the calculated level at the \( k \)th time instant, \( \Delta t \) is the time increment between samples, \( (0.5 \text{s}) \), and \( K_1 \) and \( K_2 \) correspond to the time intervals when \( PL(k) \) first and last exceeds a level which is 10 dB below the peak level \( PL \) and \( T \) is a reference time of 10 seconds.

The scales used to compute \( PL(k) \) were as follows.

(a) Overall sound pressure level, \( L \).

(b) \( A \), \( B \) and \( D \) weighted sound pressure levels, \( L_A \), \( L_B \) and \( L_D \). Values of the weighting functions used are given in Table 1. The \( A \) and \( B \) functions approximate the weighting
networks of standard sound level meters and the D-function is the inverse of the 40 noy contour (see e.g., reference [15]).

### Table 1

*Table of weighting functions*

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>$L_A$</th>
<th>$L_B$</th>
<th>$L_D$</th>
<th>$L_{NN}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>12</td>
<td>12</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>26</td>
<td>10</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>22</td>
<td>7</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>19</td>
<td>6</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>125</td>
<td>16</td>
<td>5</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>160</td>
<td>13</td>
<td>3</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>200</td>
<td>11</td>
<td>2</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>250</td>
<td>8</td>
<td>2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>315</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>400</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>500</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>630</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>800</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1250</td>
<td>-1</td>
<td>0</td>
<td>-2</td>
<td>-1</td>
</tr>
<tr>
<td>1600</td>
<td>-1</td>
<td>0</td>
<td>-6</td>
<td>-2</td>
</tr>
<tr>
<td>2000</td>
<td>-1</td>
<td>0</td>
<td>-8</td>
<td>-4</td>
</tr>
<tr>
<td>2500</td>
<td>-1</td>
<td>0</td>
<td>-10</td>
<td>-8</td>
</tr>
<tr>
<td>3150</td>
<td>-1</td>
<td>0</td>
<td>-11</td>
<td>-15</td>
</tr>
<tr>
<td>4000</td>
<td>-1</td>
<td>0</td>
<td>-11</td>
<td>-20</td>
</tr>
<tr>
<td>5000</td>
<td>-1</td>
<td>1</td>
<td>-10</td>
<td>-20</td>
</tr>
<tr>
<td>6300</td>
<td>0</td>
<td>2</td>
<td>-9</td>
<td>-15</td>
</tr>
<tr>
<td>8000</td>
<td>1</td>
<td>3</td>
<td>-6</td>
<td>-8</td>
</tr>
<tr>
<td>10,000</td>
<td>3</td>
<td>5</td>
<td>-3</td>
<td>0</td>
</tr>
</tbody>
</table>

(c) $NN$-weighted sound pressure level, $L_{NN}$.
This scale used a weighting function derived earlier [10] from subjective tests performed in the same progressive wave chamber. Although this function, which is also listed in Table 1, attributes an unusually large degree of importance to frequencies in the region of 4 kHz its form was confirmed in a subsequent independent experiment [21].

(d) Stevens’ loudness level, Mk VI, $LL_S$.
This method, described in reference [22], was discussed in the previous section. For a detailed description of the procedure, the reader is referred to Stevens’ paper. It should be noted that this procedure has now been supplemented by a much refined Mk VII version [23].

(e) Zwicker’s Loudness Level, $LL_Z$ (Approximate Method).
Zwicker’s loudness level has been computed by a subroutine developed on the basis of the “graphical evaluation charts” of reference [24]. However, the method adopted is only an approximation of the procedure since the effect of masking has been ignored. This was justified on the grounds that for the majority of aircraft noise spectra, the contributions to loudness of the spread of masking is probably small. This assumption was certainly confirmed by a small number of comparisons of computed and graphically obtained levels.
where the errors were less than 0.5 phon. However, significant discrepancies could occur in
the presence of spectral "spikes" and it must be admitted that this approximate approach
may not be as accurate as the rigorous procedure. Also, the parameters programmed are
only as accurate as readings which could be made from published documents.

(f) Perceived Noise Level, PN L.
The particular version of perceived noise level utilized was originally described by Kryter
and Pearsons [15]. The equations programmed are based on mathematical formulations
of the nay tables developed by Pinker [25].

(g) Tone-corrected Perceived Noise Level, PNLc.
This scale corrects the basic PN L for the presence of pure tones in the spectrum. Of several
alternatives, the method used is that adopted in references [18], [19] and [26], which in
turn was based upon an original recommendation by Little [14]. The integrated version
of this scale, EPN Lc, is that specified by the U.S. Federal Aviation Administration [18]
and the Air Registration Board [19] for aircraft noise certification measurements.

4. RESULTS AND DISCUSSION

4.1. STATISTICAL ANALYSIS

It has been described how the perceived level of each sound was measured in two ways:
(a) subjectively by analyzing the responses of a group of people exposed to the sounds, and
(b) objectively by physical analysis of the acoustic signals themselves. Before proceeding to
compare the two sets of values obtained, it is necessary to establish precisely what we
expect of the various rating scales.

Ideally, the objective level should be identical to the subjective level in every case. In
reality, this ideal is rarely achieved and the two will differ by an amount which varies from
sound to sound and from scale to scale. These differences represent the cumulative effects
of several errors:

(a) the error due to inherent subjective variability; i.e., errors of replication;
(b) the experimental error associated with inaccurate measurements or recording of
the subjective and objective data;
(c) the error due to the inability of the scale to accurately account for all the charac-
teristics of the sound which are subjectively important.

Little can be done about errors (a) and (b) beyond taking all normal precautions to avoid
errors and to ensure that the test panel is fully trained in its task. The third error is the
quantity we are trying to measure and the problem at hand is to distinguish between this
error and the other two which are always present.

The subjective variability (a) can be estimated in a very approximate manner from the
errors associated with the zero-crossing intercepts of the paired comparison results. An
analysis of a large set of zero-crossing data [20] showed that the average range between
the 20th and 80th percentile intercepts was approximately 9.5 dB. If this is assumed to be
approximately equal to two standard deviations, a standard t-test would indicate the 95%
confidence intervals associated with the zero crossing of the mean judgement curve to be
±1.7 dB. That is, in many repetitions of any test, we may be confident that 95% of the
median intercepts would lie within the range ±1.7 dB. This is the replication error asso-
ciated with the group of 32 subjects. For the smaller group, the error will be a little larger.
Note that these figures correspond to the group behavior of the test panel. It may also be
inferred that any individual would perform with this degree of consistency. However, for
individual results, repeatedly picked at random from the group, the equivalent confidence interval would be \( \pm 6.5 \text{ dB} \), a range which is, therefore, more representative of the variation to be expected between individuals with normal hearing picked at random from any population.

A measure of the total experimental error which effectively includes both (a) and (b) has been obtained from the 20 comparisons of the Level 1 and Level 2 reference sounds included in the tests and described in section 3, where the standard deviation of the distribution of results was 1.5 dB. Also, in the ten 1–2, 2–3, 3–1 triangulation checks also described in section 3, the average magnitude of the ten errors in closing the loops was 0.9 dB. Thus, it seems that subjective variability is the greatest source of experimental scatter whose r.m.s. value lies between 1 and 2 dB.

When it is accepted that the probable r.m.s. experimental error is of the order 1–2 dB, it may be assumed that deviations between the calculated and judged perceived level which exceed this range may be attributed to the inadequacies of the perceived level scale. In any event, there is no way in which the two sources of error can be separated, and they can be analyzed only in combination.

The value of a rating scale, of course, rests with its ability to accurately and consistently estimate the perceived noise level of aircraft flyover noise. In the present context, “accuracy” could be used to describe the absolute agreement between the judged and calculated levels, whereas “consistency” might refer to the dispersion of the errors about some central value. Thus, for example, the scale which repeatedly yields calculated levels of 90 dB for a number of sounds which are all subjectively rated at 100 dB, is not at all accurate, but very consistent. If we assign the variables \( x \) and \( y \) to the calculated and judged perceived levels of a sound, as will be done through the remainder of this report, we could determine accuracy and consistency from the distribution of the error \( (x - y) \). Accuracy is related to the mean error,

\[
\overline{z} = \frac{\sum_{i=1}^{N} (x_i - y_i)}{N},
\]

whereas consistency is reflected by the sample standard deviation of the error, which is given by

\[
s = \left( \frac{\sum_{i=1}^{N} (x_i - y_i)^2 - \left[ \sum_{i=1}^{N} (x_i - y_i) \right]^2}{N - 1} \right)^{1/2},
\]

where \( N \) is the number of samples, and \( x_i \) and \( y_i \) are the objective and subjective levels associated with the \( i \)th sample.

The practical distinction between the terms “accuracy” and “consistency” depends upon how the judged levels are measured and in this regard it is essential to recognize the importance of a reference point. In writing down equations (5) and (6), it has been assumed that the levels \( x_i \) and \( y_i \) are, in fact, available in an “absolute” sense. By absolute we mean that both calculated and judged levels, in dB, are related to some form of standardized reference pressure. However, although the subjective levels are expressed in terms of the equivalent level of a particular octave band of noise, not all scales are related to any particular definition of perceived level. The loudness and noisiness scales, \( LL_s, LL_z \), and \( PNL \), are, of course, linked to specific reference sounds, but hazy definitions of these (e.g., the type of random noise, signal level, time histories, listening conditions), together with the practical difficulties of generating those sounds for experimental purposes, tend to
obscure the precise meaning of “absolute” perceived levels. In the case of the weighted sound pressure level sounds, no similar subjective definitions exist.

Thus, in order to make the most meaningful comparisons of absolute accuracy, the corrected mean error, Δ, will be introduced. This is similar to 2, except that the subjective level y is expressed in the same units as the calculated level x. Thus, for example, the error Δ for the PNL scale is the mean difference between PNL for the aircraft sounds and PNL for the reference sound. On the basis of this parameter “accuracy” really pertains to the ability of the scale to rate consistently both aircraft sounds and narrow band noise.

The sample standard deviation, s, similarly expresses the consistency with which any scale might be expected to rate the relative perceived levels of different aircraft sounds. However, it may not provide a fair test of all scales since it is important to consider the possibility that perceived noisiness “grows” at different rates on the judged and calculated scales: i.e., y is proportional to bx, where b is a constant other than unity.

To allow for this possibility an alternative to the use of the statistic s is to fit the best straight line to the plot of y against x and to measure the dispersion of the data about it. Methods for computing the regression coefficients B0 and B1 in the equation

$$y = B_0 + B_1x,$$

which are based on minimizing the mean square error of the points (x,y) about the line, are described in most statistics texts. It is common practice for the above equation to represent the regression of y on x in which the error is minimized in the y-direction. Alternatively, a regression of x on y can be performed to determine the coefficients in the equation

$$x = B_0' + B_1'y,$$  

where the error is minimized in the x-direction. It might be expected that the lines given by equations (7) and (8) are the same so that the slopes B1 and B1' are reciprocal. In the event that all the points (x,y) lie on a straight line, this is indeed the case; otherwise, as Figure 3 clearly shows, B1 and 1/B1' can differ, by an amount which depends on the scatter. The geometric mean of the slopes B1 and B1' is called the product moment correlation coefficient Re, where

$$R_e = \sqrt{B_1B_1'}.$$  

This is a useful parameter which describes the correlation between two sets of variables without actually specifying the constant of proportionality. If all the points fall on a line, R_e = ± 1. Otherwise, |R_e| < 1 and if the x and y are completely uncorrelated, R_e = 0. In practice, a finite value of R_e may be computed, although tests for the significance of its deviation from zero can be applied. Unfortunately, for the reason that the product B1B1' is not, in general, equal to unity, the appropriate constant of proportionality linking x and y is not obvious. Reference to Figure 3, for example, shows that one slope is high and the other is low relative to the line which would probably be fitted by eye. Although either line is a perfectly valid least squares fit through the data, it seems that one lying somewhere between the two should be used as a general relation relating x and y. Thus, for present purposes, the geometric mean of B1 and 1/B1' will be used to indicate the “natural” slope of the data. This is defined as B1, where

$$B_1 = \sqrt{B_1B_1'} = B_1/R_e.$$  

These mean lines are also included in Figure 3 for comparison. Note that all three lines intersect at the centroid of the data.
Because it is difficult to relate the correlation coefficient to a physical measurement of scatter, the parameter $s^2_{yy}$, the standard deviation of the data in the $y$-direction about the regression of $y$ on $x$, will also be discussed in subsequent sections. This is sometimes referred to as the "standard error of estimate". Note that this is the scatter about the line with the slope $B_1$, and not that with the mean slope $B$.  

It is, of course, possible that the relationship between $x$ and $y$ is not linear and that errors could be further minimized by fitting higher order curves to the data. However, this possibility has not been investigated in this study, and all analyses have been based upon the assumption of linearity.

4.2. EXPERIMENTAL RESULTS

Figure 3 shows all 120 values of judged perceived level (JPL) plotted against the corresponding (a) peak overall levels ($L$), and (b) the integrated perceived noise levels ($EPNL$). These objective scales were expected to typify the worst and best scales respectively so that the graphs give some idea of the total range of scatter to be expected. It is certainly obvious from the figure that the $EPNL$ scale is significantly more consistent than the $L$ scale. It may also be noticed in both plots that one point in each, identified by a different symbol, lies apart from the remainder of the data. A detailed examination of this sound and all analyses associated with it, revealed no reason why it should appear so different to the remaining data points, particularly those describing similar aircraft and flight conditions. However, because it seems to suffer from a serious, if unexplained error, it has been omitted from the numerical analyses.

The results of the data analysis are presented in Table 2, which lists the various statistics described above for all 18 scales studied. The analysis has been applied to the total set of all 119 sounds, and also to the four subsets of data corresponding to aircraft in the different propulsion system categories. The five parts of Table 2 thus correspond to

(a) all 119 sounds,
(b) 34 turbojet or turbofan aircraft (jets),
(c) 31 propeller turbine aircraft (turboprops),
(d) 28 piston engined propeller driven aircraft (pistons),
(e) 26 helicopters.
TABLE 2

Correlation analysis

<table>
<thead>
<tr>
<th>Category</th>
<th>Statistic†</th>
<th>L</th>
<th>L_A</th>
<th>L_B</th>
<th>L_D</th>
<th>L_NN</th>
<th>L_L</th>
<th>PNL</th>
<th>PNL</th>
<th>EL</th>
<th>EL_A</th>
<th>EL_B</th>
<th>EL_D</th>
<th>EL_NN</th>
<th>EL_L</th>
<th>EPNL</th>
<th>EPNL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>4.5</td>
<td>-2.1</td>
<td>1.5</td>
<td>1.8</td>
<td>3.5</td>
<td>7.5</td>
<td>9.5</td>
<td>9.8</td>
<td>13.0</td>
<td>0</td>
<td>-6.7</td>
<td>-3.2</td>
<td>-2.3</td>
<td>-1.9</td>
<td>3.0</td>
<td>5.2</td>
</tr>
<tr>
<td>(a)</td>
<td></td>
<td>3.9</td>
<td>4.2</td>
<td>1.5</td>
<td>0.8</td>
<td>8.0</td>
<td>7.5</td>
<td>3.0</td>
<td>3.8</td>
<td>7.0</td>
<td>4.0</td>
<td>-2.7</td>
<td>0.8</td>
<td>0.7</td>
<td>1.6</td>
<td>4.0</td>
<td>2.7</td>
</tr>
<tr>
<td>All 119 sounds</td>
<td></td>
<td>0.822</td>
<td>0.881</td>
<td>0.855</td>
<td>0.881</td>
<td>0.898</td>
<td>0.907</td>
<td>0.912</td>
<td>0.905</td>
<td>0.876</td>
<td>0.821</td>
<td>0.871</td>
<td>0.869</td>
<td>0.929</td>
<td>0.936</td>
<td>0.917</td>
<td>0.913</td>
</tr>
<tr>
<td>(b)</td>
<td></td>
<td>4.5</td>
<td>-2.1</td>
<td>1.5</td>
<td>0.8</td>
<td>3.0</td>
<td>4.5</td>
<td>3.0</td>
<td>3.9</td>
<td>4.5</td>
<td>4.0</td>
<td>3.6</td>
<td>3.5</td>
<td>2.8</td>
<td>3.3</td>
<td>2.8</td>
<td>2.7</td>
</tr>
<tr>
<td>34</td>
<td></td>
<td>2.7</td>
<td>3.5</td>
<td>1.5</td>
<td>1.3</td>
<td>4.2</td>
<td>3.8</td>
<td>3.5</td>
<td>4.4</td>
<td>7.1</td>
<td>3.4</td>
<td>-2.3</td>
<td>0.8</td>
<td>1.3</td>
<td>2.5</td>
<td>4.1</td>
<td>3.1</td>
</tr>
<tr>
<td>Jet</td>
<td></td>
<td>0.891</td>
<td>0.935</td>
<td>0.895</td>
<td>0.952</td>
<td>0.927</td>
<td>0.947</td>
<td>0.951</td>
<td>0.954</td>
<td>0.940</td>
<td>0.820</td>
<td>0.912</td>
<td>0.849</td>
<td>0.941</td>
<td>0.961</td>
<td>0.938</td>
<td>0.921</td>
</tr>
<tr>
<td>sounds</td>
<td></td>
<td>3.2</td>
<td>2.5</td>
<td>3.2</td>
<td>2.2</td>
<td>2.7</td>
<td>2.3</td>
<td>2.2</td>
<td>2.1</td>
<td>2.4</td>
<td>4.1</td>
<td>2.9</td>
<td>3.7</td>
<td>2.4</td>
<td>2.0</td>
<td>2.5</td>
<td>2.8</td>
</tr>
<tr>
<td>(c)</td>
<td></td>
<td>6.4</td>
<td>-0.5</td>
<td>3.4</td>
<td>4.5</td>
<td>4.8</td>
<td>9.2</td>
<td>10.0</td>
<td>11.5</td>
<td>15.0</td>
<td>-0.1</td>
<td>-6.9</td>
<td>-3.2</td>
<td>-2.0</td>
<td>-1.3</td>
<td>3.2</td>
<td>4.9</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td>4.5</td>
<td>-0.5</td>
<td>3.4</td>
<td>3.5</td>
<td>4.3</td>
<td>6.2</td>
<td>3.4</td>
<td>3.5</td>
<td>6.5</td>
<td>10.0</td>
<td>-0.9</td>
<td>0.8</td>
<td>1.0</td>
<td>2.2</td>
<td>4.2</td>
<td>4.2</td>
</tr>
<tr>
<td>Turbo</td>
<td></td>
<td>0.810</td>
<td>0.839</td>
<td>0.827</td>
<td>0.890</td>
<td>0.872</td>
<td>0.910</td>
<td>0.875</td>
<td>0.897</td>
<td>0.864</td>
<td>0.784</td>
<td>0.849</td>
<td>0.838</td>
<td>0.921</td>
<td>0.915</td>
<td>0.925</td>
<td>0.901</td>
</tr>
<tr>
<td>prop</td>
<td></td>
<td>0.670</td>
<td>0.674</td>
<td>0.629</td>
<td>0.666</td>
<td>0.686</td>
<td>0.776</td>
<td>0.662</td>
<td>0.662</td>
<td>0.724</td>
<td>0.714</td>
<td>0.738</td>
<td>0.800</td>
<td>0.729</td>
<td>0.850</td>
<td>0.983</td>
<td>0.817</td>
</tr>
<tr>
<td>sounds</td>
<td></td>
<td>3.8</td>
<td>3.5</td>
<td>3.6</td>
<td>2.9</td>
<td>3.1</td>
<td>2.7</td>
<td>3.1</td>
<td>2.8</td>
<td>3.2</td>
<td>4.0</td>
<td>3.4</td>
<td>3.5</td>
<td>2.5</td>
<td>2.6</td>
<td>2.4</td>
<td>2.8</td>
</tr>
<tr>
<td>(d)</td>
<td></td>
<td>0.827</td>
<td>0.803</td>
<td>0.761</td>
<td>0.748</td>
<td>0.661</td>
<td>0.754</td>
<td>0.887</td>
<td>0.738</td>
<td>0.669</td>
<td>0.923</td>
<td>0.959</td>
<td>0.881</td>
<td>0.869</td>
<td>0.796</td>
<td>0.919</td>
<td>1.091</td>
</tr>
<tr>
<td></td>
<td>(d)</td>
<td>(e)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\bar{z}$</td>
<td>6.0</td>
<td>3.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta$</td>
<td>6.0</td>
<td>3.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$s$</td>
<td>3.0</td>
<td>5.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_e$</td>
<td>0.927</td>
<td>0.653</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$B_i$</td>
<td>0.855</td>
<td>0.410</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$s_{sp}$</td>
<td>2.8</td>
<td>3.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\bar{B}_i$</td>
<td>0.922</td>
<td>0.628</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>$-1.8$</td>
<td>$-3.8$</td>
</tr>
<tr>
<td>$b$</td>
<td>$2.7$</td>
<td>$-0.3$</td>
</tr>
<tr>
<td>$c$</td>
<td>$-0.1$</td>
<td>$0.0$</td>
</tr>
<tr>
<td>$d$</td>
<td>$3.9$</td>
<td>$5.2$</td>
</tr>
<tr>
<td>$e$</td>
<td>$8.5$</td>
<td>$8.2$</td>
</tr>
<tr>
<td>$f$</td>
<td>$10.1$</td>
<td>$7.5$</td>
</tr>
<tr>
<td>$g$</td>
<td>$14.3$</td>
<td>$10.3$</td>
</tr>
<tr>
<td>$h$</td>
<td>$0.4$</td>
<td>$0.5$</td>
</tr>
<tr>
<td>$i$</td>
<td>$-7.7$</td>
<td>$-6.5$</td>
</tr>
<tr>
<td>$j$</td>
<td>$-3.0$</td>
<td>$-3.3$</td>
</tr>
<tr>
<td>$k$</td>
<td>$-2.6$</td>
<td>$-2.9$</td>
</tr>
<tr>
<td>$l$</td>
<td>$-2.4$</td>
<td>$-2.9$</td>
</tr>
<tr>
<td>$m$</td>
<td>$2.7$</td>
<td>$2.6$</td>
</tr>
<tr>
<td>$n$</td>
<td>$4.6$</td>
<td>$5.5$</td>
</tr>
<tr>
<td>$o$</td>
<td>$4.9$</td>
<td>$4.6$</td>
</tr>
<tr>
<td>$p$</td>
<td>$4.8$</td>
<td>$4.6$</td>
</tr>
<tr>
<td>$q$</td>
<td>$4.2$</td>
<td>$4.6$</td>
</tr>
<tr>
<td>$r$</td>
<td>$4.2$</td>
<td>$4.5$</td>
</tr>
<tr>
<td>$s$</td>
<td>$3.8$</td>
<td>$3.8$</td>
</tr>
<tr>
<td>$t$</td>
<td>$3.4$</td>
<td>$3.4$</td>
</tr>
<tr>
<td>$u$</td>
<td>$3.5$</td>
<td>$3.5$</td>
</tr>
<tr>
<td>$v$</td>
<td>$3.3$</td>
<td>$3.5$</td>
</tr>
<tr>
<td>$w$</td>
<td>$3.5$</td>
<td>$3.6$</td>
</tr>
<tr>
<td>$x$</td>
<td>$0.628$</td>
<td>$0.608$</td>
</tr>
<tr>
<td>$y$</td>
<td>$0.625$</td>
<td>$0.624$</td>
</tr>
<tr>
<td>$z$</td>
<td>$0.647$</td>
<td>$0.661$</td>
</tr>
<tr>
<td>$A$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$B$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$C$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$D$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$E$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$F$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$G$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$H$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$I$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$J$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$K$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$L$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$M$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$N$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$O$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$P$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$Q$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$R$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$S$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$T$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$U$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$V$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$W$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$X$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$Y$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$Z$</td>
<td>$7.0$</td>
<td>$7.0$</td>
</tr>
</tbody>
</table>

$\bar{x}_i$ = calculated perceived level of $i^{th}$ sound
$y_i$ = sound pressure level of reference sound judged equally noisy
$\bar{z}$ = mean value of error $x_i - y_i$
$s$ = standard deviation of $x_i - y_i$
$R_e$ = correlation coefficient (between $x_i$ and $y_i$)
$B_i$ = best mean slope = $B_i/R$
$y_i'$ = calculated perceived level of reference sound judged equally noisy
$\Delta$ = mean value of $x_i - y_i$
An inspection of Table 2 reveals that \( x_i - y_j \) and \( s_{xy} \), the standard deviation of the error, which lie in the range 1·8 to 5·3 dB, are generally somewhat larger than the estimated experimental error, which lies in the range 1·2 db. Thus, we can be fairly confident that the differences do, in fact, reflect true differences in the performance of the scales. On the other hand, the total variation of these statistics for any aircraft category—perhaps seems—rather small; being a mere 65% for the analysis of all sounds for example. These values are, in fact, typical of previous experimental results. However, because in the present study the data sample is very much larger, the confidence with which these differences can be evaluated is very much greater.

A standard method for evaluating the difference between two measurements of data scatter is the F-test. This test utilizes the F-distribution, a mathematically derived function which assigns a probability to the likelihood that the difference in two variances occurred by chance. It is based on the assumption that the variances are computed for two samples independently and randomly selected from a normally (Gaussian) distributed population. An inspection of the histograms of the present error distributions gave no reason to suspect that they are not approximately normal. However, it does not seem that the samples can be considered “independent”, since each set of objective perceived levels \( x \)-variables are computed from the same set of 1/3-octave band level arrays. Indeed, the calculated perceived level distributions are highly correlated with each other. Consequently, the validity of the F-test for comparing the variances associated with different scales is somewhat obscure. Nevertheless, there can be no question that the smaller the variance, the more consistent the scale, and there seems little reason why the F-test cannot be used as a framework for comparing scales, provided the results are interpreted in a relative sense.

When attention is confined initially to the results in Table 2 (a) for the complete 119 aircraft set two facts are immediately apparent. The first is that the errors \( s_{xy} \) about the regression lines are substantially less than the standard deviations \( s \), leading to a significant improvement in consistency when a “floating slope” is allowed. Secondly, the integrated duration corrected scales, the “effective” perceived levels, are significantly more consistent than the peak levels.

With regard to the first of these, it seems that the slope of the line is of first order importance to the entire problem of aircraft noise rating. For the complete data set, mean slopes \( B_i \) between 0·665 (for \( L_{NN} \)) and 1·027 (for \( ELL_z \)) have been computed, and it is necessary to examine possible explanations for this in some detail. The slope \( B_i \) compares the average rate of growth of judged level to that of calculated perceived level over the experimental sound pressure level range (from 84 to 115 dB, peak). Here, the judged perceived level is the sound pressure level of the octave band standard reference, which is judged equal in perceived magnitude. Since, for the \( L \) scale, \( B_i = 0·845 \), we see that the actual perceived level is proportional to 0·845 times the peak intensity of the aircraft sounds. In other words, the perceived level of aircraft noise does not grow with overall sound pressure level at the same rate as that of the octave band of noise at 1000 Hz. We saw in section 2 that for wideband spectra, this effect is, in fact, predicted by Zwick’s loudness scale \( LL_z \). However, \( PNL \) and the sound pressure level scales predict equal growth (for relatively non-changing spectra) at all moderate and high levels. It is not surprising, therefore, that \( LL_z \) yields the highest slope \( (B_i = 0·885) \), indicating that this scale predicts a lower growth of perceived level with intensity than do the other methods.

A major influence upon the results of this study may have been the decision to play sounds to the subjects at realistic levels: i.e., as close as possible to the original levels at the time of recording. This maintained realistic relationships between overall signal intensity, spectral content and sound duration, which appear to have an important bearing upon the practical applicability of the noise rating methods.
In the first place, ignoring variations in aircraft size and power, an increase in level implies a reduction in distance between the source and the observer. This, in turn, is accompanied by a reduction in signal duration, and because of reductions in atmospheric sound absorption at higher frequencies, an increase in the high frequency content. The level-duration relationship was, in fact, investigated by computing the correlation between $PNL$ (x-variable) and its "duration correction," which is equal to $EPNL-PNL$ (y-variable). The correlation coefficient $R$, and slope $B$, were, respectively, $-0.474$ and $-0.133$, showing the expected reduction in duration with increase of signal level and a significant correlation. The frequency effect can be inferred by comparing results for the scales $L$ and $L_{NN}$. The latter gives substantially more weight to high frequencies and less to the low frequencies than a linear weighting function, so that the difference between the two levels gives a good indication of the distribution of energy between low and high frequencies. Since the slope $B$ is substantially less for $L_{NN}$ ($0.665$) than for $L$ ($0.845$), it is clear that the difference between $L_{NN}$ and $L$ increases fairly rapidly with level, indicating a shift of emphasis to higher frequencies as intensity increases. It is interesting that scales which give increasing emphasis to high frequencies progress to ever decreasing slopes ($0.845 \rightarrow 0.802 \rightarrow 0.776 \rightarrow 0.665$).

Because of the level-duration relationship, the application of a duration correction causes an increase of slope, which may be seen for all the effective scales listed in Table 2 (a). In particular, $ELL_2$ exhibits a constant of proportionality which is very close to unity. Also, the integration has caused an improvement of consistency in many of the scales. When comparing $s_y$ values, this is particularly noticeable in $EL_{NN}$, $EPNL$, and $EPNL_r$. These procedures are the ones which are most sensitive to high frequencies and this observation thus suggests that the duration correction (negative) is tending to compensate for what is perhaps an excessive emphasis upon high frequency content (positive).

Note that just as $s$ and $s_y$ tend to equalize as $B$ approaches unity, $s$ shows a marked improvement from the "peak" to the "effective" scales.

A standard F-test comparison of the $s$ values for the various procedures in Table 2 (a) suggests that if importance is attached to unit slope the three effective perceived level scales $ELL_2$, $ELL_2$, and $EPNL$, together with $EL_1$ and $LL_2$, are significantly better than the remainder. Reasons for this can be traced to either high slope ($LL_2$, $ELL_2$), low scatter ($EPNL$, $EL_1$) or both ($Ell_2$). If, on the other hand, slope is ignored, we see by comparing the standard errors of estimate $s_y$ that the same procedures rank highly, but that all have been overtaken by $EL_{NN}$ with a very low error of 2.3 dB. It seems that emphasis upon high frequencies markedly improves consistency, but that because of the particular relationship between intensity and frequency distribution, this step has caused an excessive increase in the calculated perceived level growth rate ($B = 0.778$).

Comparing further standard deviations $s$ in Table 2 (a), we see that the next group of scales includes the remainder of the duration corrected versions (except $EL$). Significantly lower again are the peak level scales $LL_{NN}$, $PNL$, $L_{D}$, $L_1$ and finally, at the bottom of the list, $L$, $L_{NN}$, and $PNL_r$. For exactly opposite reasons for which the best scales are superior, the poorest ones have low slopes, high scatter, or a combination of both. It will be noticed, for example, that in terms of $s_y$, the peak scales have exchanged places with the effective scales, so that the uncorrected versions of the more elaborate scales remain superior to the corrected versions of the poor sound pressure level scales.

The question of accuracy, which has been related to the ability of the rating procedures to accurately scale both narrow and broadband noise, may be examined by comparing the mean errors $\overline{z}$ and $\Delta$ for the 119 sound set. It will be remembered that $\overline{z}$ is simply the mean difference between the calculated levels and the average sound pressure levels of the equivalent standard reference sounds, whereas the increment $\Delta$ is based upon the calculated per-
ceived level of the reference. The increment \( z \) thus gives the direct differences between mean levels calculated on the different scales. We see, for example, an increment of 8 dB between \( L_D \) and \( PNL \). Also note that \( LL_z \) generates levels which are 2 phons higher than \( LL_s \), but that \( LL_z \) and \( PNL \) are very close.

These differences reveal the different ways in which the scales account for the increase in perceived level with bandwidth, remembering that for a narrow band signal centered at 1000 Hz, the levels would all agree to within 1 dB.

When attention is confined to the more meaningful increment \( \Delta \), it seems that for the average aircraft flyover sound, the complex loudness/noisiness procedures, with or without duration allowances, overestimate perceived level (with respect to that of a 1000 Hz reference) by around 4 dB. The tone correction increases this discrepancy by a further 3 dB. The sound pressure level errors, on the other hand, range between 4·5 dB and 2·7 dB, in each case reflecting the net attenuation introduced by the weighting function. Thus, the linear scales \( L \) and \( EL \) overestimate by 4·5 and 4·0 dB, respectively, whereas \( LA \) and \( ELA \) underestimate by 2·1 and 2·7 dB.

4.3. DIFFERENCES BETWEEN AIRCRAFT CATEGORIES

The scatter diagrams relating the judged and calculated levels for each of the four aircraft categories are presented in Figure 4. Apparently confirming the significance of subjective differences in the acoustic characteristics of aircraft with different propulsion systems, the figure and the results in Table 2 reveal that clear differences do indeed exist between the consistency of the scales as applied to the different data sets. Unfortunately, because of the smaller samples, distinctions between the scales are less clear but it is obvious that, on average, the scales are most consistent for the piston sounds followed by the jets, the turbos and the helicopters, in that order.

The results for the helicopters are remarkable in that (a) all scales are poor, and (b) in terms of standard error of estimates \( s_{xy} \), there is practically no difference between any of the scales. Although \( L \) and \( EL \) appear inferior, the differences are not significant. However, reference to the standard deviation \( s \) does help to discriminate between the scales because there is a large variation of \( B_1 \) (Table 2(c)). On the scale of \( s \), the methods can be divided into two basic categories, moderate and poor, with the effective perceived level scales being superior to a group containing all the peak scales plus \( EL, EL_A, \) and \( EL_B \). The duration correction is particularly beneficial, probably because of the long durations associated with some of the very low speed flyovers. The reason for the consistently poor performance of the scales is probably related to the domination of the helicopter sounds by low frequency energy of a pulsatile nature. We have seen that attention to high frequencies is one of the major factors which, in general, discriminates between the better and poorer scales. The fact that the helicopter sounds contain little high frequency energy, therefore, serves to explain the small range of \( s_{xy} \) values. The fact that all scales are poor suggests that the subjective effects of low frequency pulsatile sounds require further investigation.

At the other end of the range, most of the better scales perform remarkably consistently for the sounds of piston engined aircraft, with \( s_{xy} \) errors of only 1·8 dB which are probably as low as possible. However, the slopes \( B_2 \) are consistently low, and further, they are not significantly increased by the application of a duration correction. This is possibly because sound pressure level is more strongly related to aircraft size than to distance for the piston group. Because the slopes are low, the deviations \( s \) are substantially bigger than the deviations about regression. Of particular interest is that although the duration correction has a very marked effect upon \( L_D \) and \( LNN \) it does not improve \( LL_s \) or \( LL_z \) in terms of \( s_{xy} \). Again, the only explanation which can be offered is that the (negative) duration correction is
counteracting some harmful effect of emphasizing the high frequencies (which \( LL_z \) and \( LL_{z} \) do to a much lesser extent than either \( L_B \) \( L_{NS} \) or, for that matter, \( PNL \)).

