AD-A083 075 WYLE LABS EL SEGUNDO CALIF F/G 20/1
CORRECTION PROCEDURES FOR AIRCRAFT NOISE DATA. VOLUME IV. TONE --ETC(U)
FEB 80 D N MAY, E E WATSON DOT-FA-78WA-4143
UNCLASSIFIED WR-79-9-VOL-4 FAA/EE-80-1-VOL-4 NL

1 OF 1
ADA
083 075

END
DATE
FILMED
5-80
DTIC

Report No: FAA-EE-80-1, Vol. IV

LEVEL

CORRECTION PROCEDURES FOR AIRCRAFT NOISE DATA VOLUME IV: TONE PERCEPTION

D.N. May and E.E. Watson
Wyle Laboratories
128 Maryland St.
El Segundo, CA 90245

FEBRUARY 1980
FINAL REPORT

Document is available to the U.S. public through
The National Technical Information Service,
Springfield, Virginia 22161

Prepared for

U.S. DEPARTMENT OF TRANSPORTATION
Federal Aviation Administration
Office of Environment and Energy
Washington, D.C. 20591

The contents of this report reflect the views of Wyle Laboratories which is responsible for the facts and the accuracy of the data presented herein. *The contents do not necessarily reflect the* official views or policy of the Department of Transportation. This report does not constitute a standard, specification or regulation.

Technical Report Documentation Page

1. Report No. FAA-EE-80-1, Vol. IV

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle
Correction Procedures for Aircraft Noise Data, Vol. IV. Tone Perception.

5. Report Date February 1980

6. Performing Organization Code

7. Author(s) D. N. May and E. E. Watson

8. Performing Organization Report No. WR-79-9, Vol. IV

9. Performing Organization Name and Address
Wyle Laboratories
Wyle Research
128 Maryland St.
El Segundo, California 90245

10. Work Unit No. (TRAIS)

11. Contract or Grant No. DOT-FA-78WA-4143

12. Sponsoring Agency Name and Address
U. S