The most notable feature of the results for the 34 jet sounds, Table 2 (b), is that the average slope of the regression lines is rather greater than it is for the other data sets. In fact, for the effective scales, the average slope is very near to unity and \( LL_z \) here results in a \( B_1 \) of 1.218. This says that perceived level grows more rapidly with intensity for jets than it does for other aircraft. Like the sounds of other aircraft, and perhaps more so since the total power range is somewhat smaller, signal level is closely linked with aircraft proximity, and, therefore, with signal duration and frequency distribution. Also, the higher frequency energy appears in the form of compressor, fan or turbine tones. Thus, it must be conjectured that the presence of these components at the higher sound pressure levels is responsible for the relatively high growth of judged level. It is certainly not without significance that for the jets, \( EPNL_{t} \), with the tone corrections, has a particularly small value of \( s \). In general, the scales \( PNL \) and \( L_B \) fare particularly well, scales which were previously noted to require a duration correction to compensate for possible overemphasis upon high frequencies.

In terms of absolute accuracy, an inspection of Table 2 shows that the mean error \( \Delta \) for any scale does not vary significantly between aircraft categories.
4.4. BAND LEVEL SUMMATION PROCEDURES

In order to make a general comparison of the three basic band level summation procedures incorporated into (a) the weighted sound pressure level scales, (b) Stevens' loudness summation rule and (c) Zwicker's graphical integration method, it is useful to introduce the concept of "uniformly distributed noise". This is a hypothetical broadband noise with a spectrum such that each individual band, if present alone, would independently yield the same perceived level on the scale in question. Thus each scale ($L$, $LL_s$, $LL_z$, $PNL$, etc.) has its own particular noise; the spectrum shapes involved are of no direct interest for present purposes.

The first point of interest is the manner in which the perceived magnitudes of $J$ bands of noise add together to give an overall perceived level. For the weighted sound pressure levels, the sum is simply given by

$$L_w = PL_{1/3} + 10 \log_{10} J,$$  \hspace{1cm} (11)

where $PL_{1/3}$ is the perceived level of each band. Thus, the total level increases by 3 dB each time the number of contributing bands is doubled. This, of course, is simply the energy summation principle.

Equating $S_m$ and $S$ for our uniformly distributed noise in Stevens' summation rule (2), we see that

$$LL_s = 33.3 \log_{10} S [1 + F(J - 1)] + 40. \hspace{1cm} (12)$$

The quantity $33.3 \log_{10} S$ is equal to the band perceived level $PL_{1/3}$ so that we can write, for 1/3-octave bands of noise ($F = 0.15$),

$$LL_s = PL_{1/3} + 33.3 \log_{10} (0.85 + 0.15 J). \hspace{1cm} (13)$$

When $J$ is very large, the total level increases by 10 dB each time the number of admitted bands is doubled. However, the increment is less for a more realistic number of bands as shown in Figure 5, where the curve given by equation (13) is seen to cross the 3 dB per doubling line of equation (11) at between $J = 5$ and $J = 6$.

A similar analysis may be applied to Zwicker's rule. An analysis of the "specific loudness" charts similar to the one shown in Figure 1 reveals that for any given sound pressure level, the proportion of the loudness of any single 1/3-octave band of noise confined to the
sideband masking envelope is roughly equal for all bands. Thus, the proportion actually confined between the 1/3-octave band frequency limits \( f_{1,j} \) and \( f_{2,j} \) can be expressed for one band as follows:

\[
\frac{\int_{f_{1,j}}^{f_{2,j}} \frac{dS}{df} \, df}{\int_{0}^{\infty} \frac{dS}{df} \, df} = F,
\]

where the subscript \( j \) denotes the \( j \)th band. It can be seen upon inspection of Figure 1 that the sideband associated with each additional and adjacent band is masked when it is added to the first, so that the total loudness for \( j \) bands is approximately

\[
S_j = S[1 + F(j - 1)].
\]

This relationship is, of course, identical to the equation for Stevens' summation principle defined above.

Approximate average values for \( F \) for a 1/3-octave band of noise centered at 1000 Hz have been estimated from Zwicker's charts [24] as a function of level and are shown in Figure 6 to steadily decrease from around 0.6 at low sound pressure levels to a little more than 0.2 at 110 dB. Thus, at all levels the factor is greater than the value of 0.15 originally suggested by Stevens [8] for 1/3-octave band summations. However, attention is drawn to Stevens' revised and variable \( F \) function used in the Mk VII procedure [23], which is included for comparison in Figure 6. In view of the totally different derivations of the two curves, they are remarkably similar at levels above 50 dB. They do, in fact, coincide at 110 dB. The reason for the disparity below 50 dB is related to the fact that Stevens' curve takes account of departures from a power law at low levels.

The perceived level summation increment corresponding to \( F = 0.3 \) is included for comparison in Figure 5. This corresponds to the Zwicker case at band levels around 80 dB. However, variation of \( F \) between 0.4 and 0.2 at band levels above 50 dB can cause the curve to vary over a total range of 10 dB about its illustrated position. Similar variations occur in the corresponding curve for Stevens' Mk VII summation rule due to the variable \( F \)-factor. However, the difference between the curve corresponding to the lowest value \( F = 0.19 \) which occurs around 80 dB (see Figure 6) and the Mk VI curve for \( F = 0.15 \) is very small due to different sone-phon conversion factors of 33.3 (Mk V) and 30 (Mk VII).

Thus, we see that at band levels around 80 dB, the three basic procedures give different weightings to the perceived level increment caused by adding further bands of noise. For
more than 6 equal magnitude bands, the energy summation principle (sound pressure level scales) gives the smallest increment, followed by the Stevens' method and Zwicker's method. However, the differences do vary with level, and, of course, with spectrum shape. Further, the absolute differences between the Zwicker and Stevens' curves are in practice reduced by other procedural differences in the level computations.

Turning now to the experimental results, because both the band level summation technique and the frequency weighting function contribute to the performance of any particular scale it is difficult to isolate the effects of either one. The fact that the five sound pressure level scales differ only in the form of their frequency level scales differ only in the form of their frequency weighting functions allows some conclusions to be drawn regarding the independent effects of the latter and these are discussed in the next section. A significant result regarding band summation may be found in Table 2 (a) which shows that the two scales $EL_D$ and $EPNL$ are practically identical in every respect except in the mean errors $\bar{e}$ and $\Delta$. Both scales overestimate the judged level of the aircraft sounds but the error for $EL_D(\Delta = 0.7 \text{ dB})$ is rather less than that for $EPNL(\Delta = 3.9 \text{ dB})$. Since these scales utilize practically identical frequency weighting functions, this finding provides an important comparison between the different band level summation procedures incorporated in the two methods.

Reference to Figure 5 suggests that $EPNL$ may be expected to exceed $EL_D$ when the effective number of 1/3-octave bands in the signal exceeds 6 or so. To shed some further light on this, a further analysis was made of the Perceived Noise Level computations which involve the Stevens' summation rule

$$N_t = N_{\text{max}} + F(\Sigma N_j - N_{\text{max}}),$$

where $N_j$ is the noisiness of the $j$th 1/3-octave band in noys, $N_{\text{max}}$ is the noy value for the noisiest band, $F$ is a constant (0.15) and $N_t$ is the effective noisiness of the total complex signal. The summation $\Sigma$ is performed over all bands. An analysis of all 119 sounds showed that, on average,

$$N_{\text{max}} = 0.4 \cdot N_t,$$

where

$$\Sigma N = 4.5 \cdot N_t = 11 \cdot N_{\text{max}}.$$  \hspace{1cm} (17)

Thus if we revert to our concept of “uniformly distributed noise” in which all the $N_j$ are equal (and equal to $N_{\text{max}}$), such a signal would contain eleven effective 1/3-octave bands of noise. Note that Figure 5 shows a difference of 3 dB between the $PNL$ and sound pressure level curves for 11 effective bands.

Both the 4 dB discrepancy by which $EPNL$ overestimates the average perceived level of aircraft noise and the sub-unity slope of 0.875 can be corrected by introducing a variable factor $F'$ into equation (16) as recently proposed by Stevens' [23], such that

$$N_t' = N_{\text{max}} + F'(\Sigma N_j - N_{\text{max}}).$$

Since this formula must reduce the perceived level by 4 PNdB the conversion formula (2) gives the result

$$33.3 \log_{10}(N_t/N_t') = 4$$

or

$$N_t/N_t' = 1.32,$$  \hspace{1cm} (19)

i.e.,

$$\frac{N_{\text{max}} + 0.15(\Sigma N - N_m)}{N_{\text{max}} + F'(\Sigma N - N_m)} = 1.32.$$  \hspace{1cm} (20)
For a summation over 11 effective bands this yields

$$F' = 0.09.$$  

This value applies at the average level (which occurs at $JPL = 96$ dB) where $PNL \sim 100 \text{ PNdB}$. At this level the effective perceived noisiness $N_e$ is 64 noys so that, from equation (17), $\Sigma N \sim 290$ noys for the average aircraft sound. Accordingly, we may write the relationship for a unit $B_1$ slope:

$$\log_{10}N(1 + F'(J - 1)) = 0.875 \log_{10}N(1 + 0.15(J - 1)),$$

where $N$ is the effective (uniform) band noisiness and $J = 11$. Whence we obtain

$$F' = 0.4(\Sigma N)^{-1/8} - 0.1.$$  

This function, which decreases with signal intensity, is compared with Stevens' Mk VII recommendation and the function derived from Zwicker's charts in Figure 6. It is seen to be significantly smaller than both. This rather low value for $F'$, which only reaches the value 0.15 at approximately 70 PNdB, must be assumed to be characteristic of aircraft noise spectra and of course, is only known to be applicable for Effective Perceived Noise Levels in excess of 85 EPNdB. The growth of computed perceived level with bandwidth for $F = 0.09$ has been included for comparison with larger values in Figure 5. As expected, this curve is 1 dB lower than the sound pressure level curve at $J = 11$ and is 6 dB lower than the result for $F = 0.15$ for high $J$. For practical values of $J$, however, it is unlikely that the $F = 0.09$ curve will differ from the sound pressure level curve by more than one or two decibels. This result would seem to have considerable practical significance since it strongly suggests that, at least for aircraft noise in the mid-level range, the complex noisiness calculation procedure, even in its modified form, may be accurately approximated by the D-weighted sound pressure level scale.

It may be noted that, although Figure 5 suggests that Zwicker's Loudness Scale, $LL_z$, will estimate levels around 6 dB greater than $LL_5$ or $PNL$, this increment is not evidenced in Table 2. Indeed, $LL_z$ is lower by approximately 1 dB. This may be attributed to two factors. The first is that the curve in Figure 5 corresponds to a signal level of approximately 80 phons where the masking effect is reduced so that individual band contributions are greater. (At higher levels the bandwidth effect is less marked.) The second is that the neglect of the masking profiles in the computations may cause larger errors than were originally anticipated.

4.5. FREQUENCY WEIGHTING FUNCTIONS

The perceived level calculation procedures evaluated in this study utilize a very wide variety of actual or effective frequency weighting functions which were originally experimental measurements of equal perceived magnitude functions with different experimental conditions, methods and environments.

Because of different band summation procedures utilized, it is difficult to compare the sound pressure level weightings with the functions included in the perceived level procedures $PNL$, $LL_5$ and $LL_z$. Insofar as these complex scales are concerned, if it is assumed that the main summation differences lie in the growth functions discussed in the previous sections, their different frequency functions can be compared through the statistic $s_{xy}$. Based on overall performance as indicated by Table 2 (a), $EPNL$ appears a little more consistent than $ELL_5$, which in turn is a little better than $ELL_z$, although none of the differences appear significant.
In the case of the five sound pressure level scales, however, the relative merits of the weighting functions (Linear, A, B, D, and N) can be compared directly, since the same band summation method is common to each scale. This comparison, however, must take into account the three related effects upon the constant of proportionality $B_1$, the mean error $\Delta$, and the scatter, as reflected by $s_{xy}$. Trends in all three quantities may be associated with the degree of emphasis upon high frequencies. Specifically, as this emphasis is increased (from $EL_D$ to $EL_A$ to $EL_D$ to $EL_N$), the slope $B$ decreases from 0·972 to 0·778, the deviation $s_{xy}$ decreases from 3·8 to 2·3 and the mean error $\Delta$ tends to reduce although this is primarily related to the net attenuation by the weighting networks.

The high sensitivity of the slopes $B_1$ to the frequency parameter is a function of the fact that for the particular data studied, level and frequency content are correlated via the distance variable. Thus, high frequency emphasis causes greater calculated level increases at higher signal intensities than it does at lower levels, effectively introducing a change of slope. Although this change leads to higher correlation, this benefit is offset by the reduced slope. An optimum weighting function can only be defined therefore in terms of a tradeoff between scatter, slope and mean error. Certainly, of those studied, the D-network appears to most closely approach the ideal.

When used with a duration correction, $EL_D$ yields a mean error $\Delta = 0·7$ dB, a standard error of estimated $s_{xy} = 2·4$ dB and a mean slope $B_1 = 0·879$ based on the complete set of data. Although improvements are undoubtedly possible by careful attention to detailed network design, this performance is certainly good and in fact slightly better than that of $EPNL$.

4.6. DURATION EFFECTS

The results clearly show that the integrated duration correction has a beneficial effect upon the performance of the scales, both in terms of consistency and slope. An approximate correction, based on the actual time between the 10 dB-down points, was included in the study [20], but proved significantly inferior to the integrated version. It was therefore omitted from the present paper.

In order to examine the duration effect a little more closely, the correlation was computed between the subjective levels calculated according to the relationship

$$EPNL' = PNL + K(EPNL - PNL),$$

where the constant $K$ was varied over an appropriate range. This is equivalent to the equation

$$EPNL' = PNL + 10K \log_{10} T_e,$$

where $T_e$ is the effective duration computed by integration. The analysis was applied to the five different sets of data and the results are shown in Figure 7. This illustrates, for each group, the variation of $s$ and $s_{xy}$ with $K$.

In all cases, the minimum in the $s$-curve occurs near $K = 1$, suggesting that the constant of 10 is indeed optimum in all cases. However, it should be remembered that because of the low $B_1$ slopes encountered, the scatter $s$ is reduced by the increase of slope which the duration correction brings about. The curves would be somewhat different if the growth problem were remedied as discussed previously, and would, in fact, bear more resemblance to the $s_{xy}$ curves. It may be seen that the troughs in these curves, in addition to being shallower, tend to occur at fractional values of $K$. This is particularly noticeable in the case of the jet data. Nevertheless, for the complete set of 119 sounds the minimum occurs at $K = 0·9$ and because $s_{xy}$ is not particularly sensitive to $K$ anyway, the presently used 3 dB per time doubling seems to be a good choice, at least for the $PNL$ scale.
4.7. TONE CORRECTION

The same cannot be said of the tone correction which has not shown itself to be a particularly beneficial measure since, in general, its application has caused both \( PNL \) and \( EPNL \) to become less consistent evaluators of perceived level. The probable reasons for this, however, may be identified by inspection of the results for the different aircraft categories.

It does seem significant, for example, that the one case in which the tone correction proves advantageous is the application of \( EPNL \) to the jet sounds. High frequency tones could be observed in 26 of the 34 sounds used, and were "strong" in about 10 of these cases. However, tone corrections of between 1 and 5 dB were applied to all sounds, without exception, and the average increment was 2.3 dB.

It was initially somewhat surprising to find that even larger corrections with mean values of 3.2 and 2.5 dB were applied to the turboprop and the piston engine data respectively, although an inspection of individual 1/3-octave spectra quickly revealed why. Very large spikes occur at the fundamental propeller frequency, normally in the region of 100 Hz, and sometimes at its higher harmonics. Thus, even though no high frequency tone was present in the case of the piston sounds, corrections are automatically applied by the \( EPNL \) procedure. These spikes do, of course, correspond to "tones" in the spectrum. However, the quality of propeller sounds is controlled more by the higher harmonics than the fundamental. In fact, it is a well known fact that the "impulsiveness" of propeller noise, and harmonic sound in general, increases as the spectrum becomes more flat. It is thus conceivable that the tone correction, as presently constituted, works "in reverse" for the sounds of propeller aircraft, adding larger tone corrections as the harshness of the sound decreases. It should be noted that, although the maximum correction for tones
below 500 Hz is +3.3 dB, corrections as high as 5 dB have been applied in some piston aircraft cases. These must be attributed to random level fluctuations at higher frequencies. However, obvious differences between the piston and turboprop groups do exist, and, for the present, these can only be attributed to the presence of compressor and fan components in the case of the turboprops and perhaps exhaust components for the pistons. For some reason all scales, with and without tone corrections, are less consistent for the turboprops and it is possible that the tone correction might have proved advantageous for this group, as in the case of the jets, if it could have operated upon the high frequency tones only.

In view of the uncertainties regarding the subjective aspects of propeller and rotor noise, it would seem advisable for the moment to restrict the tone correction to higher frequencies, say above 500 Hz, and to ignore "tones" identified at lower frequencies.

5. CONCLUSIONS AND RECOMMENDATIONS

It has been found that significant differences do exist between scales and that they can be ranked into several strata. In terms of consistency, the better methods are essentially indistinguishable and include the three "complex" perceived level procedures, ELLs, ELLz and EPNL due to Stevens, Zwicker, and Kryter, and where the prefix E denotes the application of an integrated signal duration allowance. Also statistically indistinguishable from these for the aircraft sounds, were integrated or "effective" sound pressure levels measured on the D-scale, ELp, and peak loudness level, LLz, from Zwicker's methods. However, with the exception of Zwicker's, all methods tend to overestimate the growth of perceived level with intensity over the range of sound pressure levels investigated (84-115 dB overall).

Distinct differences were observed between the applicability of the scales to sounds in the four different aircraft categories. On average, the scales were extremely consistent for the piston engined aircraft sounds but increasingly less so for the jets, the turboprops and the helicopters, in that order. The deficiencies in the latter groups are attributed to improper account of pure tones in the turboprop spectra and low frequency harmonic sound in the case of the helicopters.

Of three alternative methods for summing contributions from different frequency bands to the total perceived level, Zwicker's adheres most closely to accepted auditory theory. It best explains the experimental observations relating to the growth of perceived levels, and possibly takes automatic account of spectral spikes. Stevens' technique, which is common to both LLs and PNL, is based upon a simpler empirical model but for practical, wideband sounds it turns out to be remarkably similar to Zwicker's calculation at lower sound pressure levels. However, both techniques overestimate the perceived level of aircraft noise with respect to that of the 1000 Hz reference sound by an average of 4 dB. At higher levels the Zwicker and Stevens' procedures differ in that Stevens' approach overestimates the growth of perceived level of wideband sounds with respect to that of narrowband sound. Based upon the experimental findings and an investigation of the relationships between the three basic summation procedures, a simple remedy for both problems has been defined for use in the EPNL calculation. This involves a smaller, but variable E-factor in the loudness/noisiness summation formula. This term, which decreases as intensity increases, simply accounts for a known increase of inter-frequency masking at higher levels.

The simple energy summation process performed by the weighted sound pressure level circuits is rather sensitive to the particular choice of weighting network and, depending on this selection, can over- or underestimate the perceived level of wideband noise relative to that of narrowband noise. Thus, a linear (flat) function overestimates, the A-weighting
SCALING AIRCRAFT NOISE PERCEPTION

underestimates, whereas the $D$-weighting, based on the inverse of the 40 nøy contour, shows a very small mean error. Otherwise the energy summation rule gives a very good approximation to the revised noisiness summation rule over a practical bandwidth range.

The procedures $LL_{s}$, $LL_{z}$ and $PNL$ directly or indirectly incorporate similar frequency weighting functions and largely for this reason tend to be equally consistent. An investigation of a set of widely differing sound pressure level weighting functions revealed an improvement in performance as emphasis shifted from low frequencies to high. However, on the basis of consistency, perceived level growth and accuracy, the $D$-weighting is the best of those studied and is probably close to optimum. For all practical purposes $L_{D}$ is at least as accurate as $PNL$ for rating aircraft noise.

Based on the assumption of a uniform duration/perceived level tradeoff allowance, the presently used correction of 3 dB per duration doubling is close to optimum for aircraft sounds in all categories. The application of this duration allowance improves the performance of the scales.

The tone correction used in the internationally accepted $EPNL$ procedure has been tested for each aircraft category. Only in the case of the jet sounds does the correction appear to perform as intended, and then the improvement is marginal, a slight improvement in consistency being offset by a further increase in the mean error. It is concluded that in the case of the piston sounds the requirement for a correction possibly exists, but that this need is not fulfilled by the selected procedure. The problem appears to lie not entirely with the form or magnitude of the correction, but in the manner by which "tones" in the spectra are detected by the computer model. It is probable that an immediate interim improvement could be made by eliminating correction based on "tones" identified at frequencies below 500 Hz.

ACKNOWLEDGMENTS

The author would like to thank B. D. Adcock and D. M. Lister for their contributions to the experimental program. He is also indebted to K. M. Eldred for many helpful discussions.

The work was supported by the Federal Aviation Administration Office of Noise Abatement, Washington, D.C. and the National Aeronautics and Space Administration, Langley Research Center, Hampton, Virginia.

REFERENCES


This describes part of an extensive study of helicopter noise perception which was initiated to help resolve the question of whether helicopter noise certification should be based upon different noise measurement scales to those adopted for fixed-wing aircraft. Previous research had raised doubts that these scales made proper allowance for the impulsiveness of helicopter noise and opinion was divided, more or less evenly, for and against the use of an impulsiveness correction.

This experimental investigation involved the largest available sample of different helicopter sounds and led to the conclusion that the previous difficulties were due, in part at least, to reliance upon an insufficient range of data. Helicopters exhibit a much wider range of noise characteristics than conventional aircraft and, probably for this reason, are less easily rated by conventional scaling techniques. However although the standard Effective Perceived Noise Level is generally less accurate for helicopters it was concluded that no special impulsiveness corrections are necessary. The study therefore lent support to the decision to adopt the standard EPNL scale for helicopter noise certification.

The basic experiment was performed four times, each with different groups of test subjects. Two involved headphone presentation of the test stimuli and two involved loudspeaker presentation. Only the main headphone experiment is described here. The repetitions confirmed the main findings of the study although there were certain differences which are summarised in the conclusions.
1.0 INTRODUCTION

Aircraft noise certification standards have been specified by the FAA and the International Civil Aviation Organization (ICAO) for subsonic jet aircraft and for both large and small propeller-driven aircraft. For the first two categories, noise limits are defined as Effective Perceived Noise Levels in EPNdB; for the latter, they are defined as Maximum A-Weighted Sound Levels, $L_A$, in dB(A).

In its deliberations to develop noise certification standards for V/STOL (vertical or short takeoff) aircraft including helicopters, Working Group B (WGB) of the ICAO Committee on Aircraft Noise (CAN) was concerned about evidence that these noise scales may be less satisfactory for rating helicopter noise than that of conventional takeoff and landing aircraft (CTOL). Much of this evidence pointed to the possibility that in the case of helicopters, the existing noise scales might not properly account for the periodic impulsiveness which characterizes the sound of rotors. It is certainly widely acknowledged that severe forms of impulsiveness, often known as "blade slap," can be particularly intrusive and annoying, and it is clearly necessary that any noise scale used for certification should properly reflect the potential of such noise components to evoke annoyance.

The history of research into suitable helicopter noise rating methods is documented elsewhere (e.g., References 4 through 7). It suffices to state here that the evidence is contradictory; some studies have suggested that standard procedures such as EPNL and $L_A$ are adequate for V/STOL and helicopter noise while others indicate that they underestimate its noisiness.

Of particular significance, WGB asked the International Organization for Standardization (ISO) to study the problem of helicopter noise and recommend a suitable noise scale. This work, some of which is described in Reference 8, was performed by Working Group 2 of ISO Technical Committee 43, Subcommittee I, and culminated in the preparation of a draft ISO standard for helicopter noise measurement. The main feature of this proposal was the adoption of a version of EPNL modified by a correction for impulsiveness (following the philosophy of the

*In this report, the standard version of Effective Perceived Noise Level which incorporates tone corrections is abbreviated EPNL to distinguish it from an alternative version EPNL which does not.

"tone correction," another EPNL⁺ modifier. The ISO impulsiveness descriptor is sensitive to large periodically occurring peaks in flyover sound pressure time history and augments EPNL⁺ by up to 6 dB.

This descriptor was subsequently tested in a field experiment⁹ at NASA's Wallops Flight Center in which two different helicopters and a propeller-driven CTOL aircraft were flown over a group of test subjects who compared their relative noisiness. When compared on the basis of EPNL⁺ (without the ISO impulse correction), the two helicopters, a Bell 204B and a Bell 0H58A, were judged equally noisy despite the fact that the 204B has a considerably more impulsive noise signature. This finding was broadly confirmed in laboratory experiments involving sound recordings made during the field trials.¹⁰ In the light of this evidence, WGB concluded that the need for an impulse correction remained unproven and both ICAO and FAA consequently framed proposed helicopter noise certification procedures around the conventional EPNL⁺ scale.¹¹,¹² The committee did, however, recognize a need for further research into the matter.

The present study was initiated during the period of deliberation in a further attempt to check the adequacy of EPNL⁺ for the practical purposes of controlling helicopter noise. The main objective was to test and compare the abilities of a number of conventional noise rating scales to predict the relative annoyance levels of a wide range of recorded helicopter sounds and to identify components and characteristics of helicopter noise which contribute to annoyance but which may not be fully accounted for in the EPNL⁺ model. Of special interest were (a) the relationships between helicopters and CTOL noise, (b) impulsiveness, and (c) the very long durations sometimes associated with helicopter flyover noise, particularly during the approach phase.

It is, of course, highly probable that many factors contribute to helicopter noise annoyance including both the acoustic qualities of the sound and nonacoustic information which the sound conveys. The precise role of each factor could only be established through extensive experiments in which each factor is varied independently of the others, either one at a time or simultaneously. The main requirements would be the correct identification and inclusion of all relevant independent variables and, as the name implies, independence of these variables.

Theoretically, single factors such as impulsiveness can be studied through relatively small scale experiments in which this factor is the only physical variable. In practice, it is often difficult, if not impossible, to vary a single factor
Independently of all others. For example, a change of impulsivity normally causes a change in the frequency spectrum. In the case of helicopter noise, impulsivity may also be associated with increased duration, as will be seen. This "confounding" of factors is difficult to unravel and the isolation of a satisfactory noise rating scale may only be possible through a trial-and-error process in which the model is evaluated and refined by testing it against new experimental data as they become available.

The basic approach to this study was to gather together a large collection of helicopter noise recordings from which a test sample could be selected to cover wide but realistic variations of at least the major variables of interest (duration, tonality, and impulsiveness). Each sound would be rated with respect to its annoyance-evoking qualities by a group of test subjects and measured on various standard scales of noise measurement including A-weighted sound level \( L_A \) and Effective Perceived Noise Level (EPNL\(_T\)). The performance of these scales as annoyance predictors could then be assessed by comparing the measured sound levels and the subjective "annoyance levels." If a sufficiently large and varied sample of sounds were available, then it would also be theoretically possible to isolate directly the independent contributions of these variables to judged annoyance by appropriate multivariate statistical methods.

Certain difficulties associated with this kind of experimentation were recognized at the outset. Foremost among them is that reliance upon available recordings of real aircraft flyover sounds imposes severe constraints upon the variations of, and relationships between, variables of importance. It might be possible to achieve a reasonable degree of decorrelation between a few primary variables but many subsidiary variables including variations of the signal with time, Doppler frequency shifts, rotor blade passing frequencies, and many others which may affect a listener's assessment of a particular event, inevitably lie beyond the control of the experimenter. As noted previously, elaborate annoyance prediction models to account for many such factors could only be synthesized on the basis of results from highly controlled experiments in which those factors are varied systematically.

Indeed, it was on systematic experiments of this kind that the foundations of EPNL\(_T\) were laid and from which emerged duration and tone corrections and more recently the ISO impulsivity correction. However, it is by no means clear that this process is entirely satisfactory when conducted in isolation. A fairly extensive test
of EPNL made by the author revealed certain deficiencies which, although of little consequence when the scale is used to compare aircraft of similar performance and acoustical characteristics, suggested that it would be unwise to place too much reliance on EPNL for the purposes of comparing the perceived noisiness of very dissimilar aircraft. The results pointed to the need for the more systematic experiments to be accompanied by practical evaluation of psychoacoustical models through tests such as those described here.

The original program plan called for the inclusion of up to 200 individual helicopter flyover recordings. These were to be evaluated in subjective tests at Loughborough using headphone presentation and subsequently at Langley Research Center using loudspeaker presentation. This very large sample of test sounds was considered practicable through the use of a fast rating scale test procedure to obtain annoyance assessments of each sound.

Because less than 200 original sound recordings were obtained and because of other difficulties, the scope of the experiments had to be curtailed. In an attempt to compensate for this to some extent, a large part of the basic experiments was duplicated in three independent tests; one at Loughborough, again using headphone presentation, and the others in two separate test facilities at Langley Research Center using loudspeaker presentation.

The use of headphones offers numerous advantages over loudspeakers: closer control over variations in sound level and frequency response, comfortable and convenient surroundings for the test subjects, and the ability to handle large numbers of subjects at a time. The disadvantages include the difficulty of accurately measuring the test stimuli and uncertainties concerning the relationships between normal free field or diffuse listening conditions and the pressure field of the headphones. A check on present headphone results using loudspeakers was therefore felt to be desirable.

In this report, the main experiment is described in detail in Sections 2, 3, and 4. This is followed in Sections 5 and 6 by a description of the duplicate experiments and the overall conclusions. Appendices contain (A) the Instructions to the Subjects, (B) a summary of the acoustic characteristics of the test sounds, (C) representative time histories and spectrum plots of some of the helicopter test sounds, and (D) a summary of basic characteristics of most of the helicopters utilized for recordings employed in this program.
3.0 DESCRIPTION OF THE MAIN EXPERIMENT

3.1 Test Tapes

The main tests involved an evaluation of 119 aircraft sounds; 89 helicopters and 30 CTOLs* which are described in Appendix B. The helicopter recordings were selected from approximately 140 available to provide the widest possible range of types and flight conditions as well as satisfying the requirements of reproduction quality. See Appendix C for representative time histories and spectra and Appendix D for general characteristics for the helicopters included.

Most of the helicopter flights were level flyovers although some recordings were made during approach descents. The CTOL's, which were included to allow direct comparison of the relative performance of the noise rating scales as applied to helicopters and fixed-wing aircraft, were recorded for this study at London (Heathrow) Airport at positions close to the nominal approach and flyover certification points.

The sounds were rerecorded in random sequence onto four test tapes. Each tape of approximately 30 minutes duration contained a total of 44 flyover sounds including eight reference sounds (the same T-28 flyover recording used in the preliminary experiments) recorded at 3 dB intervals over a dynamic range of 21 dB and the same five sounds recorded at the beginning and end of the tape (results for the first five were discarded to minimize the effects of any initial period of adjustment or adaptation by the subjects).

The test sounds were recorded on, and replayed from, a Nagra IV S tape recorder running at 7-1/2 ips. All sounds were manually "ramped" at start and finish and the interval between sounds was about 8 seconds during which a voice announcement of the next sound number was recorded (although in most test runs this was suppressed in favor of an electronically-controlled digital display).

The test tape was replayed to six subjects at a time through Koss PRO 4AA headphones driven by six specially constructed power amplifiers. A control unit mixed the test signals with a very low level broadband background sound whose function was to mask perceptible switching transients between sounds. The same unit suppressed the voice announcements and operated individual sound number

*Conventional takeoff and landing aircraft - in this case, all transport category types, mostly turbofan-powered.
displays when these were in use. This process was controlled by a 12 kHz tone recorded on the second tape recorder channel. To eliminate slight cross-talk during replay, the data channel was low-pass filtered at 8 kHz.

3.2 Test Procedures

The four test tapes were administered to between 36 and 40 test subjects, most of whom were undergraduate students in the age range 19 to 23 with roughly equal numbers of males and females.

The test subjects sat in armchairs inside a quiet test room. Written instructions read by the subjects before a test together with a score sheet are presented in Appendix A. The instructions were verbally reinforced and the broad purpose of the test was also explained. Most subjects participated in three tests on three separate occasions but prior to the first they were given a practice test comprising six typical sounds covering the sound level range to be heard subsequently. Subjects recorded their scores for each sound by marking numbers on their score sheets between 0 and 10. In most tests, the sound number was continuously presented on small LCD display units affixed to their clipboards.

3.3 Noise Levels

The sound recordings were analyzed to yield measurements on the various scales of noise level summarized in Table I, taking account of the frequency response of the headphones. Real-time one-third octave band analysis was performed on a GenRad 1921 analyzer coupled to a PDP 11/34 computer. The data reduction program incorporated a frequency response correction function which was based on the average response for the 12 individual earphones used in the tests.

To obtain this function, individual headphone output levels were measured underneath the headphone cushion on the head using a Knowles miniature microphone and a "pink noise" input. The frequency response of the miniature microphone was in turn measured by calibrating it against a pressure response condenser microphone in a flat-plate headphone coupler. The average frequency response together with the standard deviation for the 12 headphones is shown in Figure 4.

Because it was not possible to measure accurately the sound pressure inside the headphones when in normal use, the impulse correction terms were computed from the A-weighted tape recorder output. An indication of the likely effect of
Figure 4. Mean Frequency Response of Headphones (12 Ears) Showing ±1 Standard Deviation

Figure 5. Comparison of Impulse Corrections (EPNL - EPNL) As Measured from (a) Tape Recorder Output and (b) in Flat Plate Headphone Coupler
the headphones upon impulsiveness can be gained from a sample of 25 measurements made with a pressure response microphone in the flat plate coupler. Figure 5 compares the impulse correction terms \( \text{EPNL}_{t1} - \text{EPNL}_t \) as computed with the 1/2-second values measured in the two different ways. On average, the coupler values are approximately two-thirds of the direct values with a standard deviation of 0.3 dB. Repetitions of the headphone measurements for one particular flyover recording (561N1) using 10 different headphones showed very little variance in the magnitude of the average impulse correction (standard deviation = 0.1 dB).

For each basic scale, two levels were computed: (a) the maximum 1/2-second value during the event, and (b) the time-integrated or "duration corrected" value obtained by the summation process incorporated into the \( \text{EPNL}_t \) procedure which covers the upper 10 dB of the time history. Time-integrated (i.e., duration corrected) levels are denoted by abbreviations prefixed by the letter "E". It should be noted that the weighted sound pressure levels were computed from the one-third octave band level arrays using the weighting functions listed in Table 2 and plotted in Figure 6.

**Table 1**

**Noise Level Scales**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>L (EL)</td>
<td>Overall sound pressure level, dB</td>
</tr>
<tr>
<td>( L_A (EL_A) )</td>
<td>A-weighted sound level, dB(A)</td>
</tr>
<tr>
<td>( L_D (EL_D) )</td>
<td>D-weighted sound level, dB(D)</td>
</tr>
<tr>
<td>( L_E (EL_E) )</td>
<td>E-weighted sound level, dB(E)</td>
</tr>
<tr>
<td>( L_F (EL_F) )</td>
<td>&quot;F&quot;-weighted sound level, dB(F)</td>
</tr>
<tr>
<td>PNL (EPNL)</td>
<td>Perceived Noise Level, excluding tone correction, PNdB</td>
</tr>
<tr>
<td>( \text{PNL}_t (EPNL_t) )</td>
<td>Perceived Noise Level, including tone correction (EPNL_t is the Standard ICAO version)</td>
</tr>
<tr>
<td>( \text{PNL}<em>{t1} (EPNL</em>{t1}) )</td>
<td>PNL with tone correction and ISO impulse correction</td>
</tr>
<tr>
<td>( \text{PNL}<em>{tc} (EPNL</em>{tc}) )</td>
<td>PNL with tone and crest factor impulse correction</td>
</tr>
<tr>
<td>( L_{Ac} (EL_{Ac}) )</td>
<td>A-weighted sound level with crest factor impulse correction</td>
</tr>
</tbody>
</table>
Figure 6. Sound Level Frequency Weighting Curves
Table 2
Sound Level Weighting Functions

<table>
<thead>
<tr>
<th>Frequency, Hz</th>
<th>Weighting, dB</th>
<th>A</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td></td>
<td>-30.2</td>
<td>-12.8</td>
<td>-17.4</td>
</tr>
<tr>
<td>63</td>
<td></td>
<td>-26.2</td>
<td>-10.9</td>
<td>-14.5</td>
</tr>
<tr>
<td>80</td>
<td></td>
<td>-22.5</td>
<td>-9.0</td>
<td>-11.8</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>-19.1</td>
<td>-7.2</td>
<td>-9.4</td>
</tr>
<tr>
<td>125</td>
<td></td>
<td>-16.1</td>
<td>-5.5</td>
<td>-7.3</td>
</tr>
<tr>
<td>160</td>
<td></td>
<td>-13.4</td>
<td>-4.0</td>
<td>-5.3</td>
</tr>
<tr>
<td>200</td>
<td></td>
<td>-10.9</td>
<td>-2.6</td>
<td>-3.6</td>
</tr>
<tr>
<td>250</td>
<td></td>
<td>-8.6</td>
<td>-1.6</td>
<td>-2.2</td>
</tr>
<tr>
<td>315</td>
<td></td>
<td>-6.6</td>
<td>-0.8</td>
<td>-1.1</td>
</tr>
<tr>
<td>400</td>
<td></td>
<td>-4.8</td>
<td>-0.4</td>
<td>-0.3</td>
</tr>
<tr>
<td>500</td>
<td></td>
<td>-3.2</td>
<td>-0.3</td>
<td>0.1</td>
</tr>
<tr>
<td>630</td>
<td></td>
<td>-1.9</td>
<td>-0.5</td>
<td>0.1</td>
</tr>
<tr>
<td>800</td>
<td></td>
<td>-0.8</td>
<td>-0.6</td>
<td>0</td>
</tr>
<tr>
<td>1,000</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1,250</td>
<td></td>
<td>+0.6</td>
<td>2.0</td>
<td>0.7</td>
</tr>
<tr>
<td>1,600</td>
<td></td>
<td>+1.0</td>
<td>4.9</td>
<td>2.1</td>
</tr>
<tr>
<td>2,000</td>
<td></td>
<td>+1.2</td>
<td>7.9</td>
<td>4.0</td>
</tr>
<tr>
<td>2,500</td>
<td></td>
<td>+1.3</td>
<td>10.6</td>
<td>5.9</td>
</tr>
<tr>
<td>3,150</td>
<td></td>
<td>+1.2</td>
<td>11.5</td>
<td>7.6</td>
</tr>
<tr>
<td>4,000</td>
<td></td>
<td>+1.0</td>
<td>11.1</td>
<td>8.7</td>
</tr>
<tr>
<td>5,000</td>
<td></td>
<td>+0.5</td>
<td>9.6</td>
<td>9.1</td>
</tr>
<tr>
<td>6,300</td>
<td></td>
<td>-0.1</td>
<td>7.6</td>
<td>8.3</td>
</tr>
<tr>
<td>8,000</td>
<td></td>
<td>-1.1</td>
<td>5.5</td>
<td>6.5</td>
</tr>
<tr>
<td>10,000</td>
<td></td>
<td>-2.5</td>
<td>3.4</td>
<td>3.8</td>
</tr>
</tbody>
</table>
The E-weighted scale is based on Steven's generalized "perceived level" function. The F-weighting is the (nonstandard) abbreviation assigned to a curve derived by Powell from a study of the relations between impulsiveness, repetition rate, and judged annoyance of simulated helicopter sounds. Above 1 kHz, it is identical to the D-weighting. Below 1 kHz it rolls off more rapidly than the D-curve approaching the A-curve at the very lowest frequencies (see Figure 6 and Table 2).

The ISO impulsiveness correction is applied to the half-second sound level time history of the flyover sound in a manner analogous to the use of the tone correction. The correction is computed from the A-weighted sound pressure time history \( p(t) \) which is low-pass filtered at 2 kHz for anti-aliasing purposes and digitized at 5 kHz. For each half-second time period, a quantity \( X \) is computed where

\[
X = 10 \log_{10} \left[ \frac{1}{N} \sum_{i=1}^{N} \left( \frac{p_i^2 - s}{s} \right)^2 \right]
\]

\[
s = \frac{1}{N} \sum_{i=1}^{N} p_i^2
\]

and \( p_i \) are the \( N \) sampled values of \( p(t) \). The half-second impulsiveness correction is then given as follows:

- if \( X < 5.5 \) \( \Delta = 0 \)
- if \( 5.5 \leq X \leq 10.5 \) \( \Delta = 0.8 (X-3) \), dB
- if \( 10.5 < X \) \( \Delta = 6 \) dB.

The "crest factor impulse correction" is also computed from the digitized A-weighted sound pressure time history. A crest factor \( C \) is calculated for each half-second period as the ratio

\[
C = \frac{p_{\text{max}}^2}{\frac{1}{N} \sum p_i^2}
\]

where \( p_{\text{max}} \) is the largest numerical value of \( p_i \). The impulse corrected level is then given by (for \( \text{PNL}_{tc} \) for example):

\[ \text{PNL}_{tc} = \text{PNL}_t + 10 \log_{10} C - 12 \]

subject to the proviso that \( \text{PNL}_{tc} \geq \text{PNL}_t \) (subtraction of 12 ensures that \( \text{PNL}_{tc} = \text{PNL}_t \) for broadband random noise).

It must be pointed out that all noise level calculations can only be considered approximate in that (a) the weighted levels are computed from one-third octave band levels, (b) although the time integration periods are nominally 0.5 second, they were in practice controlled by the cycle time of the GR 1921 analyzer which is slightly less than this (a difference which is, of course, accounted for in the integration process), and (c) the impulse correction is also nominal rather than actual because it does not allow for unmeasurable differences caused by the headphone response. Although these approximations mean that all calculations strictly are "nonstandard," the effects of (a) and (b) are considered to be negligibly small. The magnitude of the error due to (c) which is significant cannot be estimated with any precision although we may be confident that in general the true impulsiveness will be somewhat less than the nominal value.

### 3.4 Annoyance Levels

The mean subjective score \( SS \) (and standard deviation) for each sound were calculated across all subjects. For each test, the value of \( SS \) was plotted against measured levels \( L_A \) and \( \text{EPNL}_t \) for the eight repetitions of the reference sound and the regression lines were then used to convert \( SS \) for each test sound to Annoyance Levels \( NL \) and \( \text{NLE} \), in dB(A), and EPNdB. In other words, the Annoyance Levels, \( NL \) and \( \text{NLE} \), of any sound are the levels (in dB(A) and EPNdB) of the standard reference sound which would be equally annoying. \( \text{NLE} \) was included to make suitable allowance for possible nonlinearity between dB(A) and EPNdB over the wide dynamic range of the tests. (In fact, the relationship was entirely linear for the reference sound with the relationship \( \text{NLE} = \text{NL} + 9.0 \).)

### 3.5 Accuracy Considerations

The accuracy of the experimental method can be assessed in two ways. The square root of the grand average inter-subject variance (averaged across all test
sounds) is 1.5 annoyance scale units which yields a standard error of 0.25.* Since one annoyance unit translates to approximately 4 dB on the NL scale, this may be interpreted as an average standard error of approximately 1 dB, i.e., the 95 percent confidence interval associated with any individual NL is about ±2 dB.

A check on this is provided by the annoyance scores for the standard reference sound which is repeated through the main part of the tests 32 times (albeit at different levels). The average standard error of estimate about the regression lines** may therefore be taken as a measure of the variability of individual NL values. This has a value of 1.4 dB. This is a little larger than the standard error computed above but the difference could be explained by the small sample size.*** These considerations suggest that errors (i.e., the standard deviation) associated with a perfect noise rating scale would not be less than about 1.5 dB in this experiment.

---

* For 36 subjects.

** The lines used to convert from SS to NL.

*** A standard F-test shows that there is about a 1 percent probability that this difference arose by chance.
4.0 RESULTS OF THE MAIN EXPERIMENT

4.1 Analytical Considerations

Sample results are shown in Figures 7 and 8 in the form of "scatter diagrams" of measured level plotted against annoyance level (the significance of the different plotting symbols will be discussed later). The correlation between measured (y) and judged (x) levels may be expressed in various ways and a choice depends upon the criteria of assessment, especially concerning the linearity of underlying relationships. It might be supposed for example that since both ordinate and abscissa in Figure 7(a) are maximum levels, expressed in dB(A), the underlying relationship should be the line y = x. Figure 7(a) shows that this is clearly not the case.

This discrepancy suggests that $L_A$ is not a particularly good estimator of $NL$ for the test sounds in general. But the form and magnitude of the apparent error depends on the precise choice of reference sound (which itself should be assigned no more importance than any one of the individual sounds in Figure 7(a)) and the gross deviation of the data cluster from $y = x$ may depend upon special peculiarities of the one used. On the other hand, the many test sounds may vary with respect to factors of importance not accounted for in the variable $L_A$.

The ultimate purpose of the noise measurement scales under investigation is to predict average annoyance levels. In this respect, it would be more logical to reverse the axes in scatter diagrams like Figure 7(a). However, in these tests of the predictive performance of the scales, annoyance level $NL$ is the independent variable (admittedly involving a degree of experimental error) and the dispersion of the data points in the y-direction is a measure of how well (or how poorly) the noise measurement scales do their job.

Of course, any noise measurement scale for which the data points are clustered tightly about a monotonically increasing relationship between $y$ and $x$ may be considered good for the practical purposes of rating aircraft noise. However, in the context of the present tests, it is also considered desirable that the relationship between measured level and annoyance level be constrained to be linear with unit slope. This is because the only property of the reference sound which changes significantly with $NL$ or $NLE$ is that of sound level itself. Any composite noise scale which purports to take proper account of temporal and spectral variations in the test sounds should by definition incorporate the appropriate tradeoffs between their contributions and that of sound level, maintaining the relationship $y = x + c$. The constant $c$ should ideally be zero or at least small.
The best fitting straight line of unit slope (i.e., the unit slope line about which the variance in the y-direction is minimized) passes through the centroid of the scatter diagram so that the constant \( c \) is the mean value of the error \((y_i - x_i)\). The goodness of fit is inversely related to the standard deviation \( s \) of the error. In Figure 7(a) the unit slope line is \( y = x - 4.6 \) with a standard deviation \( s = 2.5 \text{ dB} \).

4.2 General Comparison of Noise Level Scales as Predictors of Annoyance Level

Scatter diagrams comprising plots of measured level against annoyance level for some of the various noise measurement scales are presented as Figures 7(a) through 8(g). Different plotting symbols are used for the subgroupings identified in Table 3.

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less Impulsive Helicopters</td>
<td>73</td>
</tr>
<tr>
<td>More Impulsive Helicopters</td>
<td>16</td>
</tr>
<tr>
<td>CTOL Approaches</td>
<td>12</td>
</tr>
<tr>
<td>CTOL Takeoffs</td>
<td>18</td>
</tr>
</tbody>
</table>

The "more impulsive" helicopter sounds are those for which the integrated impulse correction given by \( I = \text{EPNL}_{ti} - \text{EPNL}_{t} \) is greater than or equal to an arbitrary threshold value of 4 dB.

The unit slope straight line in each diagram is fitted to all 119 data points to minimize the error variance in the y-direction. Table 4 lists the overall mean prediction error and its standard deviation together with the mean and standard deviation of the displacements of each data subgroup from the overall mean line. Thus, for example, the mean error \( L_A - \text{NL} \) for all 119 sounds is -4.6 dB with a standard deviation of 2.5 dB. The 89 helicopter points lie on average 0.2 dB below this mean error line (standard deviation = 2.6 dB) and the 30 CTOL points lie on average 0.3 dB above it (standard deviation = 1.9 dB). The further breakdowns in Table 4 give the margins for "more" and "less" impulsive helicopters separately and

i.e., the line \( y = n + c \) is positioned so as to minimize the dispersion of the data points about it in the vertical (y) direction. This dispersion will be greater than that about the linear regression line (of y on x) if the slope of the latter is not unity.
Figure 7. Measured Levels Versus Judged Annoyance Levels; Maximum Level Scales

Low Level Headset Test

Mean Error = -4.6 dB
Standard Deviation = 2.5 dB
Figure 7. (Continued)
Figure 8. Measured Levels Versus Judged Annoyance Levels; Time-Integrated Scales
Figure 8. (Continued)
LOW LEVEL HEADSET TEST

MEASURED LEVEL EPNLTI

ANNOYANCE LEVEL NLE, EPNLDB

(g) $\text{EPNL}_{11}$
Mean Error = $+0.1 \text{ dB}$
Standard Deviation = 2.4 dB

Figure 8. (Continued)
Table 4

Main Experiment Annoyance Prediction Errors, dB
Mean errors for subsamples are relative to overall mean error (listed for all 119 sounds). Standard deviations in parentheses

<table>
<thead>
<tr>
<th>Scale</th>
<th>Maximum Levels</th>
<th>Time-Integrated Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>89 Helos 73 less imp.</td>
<td>16 more imp.</td>
</tr>
<tr>
<td>All 119 Sounds</td>
<td>30 CTOLs</td>
<td>12 approach 18 takeoff</td>
</tr>
<tr>
<td>-A</td>
<td>-4.6 (2.5)</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+0.3 (1.9)</td>
</tr>
<tr>
<td>D</td>
<td>+2.3 (2.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.6 (3.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+2.6 (2.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.6 (2.1)</td>
</tr>
<tr>
<td>E</td>
<td>+0.3 (2.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.9 (2.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+1.3 (2.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.7 (1.8)</td>
</tr>
<tr>
<td>F</td>
<td>+0.6 (3.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>+1.7 (3.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.0 (2.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+4.1 (2.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.0 (2.6)</td>
</tr>
<tr>
<td>L</td>
<td>+8.6 (2.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>+1.6 (1.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.6 (1.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.1 (2.1)</td>
</tr>
</tbody>
</table>
### Table 4 (Continued)

<table>
<thead>
<tr>
<th>Scale</th>
<th>Maximum Levels</th>
<th>Time-Integrated Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>89 Helos</td>
<td>73 less imp. 16 more imp.</td>
</tr>
<tr>
<td></td>
<td>12 approach 18 takeoff</td>
<td>30 CTOLs</td>
</tr>
<tr>
<td>All 119 Sounds</td>
<td>30 CTOLs</td>
<td>12 approach 18 takeoff</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PNL</strong> <strong>+10.5 (2.6)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>+0.1 (2.3)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>** ** (**)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>-1.9 (3.2)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>+2.5 (1.7)</strong></td>
<td><strong>-1.7 (1.7)</strong></td>
</tr>
<tr>
<td></td>
<td><strong>-0.9 (2.6)</strong></td>
<td><strong>-1.5 (1.0)</strong></td>
</tr>
<tr>
<td><strong>PNL_{ti}</strong> <strong>+12.7 (2.6)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>+0.3 (2.5)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>+1.3 (3.3)</strong></td>
<td><strong>+0.1 (2.4)</strong></td>
</tr>
<tr>
<td></td>
<td><strong>-0.9 (2.7)</strong></td>
<td><strong>-2.7 (1.2)</strong></td>
</tr>
<tr>
<td></td>
<td><strong>-2.0 (2.4)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>PNL_{tc}</strong> <strong>+13.1 (2.3)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>+0.2 (2.3)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>+0.0 (2.4)</strong></td>
<td><strong>+0.8 (1.8)</strong></td>
</tr>
<tr>
<td></td>
<td><strong>+0.1 (2.2)</strong></td>
<td><strong>+0.1 (2.2)</strong></td>
</tr>
<tr>
<td></td>
<td><strong>-0.9 (1.9)</strong></td>
<td><strong>-2.4 (1.2)</strong></td>
</tr>
<tr>
<td></td>
<td>** ** (**)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>-1.6 (2.0)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>L_{Ac}</strong> <strong>-1.9 (2.4)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>+0.3 (2.5)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>+0.2 (2.5)</strong></td>
<td><strong>+0.9 (2.1)</strong></td>
</tr>
<tr>
<td></td>
<td><strong>-14.8 (2.5)</strong></td>
<td><strong>-2.5 (1.7)</strong></td>
</tr>
<tr>
<td></td>
<td><strong>-0.4 (2.4)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>**-0.7 (2.0)</td>
<td></td>
</tr>
</tbody>
</table>

40
for approaching and departing (takeoff) CTOL aircraft. Table 4 lists these statistics for all measurement scales including those which are not illustrated by scatter diagrams. It should be noted that the mean prediction error for the maximum levels is referenced to the annoyance level NL in dB(A), whereas for the time integrated levels, the reference is NLE in EPNdB (where NLE = NL + 9). The absolute values of these mean errors are of little importance; it is the differences between them which are of interest.

In Table 4, asterisks within pairs of mean values indicate that the difference is statistically significant according to student's T-test (one asterisk for 5 percent significance level, two for 1 percent). Those errors paired without asterisks are not significantly different. Asterisks (in parentheses) between pairs of standard deviation figures (in parentheses) indicate that their respective error variances are significantly different according to a standard F-test (again at the 5 percent or 1 percent level of significance).

On the basis of a broad comparison between the overall error standard deviations for the maximum levels and the time-integrated levels for all sounds, it is clear that for the commonly used scales, the duration correction is generally beneficial in that the consistency with which the scales predict annoyance level is improved. The improvement is significant at the 1 percent level in the cases of LA', LD, LF, and PNL, and at the 5 percent level for PNL (without tone correction). For LF, the improvement is very large, doubtlessly because the uncorrected maximum level is a very poor performer. For LE', PNL, PNLc, and LAc', there is no significant change of this group. The maximum level, LE, is itself a good index of annoyance but the others involve impulsiveness corrections which generally appear to do little to improve the predictive accuracy of the basic scales to which they are applied. Instead, in every case, the impulsiveness corrections counter the beneficial effects of the duration allowance (compare LA and LAc', PNL and PNLc, PNL and PNLc).

*The large effect of the duration correction in the case of the F-weighting is possibly linked to a correlation between low frequency energy and duration. Note for example that the improvement is extremely large for the CTOL subsample (standard deviation falls from 3.1 dB to 1.4 dB) for which the takeoffs have longer durations and more low frequency energy than the approaches.
Examination of the subgroup results shows that the non impulse-corrected maximum level scales tend to underestimate the annoyance levels of the more impulsive helicopters relative to the less impulsive ones by around 2 dB. However, this difference nearly vanishes when the duration allowance is included (except in the case of $L_F$) implying a degree of correlation between impulsiveness and signal duration. Confining attention to the simple weighted sound level scales, it may also be noted that the mean differences between more and less impulsive helicopters tend to decrease slightly as emphasis is transferred from high frequencies to low $[1.5(EL_F) \rightarrow 0.8(EL_A) \rightarrow 0.4(EL_D) \rightarrow 0.2(EL_E)]$. This suggests a positive correlation between impulsiveness and low frequency energy in the helicopter sounds.

Many of the subgroup error deviations are considerably smaller than the overall values. This is particularly true of the CTOL sounds (for which the standard deviations are of the same order as the experimental error, i.e., as about as low as could be expected from an ideal noise rating scale). The standard deviations for the helicopters are also small in absolute terms but for all scales except $L_{AC}$ they are significantly greater than the CTOL values (i.e., practically all scales predict noise annoyance levels less consistently for helicopters than for CTOLs).

Another feature which is common to all duration corrected scales but one ($EL_F$) is that on average they overestimate annoyance levels of helicopters relative to those of CTOLs by around 2 dB. The F-weighted scale appears to overcome this deficiency by assigning relatively more weight to higher frequency energy than the other scales, thus increasing the relative levels of the CTOL sounds (this is particularly noticeable for the CTOL approach sample).

Turning now to the question of impulse corrections, it is apparent that all the conventional duration-corrected scales predict the annoyance levels of the more impulsive helicopters with rather poor consistency. (In all cases, except $EL_F$, the error variance is significantly greater than it is for the less impulsive sample at the 1 percent level.) This difference is eliminated for all impulse-corrected scales, whether they involve the ISO factor or the crest factor based term. However, this "improvement" is achieved at least as much by increasing the variance for the less impulsive sample as it is by decreasing the variance for the more impulsive sounds. Consequently, for the impulse-corrected scales, there are increases in the
variances for the combined helicopter sample and for the total sample. However, for these scales, the pooled standard deviations for the subgroups are little larger than those of the uncorrected scales; the substantial increases in the overall variances arise because the impulse corrections generate significant differences between the mean prediction errors for the two helicopter subgroups and increase the differences between helicopters and CTOL means.

This is clearly evident in Table 5 which ranks the various duration corrected scales with respect to total error standard deviation but also lists the pooled values. (The differences between the first five scales are not significant at the 5 percent level.)

Table 5

<table>
<thead>
<tr>
<th>Standard Deviations of Annoyance Prediction Errors, in dB, for Duration-Corrected Annoyance Scales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Standard Deviation</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>EPNLt</td>
</tr>
<tr>
<td>EPNL</td>
</tr>
<tr>
<td>ELD</td>
</tr>
<tr>
<td>ELA</td>
</tr>
<tr>
<td>ELF</td>
</tr>
<tr>
<td>EPNLt tc</td>
</tr>
<tr>
<td>EPNLti</td>
</tr>
<tr>
<td>ELAc</td>
</tr>
<tr>
<td>EPNLi</td>
</tr>
</tbody>
</table>

This general review of the performance of the different noise scales begins to reveal the difficulties of isolating the contributions of the various factors such as frequency distribution, tonality, signal duration, and impulsiveness to annoyance, especially when there is a degree of association between them. In general, it seems reasonable to conclude that duration is a most important factor while tonality (as measured by the tone correction in EPNLt) is of minor importance.
The two impulsiveness corrections enhance the consistency with which the noise scales predict annoyance levels of the more impulsive helicopter sounds when they are considered in isolation but, on average, the overall magnitude of the correction is too great, causing the more impulsive helicopters to be overrated with respect to the less impulsive ones. This, together with an increase in error variance for the less impulsive helicopters, causes the disadvantages of the corrections to outweigh their advantages.

To obtain a more quantitative evaluation of the roles of the various underlying factors, it is helpful to turn to multiple regression analysis which yields the coefficient in an optimum annoyance predictor formula comprising a linear combination of the variables.

### 4.3 Multiple Regression Analysis

The equivalent level $EPNL_{tj}$ of any test sound may be written:

$$EPNL_{tj} = L + D + T + I$$

where:

- Maximum Level $L = PNL$
- Duration Correction $D = EPNL - PNL$
- Tone Correction $T = EPNLt - EPNL$
- Impulse Correction $I = EPNL_{tj} - EPNL_t$

The equivalent level is thus a linear combination of these underlying variables but the relative weight attached to each of them is fixed (and equal).

Multiple regression analysis allows the relative weights to vary; the resultant regression analysis gives the best combination. Specifically, it yields the regression coefficients $a$ through $e$ in the linear prediction equation

$$NL' = aL + bD + cT + dI + e$$

The dependent variable $NL'$ is the predicted annoyance level and the regression coefficients are those for which the variance of the prediction errors $NL' - NL$ (predicted annoyance level - actual annoyance level) is minimal. The standard deviation of this error, labeled $s_{xy}$, is sometimes called the "standard error of estimate."
If the predictor variables are truly independent (uncorrelated), the regression coefficients can be isolated with complete accuracy. However, uncertainty arises when the variables are intercorrelated and in this case the computed regression coefficients have to be assigned a probable error margin (or confidence limits). Table 6 gives the matrix of intervariable correlation coefficients (Pearson's R) for the complete sample of 119 sounds and for the subsamples of 89 helicopters and 30 CTOLs. This shows that the correlation between variables is significant in all cases except (not surprisingly) between impulsiveness and the other variables for the CTOL sample.

The relation between each of these potential predictor variables and annoyance has therefore been examined by a process of "stepwise" multiple regression in which the independent variables are admitted to the analysis one at a time in descending order of importance. At each stage of the analysis, the next most important variable is that which makes the greatest contribution to explained variance. The regression equations defined below exclude variables which were not significant at the 5 percent level.

Table 6

|          | D       | T       | I       | \(|R_{crit}|\) | 1% (5%) |
|----------|---------|---------|---------|----------------|---------|
| All sounds D | -0.647  | 0.466   | -0.616  | 0.235          | (0.176) |
| n = 119 T    |         |         |         |                |         |
| All helicopters D | -0.610  | 0.399   | -0.586  | 0.269          | (0.205) |
| n = 89 T    |         |         |         |                |         |
| All CTOLs D | -0.666  | 0.593   | -0.211  | 0.449          | (0.349) |
| n = 30 T    |         |         |         |                |         |

For the complete sample of 119 sounds, the regression equation is

\[ NL' = 0.92L + 0.56D + 1.1 \quad (s_{xy} = 1.6 \text{ dB}) \ldots \]  (1)
where \( NL' \) is the predicted annoyance level and \( s_{xy} \) is the standard error of estimate (= standard deviation of residual error \( NL' - NL \)). The variables \( T \) and \( I \) are not significant predictor variables (at the 5 percent level). However, if a dummy variable \( H \) is introduced, which takes the value 1 for helicopters and 0 for CTOLs, the result is rather different:

\[
NL' = 0.89L + 0.80D + 0.74T - 1.8H + 4.4 \quad (s_{xy} = 1.4 \text{ dB}) \ldots (2)
\]

The variable \( T \) is now significant at the 5 percent level. This result confirms that helicopter sounds are less annoying than CTOL sounds (by an amount equivalent on average to 1.8 dB) and that if this difference is ignored in the predictor model, tone corrections are of little or no value.

If the helicopters \((n = 89)\) and CTOLs \((n = 30)\) are analyzed separately, the two separate regression equations become:

Helicopters:  \( NL' = 0.89L + 0.78D + 0.90T + 2.63 \quad (s_{xy} = 1.5 \text{ dB}) \ldots (3) \)

CTOLs:  \( NL' = 0.89L + 0.73D + 5.4 \quad (s_{xy} = 0.9 \text{ dB}) \ldots (4) \)

These indicate that the tone correction is an effective annoyance predictor only in the case of the helicopter sounds.

The 95 percent confidence limits for the regression coefficients in the above equations are given in Table 7.

Table 7
Confidence Range for Regression Coefficients

<table>
<thead>
<tr>
<th>Equation</th>
<th>Sample</th>
<th>95% Confidence Range for Regression Coeff. of</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( L )  ( D )  ( T )    ( H )</td>
</tr>
<tr>
<td>1</td>
<td>All sounds (119)</td>
<td>0.84-0.99 0.38-0.74 *</td>
</tr>
<tr>
<td>2</td>
<td>All sounds (119)</td>
<td>0.82-0.97 0.62-0.97 0.22-1.26 -2.5 to -1.1</td>
</tr>
<tr>
<td>3</td>
<td>Helicopters (89)</td>
<td>0.81-0.97 0.54-1.02 0.24-1.56 **</td>
</tr>
<tr>
<td>4</td>
<td>CTOLs (30)</td>
<td>0.76-1.03 0.53-0.92 *</td>
</tr>
</tbody>
</table>

* not significant
** variable not admitted
The large confidence intervals associated with the coefficients of the tone correction term \( T \) shows that in those cases where it is a significant predictor variable, it is not a particularly strong one; indeed, in both cases its inclusion reduces the standard error of estimate by a mere 0.05 dB. However, this does not necessarily imply that the tone correction is inappropriate; more probably, it reflects the fact that the term varies very little in this sample of typical aircraft and helicopter sounds (standard deviation = 0.6 dB).

The coefficients of \( L \) and \( D \) are statistically indistinguishable between the helicopter and CTOL subsamples (Eqs.(3) and (4)); i.e., the regression lines are parallel, separated by the mean difference of around 2 dB. Inclusion of the dummy variable \( H \) in the total sample regression (Eq.(2)) thus yields very similar coefficients for \( L \) and \( D \). If the variable \( H \) is not admitted, the prediction error is significantly greater and the coefficient of \( D \) changes markedly (reflecting a degree of correlation between \( D \) and \( H \); see Eq.(1)).

Table 7 shows that the coefficients of \( L \) and \( D \) do not differ substantially from the unit values effectively specified in the EPNL\(_t\) formula (EPNL\(_t\) = \( L + D + T \)). Thus, we find in Table 8 that EPNL\(_t\) is practically as good an annoyance predictor as the regression equations.

Table 8
Comparison of Annoyance Prediction Errors
EPNL\(_t\) vs Regression Model

<table>
<thead>
<tr>
<th>Sample</th>
<th>Regression Formula</th>
<th>EPNL(_t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sounds (119)</td>
<td>1.6 (1.4*)</td>
<td>1.7</td>
</tr>
<tr>
<td>Helicopters (89)</td>
<td>1.5</td>
<td>1.6</td>
</tr>
<tr>
<td>CTOLs (30)</td>
<td>0.9</td>
<td>1.0</td>
</tr>
</tbody>
</table>

*including dummy variable \( H \)
4.4 Further Analysis of Helicopter Results

A comparison of mean annoyance prediction errors for individual helicopter types reveals significant differences, for example, between the Westland Wessex, the Bell 205, and the Bell OH58A. Some of these differences are illustrated in Figure 9 which compares some mean annoyance prediction errors associated with the time-integrated noise level scales. Five specific helicopter types are selected: Wessex, S64, Puma, Bell 205, and Bell OH58. The first four of these are drawn from four distinct groups of sounds, each of which can be represented by a typical average one-third octave spectrum shape. These groups are listed in Table 9 and the spectra are shown in Figure 10. The spectra have been drawn by eye as a best fit to a superposition of the individual spectra of all members of the groups. The individual spectra are themselves average values obtained by time integrating each one-third octave band level over its own 10 dB-down duration during the flyover. The relative levels of the four spectra in Figure 10 have been adjusted to ensure that the mean prediction errors for the four groups are correctly related on the ELA scale (this choice of scale is arbitrary and it does not affect the observations which follow).

Table 9

<table>
<thead>
<tr>
<th>Group I</th>
<th>Group II</th>
<th>Group III</th>
<th>Group IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bell 204 (10)</td>
<td>Squirrel (5)</td>
<td>S76 (5)</td>
<td>Wessex (5)</td>
</tr>
<tr>
<td>Bell 205 (4)</td>
<td>Bo 105 (8)</td>
<td>Puma (6)</td>
<td></td>
</tr>
<tr>
<td>Bell 212 (4)</td>
<td>S64 (4)</td>
<td>Super Frelon (5)</td>
<td></td>
</tr>
<tr>
<td>Bell 206 (3)</td>
<td>S61 (3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Progressing from Group I to Group IV, the typical spectra show a progressive shift in energy distribution from low frequencies to high. The Group I helicopters, all members of the large two-blade helicopter family related to the military UH1, *Because the ISO and crest factor impulse corrections are highly correlated (between (EPNL - EPNL) and (EPNL - EPNL)), the correlation coefficient for all 89 helicopter sounds is 0.94), they may be regarded as equivalent measures of the same characteristic.*
Sample Sizes: Wessex 5; S64; Puma-6; 205-4; OH58A-3

Figure 9. Relative Mean Annoyance Prediction Errors for Selected Helicopter Types
Figure 10. Typical Average Spectra for Helicopter Subgroups (at equally annoying levels)
exhibit pronounced main rotor noise with a low fundamental frequency and, often, a high degree of impulsiveness. Their acoustic energy is clearly concentrated at the low end of the audible frequency range. Group IV comprises the five flyover sounds of the Westland Wessex, a turbine-powered derivative of the four-blade S58. The sound of the Wessex is perhaps best described as "nondescript" with little or no impulsiveness and with no particular sound source dominant. Its frequency spectrum is unique among the helicopters studied in that its energy is spread broadly across frequencies above about 250 Hz with little below that limit.

The OH58A, for which results are also included in Figures 9 and 11, is the military version of the ubiquitous two-blade Bell 206 Jet Ranger. Its spectrum does not fit any of the four groups but it is of special interest because it appears to be a deviant type (in respect of mean annoyance prediction error) and it was one of the two helicopters used in the Wallops Island field experiment \(^\text{10}\) (indeed, the recordings used in this study were made during that experiment).

Figure 11 compares the mean annoyance prediction errors, together with their respective 95 percent confidence intervals, for the four groups of sounds. This diagram indicates that of the four sound level weighting functions, "F" is the least appropriate for helicopter noise since it clearly separates the four results. (The differences between the group means are all significant at the 5 percent level.) The A-scale shows some improvement in that the Group II and III errors merge but Group I and IV remain significantly different. For the D-scale, the collapse is more complete with only the Group IV (Wessex) data significantly deviating (at the 5 percent level). No deviations occur in the case of the E-scale for which no differences between means are significant at the 5 percent level.

Figure 12 provides a further comparison of the four frequency weighting functions corresponding to the A, D, E, and "F" scales. Here, the reference levels of these curves have been shifted so that the average levels for all 89 helicopter sounds would be the same on each duration corrected scale. (Thus there is a 6.8 dB difference between the A-curve and the D-curve at 1 kHz. The difference between the A-curve and the "F"-curve is 4.7 dB and between A and E it is 5.1 dB.) Relative to the A-weighting, the other curves give less emphasis to the mid-frequencies (250 to 2,000 Hz) and more to the high frequencies (greater than 2 kHz). Below 250 Hz, the "F"-curve differs little from the A-weighting but the D- and E-functions give considerably more weight. Of the four weightings, the "F"-
Figure 11. Ninety-Five Percent Confidence Intervals for Helicopter Group Prediction Errors
Figure 12. Sound Level Weighting Functions with Relative Levels Adjusted to Give Equal Average Levels for All Helicopter Sounds
curve shows the greatest variation between low frequencies and high. Between 2 and 4 kHz, E and A are similar, "F" applies considerably more weight, and D is intermediate. Above 4 kHz, E becomes dominant but this range is not particularly significant for the helicopter sounds (see Figure 10).

Although it may not be immediately apparent, consideration of Figures 10, 11, and 12 suggests that results for the four groups are harmonized as less weight is given to high frequencies and more to low. It has not been possible to explore this possibility further by fully computing modified sound levels with different weightings from the time histories of one-third octave spectra. However, a realistic assessment of the likely results can be obtained by applying alternative frequency weightings to the time-averaged spectra in Figure 10. Justifications for this approximate procedure may be found in Table 10 where the relative mean prediction errors so calculated are compared with the properly computed values.

Table 10

Comparison of Mean Annoyance Prediction Errors Based On
(a) Full Calculation from Individual One-Third Octave Spectral Time Histories and
(b) Weighting the Typical Average Spectra in Figure 10

<table>
<thead>
<tr>
<th>Sample</th>
<th>Mean Annoyance Prediction Error, dB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ELa (a)</td>
</tr>
<tr>
<td>Group I</td>
<td>-7.0</td>
</tr>
<tr>
<td>Group II</td>
<td>-5.7</td>
</tr>
<tr>
<td>Group III</td>
<td>-5.6</td>
</tr>
<tr>
<td>Group IV</td>
<td>-3.1</td>
</tr>
<tr>
<td>Range</td>
<td>3.9</td>
</tr>
</tbody>
</table>

As noted previously, the simple weighted levels have been normalized by adjusting the overall levels of the four average spectra to equate the two sets of results for the A-scale - the same basis used to determine the equal annoyance levels of the four spectra in Figure 10 (the choice of base scale is arbitrary - the conclusions are unaffected by it).
The agreement in Table 10 between the accurate (a) and approximate (b) methods for applying frequency weighting is good and lends credence to the validity of the figures in the final column of this table which shows further improvement when the D-weighting is slightly modified to reduce the high frequency weighting, i.e., to transfer still more emphasis from high frequencies to low as shown in Figure 13.

Figure 11 indicates that EL_A underestimates the annoyance levels of the Group I sounds but the difference (between Groups I, II, and III) disappears when more weight is assigned to the low frequencies by EL_D. However, the same result is achieved by applying the crest factor impulse correction in EL_AC. The dilemma therefore arises as to whether the Group I sounds (the UHI family of helicopters) are being underrated because insufficient emphasis is given to low frequency energy or to impulsiveness.* In the case of EPNL, more weight is given to both factors (than by EL_A) and the Group I sounds are substantially overrated.

The question of impulsivity is considered further in the next section. The analysis in this section has clearly served to illuminate two important general points. The first is that the diagnosis of underlying relationships is hampered by the presence of intercorrelations, even though the test sample is large. The second is that it might be misleading to draw general conclusions from an experiment involving a small number of helicopter types. Figures 9 and 11 indicate, for example, that the Bell OH58A helicopter, which was used for the Wallops Island field tests of the ISO impulse correction is, perhaps, atypical of helicopters in general. On the basis of conventional noise scales, these figures show that, relative to other helicopters, the OH58A is particularly annoying and is thus perhaps an unrepresentative standard by which to gauge them. Had the Wessex been used as a reference aircraft, the case in support of the ISO correction would have been strong (but equally misleading because the Wessex appears to have a particularly inoffensive sound). These results highlight the fact that, as a group, helicopters exhibit a range of acoustic characteristics which is probably greater than for other classes of aircraft and explained why, in general, annoyance levels for helicopter noise are predicted less consistently.

* A useful index of the frequency distribution of energy in a sound is the difference between overall (linear) level and A-weighted level, a difference which increases with the concentration of energy at lower frequencies. For all 89 helicopter sounds, the correlation between this index and impulsiveness (EPNL - EPNL_l) is significant at R = 0.52.
Figure 13. Modified D-Weighting (see Table 10)
4.5 The Need for an Impulse-Correction Term

To some extent, conclusions concerning the appropriateness of the ISO impulse correction are clouded due to the correlation between impulsiveness and low frequency content in the sample of helicopter sounds studied. However, further light may be thrown on the problem by more detailed examination of some individual results.

Table II lists the annoyance prediction errors for a subsample of helicopter sounds subdivided by helicopter type. These are the types for which some of the recorded sounds exhibit rather different impulse corrections because recordings were made in both flyover and approach conditions.

This table reveals no tendency for either $EL_A$ or $EPNL$ to underestimate the annoyance levels of the more impulsive sounds. Indeed, in the case of $EL_A$, the converse is true for this particular sample (i.e., it is the less impulsive sounds which are underestimated). There is no significant difference between the two mean errors for $EPNL$.

One of the reasons why $EL_A$ and $EPNL$ are inherently sensitive to impulsiveness may be deduced from Figure 14 which shows the average one-third octave band spectra* for some of the sounds of Table II. For each helicopter, the spectra have been overlaid (by eye) so that they coincide at the higher frequencies where the band levels tend to be controlled by noise sources other than the main rotor (i.e., nonimpulsive sources). In all cases, the more impulsive sounds are characterized by significant amplification of spectrum levels in the range 125 to 500 Hz. Since this is the region where weighted band levels of helicopter noise tend to be maximal anyway, impulsiveness directly increases the measured sound levels.

A second factor was evident in Table IV where, for most of the maximum level scales, there are significant differences between the mean prediction errors for the more and less impulsive helicopter samples (i.e., the maximum measured levels tend to underestimate the judged annoyance of the more impulsive helicopters by a significant amount). This difference largely disappears when duration allowances are included, again suggesting a correlation between impulsiveness and duration.

*As in Figure 10, the average one-third octave band levels were computed by time-integration between the 10 dB-down points.
Figure 14. Average Spectra for More and Less Impulsive Recordings of Same Helicopter Types (Impulsiveness corrections $\text{EPNL}_{t_1} - \text{EPNL}_{t_2}$ in Parentheses)
Table II
Annoyance Prediction Errors for Selected Helicopter Sounds, in dB

<table>
<thead>
<tr>
<th>Sound</th>
<th>( D_t )</th>
<th>( I )</th>
<th>( E_{LA-NL} )</th>
<th>( EPNL_{t-NLE} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>S61</td>
<td>1</td>
<td>-2.9</td>
<td>3.9*</td>
<td>-5.1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-5.5</td>
<td>0.7</td>
<td>-7.5</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-3.9</td>
<td>1.2</td>
<td>-5.5</td>
</tr>
<tr>
<td>S64</td>
<td>1</td>
<td>-1.0</td>
<td>2.9*</td>
<td>-5.3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-4.7</td>
<td>0.9</td>
<td>-6.1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-4.4</td>
<td>0.9</td>
<td>-6.5</td>
</tr>
<tr>
<td>212</td>
<td>1</td>
<td>0.0</td>
<td>5.4*</td>
<td>-5.3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-1.3</td>
<td>3.7</td>
<td>-7.8</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.0</td>
<td>3.9</td>
<td>-7.9</td>
</tr>
<tr>
<td>47G</td>
<td>1</td>
<td>-1.4</td>
<td>4.7*</td>
<td>-7.5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-1.7</td>
<td>0.5</td>
<td>-5.8</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-2.3</td>
<td>0.8</td>
<td>-7.3</td>
</tr>
<tr>
<td>Gazelle</td>
<td>2</td>
<td>-2.2</td>
<td>2.7*</td>
<td>-5.5</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-3.8</td>
<td>0.8</td>
<td>-6.3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>-4.6</td>
<td>0.6</td>
<td>-6.0</td>
</tr>
<tr>
<td>Puma</td>
<td>1</td>
<td>-0.6</td>
<td>4.1*</td>
<td>-4.0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-3.4</td>
<td>1.5</td>
<td>-5.3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-2.7</td>
<td>0.9</td>
<td>-6.2</td>
</tr>
</tbody>
</table>

Mean Errors
(Standard deviations in parentheses):

6 More Impulsive Sounds(*)
-1.4 (1.1) 4.0 (1.0) -5.5 (1.1) -0.8 (0.8)

12 Less Impulsive Sounds
-3.2 (1.6) 1.4 (1.2) -6.5 (0.9) -1.1 (0.9)

\[ D_t = EPNL_t - PNL_t; \quad I = EPNL_{t-1} - EPNL_t \]
The magnitude of such a correlation cannot be measured by computing the direct correlation between duration and impulsiveness without first making allowance for the possibility of sampling bias (e.g., the more impulsive helicopters may have been flying more slowly and thus generating longer signal durations).

The approximate effects of both speed and distance (from the microphone) can, in fact, be eliminated using theory based on spherically symmetric source characteristics. It can readily be shown that in a non-dissipative medium, the duration correction for the sound exposure level of a spherically uniform source passing with speed \( V \) at a minimum distance \( S \) from an observer increases as

\[
10 \log_{10} \left( \frac{S}{V} \right)
\]

Differences between measured duration corrections \( D_t = (EPNL_t - PNL_t) \) not accounted for by this term may therefore be attributed to differences in source directivity. Thus, higher values of the duration increment

\[
\Delta = D_t - 10 \log_{10} \left( \frac{S}{V} \right)
\]

indicate increased sound radiation in forward and/or aft directions.\(^*\)

Figure 15 shows \( \Delta \) plotted against the average impulse correction \( I = (EPNL_{t1} - EPNL_t) \) for the 73 helicopter sounds for which values of \( S \) and \( V \) are known. A clear correlation between \( \Delta \) and \( I \) is apparent; the correlation coefficient is highly significant (\( p \approx 0.001 \)) at 0.62. This result is totally consistent with the fact that blade slope tends to exhibit pronounced forward directivity. Furthermore, the natural slope of the regression line (0.8) shows that due to impulsiveness, the incremental duration correction approaches the value of the ISO impulse correction.

\(^*\)It is recognized that this analysis involves an oversimplification of actual sound radiation mechanisms. For example, the actual signal durations are also affected by atmospheric sound dissipation which in turn depends upon distance and atmospheric conditions. This "excess attenuation" reduces the signal duration by an amount which increases with the minimum passby distance \( S \) of the source. However, in the present case, all but two of the relevant recordings were made at minimum distances no greater than 300 m and, since helicopter noise is dominated by low frequency sound which is less prone to dissipation than high frequency sound, this factor is considered to be of secondary importance; i.e., variations in \( \Delta \) are largely controlled by variations in fore/aft directivity.
Figure 15. Correlation Between Duration Increment $\Delta$, and ISO Impulse Correction Term $I = \text{EPNL}_t^i - \text{EPNL}_t$. 

\[ I = \text{EPNL}_t^i - \text{EPNL}_t \]
On the basis of the present results alone, it is not possible to state whether it is impulsiveness, low frequencies, duration (or indeed any other correlated variable which may have been overlooked) or some combinations of these which cause increased annoyance. However, if due to the weight of other evidence, the conventional duration allowance made by EPNL, is accepted a priori together with the standardized frequency weightings, then the results of this study indicate that there is no requirement to include further penalties for impulsiveness.
6.0 CONCLUSIONS

Approximately 140 individual helicopter flyover recordings were obtained via the members of ICAO Working Group B. Of these, 89 were of sufficient quality and sufficiently different to include in the study. This was rather less than the 200 or so originally hoped for and it was not possible to achieve the desired degree of independence between the variables of interest (duration, tones, impulsiveness, and frequency distribution). Thirty CTOL recordings, mostly of jet transport aircraft, were included for comparison, particularly to provide a standard of performance for EPNL₂* and other noise measurement scales.

The main experiment was performed using headphone presentation to the test subjects and the maximum sound levels of the 119 test sounds covered the range 69 to 93 dB(A). A large part of the experiment was duplicated three times using different subjects and different test conditions.

The test method was based on a rating scale procedure by which each sound was assigned an average annoyance score. This annoyance score was then transformed to an annoyance level defined as the sound level, in decibels, of a common reference sound effectively judged to be equally annoying. The merits of the various noise scaling procedures, including EPNL₂, were then assessed in terms of their ability to predict the measured variations in annoyance level between the test sounds.

The main experiment was intended to test the applicability of EPNL₂ to as wide a range of helicopter sounds as possible. An original objective of deducing the independent effects of specific underlying variables by multivariate analysis was only achieved to a limited extent due to an unavoidable degree of intercorrelation between the variables.

In the measurement and analysis of the acoustic variables, allowance was made for the frequency response of the test headphones but the impulsiveness correction factors could not be measured directly inside the headphones; instead, they were computed directly from the tape recordings. The true impulse corrections were therefore somewhat less than these nominal values.

The major conclusions drawn from the main experiment were as follows:

*The abbreviation EPNL₂ is used for the conventional Effective Perceived Noise Level scale used for aircraft noise certification purposes. The subscript t is used explicitly to denote the inclusion of tone corrections since the scale was used with and without these.
1. The Perceived Noise Level scale and the commonly used weighted sound level scales are equivalent in terms of their general ability to predict annoyance level for helicopters, for CTOLs, or for all sounds combined.

2. Conventional duration corrections (+3 dB per doubling of duration) improve the annoyance predicting performance of all the basic scales to which they were applied; duration is a highly significant contributor to judged annoyance.

3. On average, helicopter flyover sounds are judged equally annoying to CTOL sounds when their measured levels are approximately 2 dB higher on the time-integrated scales (EPNL_t, EL_A, etc.). In other words, at the same duration corrected levels, helicopters are less annoying than CTOLs.

4. Multiple regression analysis indicated that provided the helicopter/CTOL difference of about 2 dB is taken into account, the particular linear combination of level, duration, and tone corrections inherent in EPNL_t is close to optimum.

5. All scales of time-integrated sound level are very consistent predictors of CTOL noise annoyance levels; for these sounds, the variance of the prediction error is of the same magnitude as that of the estimated experimental error (around 1 dB).

6. All scales of time-integrated level predict the annoyance levels of helicopter noise significantly less consistently than those of CTOL noise. This is probably due to the wide range of acoustic characteristics exhibited by helicopters of different types.

7. The integrated ISO and crest factor impulse correction terms are very highly correlated and may be considered equivalent measures of impulsivity in helicopter noise.

8. Impulse corrections did not improve EPNL_t as a predictor of helicopter noise annoyance. A small but not significant reduction in error variance for the "more impulsive" sounds (defined by a nominal ISO correction of >4 dB) is more than offset by an increase in variance for "less impulsive" sounds. Furthermore, there is no significant difference between average annoyance levels of the more and less impulsive sounds when equated on any of the time-integrated scales. The impulse correction did not emerge as a significant predictor variable in the multiple regression analysis.
9. The reason that impulse corrections are not effective/not required is attributed to the fact that impulsiveness (a) increases the spectral level of helicopter noise in the frequency range 125-500 Hz, and (b) causes a significant increase in signal duration, which together adequately amplify the sound levels as measured on the conventional scales.

10. Notwithstanding conclusion I, which is based on the fairly large sample of different helicopter types, there is evidence that the averaging process (over all helicopters) masks significant differences between results for specific helicopter types. Four subgroups of helicopter sounds were classified on the basis of average spectrum shape and a comparison of the mean annoyance prediction errors for these showed clear improvements as emphasis was shifted from high frequencies to low in the sound level weighting functions (A, D, E, and "F"). This may be attributable in part to a correlation between impulsiveness and low frequency content. However, there is a strong likelihood that the conflicting conclusions of previous research into impulsiveness corrections have arisen because of such correlations when attention has been confined to a limited number of helicopter types (especially the Wessex, UH1, and OH58 helicopters).

11. It was found during preliminary experiments that the annoyance judgments of helicopter flyover sounds were unaffected by the long (up to 3 minutes) and very noticeable onset of the sound during the approach of a very impulsive helicopter (Bell 205). This was true even when subjects were specifically instructed to consider signal duration. Accordingly, the "approach component" was not included as a variable in the experiment.

Each of the duplicate experiments involved approximately three-quarters of the test sounds including all the CTOL sounds but only two-thirds of the helicopters. The first was conducted using headphones but with all sound levels nominally 15 dB higher. The second and third were performed simultaneously in the Exterior Effects Room (EER) and Interior Effects Room (IER) at the Langley Research Center using their standard loudspeaker sound replay facilities. All four experiments involved different test subjects.
There were two significant limitations to the Langley loudspeaker tests. In the IER, the signal levels were relatively low (maximum levels between 56 and 73 dB(A)) and the signal-to-background-noise level difference caused significant changes to the duration correction terms. The level range in the EER tests (70-90 dB(A)) was very close to that of the low level headphone tests but in both the IER and the EER, the sound generation systems effectively eliminated impulsiveness from the test sounds.

Taking account of these limitations, the results of all three duplicate experiments broadly agreed with those of the main experiment and thus lend strong support to the generality of the conclusions. In particular, the basic differences in the average annoyance levels of helicopter and CTOL noise was confirmed. Also, the fact that elimination of impulsiveness in the loudspeaker tests did not cause a significant difference to emerge between those subgroups of helicopter sounds which were previously classed as "more" and "less" impulsive, corroborates the conclusion that impulsiveness per se does not contribute more to annoyance than is explained by the increase in level and duration which it causes.

On the negative side, in all three duplicate experiments, the CTOL approach sounds were found to be typically 3 dB more annoying than CTOL takeoff sounds (as measured on the duration-corrected scales). No such difference was found in the main test and this anomaly, for which no plausible explanation can be offered, casts something of a shadow over what is otherwise a surprising consistency between headphone and loudspeaker tests performed with very different groups of over 150 test subjects in different countries.

The results of this study suggest that some previous studies of impulsiveness corrections for helicopter noise indices may have been confounded by interactions between frequency distribution, duration, and impulsiveness. Although this kind of multicollinearity could not be avoided here, the risky consequences of a limited selection of test signals have been minimized. It is concluded that for general prediction of the annoyance-evoking potential of helicopter noise which is not very different in character from that to which we are accustomed, the standard Effective Perceived Noise Level procedure is at least as good as other current noise measurement scales and does not require special provision to penalize impulsiveness. The presence of impulsiveness in a helicopter flyover sound increases both its level and duration to the extent that the increase in the measured time-integrated level accounts for consequent increase in annoyance potential.
This limited endorsement of EPNL is not intended to infer that it may be considered an ideal measurement scale for helicopter noise certification. Questions remain concerning the relative contributions of the underlying variables to annoyance and it was found that like other noise scales, EPNL is a less consistent predictor of noise annoyance for helicopters than for CTOLs. This is almost certainly due to the considerably wider variety in the various characteristics of helicopter noise which impose a more rigorous test of the noise scaling procedures. This alone points to potential weaknesses in the methodology but other findings reinforce the conclusion that more extensive research into helicopter noise impact is required if a truly equitable noise certification scheme is to be devised. In particular, it is disconcerting that the very long attention-arresting sound of an approaching, highly impulsive helicopter did not affect annoyance judgments in the present experiments. This suggests that in laboratory experiments of this kind, test subjects focus their attention upon the sound of the aircraft as it passes by, perhaps in an attempt to assess its total sound power output. The fact that the sound has a pronounced forward directivity may not influence such judgments. Yet the "hearsay" evidence of complainants near heliports and under helicopter flight routes indicates that the characteristically long audible duration of much helicopter noise is a particular source of aggravation. If this can be established as fact, perhaps by field survey research, the case will be made to develop improved techniques for laboratory study and, ultimately perhaps, to formulate a better concept for helicopter noise certification standards.
REFERENCES


11. ICAO Committee on Aircraft Noise, Report Document 9286 (CAN 6), Sixth Meeting, Montreal, 23 May to 7 June 1979.


REFERENCES (Continued)


APPENDIX A

Subjects’ Instructions

These tests are part of an investigation into the characteristics of aircraft noise which cause annoyance to people who live near airports. We would like you to judge how ANNOYING some aircraft and helicopter sounds are.

Through your headphones, you will hear recordings of various aircraft and helicopter sounds. The number of each sound will be announced before it begins. On your score sheet, you will find scales like the one below which you will use to record your judgment of each sound.

<table>
<thead>
<tr>
<th>Not Annoying At All</th>
<th>Extremely Annoying</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
</tbody>
</table>

After each sound there will be a break of a few seconds. During this interval, please indicate how annoying you consider 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 AT ALL ANNOYING end of the scale. On the other hand, if you judge a sound to be very annoying, then place your mark closer to the EXTREMELY ANNOYING end of the scale. A mark may be placed anywhere along the scale, not just at the numbered locations.

When making your judgment of each sound, consider how you would feel if you heard it at home on a number of occasions during the day and take into account all the characteristics of the sound. THERE ARE NO RIGHT OR WRONG ANSWERS; we are only seeking your personal opinions.
<table>
<thead>
<tr>
<th>Subject No.</th>
<th>Group</th>
<th>Session</th>
<th>Tape</th>
<th>Sound</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Not Annoying</td>
<td>Extremely Annoying</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>At All</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>Not Annoying</td>
<td>Extremely Annoying</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>At All</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>Not Annoying</td>
<td>Extremely Annoying</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>At All</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>Not Annoying</td>
<td>Extremely Annoying</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td>At All</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td>Not Annoying</td>
<td>Extremely Annoying</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td>At All</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td>Not Annoying</td>
<td>Extremely Annoying</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td>At All</td>
<td>8</td>
</tr>
</tbody>
</table>
A FURTHER SURVEY OF SOME EFFECTS OF AIRCRAFT NOISE IN RESIDENTIAL COMMUNITIES NEAR LONDON (HEATHROW) AIRPORT.

Extract from Loughborough University Report TT 7705 by J B Ollerhead and R M Edwards

This survey was part of a wider study of public reaction to aircraft noise sponsored by the Department of Trade. Dr Edwards was a member of the research team assigned to the project who performed much of the data analysis. Except for Appendix C however the report was the sole responsibility of the present author.

The extract includes the Questionnaire (Appendix A) and those parts of the report which deal with the relationships between the different noise and response variables. Although the introductory portions of the report are omitted the main objectives of the study are restated in Section 5 (p 62 et seq). Of particular interest were alternatives to the frequently used "Noise and Number Index", NNI, (defined on p 36) and Guttman Annoyance Score, GAS, (see Appendix C).

In fact, no pair of noise and reaction variables were found to be more highly correlated than NNI and GAS. Because the distributions of many of the variables were clearly non-gaussian, non-parametric statistical analysis was used throughout and the results were subsequently seen as important verification of the main results of earlier surveys. As it turned out, this particular pair of variables, GAS and NNI, exhibit a near perfect bi-variate normal distribution and it was concluded that no new noise scale is likely to emerge to provide a more reliable or convenient predictor of aircraft noise annoyance than NNI.

* Civil Aviation Authority 1981, "The Noise and Number Index", DORA Communication 7907.
NOTATION

Symbols not listed below, are used only locally in the text and are defined where they first occur.

ABC Aircraft Noise Bother Coefficient = ABS / average bother score (Q.18)
ABS Aircraft noise bother score (Q.18b)
ANA Aircraft Noise Annoyance (Q.20)
AS4 4-point aircraft noise bother score (Q.18b, sec.4.4)
CNR Composite Noise Rating
\( \bar{D} \) Average duration of aircraft sounds (seconds)
DDD,DED,DND Perceived duration of daytime, evening and night-time aircraft sounds (Q's 27,30,33) - (seconds)
L Sound or noise level (dB*)
\( L_{10}, L_{10}, L_{50} \) - (or \( L_{10}, L_{10}, L_{50} \)) - percentiles of the noise level distribution. (dB*)
\( L_{eq} \) Equivalent continuous sound level (dB*)
\( L_{NP} \) Noise Pollution Level (dB*)
\( \Sigma \) Average noise level \( ( = \log_{10} \frac{1}{N} \sum 10 \frac{L_i}{10} ) \) (dB*)
N Number of aircraft sounds
n Number of samples (data points)
NCI Ratio NCO/Income (Q.14)
NCO Noise cost (£) = \( \frac{1}{2} \) (contribution (Q.45) - compensation) (Q.44)
NCR Ratio NCO/rent (Q.11)
NDD,NED,NND Reported number of daytime, evening, night-time disturbances (Q's 26, 29, 32)
NEF Noise Exposure Forecast
NNI Noise and Number Index (When used with no subscript, or with subscript M this refers to 'Standard' form of index)
P Sample Pearson Correlation Coefficient
R Sample Spearman (rank order) correlation coefficient.
r Population Spearman coefficient
F Population Spearman coefficient assumed for purposes of statistical inference

mostly PNdB herein
WECPNL: Weighted Equivalent Continuous Perceived Noise Level.

$\Delta R$: Differences between two R values.

$\lambda$: Threshold Noise Level (dB$^*$).

$\rho$: Population Pearson correlation coefficient.

$\sigma$: Standard deviation of noise level fluctuations.

**SUBSCRIPTS**

- **A**: Average Mode
- **D**: Daytime (0700-1900hrs)
- **E**: Evening (1900-2300hrs)
- **M**: Mode, A or W
- **N**: Night-time (2300-0700hrs)
- **P**: Time period, D, E or N
- **W**: Worst Mode
- **$\lambda$**: Threshold Level

**e.g.** $\text{NNI}_{P}^{M}$: general definition of $\text{NNI}$

$\text{NNI}_{E70W}$: $\text{NNI}$ computed for the evening period (1900-2300hrs), worst mode conditions, accounting for all aircraft sounds in excess of 70 PNdB.
4. ASSOCIATION BETWEEN VARIABLES

4.1 Approach

A large number of response variables were generated from the questionnaire answers and, as described in the next section, noise exposure variables were computed from aircraft movement data. Abbreviations used to denote many of these are listed under Notation. A study of the interrelations between these many variables, for each of which there were up to 600 values, presented a formidable task, not only in terms of the computations involved but also of interpretation.

No attempt has been made to analyse the complex relationships which may exist between several mutually dependent variables. This has been attempted on several previous occasions, for example the multiple regression of annoyance upon such noise variables as level, number of events and duration, in the 1967 Heathrow Survey (5). Although multivariate analysis clearly fulfils an important role in survey studies it was felt that because of the many uncertainties associated with the application of existing methods to survey data which meets few of the requirements of parametric statistics, it would be premature to divert effort to such nebulous exercises in this pilot study. Emphasis was rather directed towards a determination of the direct association between pairs of variables, a problem which itself is not easily resolved, as will be seen.

Probably the most convenient parameter which expresses the association between two variables is the Pearson product-moment correlation coefficient \( \rho \) (see Appendix D). The computations involved are very straightforward and the square of this coefficient gives directly the fraction of the variance of one (dependent) variable which may be attributed to the other (independent variable) although it should not necessarily be inferred that finite correlation signifies a causal relationship. Thus a correlation coefficient of unity implies that one is a unique linear function of the other. Probably because
of these attractions, estimates of $p$ are widely used for the analysis of survey data.

However, considerable care should be exercised in using and interpreting estimates of this coefficient. The distribution of the sample correlation coefficient $P$ is only known for the special case of a bivariate normal distribution and, therefore, only in this case can the usual statistical inferences be made. For present purposes these should include the inference of (a) significant correlations between variables and (b) significant differences between correlation coefficients.

As the results of Section 3 clearly show, many of the sample distributions in this survey are far from normal and the relationships between variables are far from linear. Furthermore, many variables are not continuous in the sense that they can only take certain discrete values. In the case of the noise exposure variables this is not a significant constraint since the number of values is large. However, for most response variables the number of possible values is small, typically less than ten, so that the data must be regarded as extensively 'tied'. Any assumption of normality in such cases is clearly risky.

We should therefore turn to the methods of non-parametric statistical analysis which do not involve assumptions about the distributions of the variables. One non-parametric equivalent of $p$ is Spearman's Rank Correlation Coefficient $r$ (see Appendix D). This is obtained by calculating $p$ after replacing each observation by an integral number.

For each observation this represents its rank position when all observations of that variable are arranged in order of magnitude. Thus the smallest will be 1, the next smallest 2 and so on. Spearman's Coefficient thus measures the degree of correspondence between rankings instead of between actual variate values. Its value which, like that of $p$, can vary between $+1$ and $-1$ is a measure of the probability that large (or small) values of one variable are associated with large (or small) values of the other. Its value is independent of all order-preserving transformations of either variable and thus of their distributions. It is also valid for tied data. For uncorrelated
random variables the distribution of R, the sample rank correlation coefficient for large samples is gaussian with known variance so that a routine test can be made for the significance of any R.

However, although the sampling distribution of R for correlated variables also tends to normality, the associated variance is not independent of the population distributions and it is only conveniently possible to define maximum values of this variance. A practical limitation on the use of R is thus that the confidence with which the population parameter r can be estimated is relatively low. Consequently significance can only be attached to relatively large differences between values of R. A further difficulty is that the validity of these tests for tied data is not known (20).

Despite these problems it was felt that rank correlation methods provide the best available estimates of the association between two variables for present purposes. The analysis presented subsequently should simply be interpreted with these limitations in mind. Spearman's coefficients are discussed more fully in Appendix D.

Because of these limitations it was considered most inadvisable to accept the results of the correlation analysis without some knowledge of the bivariate distributions of the samples. Accordingly a computer program was developed to list and plot these distributions in each case. The range of each 'continuous' variable was divided into appropriate class intervals and the sample size for each cell of the bivariate matrix of interest (against x) printed as an absolute value and as a percentage expressed both by row and by column. (Note that many questionnaire data were coded into class intervals by the interviewer). A simple graphical presentation of the same results was obtained by plotting the 25th, 50th and 75 percentiles of the distribution of y against x and vice-versa. The percentile values are calculated on the assumption that each subsample of observations is uniformly distributed across the dimensions of the cell. Thus even in cases where a variable takes discrete integral values (e.g. annoyance on the scale 0(1)6) percentiles are computed with non-integral values.
Although the rank correlation coefficients provide a convenient measure of the association between pairs of variables it gives no clue to any functional relationships which exist. These are frequently estimated by regression analysis or curve fitting to minimise least squares errors. Since, in many instances, the variables of this study are measured on unequal interval scales (particularly in the case of many subjective response scales where the interval cannot be given dimensions) the meaning of arithmetic means and the variances becomes obscure. Accordingly, it is considered more appropriate to define central tendencies in terms of the median and to illustrate functional relationships between variables by plotting the variation of the median of the y-distribution against x. Although, as calculated, these median values involve linear interpolation between scale points (or across class intervals) the biases are considerably less than those involved in numerical averaging of all y values.

4.2 Noise Estimates

The procedures used to estimate noise exposure levels around the airport are fully described in Reference (15) and only the broad details will be repeated here.

The various noise exposure parameters were calculated from known aircraft movements over a period of four weeks prior to the date of the last interviews. A computer programme was developed to plot contours of equal noise exposure and to calculate levels at individual interview locations which were designated by 8-digit ordinance survey map co-ordinates. Input to the programme included estimated aircraft route ground tracks, the noise radiation patterns and climb performance characteristics as a function of destination range for the various aircraft types or categories, sound propagation parameters and a breakdown of aircraft traffic on each route during different periods of the day and night.

Basic Variables

In the limit the noise exposure at any point can only be
completely defined by a time history of noise level (in dBA or PNdB for example) over the entire four-week period, a clearly impractical procedure. A variety of steps can be taken to condense such information into a more tractable form and numerous existing noise exposure indices make use of some of them. Since it was desirable not to prejudge the relative merits of these techniques, noise data were generated in as general a form as possible. The following basic parameters describing outdoor noise levels were computed for each location:

\[ \log_{10}(\text{peak level}) = \lambda \]  
\[ N_{P,M} \] the number of individual aircraft sounds which exceed level \( \lambda \)  
\[ D_{P,M} \] the average duration for which the \( N_{P,M} \) sounds exceeded level \( \lambda \) (seconds)

The 'peak' level is defined as the maximum level reached during the aircraft passby event.

The suffix \( P \) refers to the diurnal period and takes one of three values, \( D \) (daytime 0700-1900), \( E \) (evening 1900-2300) or \( N \) (night-time 2300-0700).

The 'threshold' level \( \lambda \) is effectively the level below which individual sounds are ignored so that \( N_{P,M} \) decreases as \( \lambda \) increases. It was varied by 5PNdB steps in the range 65 to 90PNdB.

The suffix \( M \) defines one of two operational conditions for which the estimates were made. Average mode (A) results were obtained on the basis of the four-week average traffic on each route for the appropriate period \( P \). Worst mode (W) results represent the maximum noise exposure conditions experienced on any one day during the four-week interval. It is important to note here that of the two, worst mode results are (a) considerably more difficult to compute and (b) more error-prone. The computational problem is that whereas average mode values can be obtained from one computer run, worst mode conditions involve
different traffic route assignments for different field locations and can therefore require a large number of separate computer runs. Increased confidence intervals must also arise because of aircraft position errors and variations in sound propagation characteristics, which are not accounted for in the computations. In the average mode case such errors will be reduced by the averaging process.

There are obvious limitations to the accuracy with which actual aircraft noise exposure parameters can be predicted in this way. These may be summarised as follows:

(i) Aircraft do not follow nominal ground tracks (Minimum Noise Routes and Standard Instrument Departures) with any great degree of precision. Individual deviations can be large.

(ii) For the purposes of practical computation certain aircraft types and variants of those types must be grouped into a limited number of categories.

(iii) Even identical aircraft operating under the same nominal configuration generate different noise levels; differences in payload, fuel load, power setting and air speed increase the variations.

(iv) Sound propagation characteristics can vary significantly from day to day. The most important factor is the variation of the speed of sound with altitude which is influenced by wind and temperature gradients; this exerts a major influence upon sound propagating at low angles of elevation.

The combined effects of these sources of error are extremely difficult to estimate but it is clear that the accuracy of all predicted noise parameters decreases with increasing distances from the flight routes, that is as noise exposure decreases.

Derived Variables

From the basic parameters $L_{PA}^M$, $N_{PA}^M$ and $D_{PA}^M$ the noise
levels at the interview sites have also been calculated on the following scales:

Noise and Number Index \( \text{NNI}_{P\lambda M} \) given by

\[
\text{NNI}_{P\lambda M} = \bar{L}_{P\lambda M} + 15 \log_{10} N_{P\lambda M} - 80
\]  

(1)

The 'standard' version of \( \text{NNI}_M \) (denoted here by omitting suffices \( P \) and \( \lambda \)) is simply one member of this family, i.e.

\[
\text{NNI}_M = \bar{L}_{\text{D80M}} + 15 \log_{10} N_{\text{D80M}} - 80
\]  

(2)

Note that the constant term is 80 for all \( \lambda \). Since this constant is essentially arbitrary (although originally selected by the Wilson Committee (16) to define zero NNI as that level where 'average' annoyance was zero) this is done so that changes in \( \text{NNI}_{P\lambda M} \) are due solely to variations in \( \lambda \).

The levels exceeded for 1%, 10% and 50% of the time

\( L_{1\text{PM}}, L_{10\text{PM}} \) and \( L_{50\text{PM}} \)

These variables are of course independent of \( \lambda \).

Equivalent continuous sound level \( \text{Leq}_{PM} \)

This is rigorously defined by the equation

\[
\text{Leq}_{PM} = 10 \log_{10} \left[ \frac{1}{T_P} \int_0^{T_P} L_M/10 \, dt \right] \tag{3}
\]

where \( T_P \) is the duration of the period \( P \) and \( L_M \) is the instantaneous level in \( \text{PNdB} \). This variable is also independent of \( \lambda \).

Noise Pollution Level \( \text{LN}_{\text{NP}_{PM}} \)

defined as

\[
\text{LN}_{\text{NP}_{PM}} = \text{Leq}_{PM} + 2.56 \sigma
\]  

(4)
where $\sigma$ is the standard deviation of the fluctuations of level $L_M$ during the period $P$.

Although $NNI_P \lambda M$ may be computed according to Equation (1) as written, the remaining parameters have been derived by an approximate method, describing in detail in Reference (15). In most cases, the distribution of $D_\lambda$ (as a function of $\lambda$) is approximately normal, particularly towards higher values of $\lambda$. Accordingly, the computed distribution for each location was approximated by the best fitting Gaussian curve from which the parameters $L_1$, $L_{10}$, $L_{50}$, $L_{eq}$ and $L_{NP}$ could be determined directly. (Note that for normally distributed sound levels $L_{eq} = L_{50} + \sigma^2/8.68$ where $\sigma$ is the standard deviation.)

It is most important to note that these computations make no allowance for the background noise level for which no reliable estimates are available. Thus, although these various scales are frequently used for measuring total environmental noise exposure they should be considered here to apply only to the aircraft noise component. At high aircraft noise levels (i.e. high $\overline{L}$ and $N$) the values will closely approximate the true overall values (which will be dominated by aircraft noise). However, as $\overline{L}$ and $N$ decrease at positions more remote from the flight routes the computed values may be considerably less than actual levels which might in reality be controlled by noise sources other than aircraft.

**Relationships between Noise Variables**

It is clearly impractical to examine the correlations between all pairs of variables out of several hundred computed. However, because of the basic association between numbers of aircraft heard and the average levels we may be quite certain that all correlations will be high. Table 2 lists the rank correlation coefficients between all pairs selected from nine basic variables for average mode operations during the daytime period.
## TABLE 2  
**Rank Intercorrelations Between Various Daytime/Average Mode Noise Exposure Variables (Sample = 600)**

<table>
<thead>
<tr>
<th></th>
<th>L_1</th>
<th>L_{10}</th>
<th>L_{50}</th>
<th>L_{eq}</th>
<th>L_{NP}</th>
<th>L_{80}</th>
<th>N_{80}</th>
<th>D_{80}</th>
<th>R</th>
<th>No.of Sig. Diffs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NNI_{80}</td>
<td>.968</td>
<td>.952</td>
<td>.832</td>
<td>.951</td>
<td>.877</td>
<td>.961</td>
<td>.936</td>
<td>.952</td>
<td>.929</td>
<td>1</td>
</tr>
<tr>
<td>D_{80}</td>
<td>.919</td>
<td>.986</td>
<td>.916</td>
<td>.885</td>
<td>.772</td>
<td>.858</td>
<td>.971</td>
<td>.907</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>N_{80}</td>
<td>.885</td>
<td>.975</td>
<td>.945</td>
<td>.842</td>
<td>.726</td>
<td>.820</td>
<td></td>
<td>.888</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>L_{80}</td>
<td>.951</td>
<td>.859</td>
<td>.683</td>
<td>.960</td>
<td>.945</td>
<td></td>
<td></td>
<td>.880</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>L_{NP}</td>
<td>.933</td>
<td>.787</td>
<td>.577</td>
<td>.957</td>
<td></td>
<td></td>
<td></td>
<td>.821</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>L_{eq}</td>
<td>.991</td>
<td>.991</td>
<td>.897</td>
<td>.719</td>
<td></td>
<td></td>
<td></td>
<td>.900</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>L_{50}</td>
<td>.778</td>
<td>.931</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.798</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>L_{10}</td>
<td>.933</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.915</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>L_{1}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.920</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

* Average correlation coefficient with other variables

Out of the 36 values, 20 exceed 0.9 and only 7 are less than 0.8.

Appendix D shows that for two samples of n = 600, and an average population correlation between 0.8 and 0.9 a difference in sample coefficients of approximately 0.1 is significant at the 5% level. Thus, the difference between the highest and the 20th highest coefficients in Table 2 is not significant. The average correlation coefficient $\bar{R}$ is the arithmetic average of the eight values associated with each variable. The number from each eight which are significantly different from the highest is also tabulated. Thus, we see that at the 5% level of significance NNI_{80}, L_{10} and L_{1} are indistinguishable from 7 of the remaining variables in terms of the order in which they rank the noise exposures. At the other extreme, in the case of L_{50}, which has a significantly smaller average coefficient, 5 of the coefficients are significantly
lower than the highest.

Obviously these observations result from the particular variables selected, the high correlations being connected with those which tend to emphasise peak levels. The importance of these correlations will become evident later.

Numbers and average levels are highly correlated because aircraft converge near the airport. Thus, locations where high numbers are heard are also exposed to high average levels. This is reflected in Table 3 which lists the distributions of $L_{D80}$ and $N_{D80}$ for the 600 interview locations. Both average and, in brackets, worst mode results are included. The clustering of cells about a diagonal line is apparent, as is the shift to higher numbers and levels in the worst mode case. The subsample of 151 for $N<10$ and $L<80$, which is unfortunately large, results from the use of available 1967 noise exposure contours for the sample selection. Probably because the methods of Reference (15) used in this study involve higher excess sound attenuation rates for small air-to-ground propagation angles than the previous techniques, noise exposures to the side of the runways are somewhat lower. Consequently, noise levels in those regions to the north of the runways which were estimated from earlier maps to lie in the range 20 to 35NNI were computed later to be less than 20NNI and in many instances to be less than zero NNI, i.e. no levels exceeded 80PNdB. Thus, as noted in Section 3.1, the distribution of NNI values at the low end of the range was very poor.

The value of N is highly dependent upon the threshold level above which it is counted. Table 4 shows the distribution of $N_{D80A}$ as $\lambda$ varies from 65 to 90PNdB. Worst mode values are given in brackets.

The high correlations between $\bar{L}$, $N$ and NNI ($D80A$) are clearly illustrated in Figure 9. The three lines in each graph, as in many subsequent figures, represent the 25th, 50th and 75th percentiles of the distribution of the ordinate variable in each interval of the abscissa. Class intervals in all cases have been selected to make the cell matrix as square as possible. The figures in italics give the size of the subsample in each abscissal interval.

The correlation between $\bar{L}$ and NNI is very high ($R = 0.961$).
TABLE 3  
**Joint Distribution of Levels and Numbers of Events**
(daytime) 80PNdB Threshold, Average (and Worst) Mode

<table>
<thead>
<tr>
<th>LI_{D80A} (PNdB)</th>
<th>N_{D80A}</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 80</td>
<td>151(151)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80-83</td>
<td>- -</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>83-86</td>
<td>1 -</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>86-89</td>
<td>- -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>89-92</td>
<td>- -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>92-95</td>
<td>- -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95-98</td>
<td>- -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98-101</td>
<td>- -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>101-104</td>
<td>- -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>104-107</td>
<td>- -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>107-110</td>
<td>- -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 110</td>
<td>- -</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>152(151)</td>
<td>11(0)</td>
<td>0(0)</td>
<td>142(1)</td>
<td>153(160)</td>
<td>103(235)</td>
<td>39(53)</td>
<td>600(600)</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 4  
**Effect of Threshold Level λ on Distribution of No. of Daytime Events for Average (Worst) Mode**
(Total Sample = 600)

<table>
<thead>
<tr>
<th>ND′A</th>
<th>THRESHOLD LEVEL λ - PNdB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>65</td>
</tr>
<tr>
<td>2.5</td>
<td>48</td>
</tr>
<tr>
<td>2.5-5</td>
<td>9</td>
</tr>
<tr>
<td>5-10</td>
<td>15</td>
</tr>
<tr>
<td>10-20</td>
<td>23</td>
</tr>
<tr>
<td>20-40</td>
<td>47</td>
</tr>
<tr>
<td>40-80</td>
<td>41</td>
</tr>
<tr>
<td>80-160</td>
<td>170</td>
</tr>
<tr>
<td>160-320</td>
<td>195</td>
</tr>
<tr>
<td>320</td>
<td>52</td>
</tr>
</tbody>
</table>
with a typical interquartile range of 3PNdB. For a normal distribution this would represent a standard deviation of about 2PNdB. The correlation between $N$ and NNI is slightly smaller (0.936) although still high in absolute terms. The relations between $\bar{L}$, $N$ and NNI are clearly fairly linear but this is not so in the case of $\bar{D}$, the average duration in excess of the threshold (80PNdB), and NNI. Although $R$ remains high, duration is practically invariant with NNI above 20NNI.

It should be noted that in Figure 9, 151 values which were listed in the computer as zero NNI, but in reality were less than zero, have been omitted. However, these same 'zero-zero' values have been included in the rank correlations, which are therefore somewhat higher than if they had been excluded (see Appendix D).

Figure 10 shows that the relation between $L_{eq}$, $L_{NP}$ and NNI are fairly linear and that correlations are high. In this case the plots are based on a sample of 496 which excludes 104 cases where the average durations were too small to allow accurate estimation of the duration-based parameters $L_{eq}$ and $L_{NP}$. However, in order to compare correlation coefficients with others based on 600 points, 104 zero-zero's were included in the $R$-computations. An important point which is illustrated by Figures 9 and 10 is that although $L_{NP}$ and NNI intervals are roughly equal (i.e. a 45° slope) the level scales $\bar{L}$ and $L_{eq}$ are condensed, yielding slopes of roughly 0.6 when plotted against NNI or $L_{NP}$.

Diurnal Variation of Noise Exposure

At Heathrow approximately 80% of aircraft movements occur during the daytime period. Consequently, the evening and nighttime noise levels are rather lower than daytime values. From Figure 11, the relations between median NNI values are:

$$NNI_{E80A} \approx NNI_{D80A} - 13 \quad (R = 0.986)$$

$$NNI_{N80A} \approx NNI_{D80A} - 17 \quad (R = 0.950)$$

The correlation between $NNI_{E80A}$ and $NNI_{N80A}$ is also extremely high at $R = 0.940$. 

-41-
TABLE 5 Correlation Between Disturbances and Noise Parameters
(All Coeffs. Significant at 1% Level Unless Indicated)

<table>
<thead>
<tr>
<th>Disturbance Variable</th>
<th>Sample Size</th>
<th>Noise Variable</th>
<th>Threshold Level λ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>65</td>
</tr>
<tr>
<td>NDD</td>
<td>365</td>
<td>$\bar{D} \Lambda_{\omega}$</td>
<td>.479</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\bar{D} \Lambda_{\alpha}$</td>
<td>.491</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_{D} \Lambda_{\omega}$</td>
<td>.448</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_{D} \Lambda_{\alpha}$</td>
<td>.423</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_{NI}D \Lambda_{\omega}$</td>
<td>.490</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_{NI}D \Lambda_{\alpha}$</td>
<td>.480</td>
</tr>
<tr>
<td>NED</td>
<td>374</td>
<td>$\bar{E} \Lambda_{\omega}$</td>
<td>.500</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\bar{E} \Lambda_{\alpha}$</td>
<td>.536</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_{E} \Lambda_{\omega}$</td>
<td>.372</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_{E} \Lambda_{\alpha}$</td>
<td>.426</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_{NI}E \Lambda_{\omega}$</td>
<td>.504</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_{NI}E \Lambda_{\alpha}$</td>
<td>.497</td>
</tr>
<tr>
<td>NND</td>
<td>377</td>
<td>$\bar{N} \Lambda_{\omega}$</td>
<td>.211</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\bar{N} \Lambda_{\alpha}$</td>
<td>.173</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_{N} \Lambda_{\omega}$</td>
<td>.149</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_{N} \Lambda_{\alpha}$</td>
<td>.185</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_{NI}N \Lambda_{\omega}$</td>
<td>.207</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_{NI}N \Lambda_{\alpha}$</td>
<td>.205</td>
</tr>
<tr>
<td>DDD</td>
<td>308</td>
<td>$\bar{D} \Lambda_{\omega}$</td>
<td>.064*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$D \Lambda_{\alpha}$</td>
<td>.089*</td>
</tr>
<tr>
<td>DED</td>
<td>300</td>
<td>$\bar{E} \Lambda_{\omega}$</td>
<td>.126+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$E \Lambda_{\alpha}$</td>
<td>.147+</td>
</tr>
<tr>
<td>DND</td>
<td>148</td>
<td>$\bar{N} \Lambda_{\omega}$</td>
<td>-.196+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N \Lambda_{\alpha}$</td>
<td>.080*</td>
</tr>
</tbody>
</table>

* significant at 5%;  • not significant at 5%
4.3 Relations Between Disturbance and Noise

Responses to questions 26 through 34 yield perceived frequencies and durations of aircraft noise disturbance during each of the three periods daytime (0700-1900), evening (1900-2300) and night-time (2300-0700). No special meaning of the word 'disturbed' was elucidated to the respondents at Q.26 (which required an estimate of the number of disturbances caused during the daytime, when aircraft noise was "at its worst"); however, at questions 27 and 30 concerning the average duration of daytime and evening disturbances, it was suggested that disturbance implied distraction or interruption. The similar question 33, pertaining to night-time disturbance referred to sleep interference. The questions were administered to some 360 respondents, the remainder having been excluded by prior indications that they were at no time affected by aircraft noise.

As discussed in the Introduction, it was anticipated from the outset that such disturbance variables might be well-correlated with appropriate noise exposure variables, obvious associations being between (a) perceived numbers of disturbances and (estimated) actual numbers of events and (b) between perceived disturbance durations and average signal durations. The computed numbers of events and signal durations depend upon the threshold level below which sound is ignored, with both numbers and durations diminishing as \( \lambda \) increases. The sample rank correlation coefficients \( R \) which measure these associations are listed in Table 5 for \( \lambda \) between 65 and 90PNdB.

Number of Disturbances and Number of Events

The correlations between the reported number of disturbances \( N_{DD} \) and \( N_{ED} \) and the corresponding numbers of events \( N_{D\lambda M} \) and \( N_{E\lambda M} \) for daytime and evening periods are significant at the 0.1% level for all \( \lambda \) and both worst and average modes. For night-time, the correlations between \( N_{ND} \) and \( N_{N\lambda M} \) are significant at the 1% level and in some cases at 0.1%. However, the correlations for day and evening are significantly higher.
(at the 5% level) than those for night-time disturbances. The differences between day and evening and between worst and average mode are not significant.

With regard to the effect of changing the threshold level there is a tendency for \( R \) to increase with increasing \( \lambda \). However, the variation of \( R \) with \( \lambda \) is relatively small and generally not significant according to the tests described in Appendix D.

The distributions of NDD, NED and NND as functions of the appropriate noise variables for \( \lambda = 80 \)PNdB are illustrated in Figure 12, where the large scatter implied by the relatively low values of \( R \) is clearly evident. Here as elsewhere, the non-Gaussian features of the distributions are apparent.

The effect of \( \lambda \) on the median distributions (denoted by NDD\(_{50}\), NED\(_{50}\) and NND\(_{50}\)) may be seen in Figure 13. Since higher values of \( \lambda \) lead to smaller values of \( N_{p} \lambda M \) there is a tendency for the curve to move to the left with increasing \( \lambda \). This effect is hardly noticeable for the night-time figures since the slope of the curve is so low.

**Number of Disturbances and NNI**

In the associations between number of disturbances and number of events examined so far, the effect of varying noise levels has been considered to some extent through the threshold level \( \lambda \). Although the changes of \( R \) with \( \lambda \) are not significant the results show that for this sample disturbance is more highly related to the more intense sounds. It is thus of interest to examine the relationship between disturbance frequency and Noise and Number Index which accounts for both intensity and frequency of exposure.

The correlations are listed in Table 5. They are higher, but not significantly so, than those obtained for event numbers. Of particular interest, however, is that \( R \) is now almost independent of \( \lambda \). This can probably be explained by the fact that as the threshold \( \lambda \) increases the level term \( \bar{L} \) also increases (due to the elimination of contributions from levels below \( \lambda \)) tending to cancel decreases in the number term \( 15 \log N \) in the NNI expression. Again, there is no significant difference between results for average and worst mode noise parameters.
The variations of disturbance frequencies with NNI are plotted in Figure 14 which is of course very similar to Figure 12.

**Relationship Between Day, Evening and Night Disturbance Frequencies**

The observations made in Section 3.2 concerning the relative intrusiveness of aircraft noise during daytime, evening and night-time periods may now be extended. Figure 15 compares the median number of disturbances (from Figures 13 and 14) as functions of (a) number of events in excess of 80PNdB and (b) NNI for each of the three periods. Both noise parameters relate to average mode conditions.

The broken line in Figure 15 (a) roughly equates number of disturbances and number of events. Comparing the experimental curves with this line it may be seen that during the daytime the median respondent is disturbed by roughly 1 in 30 aircraft sounds which exceed 80PNdB. During the evening he is roughly three times more sensitive, being disturbed by approximately 1 out of 10 sounds. Perhaps the most surprising feature of Figure 15 is the night-time curve which indicates that the median respondent is extremely insensitive to night-time noise, being disturbed by a very small fraction of sounds, typically of the order 1/100. Also, whereas for day and evening, number of disturbances is roughly proportional to number of events, this does not appear to be true for night disturbances. However, it is very probable that the slope of the line is severely constrained by the limited choice of options available on Card F for those people disturbed very infrequently. Thus, as a rough guide, it appears that median noise sensitivity, expressed in terms of the fraction of sounds which cause disturbance, varies for day: evening: night in the proportion 3:10:1.

However, the very large differences between people must not be overlooked. For example, although the median respondent is very insensitive to night-time noise Figure 15 shows that 25% of people are disturbed more than once to twice a night at the highest noise levels.

Both the variations in noise exposure and the variations between people are conveniently accounted for in Figure 16 which
shows the percentage of people $y$, who during daytime are disturbed by more than $x$ percent sounds in excess of 90PNdB (average mode). Plotted on a log-linear scale this data points collapse to a straight line giving the relationship

$$y = 89 - 42.5 \log_{10} x$$

Thus 4% of respondents are disturbed by all the sounds, 46.5% by more than 10% of the sounds and 89% by more than 1% of the sounds which exceed 90PNdB. However, although similar relationships can be established for evening and night and for different threshold levels the data collapses in the other cases are not so impressive and the simplicity of Equation (5) may be fortuitous.

**Number of Disturbances and Other Noise Variables**

The sample correlation coefficient $R$ relating the number of daytime disturbances to other noise exposure variables including $L_1$, $L_{10}$, $L_{50}$, $L_{eq}$ and $L_{NP}$ are listed in Table 6. Only $L_{50}$ (average mode) yields a coefficient which is significantly lower than some of the others and then only at around the 20% level of significance.

**TABLE 6 Correlation Between Number of Disturbances and Various Noise Exposure Scales**

<table>
<thead>
<tr>
<th>Disturbance Variable</th>
<th>Sample Size</th>
<th>Mode (M)</th>
<th>Noise Variable</th>
<th>$L_1$</th>
<th>$L_{10}$</th>
<th>$L_{50}$</th>
<th>$L_{eq}$</th>
<th>$L_{NP}$</th>
<th>NDD90M</th>
<th>NNTD80M</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDD</td>
<td>365</td>
<td>Worst</td>
<td></td>
<td>.483</td>
<td>.504</td>
<td>.441</td>
<td>.480</td>
<td>.461</td>
<td>.485</td>
<td>.487</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Avge.</td>
<td></td>
<td>.489</td>
<td>.411</td>
<td>.313</td>
<td>.495</td>
<td>.442</td>
<td>.472</td>
<td>.476</td>
</tr>
</tbody>
</table>

**Duration of Disturbance**

The correlations between reported disturbance durations for day, evening and night and $D_{DA\ M}$ are listed in Table 5, it being assumed that the physical duration parameters for evening and night
are very similar to the daytime values. (Due to computational difficulties evening and night-time values of $D$, $L_1$, $L_{10}$, $L_{50}$, $L_{eq}$ and $L_{NP}$ were not available.) In general the coefficients are all low, only a few being significant at the 0.1% level. For all three periods there is a tendency for $R$ to increase with $\lambda$ up to $\lambda = 80$PNdB. For $\lambda > 80$PNdB the correlations are not significant at the 5% level with the exception of $R = 0.215$ for DND v. $D_{D85A}$ which is significant at the 1% level. Again the differences between the worst mode and average mode results are not significant at the 5% level.

The distributions of the disturbance durations are plotted against $D_{D80A}$ in Figure 17. The graphs for day and evening are very similar but although a much larger number of people report no disturbance at night (compared with those who are never disturbed during the day and/or evening) those who are disturbed claim that the disturbances last longer at least at the higher signal durations. This is clarified in Figure 18 (a) which compares the median disturbance times for the three periods.

The correlations between disturbance durations for the three periods are

\[
\begin{align*}
\text{DED v DDD} & \quad R = 0.904 \quad n = 273 \\
\text{DND v DDD} & \quad R = 0.546 \quad n = 132 \\
\text{DND v DED} & \quad R = 0.458 \quad n = 134
\end{align*}
\]

The very high correlation between DED and DDD suggests that the disturbance mechanisms for day and evening are very similar, i.e. that similar activities are disturbed leading to very similar duration estimates. The correlations of both DDD and DED with DND are very significantly lower indicating a dissimilar disturbance mechanism at night. This result is not surprising of course - but it does serve to support the validity of the responses to the disturbance questions.

The reasons for the sharp reduction in $R$ for $\lambda > 80$PNdB is somewhat obscure. The results of Table 5 suggest that disturbance durations are, at least for this sample, most closely related to signal durations in excess of 80PNdB. However, Figure 18 (a) shows that median disturbances last for rather less
than $D_{80}$, in turn indicating that a higher threshold is appropriate. Figure 18 (b) shows the effect of $\lambda$ on the median daytime disturbance duration. As might be expected this indeed approaches $D_{D\lambda A}$ as $\lambda$ increases. The most probable explanation for the low correlations for $\lambda = 85$ and 90 is that in a large proportion of cases the peak levels do not reach these thresholds (i.e. $D = 0$). The ability of respondents to accurately estimate durations may be assessed from their responses to Q.35 concerning the actual duration of the signal (RAD). The correlations between RAD and $D_{D\lambda A}$ are as follows:

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>65</th>
<th>70</th>
<th>75</th>
<th>80</th>
<th>85</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>.218</td>
<td>.241</td>
<td>.265</td>
<td>.315</td>
<td>.261</td>
<td>.165</td>
</tr>
</tbody>
</table>

The sample size is 363 for which correlation coefficients in excess of 0.17 are significant at the 0.1% level. Figure 19 shows that perceived and physical durations are very similar for $\lambda = 70$ PNdB, i.e. the typical respondent is reporting the duration of the aircraft sound in excess of this threshold.

4.4 Relations Between Indirect Response and Noise

Aircraft noise annoyance is defined here as an indirect effect of the noise, i.e. a feeling of displeasure evoked by the direct or disturbing effects discussed in Section 4.3. The indirect response variables measured or generated from the results of the survey are listed in Table 7.

The relationships between these variables and noise will be examined in terms of the noise variable NNI ($NNI_{D80A}$). As will be seen, the differences between noise variables of which NNI is representative are less significant than differences between response scales.

As discussed in Section 3.2, the 7-point bother scales, of which ABS is just one, yielded suspicious results in that in practically all cases the distributions of responses were very bimodal. This tendency in the case of ABS is clearly evident in Figure 20 which shows highly skewed distributions at both high and low NNI, with a consequently S-shaped relation between median ABS and NNI. The correlation coefficient
### TABLE 7 Indirect Response Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Abbreviation</th>
<th>Scale Points</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>A/c noise bother score</td>
<td>ABS</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>A/c noise bother coefficient</td>
<td>ABC</td>
<td>-</td>
<td>9,17</td>
</tr>
<tr>
<td>4-pt. A/c noise bother score</td>
<td>AS4</td>
<td>4</td>
<td>17</td>
</tr>
<tr>
<td>A/c noise annoyance</td>
<td>ANA</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>Guttman Annoyance Score</td>
<td>GAS</td>
<td>7</td>
<td>20,21</td>
</tr>
<tr>
<td>Noise Costs*</td>
<td>NCO</td>
<td>-</td>
<td>44,45</td>
</tr>
<tr>
<td>Noise Cost (Rates + Rent)</td>
<td>NCR</td>
<td>-</td>
<td>44,45,11,12</td>
</tr>
<tr>
<td>Noise Costs/Income</td>
<td>NCI</td>
<td>-</td>
<td>44,45,14</td>
</tr>
</tbody>
</table>

* = (contribution + compensation)/2

Table 8 gives the rank correlations between some of these variables.

### TABLE 8 Rank Correlations Between Indirect Response Variables

<table>
<thead>
<tr>
<th></th>
<th>NCI</th>
<th>NCR</th>
<th>NCO</th>
<th>GAS</th>
<th>ANA</th>
<th>AS4</th>
<th>ABC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.815</td>
</tr>
<tr>
<td>ABC</td>
<td></td>
<td>.565</td>
<td>.668</td>
<td>.802</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AS4</td>
<td>.303</td>
<td>.811</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANA</td>
<td>.821</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAS</td>
<td>.173</td>
<td>.065</td>
<td>.249</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
is unfortunately not available.

The favourable effect of normalising ABS with respect to the average bother score determined at Q.9. may be seen in Figure 20 where the resulting coefficient ABC is plotted against NNI. There is no evidence of skewing at high NNI and the median relationship is relatively linear.

The critical value ABC = 1.0, the point at which aircraft noise bother becomes equal to the average bother, is reached by the median at 32NNI.

Inspection of the distributions of ABS suggested that a more valid scale might be constructed by combining the scores 2 and 3, 4 and 5, 6 and 7 to construct a four-point bother scale AS4. In fact, this yields no reduction in skewness at high and low NNI but it is of interest that AS4 correlates highly (R = 0.821, n = 490) with the four-point annoyance scale ANA (Q.20). This at least indicates that subjects were responding consistently when faced with different multi-point response scales. However, their respective rank correlations with NNI are 0.239 (AS4, n = 597 including 104 exclusions prior to Q.20 as AS4 = 1) and .519 (ANA, n = 493) where the difference is clearly significant at the 5% level.

The Guttman Annoyance Score (GAS) correlates highly with NNI (R = 0.597, n = 496), a fact which is reflected by the smaller interquartile range (relative to the total variation of GAS with NNI) in Figure 20. A comparison of ANA and GAS gives the following relationship; results from the 1961 and 1967 surveys are

<table>
<thead>
<tr>
<th>ANA (degree of bother or annoyance)</th>
<th>Average GAS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Present</td>
</tr>
<tr>
<td>Not at all</td>
<td>0.30</td>
</tr>
<tr>
<td>A little</td>
<td>1.80</td>
</tr>
<tr>
<td>Moderately</td>
<td>2.84</td>
</tr>
<tr>
<td>Very much</td>
<td>4.45</td>
</tr>
</tbody>
</table>
The response variables ABC and GAS are compared in Figure 21 (a) which shows the variation of ABC with respect to GAS and 21 (b) which is vice versa. The rank correlation is $R = 0.565$ ($n = 493$). The graphs suggest the reason for the relatively low $R$ is that although the correlation is high for low values of both variables, they become practically independent at high values. However, the point $ABC = 1.0$ is within the correlated range and at this value, the median GAS lies between 2 and 3. The value GAS = 3.5, accepted from the results of previous surveys as a critical annoyance level (5,7), is within the uncorrelated region and corresponds to a median ABC of more than 3.

The three 'cost' variables NCO, NCR and NCI are plotted against NNI in Figure 22. The correlations are very low; 0.235 for NCO which is significant at the 5% level, 0.137 for NCR which is not significant and 0.220 for NCI which is just significant at the 5% level. Clearly, an association between Noise Costs and noise levels, which is low to begin with, is not improved by relating the named compensation or contribution figures to either rents or incomes.

To summarise, it is evident that none of the indirect response variables are more highly correlated with noise than is the Guttman Annoyance Scale, although this finding is clearly influenced by the poor performance of the 7-point bother scale which casts some doubt on the validity of ABS, ABC and AS4.

Not only are people disinclined to think of noise in monetary terms but also the sums they reluctantly suggest, whether as reasonable compensation for suffering noise nuisance or as payments to get rid of it, are only weakly related to noise exposure. Perhaps the main value of Figure 22 is to illustrate the difficulty of costing noise nuisance for planning and other purposes.

The fact that the rank correlations between response variables themselves vary substantially suggest that the measurement of "annoyance" is a precarious task, in that results will depend very highly upon the form of the chosen scale. If the test of validity of a noise annoyance scale is that it
should correlate highly both with other annoyance scales and with noise exposure, then the validity of the scales tried herein has not been proved.

The correlations between the annoyance variables and \( \text{NNI}_{D80A} \) are summarised in Table 9.

### TABLE 9 Correlations Between Indirect Response and Noise

<table>
<thead>
<tr>
<th>'Annoyance' Variable</th>
<th>Rank Correlation with ( \text{NNI}_{D80A} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n )</td>
</tr>
<tr>
<td>ABC</td>
<td>493</td>
</tr>
<tr>
<td>AS4</td>
<td>597</td>
</tr>
<tr>
<td>ANA</td>
<td>493</td>
</tr>
<tr>
<td>GAS</td>
<td>496</td>
</tr>
<tr>
<td>NCO</td>
<td>141</td>
</tr>
<tr>
<td>NCR</td>
<td>126</td>
</tr>
<tr>
<td>NCI</td>
<td>86</td>
</tr>
</tbody>
</table>

For \( n_1 = n_2 = 500 \) and \( r = 0.5 \), Appendix D gives 0.187 as the significant difference between coefficients at the 5% level, 0.157 at 10% and 0.122 at 20%. It seems reasonable to conclude that although the difference between ANA and GAS is not significant, that between GAS and the remaining scales is sufficient to justify concentrating upon GAS in a more detailed analysis of the relationship between noise and annoyance.

### 4.5 Relation Between Annoyance (GAS), Disturbance and Noise

Table 10 lists the correlation coefficients between GAS, the basic noise parameters \( \bar{L}_p \) and \( N_p \) and the combined parameter \( \text{NNI}_{P,M} \) for all periods, threshold levels and the
two modes. The sample size in all cases is 600, obtained by assuming that the 104 people excluded from further questioning prior to question 20 effectively scored GAS = 0. The addition of a large fraction of zero scores, particularly since many of them probably emerge from zero NNI areas, has the effect of increasing R (see Appendix D). However, since the effect is very similar in all cases, it in no way affects the comparison of results in Table 10.

In fact, the Table has a very similar appearance to Table 5 which concerns numbers of disturbances. There are similar tendencies for R to be maximised between $\lambda = 80$ and $\lambda = 85\text{PNdB}$ and for average mode noise levels to yield slightly higher correlations than worst mode. However, no differences are very significant.

A major dissimilarity between Tables 5 and 10 is that the correlation between GAS and night-time NNI is high. Since daytime, evening and night-time NNI's are highly correlated with each other ($R \geq .940$) this is to be expected. In Table 5 night-time noise exposures are correlated with night-time disturbances and the low coefficients must reflect a basically low association between night noise and night disturbance.

When comparing Tables 5 and 10, the difference in sample sizes must be remembered. Approximately 400 respondents were asked the disturbance questions, 108 having been eliminated before the GAS questions and a further 88 having been excluded on the grounds of zero GAS (which itself implies no disturbance at any time). Although the different samples may be accounted for to some extent in the statistical tests of significance (Appendix D), the inclusion of a large number of 'zero-zeros' in the GAS correlations definitely increases the coefficients. Table 11 gives further GAS/noise correlations for $n = 496$ (i.e. those who were asked the GAS questions and $n = 408$ (those who were asked the disturbance questions). Only average mode results are included. Table 12 summarises equivalent disturbance v. average mode results for comparison (data taken from Table 5). For sample sizes of 400 and 600 and $r = 0.5$ Appendix D gives $\Delta R \approx 0.197$ at the 5% level, $\approx 0.165$ at 10% and $\approx 0.129$ at 20%. Thus, there may be no significant difference between the correlations for annoyance/noise and
<table>
<thead>
<tr>
<th>Period</th>
<th>Noise Variable</th>
<th>Threshold Level</th>
<th>65</th>
<th>70</th>
<th>75</th>
<th>80</th>
<th>85</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Day</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$L_{DAW}$</td>
<td></td>
<td>.627</td>
<td>.625</td>
<td>.631</td>
<td>.634</td>
<td>.627</td>
<td>.630</td>
</tr>
<tr>
<td></td>
<td>$L_{DA}$</td>
<td></td>
<td>.641</td>
<td>.645</td>
<td>.649</td>
<td>.651</td>
<td>.635</td>
<td>.625</td>
</tr>
<tr>
<td></td>
<td>$N_{DAW}$</td>
<td></td>
<td>.550</td>
<td>.574</td>
<td>.603</td>
<td>.597</td>
<td>.630</td>
<td>.628</td>
</tr>
<tr>
<td></td>
<td>$N_{DA}$</td>
<td></td>
<td>.588</td>
<td>.596</td>
<td>.598</td>
<td>.618</td>
<td>.651</td>
<td>.663</td>
</tr>
<tr>
<td></td>
<td>$NNI_{DAW}$</td>
<td></td>
<td>.627</td>
<td>.625</td>
<td>.629</td>
<td>.634</td>
<td>.637</td>
<td>.638</td>
</tr>
<tr>
<td></td>
<td>$NNI_{DA}$</td>
<td></td>
<td>.653</td>
<td>.654</td>
<td>.658</td>
<td>.663</td>
<td>.667</td>
<td>.654</td>
</tr>
<tr>
<td><strong>Evening</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$L_{EAW}$</td>
<td></td>
<td>.636</td>
<td>.641</td>
<td>.653</td>
<td>.648</td>
<td>.623</td>
<td>.642</td>
</tr>
<tr>
<td></td>
<td>$L_{EA}$</td>
<td></td>
<td>.680</td>
<td>.612</td>
<td>.691</td>
<td>.686</td>
<td>.680</td>
<td>.637</td>
</tr>
<tr>
<td></td>
<td>$N_{EAW}$</td>
<td></td>
<td>.573</td>
<td>.609</td>
<td>.631</td>
<td>.682</td>
<td>.684</td>
<td>.606</td>
</tr>
<tr>
<td></td>
<td>$N_{EA}$</td>
<td></td>
<td>.597</td>
<td>.617</td>
<td>.638</td>
<td>.674</td>
<td>.655</td>
<td>.516</td>
</tr>
<tr>
<td></td>
<td>$NNI_{EAW}$</td>
<td></td>
<td>.660</td>
<td>.661</td>
<td>.664</td>
<td>.666</td>
<td>.671</td>
<td>.665</td>
</tr>
<tr>
<td></td>
<td>$NNI_{EA}$</td>
<td></td>
<td>.673</td>
<td>.679</td>
<td>.683</td>
<td>.688</td>
<td>.667</td>
<td>.651</td>
</tr>
<tr>
<td><strong>Night</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$L_{NAW}$</td>
<td></td>
<td>.611</td>
<td>.617</td>
<td>.615</td>
<td>.614</td>
<td>.607</td>
<td>.602</td>
</tr>
<tr>
<td></td>
<td>$L_{NAA}$</td>
<td></td>
<td>.597</td>
<td>.624</td>
<td>.635</td>
<td>.631</td>
<td>.620</td>
<td>.611</td>
</tr>
<tr>
<td></td>
<td>$N_{NAW}$</td>
<td></td>
<td>.498</td>
<td>.502</td>
<td>.540</td>
<td>.593</td>
<td>.619</td>
<td>.618</td>
</tr>
<tr>
<td></td>
<td>$N_{NAA}$</td>
<td></td>
<td>.541</td>
<td>.517</td>
<td>.513</td>
<td>.555</td>
<td>.608</td>
<td>.510</td>
</tr>
<tr>
<td></td>
<td>$NNI_{NAW}$</td>
<td></td>
<td>.613</td>
<td>.613</td>
<td>.612</td>
<td>.620</td>
<td>.620</td>
<td>.610</td>
</tr>
<tr>
<td></td>
<td>$NNI_{NAA}$</td>
<td></td>
<td>.624</td>
<td>.611</td>
<td>.608</td>
<td>.611</td>
<td>.625</td>
<td>.613</td>
</tr>
</tbody>
</table>
TABLE 11 Rank Correlations Between GAS and Noise Variables

(a) Sample n = 600 (assumes 104 excluded before Q.20 score GAS = 0)
(b) Sample n = 496 (those who were asked Q.'s 20 and 21)
(c) Sample n = 104 (excluding GAS = 0; i.e. people who were not asked disturbance questions)

<table>
<thead>
<tr>
<th>Period</th>
<th>Variable</th>
<th>(a) n = 600</th>
<th>(b) n = 496</th>
<th>(c) n = 408</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(\text{dB(PN)})</td>
<td>(\text{dB(PN)})</td>
<td>(\text{dB(PN)})</td>
</tr>
<tr>
<td></td>
<td></td>
<td>65</td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td>Day</td>
<td>(L_{\text{DA}})</td>
<td>.641</td>
<td>.645</td>
<td>.649</td>
</tr>
<tr>
<td></td>
<td>(N_{\text{DA}})</td>
<td>.588</td>
<td>.596</td>
<td>.598</td>
</tr>
<tr>
<td></td>
<td>(NN_{\text{DA}})</td>
<td>.653</td>
<td>.654</td>
<td>.658</td>
</tr>
<tr>
<td>Evening</td>
<td>(L_{\text{EA}})</td>
<td>.680</td>
<td>.682</td>
<td>.691</td>
</tr>
<tr>
<td></td>
<td>(NN_{\text{EA}})</td>
<td>.673</td>
<td>.679</td>
<td>.683</td>
</tr>
<tr>
<td>Night</td>
<td>(L_{\text{NA}})</td>
<td>.597</td>
<td>.624</td>
<td>.635</td>
</tr>
<tr>
<td></td>
<td>(N_{\text{NA}})</td>
<td>.541</td>
<td>.517</td>
<td>.513</td>
</tr>
<tr>
<td></td>
<td>(NN_{\text{NA}})</td>
<td>.624</td>
<td>.611</td>
<td>.608</td>
</tr>
</tbody>
</table>
disturbance/noise (for day and evening). This is certainly true when we compare the GAS results for n = 408 where the correlations coefficients are very similar to those of Table 12.

TABLE 12  Rank Correlations Between Disturbance and Noise Variables (Extracted from Table 5)

<table>
<thead>
<tr>
<th>Disturbance Variable</th>
<th>Sample Size</th>
<th>Noise Variable</th>
<th>Threshold Level ( \lambda ), dB(PN)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Noise Variable</td>
<td>65</td>
</tr>
<tr>
<td>NDD</td>
<td>365</td>
<td>( \overline{I}_D \text{ vs. } A )</td>
<td>.491</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_D \text{ vs. } A )</td>
<td>.423</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{NNI}_D \text{ vs. } A )</td>
<td>.480</td>
</tr>
<tr>
<td>NED</td>
<td>374</td>
<td>( \overline{I}_E \text{ vs. } A )</td>
<td>.536</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_E \text{ vs. } A )</td>
<td>.426</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{NNI}_E \text{ vs. } A )</td>
<td>.497</td>
</tr>
<tr>
<td>NND</td>
<td>377</td>
<td>( \overline{I}_N \text{ vs. } A )</td>
<td>.173</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_N \text{ vs. } A )</td>
<td>.185</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{NNI}_N \text{ vs. } A )</td>
<td>.205</td>
</tr>
</tbody>
</table>

Figure 23 shows the relationships between number of disturbances and GAS. The correlations (day 0.644, evening 0.638, night 0.450) are high by comparison with those relating the different indirect response variables (Table 8) and in some cases significantly higher at the 20% level than those between either disturbance or annoyance and noise (for day and evening periods).

It may be observed from Tables 11 and 12 that the combination of level and number variables into NNI does not necessarily lead to an increase in correlation with either disturbance or annoyance. In most instances the correlation with level \( \overline{L} \) is higher than that with \( N \). However, none of these changes is significant and indeed in previous surveys correlations between annoyance and \( N \) have been higher than between annoyance and \( \overline{L} \).
As before, the question of the 'trade-off' between $\bar{L}$ and $N$ is confused by the high correlation between these two variables. Figure 24 shows the variation of $R$ with $K$ in the rank correlation between GAS and the modified NNI-type formula.

$$\bar{L}_{D80A} + K \log N_{D80A}$$

for $K$ in the range 1 to 500 (note that for $\bar{L}$ alone, $K = 0$, $R = .651$ and for $N$ alone, $K = \infty$, $R = .618$). The variation of $R$ is extremely small, reaching a maximum at $K = 15$. (For other periods and threshold levels the curve may have a characteristically different shape.) Therefore the conclusion is that, all combinations of $\bar{L}$ and $N$ will provide response predictor formulae of equivalent reliability. This is a consequence of the high correlation between $\bar{L}$ and $N$.

For similar reasons, no significant changes in $R$ result when disturbance or annoyance is correlated with any other noise variable studied herein. Table 13 lists a sample of rank correlation coefficients relating NDD and GAS and various noise variables including four composite indices which are as follows:

$$\text{NNI}_{DE} = \bar{L}_{D80A} + 15 \log_{10} (N_{D80A} + 3N_{E80A})$$
$$\text{NNI}_{DN} = \bar{L}_{D80A} + 15 \log_{10} (N_{D80A} + N_{E80A} + 10N_{N80A})$$
$$\text{NNI}_{DEN} = \bar{L}_{D80A} + 15 \log_{10} (N_{D80A} + 3N_{E80A} + 10N_{N80A})$$
$$\text{NNI}_{24} = \bar{L}_{D80A} + 15 \log_{10} (N_{D80A} + N_{E80A} + N_{N80A})$$

These first three of these are similar in form to the variants of the noise index WECPNL proposed by ICAO for aircraft noise rating purposes(2). Like other noise indices used in the USA (NEP, $L_{dn}$, CNR) they give different weights to daytime, nighttime and sometimes evening events. The simplifications above assume that $\bar{L}$ is equal for all three periods, i.e. that the main differences arise through the number term.
### TABLE 13

**Rank Correlations Between Disturbance, Annoyance and Various Noise Exposure Variables**

<table>
<thead>
<tr>
<th>Resp. Vble.</th>
<th>Sample</th>
<th>Mode</th>
<th>$\overline{L}_{D80}$</th>
<th>$\overline{N}_{D80}$</th>
<th>NNI$_{D80}$</th>
<th>NNI$_{DE}$</th>
<th>NNI$_{DN}$</th>
<th>NNI$_{DEN}$</th>
<th>NNI$_{24}$</th>
<th>$L_1$</th>
<th>$L_{10}$</th>
<th>$L_{50}$</th>
<th>$L_{eq}$</th>
<th>$L_{NP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDD</td>
<td>365</td>
<td>W</td>
<td>.476</td>
<td>.415</td>
<td>.487</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.483</td>
<td>.504</td>
<td>.441</td>
<td>.480</td>
<td>.461</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A</td>
<td>.493</td>
<td>.387</td>
<td>.476</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.489</td>
<td>.411</td>
<td>.313</td>
<td>.495</td>
<td>.442</td>
</tr>
<tr>
<td>GAS</td>
<td>600</td>
<td>W</td>
<td>.634</td>
<td>.597</td>
<td>.634</td>
<td>.640</td>
<td>.633</td>
<td>.636</td>
<td>.634</td>
<td>.626</td>
<td>.642</td>
<td>.662</td>
<td>.627</td>
<td>.606</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A</td>
<td>.651</td>
<td>.618</td>
<td>.663</td>
<td>.668</td>
<td>.660</td>
<td>.663</td>
<td>.664</td>
<td>.674</td>
<td>.632</td>
<td>.541</td>
<td>.676</td>
<td>.617</td>
</tr>
</tbody>
</table>
Again, as may be expected from the high correlations between day, evening and night noise exposure these complex formulae offer no advantage for response prediction purposes, there being no significant differences in their associated rank correlation coefficients.

Indeed the only noteworthy features of Table 13 are the small coefficients for $L_{50}$ (average mode) relative to $L_{50}$ (worst mode) which probably reflect large inaccuracies associated with the estimation of $L_{50}$ where $N$ is small.

However, in view of current interest in the scales $L_{eq}$ and $L_{NP}$, the relations between GAS, $L_{eq}$ and $L_{NP}$ are compared with that between GAS and NNI in Figure 25.

4.6 Comparison with Previous Survey Results

The inclusion of the Guttman Annoyance Scale in this survey allows direct comparisons to be made with earlier surveys. Figure 26 compares the relations between "group mean annoyance scores" GAS from the four surveys (Heathrow, 1961(7); 1967(5) and 1972 (present) and Gatwick 1971(17). The mean scores are the arithmetic averages of the scores of all respondents living within a band of noise exposure for which the plotted NNI value is an arithmetic average.

Although the collapse of the data in Figure 26 is good, such a diagram is somewhat misleading from a practical point of view because it disregards the considerable scatter of individual annoyance scores. Ollerhead has previously noted (1) that GAS and NNI are related in an approximately joint normal probability distribution and all four sets of data are plotted on normal probability paper in Figure 27 to illustrate this point. This graph shows that to a fair approximation (a) individual annoyance scores in any noise stratum are normally distributed and (b) all scores of any value (i.e. 1,2,3,4,5 or 6) are normally distributed with respect to NNI. The respective standard deviations are (a) 2 annoyance units and (b) 20NNI (reading from the idealised straight line fits).

This rather neat relationship is obviously attractive from a mathematical point of view and indeed might be considered ideal. With hindsight it might even be claimed that a desirable
feature of scales of both noise exposure and response is that they should exhibit these Gaussian characteristics. However, it can only be concluded that the satisfaction of these requirements is fortuitous, if not as a pure accident then as an accidental by-product of the search for scales which meet other validity criteria. In passing it should perhaps be noted that although the data points for GAS scores from 1 to 5 fit the lines reasonably well, scores of 6 fall some way below the 'ideal' line suggesting that the interval between scores of 5 and 6 represents a greater annoyance interval than the other four (on the assumption that the numerical scale of annoyance should be normally distributed throughout the population by definition).

The agreement between the four sets of data is extremely good indicating that the relationship between GAS and NNI is stable with time (at least over eleven years) and is the same for Gatwick and Heathrow airports. Whether or not it remains true for other airports cannot be ascertained from present results; a much more complete data matrix of noise and number combinations is required to test this. Gatwick, although handling only a fraction of the traffic volumes at Heathrow, must still be categorised as a major airport and further conclusions must await data from substantially smaller airports.

Figure 27 should be transformable into similar joint-Gaussian plots for any noise variables which are linearly related to NNI. Figure 10 indicates fair linearity between NNI, $L_{eq}$ and $L_{NP}$ and Figures 28 and 29 show the results for $L_{eq}$ and $L_{NP}$ for present (1972) survey data. The standard deviation of annoyance scores (noise constant) remains fixed at 2 units and the standard deviation with respect to $L_{eq}$ is 12PNdB which is consistent with a slope of 0.6PNdB/NNI in Figure 10. The fit of the data to the ideal lines is equally as good as that for NNI and it may be concluded that on this basis $L_{eq}$ constitutes a perfectly adequate substitute for NNI for scaling aircraft noise, using the transformation

\[ L_{eq} = 0.6 \text{NNI} + 56, \text{PNdB} \]

The standard deviations in Figure 29 for $L_{NP}$ are two
annoyance units and 20PNdB ($L_{NP}$) the latter being consistent with the unit slope of Figure 10. However, the agreement between the lines and the data is less impressive and although this data yields the transformation

$$L_{NP} = NNI + 70, \text{ PNdB}$$

it is unlikely that the scale of $L_{NP}$ will demonstrate the attractive features of NNI and $L_{eq}$ for predicting the community impact of aircraft noise.

4.7 The Role of Background Noise

On both intuitive grounds and upon the basis of laboratory experimentation it is reasonable to suppose that reaction to noise intrusion is sensitive to the masking affect of background noise. Although no background noise information was acquired in this study, it may also be supposed that road traffic is a major contributor. The results of previous Loughborough University study(10) suggested that annoyance reactions are sensitive to an interaction between aircraft noise and traffic noise. To examine this possibility, the traffic noise estimates described in Section 3.1 have been used to separate respondents into high ($L_{50} \geq 55 \text{ dB(A)}$) and low ($L_{50} < 55 \text{ dB(A)}$) traffic noise categories.

Figure 30 shows the median aircraft noise bother coefficient (ABC) plotted against NNI for the two groups. Although the reaction to aircraft noise is higher at higher NNI for the low traffic noise group, the difference is not sufficiently significant to support the linear findings. However, the unreliable nature of the traffic noise estimates may, of course, observe possible real differences.
5.0 CONCLUSIONS AND RECOMMENDATIONS

Six hundred residents of suburban communities near London (Heathrow) Airport were interviewed in a pilot survey designed to investigate further aspects of aircraft noise assessment.

The questionnaire included questions designed to scale human reaction to noise in various ways whilst aircraft noise exposure variables at each residence were estimated from data describing aircraft movements by type, route and time during a period of four weeks prior to the interviews. The principal objectives of the study were:

(1) to seek some indication that the correlation between physical noise exposure and people’s reaction to it, generally recognised to be low, might be increased through better definitions of both noise and reaction variables.

(2) to test alternative scales of reaction which have more quantitative meaning for planning purposes;

(3) to relate aircraft noise to other sources of community nuisance;

(4) to examine the variation of community noise sensitivity with time of day;

(5) to define the requirements for further full-scale survey research.

Although various elements of the study were concerned with more than one of these objectives, the main conclusions are summarised below under the headings to which they are most relevant.

5.1 Correlation between noise and response

It was initially hypothesised that the directly disturbing or intruding effects of noise (disturbance) would be more highly correlated with noise exposure than would the indirect (annoyance) effects which are probably more sensitive to intervening psychological and sociological factors.

To test this hypothesis, direct effects were measured through questions concerning the frequency and duration of
disturbances (which were undefined) whilst numerous indirect effects were scaled through questions related to annoyance, bother and monetary costs. For comparison with previous surveys a series of questions which form the basis for a previously used Guttman Annoyance Scale (GAS), was included.

The relationship between GAS score distributions and noise exposure, expressed in NNI, agreed closely with those found in three earlier surveys. Furthermore the combined results show that to a good approximation GAS scores are normally distributed at any constant NNI and also that those people scoring any particular GAS are normally distributed with respect to NNI. The mathematical convenience of this relationship has obvious appeal for planning purposes as well as satisfying an intuitive feeling that annoyance and noise should be associated in this way.

Other indirect response variables were less highly correlated with noise than was GAS. Of some concern also is that the different response variables were rather poorly correlated with each other suggesting that they are not associated with the same subjective dimension. However a scale of 'bother', which might have been expected to show high correlation with annoyance, was a generally poor performer and it cannot be used in any way to test the validity of GAS.

The hypothesis that disturbance is more highly correlated with noise than indirect effects could not be verified; the correlation between disturbance and noise was not significantly different from that between noise and annoyance. Whether the hypothesis is incorrect or whether the disturbance measures were inadequate (or both) cannot be established from present results. However, it is likely that disturbance, as defined, and annoyance are not separate responses.

5.2 Scales of subjective response

The Guttman scale of aircraft noise annoyance again satisfied cumulative attitude scaling and reproducibility criteria and correlated reasonably well with noise exposure estimates. This, together with the close agreement of the GAS-NNI relationships with those obtained in earlier surveys suggest that GAS remains a convenient and reliable response
scale. However, although the scale values are unaffected, it has been noted that the hierarchical sequence of responses to the questions involved vary from survey to survey, and thus warn against any general acceptance of the measuring and validity of this particular scale. The disturbance frequency measurements, although no more highly correlated with noise than annoyance, were no less so and offer a viable alternative or complement to annoyance measurements. Furthermore, disturbance frequency provides a meaningful and quantitative response scale which can be more precisely related to specific events or time periods.

Disturbance duration was poorly correlated with estimated sound duration but this is probably explained by a very small variation of the physical variable. Subjective estimates of signal duration appear to be quite realistic.

There is little evidence that people think about noise nuisance in monetary terms, either spontaneously or after persuasion to do so. Only 1% of respondents spontaneously suggested that monetary compensation might provide a solution to the aircraft noise problem and many respondents could not be prompted to suggest a suitable payment. Upper and lower bounds to nuisance costs were obtained as estimates of (a) compensation and (b) payments to eliminate the noise. Of those who answered the questions the median respondent was prepared to pay roughly one third of what he felt to be fair compensation. The correlation between the average of (a) and (b) and noise was significant but very low and the correlation could not be improved by normalising these costs with respect to either rent or income.

5.3 Aircraft noise exposure scales

Several basic noise exposure parameters were computed for the purposes of this study in relation to a variable threshold level $\lambda$. These were, for each of the 600 residential locations (and for day, evening and night),

$N$, the number of aircraft sounds which exceed the threshold $\lambda$
$\overline{L}$, the (energy) average peak level of these sounds and $\overline{D}$, the average duration of these sounds in excess of $\lambda$
Each value was determined for both average mode (i.e. average daily aircraft traffic) and worst mode (heaviest traffic encountered during the four week period) operations.

From various composite scales such as NNI and the duration-based parameters \( L_1, L_{10}, L_{50}, L_{eq} \) and \( L_{NP} \) were computed. The latter can only be considered valid for the aircraft noise component of the total noise environment since no background component was included. However, since at Heathrow general background noise levels may be expected to accompany high aircraft noise levels, this is not thought to be a stringent limitation upon the following conclusions.

Because the parameters \( N, L \) and \( D \) are highly correlated with each other within the survey area, present results provide no more guidance to the relative merits of different noise scales as response predictors than do those of previous surveys. In a region where this high correlation exists most noise scales will be equivalent. Thus for example, the correlation between annoyance and the composite scale given by \( L + K \log_{10} N \) does not change significantly as \( K \) varies from zero to infinity. Therefore, pending the availability of further information, a choice must be based upon other considerations such as convenience, linearity and stability.

For aircraft noise the currently favoured scale, NNI, satisfies many of these requirements. It is convenient to calculate or measure and its relationship with response, at least on the GAS scale, may be regarded as ideal (see 5.1). The effect of the threshold level \( \lambda \) on its correlation with response is insignificant but the fact that other correlations are affected when \( \lambda \) exceeds 80PNdB suggests that this currently accepted value may be an optimum choice (as \( \lambda \) decreases, computations become more complex and unreliable). The difference in correlation for average and worst mode conditions is also insignificant and, since worst mode results are considerably more difficult to compute, the practical advantage rests with average mode conditions.

The alternative scale of \( L_{eq} \) (the equivalent continuous noise level which would represent the same total energy as the real variable signal) exhibits the same favourable relationship with response as NNI to which it is related by the practical
formula

$$L_{eq} \approx 0.6\text{NNI} + 56, \text{PNdB}$$

Also, by definition, it is independent of threshold level $\lambda$ (although one could be incorporated for ease of measurement or calculation or to modify sensitivity). A disadvantage that its calculation is more complicated is offset by the fact that it is readily applied to complex time histories of noise (for which reason it is rapidly gaining international favour as a universal noise scale).

The more elaborate scale of Noise Pollution Level ($L_{NP}$) is rather less convenient than $L_{eq}$ due to difficulties associated with both measurement and calculation and its relationship with annoyance was found to be rather less neat than was the case with NNI and $L_{eq}$. On the other hand it must be admitted that the available input data was not really adequate for the computation of the $L_{NP}$ values (which are more sensitive to background noise than either NNI or $L_{eq}$).

A very approximate estimate of the background noise level at each location was made on the basis of the interviewer's observations of traffic noise sources and this was used to examine possible effects of traffic noise on aircraft noise annoyance.

No effect could be identified and no real conclusion can be drawn due to the unreliable nature of the traffic noise estimates.

5.4 Aircraft noise in relation to other sources of nuisance

A semantic differential scale of 'bother' was used to obtain reactions to a wide range of sources of dissatisfaction with local living conditions. Although the validity of the scale was suspect due to skewed and bimodal response distributions the results provide some indication of community perception of various undesirable agents. Of 19 bothersome items investigated, noise ranked second to the risk of road accidents and was followed by atmospheric pollution. The predominant sources of noise nuisance, in rank order, were aircraft, road traffic, children, and neighbours. Surprisingly, although 13% of the respondents lived within 200 yards of a railway line
and 43% within 1/2 mile, trains caused substantially lower bother than any other nuisance, acoustic or otherwise.

5.5 Variation of noise sensitivity with time of day

The disturbance measurements provide unique insight into the variation of community noise impact with time of day. Three periods were defined: daytime (0700-1900), evening (1900-2300) and night-time (2300-0700). For each period both noise exposure estimates and reported disturbance frequencies and durations were obtained.

Although inter-subject variations were inevitably large, the results clearly show that aircraft noise is most intrusive during the evening, followed by the daytime and the night. The median respondent was at worst disturbed by roughly one out of ten evening sounds (which exceed 80PNdB), one out of thirty daytime sounds and one out of one-hundred night-time sounds. Thus the variation in median sensitivity is roughly (for day; evening; night) 3:10:1.

It must be emphasised that individual susceptibilities vary widely and that the only indication of perceived severity of disturbance is its reported duration. In this regard daytime and evening disturbances were rated equal (median = 40 sec.) whereas night-time durations were rather longer (median = 2 min.).

The results of the pilot study thus tend to support the use of a weighting for evening sounds with a multiplying factor of 3 applied to evening numbers being appropriate for NNI-type noise indices. On the other hand the use of a large night-time weighting cannot be justified on the same basis.

Yet the disturbances of sleep can only be viewed with considerable concern and it is never likely to be considered desirable to reschedule flights from day or evening to the night-time period. Thus some rationale other than disturbance frequency or duration must form the basis for night-time noise assessment.

That the night-time problem must be treated separately is evident from the fact that the correlation between night-time noise and disturbance is significantly less than between day or evening noise and disturbances. Similarly the correlation between day and evening disturbance is significantly higher.
than that between night-time disturbance and either day or evening disturbance.

These results are in contrast to the relations between annoyance (GAS) and noise exposures during the three periods for which the correlations are not significantly different. This illustrates a major limitation with GAS (or any other scale of 'Chronic' annoyance); that it is totally insensitive to time of day variations.

5.6 Recommendations

(a) The relationship between annoyance (GAS) and NNI appears to be stable and convenient for major airport planning purposes and there is no evidence to suggest that current techniques should be discontinued at present; rather that they should be reinforced by the findings of studies such as this. At the same time GAS exhibits certain weaknesses and there is a need for a new scale of annoyance which is less dependent (and ideally has no dependence) upon the sources of that annoyance.

(b) Equivalent continuous noise level $L_{eq}$ appears to be an alternative scale for aircraft noise exposure which, on balance, is equal to NNI as a response predictor. In view of current needs for a degree of unification in noise scaling methodology there is a strong case to develop experience in the use of $L_{eq}$ for aircraft noise assessment.

(c) Simple techniques for the measurement of disturbance show promise for investigating public reaction to noise. Although the results presented provide a tentative basis for noise impact assessment, the survey was of a preliminary nature and a full-scale survey is desirable. Methods for more detailed disturbance analysis should also be investigated.

(d) In the short-term the applicability of the GAS-NNI relations, which appear to be stable for large airports, should be checked for smaller, regional airports.

(e) In the longer term, the remaining dilemma concerning the trade-off between noise and numbers must be resolved through a widespread survey with a sample for which
these variables are decorrelated, possibly by simultaneous studies at several airports. Perhaps concurrently, closer attention should be paid to the question of background noise, although this will inevitably involve a very laborious noise measurement and/or estimation exercise.
FIGURE 9  RELATIONS BETWEEN MEAN PEAK LEVELS, NUMBERS, DURATIONS AND NNI

* 151 'zero-zeros' excluded from plots
FIGURE 10

RELATIONS BETWEEN $L_{eq}$, $L_{NP}$ and NNI

- 'zero-zeros' excluded from plots
FIGURE 11 RELATIONS BETWEEN DAY, EVENING AND NIGHT NNI
FIGURE 12  RELATIONS BETWEEN NUMBER OF DISTURBANCES AND NUMBER OF EVENTS
FIGURE 13  EFFECT OF THRESHOLD LEVEL $\lambda$ (PNdB) ON RELATIONS BETWEEN
MEDIAN NUMBER OF DISTURBANCES AND NUMBER OF EVENTS
FIGURE 14  NUMBER OF DISTURBANCES AS FUNCTIONS OF NNI (FOR DAY, EVENING AND NIGHT)
FIGURE 15 VARIATION OF SENSITIVITY TO DISTURBANCE WITH TIME OF DAY

(a) Median Number of Disturbances

(b) Median Number of Disturbances

Number of disturbances ≈ Number of events

EVENING

DAY

NIGHT

Number of Events in Excess of 80 PNdB

NNI (Average Mode; day, evening, night)
FIGURE 16  PERCENTAGE OF PEOPLE DISTURBED BY MORE THAN GIVEN PERCENTAGE OF SOUNDS

\[ Y = \text{Percentage of people disturbed by } x \times \text{percentage of sounds or more} \]

\[ x = \text{Percentage of sounds with peak levels } > 90 \text{ PNdB} \]
FIGURE 17  RELATIONS BETWEEN DISTURBANCE DURATIONS AND SIGNAL DURATIONS (> 80 PNdB)
FIGURE 18  FURTHER RELATIONS BETWEEN DISTURBANCE DURATIONS AND SIGNAL DURATIONS
FIGURE 19  PERCEIVED (MEDIAN) DURATION VS. ACTUAL SIGNAL DURATION IN EXCESS OF
FIGURE 20  RELATIONS BETWEEN ANNOYANCE INDICES AND NNI

n = 600
R = 0.663

n = 597
R = 0.451

n = 493
R = 0.451
FIGURE 21. RELATIONS BETWEEN ABC AND GAS

n = 493
R = 0.565
FIGURE 22  NOISE 'COSTS' IN RELATION TO NOISE EXPOSURE
FIGURE 23  RELATIONS BETWEEN NUMBER OF DISTURBANCES AND ANNOYANCE
FIGURE 24  RANK CORRELATION BETWEEN GAS  
AND $\bar{L} + k \log_{10} N$
FIGURE 25  DISTRIBUTIONS OF ANNOYANCE Vs. NNI, $L_{eq}$ and $L_{NP}$
FIGURE 26 VARIATION OF AVERAGE ANNOYANCE GAS WITH NNI - FOUR SURVEYS.
FIGURE 27  DISTRIBUTION OF ANNOYANCE SCORES
(GAS) Vs. NNI - FOUR SURVEYS

-96-
FIGURE 28 DISTRIBUTION OF ANNOYANCE SCORES (GAS) Vs. $L_{eq}$
FIGURE 29  DISTRIBUTION OF ANNOYANCE SCORES (GAS) Vs. $L_{NP}$
FIGURE 30  AIRCRAFT NOISE 'BOTHER' REACTIONS IN HIGH AND LOW ROAD TRAFFIC NOISE AREAS
### Classification Page

Total no. of interviews - 600  Figures in parentheses refer to missing data or 'don't knows'

<table>
<thead>
<tr>
<th>Sex</th>
<th>Male</th>
<th>286</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Female</td>
<td>311</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0-9 years</td>
<td>24</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>10-19 years</td>
<td>115</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>20-29 years</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>30-39 years</td>
<td>24</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>40-49 years</td>
<td>109</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>50-59 years</td>
<td>144</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>60 years or over</td>
<td>missing</td>
<td>(2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relationship to head of household</th>
<th>Head</th>
<th>2.77</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spouse</td>
<td>254</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Immediate family</td>
<td>63</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>(6)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Social grade</th>
<th>A</th>
<th>17</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>62</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>C1</td>
<td>167</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>197</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>107</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>38</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>(12)</td>
<td></td>
</tr>
</tbody>
</table>

#### BY OBSERVATION - DO NOT ASK

<table>
<thead>
<tr>
<th>No evidence of hearing deficiency?</th>
<th>571</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some</td>
<td>24</td>
<td>2</td>
</tr>
<tr>
<td>Considerable</td>
<td>missing</td>
<td>(2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>During the interview did you hear?</th>
<th>Aircraft</th>
<th>248/69</th>
<th>51</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Road Traffic</td>
<td>216/11</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>Industrial</td>
<td>2/0</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>Constructions</td>
<td>6/0</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Trains</td>
<td>11/4</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Household</td>
<td>16/2</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>14/8</td>
<td>57</td>
</tr>
</tbody>
</table>

| If Other (State) | 58 | 59 |

<table>
<thead>
<tr>
<th>Source of Road Traffic noise</th>
<th>Less than 20 yards</th>
<th>60 Trunk Major Minor Access</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ask CASUALLY - Had you heard anything about this enquiry before?</th>
<th>64</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>6</td>
</tr>
<tr>
<td>No</td>
<td>missing</td>
</tr>
</tbody>
</table>

-101-
I am carrying out a survey for BMRB on behalf of Loughborough University and I would like to ask you a few questions about local living conditions. May I come in?

Show Identity Card

I am working on a survey to find out people's opinions on various factors which affect their general living conditions. Any views you give us will be kept strictly confidential. For the moment this survey is to be restricted to people with normal work and rest periods. Are you by any chance a night worker?

If IRREGULARLY — During the last few weeks?

If YES — CLOSE INTERVIEW and do not record anything.

1 How long have you lived in this particular neighbourhood? (If necessary define as within ½ mile or so of here)

   less than 4 weeks*..........................1
   1-5 months...............................2
   6-11 months...............................3
   1-2 years...................................4
   3-5 years...................................5
   6-10 years..................................6
   11-20 years.................................7
   More than 20 years........................8
   Don't know/missing (0)...................0

2 And have you actually been here for the last few weeks?

   Yes..............................559
   No - less than 2 weeks.............40
   No - more than 2 weeks*...........missing (1)

3 On the whole how do you like living in this neighbourhood? Do you rate it as an excellent, good, fair, poor, or very poor place to live in?

   Excellent..............................52
   Good....................................283
   Fair.....................................198
   Poor.....................................38
   Very poor................................29
   Don't know/missing (0)...........0

4 At the present time, what are some of the things you particularly like about living in this neighbourhood — things you feel are advantages and make this a good place to live?

   Area quiet..............................68
   Other environmental advantages.....69
   Convenient location...................70
   Other.....................................71
   Nothing/vague/don't know...........72

5 What about the things you DON'T like about living in this neighbourhood? What are the disadvantages — things you feel are unpleasant or undesirable?

   Aircraft noise...........................73
   Other noise................................74
   Other environmental disadvantages...75
   Inconvenient location...................76
   Other.....................................77
   Nothing/vague/don't know............78

6 What is the ONE thing you would most like to CHANGE in this neighbourhood?

   Aircraft noise.........................66
   Other noise.............................26
   Other environmental disadvantages....30
   Inconvenient Location..................20
   Other.....................................150
   Nothing/vague/don't know............(223)
7 (a) Have you ever felt like moving away from here?
Yes............................................. 337  
No................................................. 263  
Don't know.................................. (0) 

If YES
(b) Why did you feel like moving?
Aircraft noise.................................. 56  
Other noise..................................... 14  
Other environmental disadvantages...... 76  
Inconvenient location....................... 18  
Other............................................. 150  
Nothing/vague/don't know................ (48) 

Show CARD A
I want to find out how you feel about a number of things which affect many people's enjoyment of life. To do this I would like you to use this scale to indicate how much they bother you personally. At the top (POINT TO SEVEN) is Extremely bothered and at the bottom (POINT TO ONE) is Not at all bothered. If you ARE bothered, you would choose a number somewhere between one and seven depending on how bothered you feel.

Show CARD A
8 As an example let us consider the quality of TV programmes. Please look at the card and give me a number which most nearly represents how strongly you normally feel about violence and immorality on TV.

Ensure respondent understands use of CARD A before continuing.

9 Now that you understand what is required I am going to read you a list of things you may feel concerned about. For each one, please look at the card and tell me how much it normally bothers you in this particular neighbourhood.

For EACH item below ask: How bothered are you by ...................... ?  
(a) Bad traffic conditions and the risk of road accidents around here ......... (1)  
(b) Poor street lighting .................................................. (4)  
(c) Dirt and litter in the street around here ................................... (3)  
(d) Air pollution, fumes, dirt and smells ........................................ (4)  
(e) The sort of people who live around here .................................... (8)  
(f) Noise .............................................................................. (4)  
(g) Neglect of the neighbourhood appearance ............................... (7)  
(h) Poor public transport .......................................................... (7)  
(i) Poor on inconvenient shops ................................................. (16)  
(j) Lack of entertainment and recreational facilities ......................... (15)  
(k) Being overlooked or cramped and general lack of privacy .......... (2)  
(l) Poor parking facilities ...................................................... (18)  
(m) The worry of theft or other threats to persons and property .......... (8)  
(n) Poor planning in this district .............................................. (40)  
(o) The rate at which this area is growing ...................................... (22)  
(p) Poor quality housing ....................................................... (23)  
(q) The weather .................................................................... (39)  
(r) Poor maintenance and repair of public property (roads, paths, fences etc)........ (13)  
(s) People's lack of interest in local affairs .................................. (38)  

10 Summing up, how bothered are you in general by things you dislike in this neighbourhood?  

Score (Don't know=0)

-103-
11 (a) What about your home? Is this (house/flat) owned or rented by you or do you live here rent free?

If OWNED ONLY

Less than £2,000 .................................. 8
£2,000-£2,999 .................................. 8
£3,000-£4,999 .................................. 65
£4,500-£6,999 .................................. 136
£7,000-£9,999 .................................. 49
£10,000-£14,999 .................................. 202
£15,000-£19,999 .................................. 52
£20,000 or more .................................. 11
Don't know ........................................ 5

Results from Q.s 11(b) and 11(c) combined on assumption that house values are equivalent to rents with same code at 11(c)

12 How much do you pay in rates each year?

None or rates inclusive .................................. 208
Less than £20 .................................. 3
£20-£29.99 .................................. 9
£30-£44.99 .................................. 4
£45-£69.99 .................................. 3
£70-£99.99 .................................. 140
£100-£149.99 .................................. 82
£150 or more .................................. 8
Refused .................................. 5
Don't know .................................. 80

13 How many people are there in your household?

Resident family ONLY (excludes lodgers, paying guests etc.) .................................. 1
less than £600 p.a. .................................. 29
£600-899 .................................. 32
£900-1299 .................................. 50
£1300-1999 .................................. 149
£2000-2999 .................................. 110
£3000-4499 .................................. 56
£4500 or more .................................. 17
Refused .................................. 60
DK/missing .................................. 97

SHOW CARD B

14 Here is a card showing typical household incomes. Which category most nearly represents your total household income — from all sources and before taxes?

(If required define household as all members supported by stated income)
15 Are you personally working nowadays for pay either for yourself or somebody else?

- Yes - full-time ........................................ 245
- part-time ........................................ 73
- No - retired* ........................................ 56
- housewife* ........................................ 203
- student* ........................................ 15
- Other* ........................................ 4
- missing(4)

If Other (state) ........................................

II WORKING
16 What is your main occupation?

Probe if necessary: Tell me a little more about what you do exactly.

17 I'd like to find out how much time you spent here during the last four weeks.
(Note: EXCLUDING holiday periods)

For each of the following periods, will you tell me whether you were usually here, or usually away from here?

18 (a) What different kinds of noise have you heard around here during the last four weeks?
Tick Box A for noises MENTIONED SPONTANEOUSLY-do not probe or prompt.

Show CARD A

For each noise MENTIONED ask-How bothered are you by the...........noise?
Enter score in Box B (Don't know=0)

(c) For the REMAINDER of list ask- Have you heard any...........during the last four weeks?
If NO- Score 8 in Box B.

OTHERWISE-Show CARD A and ask- How bothered are you by the...........noise?
Enter score in Box B (Don't know=0)

19 (a) If HOUSEHOLD noise not mentioned spontaneously at Other in Q18(a) ask- What about the noises from inside your home? Do you have any noisy appliances or equipment?

If NO- Score 8 in box at (b) over page and go to filter before Q20

If YES- State.
(b) How bothered are you by the amount of noise made by (all) these?

If Aircraft NOT heard at Q18 (i.e. Score 8 in Box B) TERMINATE INTERVIEW NOW.

OTHERWISE - Show CARD C

I would like to ask you some more questions about the Aircraft noise. Please look at this different card and tell me how much the aircraft noise has bothered or annoyed you during the last few weeks.

Ask for EACH item below

21 (i) During the last few weeks did the aircraft ever.............?

(ii) If YES: Show CARD C and ask:
When they.............how annoyed did this make you feel?

(a) Startle you? (b) Wake you up? (c) Interfere with LISTENING to TV, radio or records? (d) Make the TV picture flicker? (e) Make the house vibrate or shake? (f) Interfere with conversation? (g) Interfere with or disturb any other activity?

If YES - State ONE

(h) Bother, annoy or disturb you in any other way?

If YES - State ONE

If NO to Q21(h) - REPEAT

22 What would you say is the MOST disagreeable thing about Aircraft noise? Exactly what is it about the noise you dislike most?

If NO numbers 1, 2, or 3 have been ringed in Qs 20 and 21 TERMINATE INTERVIEW.

88 terminations here (409 interviews for remainder of questionnaire)
24. Do you find any particular kind of aeroplane to be noticeably quieter than average? Which?

   No: ........................................... 164  
   Yes-Jumbo (747): .................. 106  
   Other: .................................. 54  
   Don't know/vague: ................. (75)  

If respondent appears to identify Jumbo (747) remind him of name if necessary.

If Jumbo (747) QUIETER ask:

25. How many Jumbo (747) flights do you think it would take to cause as much disturbance as one ordinary jet flight?

   One: ...................................... 3  
   Two: ..................................... 32  
   3-5: ...................................... 44  
   6-10: .................................... 5  
   11-20: .................................... 2  
   21-50: .................................... 0  
   More than 50: ......................... 2  
   Don't know: .................................. (18)  

Show CARD D

26. On the occasions when you have been here during the last four weeks, and when the aircraft noise has been at its worst, how many times would you say you have actually been DISTURBED by aircraft during the DAYTIME hours between 7am and 7pm?

   None*: .................................. 69  
   Once or twice altogether........... 42  
   Once or twice a week................ 49  
   Once or twice per day............. 43  
   3-5 times per day................... 60  
   6-10 times per day................... 42  
   10-20 times per day.................. 34  
   21-50 times per day................... 23  
   More than 50 times per day........ 24  
   Don't know: .................................. (43)

   Total: 366

Final attempt  
2nd Attempt

27. Considering all the effects of noise, how long would you say each of these daytime disturbances lasts on average? For how long are you actually distracted or interrupted? (Should have been asked 409-69=340 times)

   ..x...x...x...x...x...x...x...x...x... Less than 10 seconds............. 33  
   ..x...x...x...x...x...x...x...x...x... 10 seconds to less than 20 seconds 44  
   ..x...x...x...x...x...x...x...x...x... 20 seconds to less than 45 seconds 31  
   ..x...x...x...x...x...x...x...x...x... 45 seconds to less than 1½ minutes 51  
   ..x...x...x...x...x...x...x...x...x... ½ minutes to less than 3½ minutes 63  
   ..x...x...x...x...x...x...x...x...x.... 1½ minutes to less than 3½ minutes 5  
   ..x...x...x...x...x...x...x...x...x.... 3½ minutes to less than 8 minutes 15  
   ..x...x...x...x...x...x...x...x...x.... 8 minutes to less than 18 minutes 2  
   ..x...x...x...x...x...x...x...x...x.... 18 minutes to less than 40 minutes 4  
   ..x...x...x...x...x...x...x...x...x.... 40 minutes or more or continuous 3  
   Don't know: .................................. (32)

   Total: 308

Total combinations 358

If percentage NOT FOUND in circled column go to Q29

If percentage FOUND continue:

28. Your answers indicate that during the daytime you are actually disturbed for more than ....... per cent of the time (if MORE than 100% - i.e. continuously).
   Do you think that is correct?

   Yes*: ........................................ 3  
   No: ......................................... 28  
   Unsure/don't know: ....................... 0  

If NO or DON'T KNOW ask: Then will you please reconsider your answers to the question...................... (REPEAT Qs26 and 27)
29 What about the evening hours between 7pm and 11pm? On the occasions when you have been here during the last four weeks, how many times on average would you say you have been disturbed by aircraft noise during the evenings, again when it has been at its worst?

- None
- Once or twice altogether
- Once or twice per week
- Once or twice per evening
- 3-5 per evening
- 6-10 per evening
- 11-20 per evening
- More than 20
- Don’t know

Total 375

30 Considering all the effects of the noise, how long would you say EACH of these EVENING disturbances lasts? Again, for how long are you actually distracted or interrupted? (Should have been asked 409 – 80 = 329 times)

- Less than 10 seconds
- 10 seconds to less than 20 seconds
- 20 seconds to less than 45 seconds
- 45 seconds to less than 1 minute
- 1 minute to less than 3 minutes
- 3 minutes to less than 8 minutes
- 8 minutes to less than 18 minutes
- 18 minutes to less than 40 minutes
- 40 minutes or more

4 5 6 7 8

Don’t know

Total 301

Total combinations 360

If percentage NOT FOUND in circled column go to Q32
If percentage FOUND continue.

31 According to your answers you were actually disturbed for more than _______ per cent of the evening period (if MORE than 100% - i.e. continuously). Do you think this is correct?

Yes
No
Unsure/don’t know

Total 301

32 Again during the last four weeks when the noise has been at its worst, how often have you been disturbed at NIGHT, between the hours of 11pm and 7am?

- None
- Once or twice altogether
- Once or twice per week
- Once or twice per evening
- 3-5 per day
- 6-10 per day
- 11-20 per day
- More than 20 per day
- Don’t know

Total 378
33 And how long would you say each night-time disturbance lasts on average?

For how long are you prevented from sleeping each time you are disturbed?

(Should have been asked 409-229 = 180 times)

<table>
<thead>
<tr>
<th>Duration</th>
<th>18</th>
<th>22</th>
<th>28</th>
<th>33</th>
<th>37</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-20 sec</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-45 sec</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-1½ min</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1½-3½ min</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3½-8 min</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8-18 min</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-40 min</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-90 min</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90 min+</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Don't know</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total combinations 365

If percentage NOT FOUND in circled column go to Q35

If percentage FOUND continue...

34 According to your answers you were certainly kept awake for more than...... per cent of the night (if MORE than 100% - in other words you didn't sleep at all). Do you think this is correct?

Yes*.......................................................... 1
No.............................................................. 12
Unsure/don't know ................................ 0

*Go to Q35

35 Regardless of the time of day, how long do you think you could actually hear each aircraft for, on average, if you listened and paid close attention to it?

<table>
<thead>
<tr>
<th>Duration</th>
<th>14</th>
<th>24</th>
<th>31</th>
<th>37</th>
<th>57</th>
<th>119</th>
<th>54</th>
<th>41</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 10 sec</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 sec to less than 20 sec</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 sec to less than 45 sec</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45 sec to less than 1½ min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1½ min to less than 3½ min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3½ min to less than 8 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 min to less than 18 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 min to less than 45 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More than 45 min</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Don't know</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

42

36 When do you find the noise of an aircraft MOST disturbing around here?

<table>
<thead>
<tr>
<th>Time</th>
<th>42</th>
<th>133</th>
<th>172</th>
</tr>
</thead>
<tbody>
<tr>
<td>During the night when you are trying to sleep</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>During the evening or</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>During the daytime</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Don't know/not disturbed</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

43

Code and ALSO circle response at Q38-A (next page)

37 And during which period would you say the noise of an aircraft is LEAST disturbing?

<table>
<thead>
<tr>
<th>Period</th>
<th>233</th>
<th>27</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>Night-time</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evening</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daytime</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Don't know</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

44

Total combinations 365

Code and ALSO circle response at Q38-B (next page)
43·

..42

As you probably realise, air travel is

41

would

40

Mentioned

something about the

39

country's economy and our way of life,

38

becoming increasingly important to the

flight

this year than in the past, or have you

37

would be as bad as one (A) flight?

36

some idea of the difference, could you say

35

If

34

when you hear the aircraft fly overhead

33

and will doubtless continue to expand.

32

Considering this, what in your opinion

31

there is a danger they

30

will you be prepared to accept in

29

exchange?

28

in other words, if you could swap each (A)

27

flight for a number of (B) flights, how many

26

would you be prepared to accept in

25

exchange?

24

On the whole would you say that in general

23

you have been more bothered by aircraft

22

this year than in the past, or have you

21

become more used to them?

20

When you hear the aircraft fly overhead

19

do you EVER feel there is a danger they

18

might crash nearby?

17

If YES

16

Would you say you feel this

15

Read

14

Very often

13

Fairly often

12

Only occasionally

11

Don’t know

10

As you probably realise, air travel is

9

becoming increasingly important to the

8

country's economy and our way of life,

7

and will doubtless continue to expand.

6

Considering this, what in your opinion

5

should be done about the noise problem?

4

Mentioned SPONTANEOUSLY

3

Other

2

If Other (state)

1

Don’t know

0

If the authorities really wanted to do

9

something about the problem yet could not

8

reduce the noise itself, which alternative

7

would you choose if the authorities offered?

6

Select ONE only

5

Code for Serial Number

4

Odd

3

Even
Suppose that the payment of compensation turned out to be the only practical solution to the noise problem, how much PER YEAR would be a fair and satisfactory payment for this household, for the amount of noise suffered here, either in the form of a direct payment or indirectly as a reduction of rates or rent?

If DON'T KNOW or VAGUE answer ask —
(ODD Serial Numbers start at £2 and proceed up)
(EVEN Serial Numbers start at £200 and proceed down)

Do you think £ ............... per year would be fair and satisfactory?

* ODD start here and code first YES
  * £2 per year.............................
  * £5 per year..........................
  * £10 per year..........................
  * £20 per year.........................
  * £50 per year.........................
  * £100 per year.......................

** EVEN start here and code first NO
  ** £200 per year.....................
  ** More than £200 per year........

If NO to £200, determine amount and code 8, 9, or 0

State (if possible) £ .............
Compensation not acceptable ....
Don't know..........................

Total ane\$ odd even

spont prompted

---

Unfortunately the cost of solving the Aircraft noise problem could run into hundreds of millions of pounds. This money could be raised by the Government or the air transport industry, but we would all be paying in one way or another through higher taxation or living costs. To give me some idea of how much you value peace and quiet yourself, would you tell me how much more per year you would be willing to pay to keep this area COMPLETELY free of Aircraft noise?

If DON'T KNOW OR VAGUE answer ask —
(ODD Serial Numbers start at More than £200 and proceed downwards)
(EVEN Serial Numbers start at More than £2 and proceed upwards)

Would you be willing to contribute ..............

* ODD start here and code first YES
  * More than £200 per year ........
    State (if possible) £ ..........
  * More than £100 per year ........
  * More than £50 per year ..........
  * More than £20 per year ........
  * More than £10 per year ..........
  * More than £5 per year ...........

** EVEN start here and code first NO
  ** More than £2 per year .........
  ** Less than £2 per year .........

If NO to £2 determine amount and code 8, 9, or 0

Nothing..............................
Don't know..........................

Was answer prompted?

Spontaneous reply............
Prompted reply ..............

Thank you for your co-operation.

NOW COMPLETE PAGE ONE.
APPENDIX C

GUTTMAN ANNOYANCE SCALE

Guttman Scaling is a device widely used in psychological scaling problems. A full description of it is not appropriate here - the reader must be referred to Reference 18. The basis of the procedure is a series of questions each of which can be arranged to have a yes-no answer. The questions form a Guttman scale if it is possible to tell to which questions a person answered yes in merely by counting his number of yeses. This is the same as saying that the questions can be rank ordered, such that a person will answer yes to all the questions above a certain rank, and no to those below. This order will be called the hierarchical order of the questions.

The questionnaire that forms the subject of this report uses virtually the same Guttman Annoyance Scale devised by McKennell and used in the Heathrow surveys of 1961 and 1967 (7,5). The Guttman Scale questions are numbers 20 (Bother); 21(b) (Wake Up); 21(c) (T.V. Listen); 21(e) (House Shakes); 21(f) (Conversation); 21(g) (Other).

A respondent can score 1 or 0 on each of these questions, and can thus score a total between 0 and 6. The scoring is as follows:

<table>
<thead>
<tr>
<th>item</th>
<th>response code</th>
<th>score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bother</td>
<td>1 2 3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4 0</td>
<td>0</td>
</tr>
<tr>
<td>Wake Up</td>
<td>1 2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3 4 0</td>
<td>0</td>
</tr>
<tr>
<td>T.V. Listen</td>
<td>1 2 3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4 0</td>
<td>0</td>
</tr>
<tr>
<td>House Shake</td>
<td>1 2 3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4 0</td>
<td>0</td>
</tr>
<tr>
<td>Conversation</td>
<td>1 2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4 0</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>1 2 3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4 0</td>
<td>0</td>
</tr>
</tbody>
</table>
The crucial tests of the performance of the scale are twofold. First the Coefficient of Reproducibility (C of R), which is a percentage measure of the proportion of true scale responses (as opposed to error responses that cannot be scaled properly) to total responses. Guttman suggests that the C of R should exceed 90% for a good scale (18). Secondly, we may examine the hierarchical order of the Guttman questions that must be used in order to give this highest C of R. In the present survey the C of R is 91% when this order is used:

Bother, T.V. Listen, Conversation, House Shake, Wake Up, Other.

This is quite different from the best order found in the 1967 survey:

Bother, Other, House Shake, T.V. Listen, Conversation, Wake Up.

It also differs from the orders found in other surveys performed by Loughborough University (4,17) as indicated by the following table.
<table>
<thead>
<tr>
<th>Reference</th>
<th>Maximum C of R (with source)</th>
<th>Hierarchical Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>96% (McKennell quote)</td>
<td>Bother: W. Up, Convers., T.V., H. Shake, Other</td>
</tr>
<tr>
<td>5</td>
<td>85% (Sample 200)</td>
<td>Bother: Other, H. Shake, T.V., Convers., W. Up</td>
</tr>
<tr>
<td>This Study</td>
<td>91% (Sample 136)</td>
<td>Bother: T.V., Convers., H. Shake, W. Up, Other</td>
</tr>
<tr>
<td>4</td>
<td>89% (Sample 100)</td>
<td>Bother: T.V., Convers., Other, W. Up, H. Shake</td>
</tr>
<tr>
<td>17</td>
<td>92% (Sample 100)</td>
<td>Bother: W. Up, Convers., T.V., H. Shake, Other</td>
</tr>
</tbody>
</table>
1) **Pearson Product-Moment Correlation**

For a population of variable pairs $x_i, y_i$ the correlation coefficient is given by

$$
\rho = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\left( \sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2 \right)^{1/2}} \tag{D1}
$$

where $\bar{x}$ and $\bar{y}$ are the mean values of $x_i$ and $y_i$. The sample coefficient is

$$
P = \frac{\sum x_iy_i - \sum x_i\sum y_i/n}{S_x S_y} \tag{D2}
$$

where $n$ is the number of pairs of data $x_i, y_i$, and $S_x, S_y$ are the sample standard deviations of $x_i$ and $y_i$. This is invariant under changes of scale and location in $x$ and $y$.

**Confidence Interval for $\rho$**

To test the hypothesis that the population correlation coefficient $\rho = 0$, we may use the sampling distribution of $P$ for the case $\rho = 0$. This is tabulated in most statistics texts.

However, this table cannot be used to test the hypothesis that $\rho$ equals some non-zero number. In this case, we may use the transformation

$$
w = \frac{1}{2} \log_e \frac{1 + P}{1 - P} \tag{D3}
$$

which has a near-normal sampling distribution with mean
\( \frac{1}{2} \log \left[ \frac{1+\rho}{1-\rho} \right] \) and standard deviation \( S = (n-3)^{-\frac{1}{2}} \).

To test the hypothesis that two samples, with correlation coefficients \( P_1 \) and \( P_2 \) are drawn from the same population we may then use the statistic

\[
z = \frac{w_1 - w_2}{\left( S_1^2 + S_2^2 \right)^{\frac{1}{2}}}
\] (D4)

which is normally distributed with unit standard deviation.

The above tests are only valid for samples with a bi-variate normal distribution; e.g. for bi-variate normal populations the result \( \rho = 0 \) indicates zero correlation. For unknown non-normal distributions we require statistics which are distribution free, such as Spearman's Rank Correlation Coefficient.

2) **Spearman Rank Correlation**

Spearman's Rank order sample correlation coefficient for tied data is (19)

\[
R = \frac{n(n^2-1) - 6 \sum_{i=1}^{n} D_i^2 - 6(t' + u')}{\left[ \left( n(n^2-1) - 12t' \right) \left( n(n^2-1) - 12u' \right) \right]^{\frac{1}{2}}}
\] (D5)

where \( n \) = total number of pairs of data \( (x_i, y_i) \)

\[ u' = u(u^2-1)/12 \]

\( n \) is the number of tied observations in each group and the summation is extended over all sets of tied ranks in the \( Y \) sample.

\( t' \) is the corresponding sum for the \( X \) sample

\[ D_i = R_i - S_i \quad \text{where} \quad R_i = \text{rank} (x_i) \]

\[ S_i = \text{rank} (y_i) \]

In any group of tied data, each observation is assigned the same rank. In each group of \( u \) tied observations, which, if not tied would be assigned the ranks \( p_k + 1, p_k + 2, \ldots, p_k + n \), the rank assigned to all is
\[
\sum_{i=1}^{u} \frac{p_k + 1}{u} = p_k + \frac{u + 1}{2}
\]

Gibbons (19) states that unless the ties are extremely extensive they will have little effect on the value of R. Thus, in practice the common expression

\[
R = 1 - \frac{6 \sum D_i^2}{n(n^2 - 1)}
\]  \hspace{1cm} (D6)

(to which Equation D5 reduces when \(u' = t' = 0\)) is often used without correction for ties. Note that the effect of correlation is to decrease value of R, i.e. a negative R is closer to \(-1\), not to zero.

3) **Sampling Distribution of R**

(a) **Null distribution** (Samples from uncorrelated variables)

For large \(n\) \((n > 20\) or so\) R is normally distributed with variance

\[
\frac{1}{n - 1}
\]

For small \(n\), the variance is very difficult to ascertain and has only been worked out for \(n\) up to 13(20).

For large \(n\), the above result can be used to test the null hypothesis that variables are uncorrelated to any desired degree of significance. The probability that a particular value of R arose by chance (for uncorrelated variables) is indicated in Figure D1.

(b) **Non-null case** (Samples from correlated variables)

For any present population the distribution of R tends to normality as \(n\) increases provided \(r\) is not too near
unity. But variance of $R$ depends upon (unknown) quantities other than $r$ and it can only be stated that:

$$\text{variance of } R \leq \frac{3}{n} \left( 1 - r^2 \right) \quad \text{(D7)}$$

4) **Test for significant difference between two (sample) rank correlation coefficients $R_1$ and $R_2$**

We wish to test the hypothesis that the two populations 1 and 2 have the same means (variances not equal). We have estimates $R_1$ and $R_2$ of the two population Spearman Rank Correlation coefficients $r_1$ and $r_2$ and maximum values of the variances $S_1^2$ and $S_2^2$ of the sampling distributions of $R_1$ and $R_2$ given by Equation D7. The hypothesis can be tested at any level of significance using the statistic

$$z = \frac{R_1 - R_2}{\left[ S_1^2 + S_2^2 \right]^{1/2}} \quad \text{(D8)}$$

which is normally distributed with mean 0 and variance 1. Note that this will be a conservative estimate since the true values of $S_1^2$ and $S_2^2$ may be less than the maximum values. Thus, we minimise the risk of attaching significance to non-significant results, though we may, in some cases, fail to discern significance where it really exists.

Figure D2, based on Equations D7 and D8, shows the differences $\Delta R$ between two sample correlation coefficients $R_1$ and $R_2$ which might arise by chance (at various levels of significance) from a sample of pairs of variables taken from a population with correlation coefficients $\bar{r}$.

5) **Effect of (0,0) pairs on $R$**

Spearman's Rank Correlation Coefficient for tied data is given by Equation D5 and for no ties by Equation D6. For large $n$ we may rewrite Equation D6
\[ R \approx 1 - \frac{\Delta}{n^3} \quad \text{(D9)} \]

where \( \Delta = 6 \sum_{i=1}^{n} D_i^2 \)

If we add \( n' \) observations at \( x = y = 0 \), Equations D5 gives:

\[
R \approx \frac{(n+n')^3 - \Delta - 6 \left\{ \frac{n'^3}{12} + \frac{n'^3}{12} \right\}}{\left[ \left\{ (n+n')^3 - 12 \frac{(n')^3}{12} \right\} \right]^2}
\]

\[ = 1 - \frac{\Delta}{n^3 + 3n^2n' + 3nn'^2} \quad \text{(D10)} \]

where \( \Delta \) remains unchanged.

Combining D9 and D10 we obtain

\[
R' = R + 3\frac{\Delta}{n^3} \left\{ \frac{nn' + n'^2}{n^2 + 3nn' + 3n'^2} \right\} \quad \text{(D11)}
\]

from which it may be seen that as \( n' \) increases, \( R \) tends to unity. Figure D3 shows this tendency for

\[ R_n = .2, .4, .6 \text{ and } .8 \]

The ordinate \( K \) is the variable in

\[ n' = Kn \]

using which, D11 becomes

\[
R' = R + \frac{\Delta}{n^3} \left[ \frac{3(K + K^2)}{1 + 3(K + K^2)} \right] \quad \text{(D12)}
\]
FIGURE D1  PROBABILITY (LEVEL OF SIGNIFICANCE) THAT A GIVEN R MAY ARISE BY CHANCE FOR A SAMPLE OF PAIRS OF UNCORRELATED VARIABLES
FIGURE D2
DIFFERENCE $\Delta R$ BETWEEN TWO SAMPLE RANK CORRELATION COEFFICIENTS WHICH ARISES WITH PROBABILITY $p$ FROM VARIABLES WITH POPULATION COEFFICIENT $\tau$
FIGURE D3  EFFECT OF ADDING 0,0 POINTS UPON SPEARMAN RANK CORRELATION COEFFICIENT R

R = Correlation for n data pairs with no 0,0 points
R' = Correlation with Kn 0,0 points added
REFERENCES


VARIATION OF COMMUNITY RESPONSE TO AIRCRAFT NOISE WITH TIME OF DAY


This paper compares some results of the previous chapter with those of other research and relates them to current planning practice.
Variation of Community Response to Aircraft Noise with Time of Day

Existing composite noise indices add a 10-dB penalty to sound heard during the night, on the assumption that this reflects a nighttime increase in public sensitivity to noise. Quantitative evidence to support the use of this penalty is fragmentary and the data reviewed by J. B. Ollerhead lead to the conclusion that for predicting aircraft noise nuisance, it would be more logical to apply a smaller weighting to an extended evening period only.

Effective environmental planning and noise control rely on an ability to measure noise in terms which relate to its acceptability to the public. In the USA, this requirement has led to a variety of composite noise indices, such as Noise Exposure Forecast (NEF), Community Noise Equivalent Level (CNEL), and Day-Night Sound Level (L_{dn}), which take account of variables believed to contribute to the adverse effects of long-term noise. One of these variables is that community sensitivity to noise varies with the time of day; accordingly, each of the three indices applies a 10-dB penalty to any noise which occurs during the night. Logically, this implies that at night, sound is judged as though it were 10 dB more intense than it really is, or (since the three scales are based on the principle that noise intrusion is a function of total noise energy) ten times as long or ten times as frequent.

In theory, the implications of this penalty are far-reaching. For example, according to \( L_{dn} \) methodology, one aircraft departing at 11:05 PM is as bad for the airport neighbors as ten aircraft departing between 10:45 and 10:59 PM. In practice, noise exposure tends not to follow such a precise pattern in any regular way and, largely because nighttime noise exposure levels are usually fairly low anyway, the 10-dB penalty is never put to a severe test. However, this does not mean that it never will be; since very long term outcomes of planning decisions may well depend on the broad validity of composite noise indices, it is important that although built-in penalties cannot be based soundly on scientific principles, they at least represent the best possible guesses.

The "Levels Document" states that the adoption of a 10-dB nighttime weighting in the \( L_{dn} \) formula "was predicated on its extensive prior usage, together with an examination of the diurnal variation in environmental noise." More specifically, von Gierke identifies three factors which support the view that intrusive noise events are more disturbing at night:

1. Most community response and public opinion surveys reveal that the same noise environment is considered more disturbing during nighttime than daytime.
2. Not only do the requirements for undisturbed sleep and relaxation make a lower noise level desirable, but the exterior background noise levels drop by 10 dB or more during the night in most communities.
3. The reduced activity inside homes contributes to the general lowering of noise levels there.

That the second and third factors combine to make intruding sound more noticeable cannot be disputed. With regard

*Received 9 June 1977; revised 30 June 1978
†Department of Transport Technology, Loughborough University, Leicestershire LE11 3TU, England

NOISE CONTROL ENGINEERING / September-October 1978
to the first. Galloway observed that although there is no strong evidence to contradict a 10-dB penalty. "solid data to support actual choice of numbers is hard to come by."³

As will be seen, the development of composite noise indices has been more of an art than a science, mainly because of the very large differences in individual reactions to noise but also because of the difficulties of putting numbers to these reactions. The day-night question is very important; unfortunately, it has been particularly difficult to solve. In the UK, engineers have fought shy of the problem; there is not yet any kind of weighted index in common use, and daytime noise and nighttime noise are treated as separate issues. This practice will probably continue until more convincing evidence on day-night differences becomes available.

In what follows, the words annoyance and disturbance, which are frequently used, have different connotations. As is usual, annoyance refers to the general adverse feeling of displeasure evoked by noise, whereas disturbance relates to the interruption of or distraction from some activity (including relaxation and sleep). Such disturbances need not cause annoyance, but it can be hypothesized that annoyance is caused by disturbance (although a casual relationship may not be found in practice because of shortcomings in the procedures used to make the subjective measurements).

The data reviewed come from studies of aircraft noise nuisance, and it would therefore be wrong to generalize the conclusions. At the same time, it should be recognized that much of our knowledge of community noise impact stems from aircraft noise research and existing noise indices have been strongly influenced by the resultant data.

### Origins of the 10-dB Weighting

The concept of a community noise index was pioneered by Rosenblith and Stevens in 1952, and their Composite Noise Rating (CNR) was described in the open literature in 1955.⁴ This was the basis of an ambitious scheme for predicting likely origins of both acoustic and sociopsychological factors. With regard to the day-night question, it was stated that "During the day-time many people are away from their residences and do not hear the noise. Residents who stay near their homes are often engaged in activities that are not greatly disturbed by moderate noise levels. In the evening and at night however, the noise tends to interfere with relaxation and sleep. We expect therefore that a noise of a given level rank will produce a more severe response if it occurs at night than if it occurs only in the day-time."

Based on empirical evidence involving a number of case histories of community noise problems, a reduction of 5 dB in the effective noise level was assumed if the noise was heard during the daytime only. The adoption of this nighttime weighting (or rather, negative daytime weighting) appears to have become the main foundation for a series of similar devices which are directly descended from it.

The 10-dB weighting appeared in 1957, when Stevens and Pietersanta adapted the CNR procedure for application to the rapidly growing problem of aircraft noise.⁶ Among various other modifications, the 24-hour day was split into day, evening, and night periods, with assigned weightings of -5, 0, and +5 dB respectively. No explanation was offered for this apparently arbitrary decision. When the modified version of CNR was subsequently revised to its existing form, the time-of-day effect was again restricted to two categories, day and night, but the 10 dB day-night penalty was retained.⁷ It was also retained in essentially unchanged form in the alternative procedure known as Noise Exposure Forecast (NEF), developed for the US Federal Aviation Administration (FAA) in 1967 as a guide for land-use planning and zoning around airports.⁸

On the other hand, the state of California based its aircraft noise legislation on a new index, Community Noise Equivalent Level (CNEL), which again split the day into day, evening, and nighttime periods with effective weightings of 0, 5, and 10 dB respectively.⁹ The justification for including these weightings was "to account for the increased need for quiet in residential areas at night."¹⁰

Outside the USA, such countries as France, Germany, the Netherlands, and South Africa developed similar aircraft noise rating procedures which accounted for diurnal variations in public reaction to noise, and a generalized method was eventually recommended by the International Civil Aviation Organization (ICAO).¹¹ However, it seems likely that most of these procedures were based more on the prototype American indices than on new experimental evidence that weightings were required.

In the USA, the selection of a relatively simple day-night model for the general noise index $L_{eq}$, was based on the conclusion that for typical time variations of environmental noise level, the two-period and three-period models give results which agree to within fractions of a decibel. The retention of a 10-dB nighttime penalty was supported by an analysis of "55 community reaction cases presented in the EPA report to Congress of 1971."¹² It was stated that these data have a standard deviation of 3.3 dB when a 10-dB nighttime penalty is applied, but that the correlation worsens (standard deviation = 4.0 dB) when no nighttime penalty is applied. It was also pointed out that little difference was observed among values of weighting between 8 and 12 dB. This is one of the few fragments of numerical support for the 10-dB penalty, but it should be noted that according to a standard F-test, the difference between the above two standard deviations is not significant at the 5% level.

### Social Survey Annoyance Data

**London airports.** It has been suggested that the 1961 survey of aircraft noise nuisance at London (Heathrow) Airport
supports the use of a substantial night weighting by showing that noise exposure from nighttime aircraft operations needs to be some 17 dB lower than daytime levels to achieve a comparable level of community response. In fact, this increment was the estimated difference between typical daytime (0800 to 2300) and nighttime (2300 to 0800) values of Noise and Number Index (NNI) then in existence. Since 28% of the respondents said that they were most bothered by aircraft noise at night and 24% were most bothered during the day (the remainder either were not bothered at all or did not directly discriminate between these two periods), it was assumed that daytime and nighttime noise exposures made roughly equal contributions to evoked annoyance. Superficially, this seems logical enough. However, one might wonder what conclusion would have been reached if the survey had been performed at an airport where the noise exposure difference was not 17 NNI and/or where the day-night percentages were very different. Certainly, it is not possible to infer from these results what percentage of people would be most bothered at night if, for example, the day and night NNI values were equal.

It must also be pointed out that of the 24% most bothered during the "daytime," 19% were referring specifically to the evening period between 6:00 and 11:00 PM (that is, only 5% were most bothered during the day between 8:00 AM and 6:00 PM).

In two subsequent surveys, attempts were made to throw more light on the day-night problem. In the 1967 Heathrow study and in an unpublished survey performed at London (Gatwick) Airport in 1971, respondents were shown cards displaying the integers between the endpoints 1 and 7, where these were labeled "not at all bothered" and "very much bothered," respectively. In order to separate annoyance components corresponding to different times of the day, they were then asked

"Look at this scale and pick out the number which indicates how bothered or annoyed you feel during the morning... and during the afternoon/evening/night."

The mean responses and standard deviations listed in Table I show that both surveys gave similar results. The differences between morning and afternoon means and between evening and night means are statistically indistinguishable; however, the differences between the means for morning or afternoon and evening or night are highly significant (at the 0.05% level). These results appear to confirm that people are more annoyed by aircraft noise at night than during the day, but that the noise is equally bothersome during the evening and night.

Unfortunately, it is not clear whether respondents were describing the cumulative effects of noise during each of the time periods or the response evoked by individual aircraft sounds as and when they intrude. Although it seems more likely that they would be expressing some general or continuing level of agitation (corresponding to the second response), they were given no instructions concerning this important distinction.

<table>
<thead>
<tr>
<th>Table I</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Annoyance Scores for Different Times of the Day</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Heathrow</td>
</tr>
<tr>
<td>(4699 respondents)</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Standard Deviation</td>
</tr>
<tr>
<td>Gatwick</td>
</tr>
<tr>
<td>(1030 respondents)</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Standard Deviation</td>
</tr>
</tbody>
</table>

No traffic details were reported, but it may be assumed that hourly aircraft movements were distributed roughly in the typical ratio 5:4:1 for the three periods of day, evening, and night. Thus, if the second response is true such that annoyance is dependent upon the number of flights per hour, then in terms of their capacity to evoke annoyance, aircraft are over four times more effective at night than during the daytime or evening. If, on the other hand, the cumulative effects are more relevant, the total number of flights per period (morning, afternoon, evening, or night) may be the controlling factor, and these are distributed in the approximate ratio 4:4:2:1. This would imply that one nighttime aircraft is only as annoying as two evening aircraft but more annoying than four daytime aircraft. How much more depends upon the meaning of the differences between mean annoyance scores, which cannot be ascertained.

**Los Angeles Airport.** An alternative approach was demonstrated by Fidell and Jones, who studied the effects upon community annoyance of a dramatic reduction in nighttime aircraft noise. This was an experiment in which night approaches to Los Angeles International Airport were diverted from a westerly to an easterly direction — a measure which between 11:00 PM and 6:00 AM reduced equivalent continuous noise levels by 25 to 30 dBA in residential areas to the east of the airport. Surveys of responses were made during the week prior to the change, immediately after the change, and four to six weeks after the change. In each survey the following question was asked:

"Does aircraft noise annoy you more when you are trying to sleep at night or does it annoy you more at other times during the day?"

The percentages of respondents who reported more annoyance during the daytime were (for the three surveys) 68%, 66%, and 72%. No significance can be attached to the differences between these three results and they appear to reveal a total lack of public sensitivity to the nighttime noise

*Taking morning and afternoon as six hours, evening as four hours, and night as eight hours.
TABLE II
VARIATIONS IN ANNOYANCE AT NEW YORK (JFK) AIRPORT

<table>
<thead>
<tr>
<th>Average Hourly Aircraft Movements</th>
<th>Average Annoyance Score</th>
<th>Annoyance/Movement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>6.12</td>
<td>2.24</td>
</tr>
<tr>
<td>Evening</td>
<td>8.30</td>
<td>2.83</td>
</tr>
<tr>
<td>Night</td>
<td>2.92</td>
<td>1.93</td>
</tr>
</tbody>
</table>

*(which was removed). This is supported by the responses, in the third survey, to the question.*

"Have you noticed any increase or decrease in number of flights near your home in the last month?"

These responses were distributed as follows: none noticed (or can't know) — 60%; decrease — 20%; and increase — 20%.

It is interesting that a frequent comment among responses to the latter question was "How could I have noticed, I was asleep?" Whatever the explanation for the nighttime insensitivity, these results seem to be at odds with the London results, which do indicate some increase in annoyance at night. However, results from Borsky's survey of residents living near New York's John F. Kennedy Airport lie somewhere between the two. 18

**New York (JFK) Airport.** As in the London study, respondents to this survey were asked to quantify separately their feelings of annoyance during day (0700 to 1900), evening (1900 to 2300), and nighttime (2300 to 0700) periods. The average responses, in relation to average aircraft movements, are listed in Table II.

On the basis of the figures for annoyance/movement, Borsky suggested that "each nighttime flight has the equivalent annoyance effect of two day or evening flights" and that a widely used 10-dB penalty, which implies a 10:1 ratio, is therefore "much too high."

**Discussion.** It is difficult to draw conclusions from a comparison of the annoyance data examined. The London studies (Table I) indicate that annoyance is higher during the evening and night than during the day. The Los Angeles experiment suggests that there is very little awareness of nighttime aircraft traffic at all. Finally, the New York figures in Table II show lower annoyance at night, but they can be interpreted to infer a higher annoyance at night.

Two uncertainties hamper interpretation of these confusing observations. The first is that annoyance, depending upon how it is measured, may or may not be a cumulative reaction; that is, it may represent a degree of reaction evolved by a single event or it may be some kind of overall reaction to all events during the time period in question. Indeed, the reaction may be influenced by the experience of events over a much longer period of time altogether (including different periods of the day). The assumption underlying composite noise indexes of the $L_{eq}$ type — that annoyance can be related to a time average of the noise — must itself presuppose that annoyance too can be time-averaged. For our purposes, therefore, it is instantaneous (or short-time average) annoyance $A(t)$ which we seek. Fidell and Jones' question, "Does aircraft noise annoy you more at night ... or at other times ... ?" may tap time-integrated annoyance $\int A(t)dt$, while the London survey question referring to "how bothered or annoyed you feel during the morning ... " taps the ongoing annoyance $A(t)$. This would explain the apparent lack of nighttime reaction in Los Angeles, since the duration of nighttime annoyance could be very small.

Borsky's interpretation of the New York results (Table II) implies that the measured annoyance is either cumulative and directly proportional to the number of aircraft, or an average value which is proportional to frequency of events. In either case, a linear relationship seems somewhat doubtful and is inconsistent with the common assumption that annoyance is proportional to the logarithm of number or frequency; with regard to this discussion, the matter is of secondary importance.

The second uncertainty concerns the time divisions chosen to represent day, evening, and nighttime periods. It would be difficult to propose a more logical three-way division — assuming that the periods selected are appropriate to local customs and habits — since three distinctly different activities can be associated with them (work, leisure, and sleep). Yet the mechanisms of noise intrusion change markedly with the onset of sleep and, in particular, maximum noise sensitivity might coincide with that relatively brief period when people are trying to get to sleep. Some of the differences in the survey results may well be attributable to slight differences in noise exposure patterns during that period which barely affect the noise variables but which have a profound influence upon reported annoyance.

**A Further Study at London (Heathrow) Airport**

**Annoyance reactions.** Further evidence concerning the time-of-day question was sought in a later (pilot) survey performed in the environs of London (Heathrow) Airport in 1972, during which a questionnaire was administered to 600 residents. 19 In addition to questions concerning the relative intrusiveness of aircraft noise at different times of the day and night, they also included questions which were used in the two previous Heathrow surveys to construct a Guttman scale of annoyance.

The main purpose of the survey was to test the questionnaire, and no attempt was made to obtain a population sample which was fully representative of any segment of the community. Instead, an adequate range of aircraft noise exposures (in terms of Noise and Number Index), road traffic noise level, and socioeconomic status was provided. It was found that the distributions of age, sex, and socioeconomic status within the sample were very similar to those achieved previously.
To estimate the most relevant aircraft noise characteristics at each residence, respondents were asked to confine their attention to conditions during the four weeks immediately preceding the interview. Aircraft noise variables estimated at the exterior of each residence on the basis of known aircraft movements during that period included Noise and Number Index, the number and average duration of aircraft sounds as a function of the level exceeded, and equivalent continuous level. These were calculated separately for the three periods of interest: day (0700 to 1900), evening (1900 to 2300), and night (2300 to 0700). Relative to some of these variables, the survey sample was distributed as shown in Tables III and IV.

For any period, Noise and Number Index is given by the formula

\[ NNI = L_{PN} + 15 \log_{10} N - 30, \]  

(1)

where \( N \) is the number of sounds whose peak levels exceed 80 PNdB and \( L_{PN} \) is the (energy) average of these \( N \) peak levels in PNdB.

The mean relationship between the daytime, evening, and nighttime NNI variables were

- NNI (evening) = NNI (day) - 13 \quad (\text{rank correlation } R = 0.99^*),

- NNI (night) = NNI (day) - 17 \quad (R = 0.95).

(It is interesting that the latter result is identical to that estimated for the 1961 survey by the Wilson Committee.\(^{14}\))

Since the evening and nighttime values of NNI are relatively small, the daytime \( L_{en} \) value is practically identical to the day-night sound level \( L_{an} \), and an approximate transformation was found to be

\[ L_{an} \approx L_{en} \text{ (day)} = 0.61 \times \text{NNI (day)} + 43, \text{ dB(A)} \quad (R = 0.95). \]  

Survey techniques used to attach numerical values to feelings of annoyance include cumulative scaling methods in which answers to a series of questions concerning various effects of noise in and around the home are analyzed to yield a numerical score. A Guttman scale is a particularly elaborate procedure of this kind. As in the earlier studies,\(^{14,19}\) the questions included in this survey concerned annoyance caused by startle, sleep disturbance, interference with conversation and television reception, and other disturbances.

Analysis showed that the pattern of responses to these questions again met Guttman scaling criteria, allowing reliable annoyance scores to be computed on a scale from 0 (no annoyance) to 6 (maximum annoyance).\(^{11}\) Respondents who stated that they did not hear aircraft noise (and therefore were not asked the relevant annoyance questions) were assigned a zero score. Fig. 1 compares group mean annoyance scores as functions of daytime NNI. A linear regression line is fitted and the product-moment correlation coefficient is 0.98. The collapse of the three sets of data indicates that reactions to aircraft noise in the vicinity of London Airport have remained fairly uniform over a period of several years (the 1972 results for 20 to 30 NNI have been omitted due to small sample size).

The vertical lines in Fig. 1 indicate the scatter of individual annoyance scores about the mean values in terms of ±1 standard deviation (even though the distributions are skewed at high and low noise levels where the lines have been truncated). That these are rather large is apparent from the fact that for the 1972 survey, the rank correlation coefficient relating daytime noise exposures (NNI) and individual annoyance scores is only 0.597. The corresponding correlations when the daytime NNI values are replaced by evening and nighttime values are 0.628 and 0.533 respectively. However, the fact that the differences between these coefficients are not significant at the 5% level and that the correlations between the day, evening, and nighttime noise variables are all in excess of 0.940 illustrates the difficulty of unraveling the relative contributions of noise exposure during the three periods to general noise annoyance.

**Disturbance reactions.** This difficulty may arise because the strength of an individual's annoyance reactions is the result of a long experience of noise nuisance and is thus only weakly related to hour-to-hour variations in the noise variables. To try to distinguish day, evening, and night effects more clearly, respondents were asked direct questions con-
cerning the frequency and duration of disturbances during the three periods on the grounds that these measures of noise nuisance would not necessarily be directly identified with feelings of annoyance. For the daytime period, the first question was:

"On the occasions when you have been here during the past four weeks, and when the aircraft noise has been at its worst, how many times would you say that you have actually been disturbed by aircraft noise during the daytime hours between 7:00 AM and 7:00 PM?"

To simplify the respondent's task, he or she was shown a card clearly marked with the following categories from which to select a response:

- None
- Once or twice altogether
- Once or twice a week
- Once or twice a day
- 3 to 5 times a day
- 6 to 10 times a day
- 11 to 20 times a day
- 21 to 50 times a day
- More than 50 times a day.

For evening (7:00 PM to 11:00 PM) and night (11:00 PM to 7:00 AM), the last two categories were replaced by "more than 20 times per evening/night."

The first question was followed by:

"Considering all the effects of the noise, how long would you say each of these daytime disturbances lasts on average? For how long are you actually distracted or interrupted?"

In this case, the accompanying card showed a series of class intervals, again arranged in a roughly geometrical progression, from less than 10 seconds to more than 40 minutes.

Questions relating to the evening and nighttime periods were very similar. However, with regard to the duration of nighttime disturbances, respondents were asked, "For how long are you prevented from sleeping each time you are disturbed?"

The distributions of the responses to these questions with respect to the noise variables \( N \) and \( NNI \) are detailed in Tables V and VI. It is clear from these tables that the variation of responses within each class interval of noise and the percentage of respondents who were not disturbed at all are both large. The relevant rank correlation coefficients are given in Table VII.

The large scatter makes it difficult to characterize the relationship between noise and response in any simple way. Two attempts to do so are shown in Figs. 2 and 3 (from which results for subsamples of less than fifteen respondents have been omitted). Fig. 2 shows the percentage of undisturbed respondents as a function of \( N \) for the three diurnal periods. This indicates that people are least disturbed by aircraft noise at night and most disturbed during the evening. This is confirmed by Fig. 3, which compares the median number of disturbances, again as functions of \( N \). Except for the special case of no disturbance, the medians have been calculated on the assumption that each subsample of observations is uniformly distributed across the dimensions of the cell; otherwise, the disturbance variable is treated as an ordinal scale. That is, no attempt is made to depict the departures from a true geometric progression (which are greatest at the low end of the scale).

Unfortunately, the sample is rather badly distributed with respect to the class intervals of \( N \) (see Table V). Also, a downward extension of the disturbance scale (to rates of less than once or twice per month) may have helped to reveal differences between the three periods at the lower noise exposures, where all median responses converge upon zero. However, Fig. 3 does indicate that during the evening (at least at the higher flight frequencies), aircraft sound is some three or four times more likely to cause disturbance than during the daytime. During the evening, less than one out of every ten sounds disturbs the median respondent; during the day, the ratio is nearer to one in forty.

Defining the noise variable in terms of the number of peaks in excess of 80 PNdB tends to obscure the probable role of sound level itself. As \( N \) increases, so does the average level and presumably, in consequence, the probability that any sound will cause disturbance. This may be investigated by regrouping the data with respect to \( NNI \), which takes both level and number of events into account.

Fig. 4 shows that the percentages of respondents who are undisturbed during each of the three periods are more clearly
### TABLE V
**DISTRIBUTION OF 600 SURVEY RESPONDENTS BY NUMBER OF EVENTS \(N\) AND FREQUENCY OF DISTURBANCE**

<table>
<thead>
<tr>
<th>Period</th>
<th>Number of Events (N)</th>
<th>Number of Disturbances per Period per:</th>
<th>Month</th>
<th>Week</th>
<th>Day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>None</td>
<td>1 to 2</td>
<td>1 to 2</td>
</tr>
<tr>
<td><strong>Day</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0700 to 1900)</td>
<td>Under 3</td>
<td>136</td>
<td>15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3 to 5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>6 to 16</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>17 to 40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>41 to 100</td>
<td>105</td>
<td>16</td>
<td>15</td>
<td>23</td>
</tr>
<tr>
<td>Over 100</td>
<td>55</td>
<td>11</td>
<td>3</td>
<td>17</td>
<td>28</td>
</tr>
<tr>
<td><strong>Evening</strong></td>
<td>(1900 to 2300)</td>
<td>Under 3</td>
<td>147</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>3 to 5</td>
<td>54</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>6 to 16</td>
<td>48</td>
<td>12</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>17 to 40</td>
<td>40</td>
<td>8</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>41 to 100</td>
<td>20</td>
<td>4</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td><strong>Night</strong></td>
<td>(2300 to 0700)</td>
<td>Under 3</td>
<td>215</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>3 to 5</td>
<td>48</td>
<td>5</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>6 to 16</td>
<td>88</td>
<td>14</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>17 to 40</td>
<td>102</td>
<td>27</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

*Possible responses to evening and night questions excluded 51+ category and replaced 21 to 50 range by 20+.

### TABLE VI
**DISTRIBUTION OF 600 SURVEY RESPONDENTS BY NNI AND FREQUENCY OF DISTURBANCE**

<table>
<thead>
<tr>
<th>Period</th>
<th>NNI</th>
<th>Number of Disturbances per Period per:</th>
<th>Month</th>
<th>Week</th>
<th>Day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>None</td>
<td>1 to 2</td>
<td>1 to 2</td>
</tr>
<tr>
<td><strong>Day</strong></td>
<td>Under 20</td>
<td>138</td>
<td>15</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>20 to 30</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>30 to 40</td>
<td>84</td>
<td>15</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>40 to 50</td>
<td>56</td>
<td>7</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>50 to 60</td>
<td>11</td>
<td>4</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>60 to 70</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Evening</strong></td>
<td>Under 20</td>
<td>196</td>
<td>17</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>20 to 30</td>
<td>55</td>
<td>8</td>
<td>18</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>30 to 40</td>
<td>46</td>
<td>10</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>40 to 50</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>50 to 60</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Night</strong></td>
<td>Under 20</td>
<td>248</td>
<td>11</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>20 to 30</td>
<td>111</td>
<td>20</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>30 to 40</td>
<td>70</td>
<td>16</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>40 to 50</td>
<td>22</td>
<td>7</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>
The improved distribution of the sample also provides a slightly more informative plot of the median disturbance rate in Fig. 5. In this case, the scale has been extended to include a curve showing the median values of $N$ associated with the NNI groups. Also shown (right-hand scale) is the median peak level $L_{PN}$. Both $N$ and $L_{PN}$ are highly correlated with NNI ($R = 0.94$ and $0.96$ respectively). The fact that the disturbance curves (for day and evening) have greater slopes than the curve of $N$ confirms that the probability of disturbance (number of disturbances $N$) increases with average noise level. Indeed, assuming that the ordinal scale is a true geometric progression, approximate equations for the two curves drawn through the disturbance data are

$$10 \log_{10} M_D = L_{PN} + 10 \log_{10} N_D - 112.1$$

and

$$10 \log_{10} M_E = 10 \log_{10} M_D + 6.$$  

where $M$ is the median number of disturbances, $N$ is the number of sounds, and the subscripts $D$ and $E$ refer to day and evening periods. The difference of $6$ dB of course implies a ratio of four between the daytime and evening disturbance rates, reflecting the fact that people are more likely to be disturbed by noise when reading, watching television, or generally relaxing in the evening than when involved in the busy activities of the daytime.

The median disturbance rate for the nighttime period is zero at all noise levels, so no similar ratio can be established for the night. However, if the convergence of the three curves at zero disturbance is a consequence of truncating the lower end of the response scale, it may be concluded that median disturbance rate at night is possibly two or more orders of magnitude less than during the day. Indeed, any method of comparing day and night response data in Tables V and VI points to a remarkable insensitivity to nighttime noise disturbance. Of course, it is important not to lose sight of the fact that the variability is large — more than $25\%$ of respondents in the noisiest areas report being disturbed at least once or twice per night.

The survey respondents were given the opportunity to make a direct comparison between the effects of noise during three periods in response to the questions

"When do you find the noise of an aircraft most disturbing around here: during the night when you are trying to sleep, during the evening, or during the daytime?"

"During which period would you say that the noise of an aircraft is least disturbing?"

There is a marked similarity between the first question and the one concerning annoyance included in the Los Angeles survey. It is interesting that despite the disturbance/annoyance difference, the results (presented in Table VIII) are also similar to the American results. In Los Angeles $68\%$, $66\%$, and $72\%$ of respondents reported more annoyance during times other than when they were trying to sleep. From Table VIII we see that $72\%$ of the London respondents were most disturbed during the day and evening. However, the table shows that daytime noise is worse than evening noise, which is a reversal of the order displayed by the associated data plotted in Figs. 2 through 5. Table VIII also appears to

<table>
<thead>
<tr>
<th>RANK CORRELATION BETWEEN NUMBER OF DISTURBANCES AND NOISE VARIABLES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>Number of sounds ($N$)</td>
</tr>
<tr>
<td>Number Need for Insertion ($NNI$)</td>
</tr>
</tbody>
</table>
It is very difficult to separate diurnal variations of noise sensitivity from social surveys of noise annoyance. The data reviewed are not consistent; some indicate higher annoyance at night; some lower. A possible explanation is that nighttime annoyance is very sensitive to events in that short but critical period when people are trying to fall asleep. No special attention has been given to this period in the surveys, so it can only be conjectured that significant differences in airport activity during this phase, which have little or no impact in the noise exposure variables measured, cause significant variations in annoyance reactions.

An alternate approach to the problem is to question survey respondents about the frequency and duration of the disturbances they experience, rather than the annoyance they feel. The technique used allowed the effects of noise to be expressed in physical terms and provided direct comparisons of the intrusiveness of noise during day, evening, and nighttime periods.

The following hypotheses seem to explain the various observations from the different surveys, which at first sight appear rather inconsistent. In terms of disturbance or annoyance, aircraft noise is considered to be worse during the evening than during the day. As a rough quantification, one evening aircraft is equivalent to four daytime aircraft.

### Conclusions

 día, evening, and 180 for night (of whom only 308, 301, and 148 respondents were able to quantify the disturbance durations). The scatter of responses is again large, but Fig. 6 clearly shows that those people who are disturbed at night are aware of the disturbance for considerably longer periods than those who are disturbed during the day and evening, the mean durations being about two minutes during the day and evening and about eleven minutes at night. Thus, in terms of the duration of the intrusion, one nighttime disturbance is on average equivalent to more than five day or evening disturbances, and it does not seem unreasonable to suppose that individual disturbances will contribute to general feelings of annoyance or vexation to a degree which depends upon their duration.

### Table VIII

<table>
<thead>
<tr>
<th>Periods When Aircraft Noise is Found to Be Most and Least Disturbing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Most Disturbing</strong></td>
</tr>
<tr>
<td>Night: 10.0%</td>
</tr>
<tr>
<td>Evening: 31.3%</td>
</tr>
<tr>
<td>Day: 40.8%</td>
</tr>
<tr>
<td>Don't know/no response: 17.9%</td>
</tr>
</tbody>
</table>

---

**Figure 4** — Percentage of respondents who were not disturbed in relation to NNI

**Figure 5** — Number of disturbances in relation to NNI

Contradict the earlier Heathrow finding that nighttime annoyance levels were high. These comparisons point to the possibility that despite the use of the expression “the noise of an aircraft” (which was an attempt to ascertain relative intrusiveness — that is, the probability that a single aircraft sound would cause a disturbance), the question, like its Los Angeles counterpart, has tapped feelings of cumulative noise impact. This again suggests that although nighttime noise disturbance may be particularly annoying, the probability of disturbance is low.

A clue to the magnitude of the higher annoyance response comes from answers to the questions concerning the durations of the disturbances, the distributions of which are shown in Fig. 6. It is stressed that these distributions relate only to those people who were actually disturbed. Out of 600 survey respondents, these numbered 340 for the day, 329 for the
Overall, aircraft noise causes little or no disturbance to most people at night, presumably because they sleep through it. Thus, when specifically asked to compare different time periods, most people say that they are more bothered by aircraft noise when they are up and about than when they are in bed, either asleep or trying to get to sleep. However, people who are disturbed at night consider the disturbance to be more severe and more annoying than during the waking hours. This increase is difficult to quantify, although it may be associated with a five-fold increase in disturbance duration.

With regard to the structure of composite noise scales, an application of these hypotheses is that for predicting community annoyance, an evening weighting of around 5 or 6 dB (or whatever is the appropriate trade-off for a four-fold increase in number in the particular scale in question) is a clear requirement. For night, the commonly used weighting of 10 dB is probably too large and extends over too long a period of time.

Indeed, it would seem more appropriate to extend the evening period to perhaps 1:00 AM, to cover the critical "falling asleep" phase and to apply a zero weighting for the remainder of the night, effectively including it with the unweighted daytime period.

These conclusions are based upon surveys of aircraft noise annoyance. Whether or not they would apply to other kinds of noise and whether a zero night weighting is commensurate with good sleeping conditions is perhaps open to question. From the standpoint of signal-to-noise ratio, aircraft noise may be considered particularly intrusive; if, as the evidence suggests, the great majority of people sleep through this, it seems likely that they would sleep equally well through a more uniform noise climate of similar equivalent level. Nevertheless, the argument that the typically lower background noise levels at night impose a general need for lower noise emissions is likely to be a compelling point, at least to the general public. Coupled with the fact that the mechanisms of noise disturbance during the night and during the day/evening are very different, this suggests that a practical compromise would be to exclude the (shorter) night period altogether when calculating composite noise levels. If nighttime noise levels are likely to be particularly high (as is rarely the case), then this matter would best be treated as an entirely separate issue.

Acknowledgments

Some of the data presented herein stem from research sponsored by the Civil Aviation Authority. The paper is a substantially expanded version of a paper, "Variation of Community Noise Sensitivity with Time of Day," presented at Inter-Noise 77 in Zurich. The author is grateful to one of the referees for drawing his attention to Dr. Borsky’s results.

References

9. California Department of Aeronautics, Title 4, Register 70, No. 45-11-28-70, Subchapter 6, Noise Standards.
16. J. B. Oliverhead, unpublished data.
This paper arose from the author's membership of a Noise Advisory Council Working Group on Noise Units which was concerned with the prospects for a more unified approach to noise impact analysis. The Group recommended that the scale of Equivalent Continuous Sound Level (Leq) be adopted as a general measure of long-term noise exposure but recognized that difficulties might arise when using it in the assessment of noise from mixed sources. The model proposed here provides a possible procedure for dealing with combinations of sounds which, individually, provoke different degrees of reaction.

* Noise Advisory Council, 'Noise Units', HMSO 1975
PREDICTING PUBLIC REACTION TO NOISE FROM MIXED SOURCES

J.B. Ollerhead

Loughborough University of Technology,

Separate study of community responses to noise from different sources, especially aircraft and road traffic, has led to the development of different indices of noise exposure for predicting public annoyance reactions. Yet because inherent human variability leads to low correlation between individual annoyance and noise exposure, many different noise variables tend to be equivalent in their ability to predict response. Evidently one may as well select a scale on grounds of practical convenience and then establish appropriate dose-response relationships for that particular scale. At the same time, for many practical reasons, it is desirable if possible to adopt a common scale for assessing noise problems of all kinds. In the U.S.A. the index $L_{dn}$ has been proposed for this purpose [1]; in the U.K. it has been suggested that $L_{eq}$ should be used, but in a rather more flexible way [2].

This flexibility was felt to be important partly because of a concern that with regard to annoyance there may be no unique dose-response relation. Evidence to support this view is shown in Figure 1 which compares reactions obtained from three social surveys of transportation noise [3,4,5]. These studies are comparable since, in all three, (a) respondents were asked to express
their dissatisfaction with noise as a number between 1 and 7 and (b) information was provided which allows noise to be expressed in $L_{eq}$. The three lines, which were obtained by linear regression of ground response data reveal significant differences in the responses to noise from different sources; in this case trains, road traffic and aircraft. It would appear that if $L_{eq}$ is to be used as a response predictor, different interpretations must be applied in the three cases.

**A MIXED-SOURCE REACTION MODEL**

Let us suppose that over a period of time the equivalent continuous sound level $L$, in dB(A), is the result of simultaneous contributions $L_i$ from several independent sources such that

$$L = 10 \log_{10} \left( \sum_{i} 10^{L_i/10} \right)$$  \hspace{1cm} (1)

Let us further suppose that if the $i$th source were heard alone, the mean annoyance reaction could be expressed by the linear relationship

$$R_i = aL_i + b_i$$  \hspace{1cm} (2)

For simplicity it has been assumed that the constant of proportionality $a$ is the same for all sources (even though at first sight this appears to contradict the evidence of Figure 1) but that the level of reaction (embodied in $b_i$) can vary from source to source. The problem is to define the resultant reaction $R$ which would be evoked by all sources when heard simultaneously.

In seeking a suitable model we may be guided by the boundary condition that if the $i$th source masks all other sources such that $L - L_i$ is very small, then by definition $R = R_i$. A relation which satisfies this condition is

$$R = aL + \sum b_i 10^{(L_i - L)/10}$$  \hspace{1cm} (3)

Alternatively it is convenient to introduce the transformation $b_i = aD_i + b$ so that the response is only dependent upon the "effective level" $L_{eff}$, i.e.

$$R = aL_{eff} + b$$  \hspace{1cm} (4)

where

$$L_{eff} = L + \sum D_i 10^{(L_i - L)/10}$$  \hspace{1cm} (5)

and $D_i$ is the effective level increment associated with the $i$th source.
EXPERIMENTAL SUPPORT FOR THE MODEL

Langdon's Traffic Noise Survey [3] involved interviews with 2933 respondents exposed to varying levels of road traffic noise. At each of 53 sites both noise levels and the fraction of traffic $f$ made up of heavy goods vehicles (HGV) were determined. Linear regression for the 53 grouped responses gave the relation between median dissatisfaction $R$ and 24-hour $L_{eq}$ (dB(A)) as

$$R = 0.124 L_{eq} - 3.54 \quad (6)$$

with a correlation coefficient of 0.509.

By introducing the variable $\log_{10} f$ into a multiple linear regression, Langdon computed the regression equation

$$R = 0.078 L_{10} + 3.34 \log_{10} f + 2.18 \quad (7)$$

where $L_{10}$ is the level exceeded for 10% of the time in dB(A). Use of equation (7) to predict median dissatisfaction raises the correlation with measured values to 0.703, clearly suggesting that people respond to the presence of HGV's to an extent not fully explained by the additional noise they generate.

Langdon's data may be re-analysed in terms of the proposed mixed source model to determine whether a difference in sensitivity (to cars and trucks) would explain the observed results. Making the assumption that, typically, HGV's are 10dB noisier than cars, equations (1), (4) and (5) may be combined to yield the theoretical result

$$R = a \left[ L + \frac{10f}{1 + 9f} D \right] + b \quad (8)$$

where $D$ is the difference in the "effective" level of trucks and cars. Use of equation (8) as a basis for the re-analysis leads to the regression equation

$$R = 0.070 L + 5.51 \frac{10f}{(1+9f)} - 3.14 \quad (9)$$

with a multiple correlation coefficient of 0.698.

A striking similarity between equations (7) and (9) in the range of $f$ encountered (0.05 to 0.32) may be seen in Figure 2 where the equations are compared for $L_{eq} \approx L_{10} - 3 = 70$dB(A). However whereas the variable $\log_{10} f$ appears to be an arbitrary choice, the term $10f/(1+9f)$ emerges as a natural consequence of the mixed source model. Figure 3 compares the response predicted by equation (9) for $f = 0.05$, 0.10 and 0.15 with the regression lines fitted to Langdon's raw data and to the train noise data of Walker and Fields [4]. This shows that the slopes are very similar when the data are 'corrected' to fixed source composition and
the response to train noise appears equivalent to the response to traffic noise with about 6% heavy vehicles.

Bottom's Survey of Mixed Aircraft and Traffic Noise [5] was designed to investigate the influence of background road traffic noise on the annoyance caused by aircraft noise. Nine sites near London Airport were selected where each of three road traffic flow conditions were combined with each of three aircraft noise conditions and 35 people were interviewed at each site. Figure 4 in which the mean dissatisfaction scores at the different sites are plotted against aircraft noise in Noise and Number Index (NNI) suggests that increased levels of traffic noise cause a reduction in dissatisfaction with the overall noise environment (which respondents were asked to consider). It was concluded that this lent support to the Noise Pollution Level concept [6] that reactions increase with fluctuations of level as well as with average sound level. However, again we may examine the possibility that equal levels of aircraft noise and traffic noise evoke different levels of dissatisfaction. To do this the component levels in

FIG 2 REGRESSION MODELS
FIG 3 EFFECT OF HGVs ON INCLUDING FRACTION OF HGVs DISSATISFACTION
FIG 4 DISSATISFACTION WITH MIXED AIRCRAFT AND ROAD-TRAFFIC NOISE.
L_{eq} have been re-estimated from the original traffic data.

The regression of mean dissatisfaction on equivalent level L (obtained by combining aircraft and traffic noise components) is

$$ R = 0.161 L - 6.89 $$

(10)

with a correlation coefficient 0.906. Use of the mixed source model (equations (4) and (5)) leads to the multiple linear regression equation

$$ R = 0.124[ L + 9.45 \times 10^{(L_a - L)/10} ] - 4.98 $$

(11)

(wher$L_a$ is the aircraft component) with a correlation coefficient of 0.955. The improved correlation indicates that this model provides an equally plausible explanation of the observed results with the implication that in terms of its capacity to induce dissatisfaction aircraft noise is effectively some 9.45dB(A) noisier than road traffic noise at the same physical level. Figure 5 compares equation (10) for mixed noises with the mean dissatisfactions given by equation (11) for (a) aircraft noise and (b) road traffic noise heard in isolation. Also shown is Langdon's result (equation 6) for road traffic noise alone.

The fact that the two traffic noise lines have practically identical slopes seems to be a logical outcome since (presumably) in both cases, increasing noise tends to be associated with higher fractions of heavy vehicles. The identical slope in the aircraft case is due to the assumption of equality in equation (3). Although it was not possible due to lack of data, a more appropriate test of the model would have allowed three separate source types - aircraft, cars and heavy vehicles. The fact that the two road traffic curves are separated by 11.5dB(A) is probably attributable to the fact that Langdon's survey was performed in the summer and Bottom's in the winter.
CONCLUSIONS

In the light of the evidence that there are significant differences in people's reactions to noise from different sources a simple model has been proposed to describe how these may be aggregated in situations where several sources are heard simultaneously. Although the analysis described is of a very preliminary nature the results appear sufficiently encouraging to warrant further studies of the model as a means for extending the usefulness of 'unified' noise indices such as $L_{eq}$ and $L_{dn}$.

REFERENCES


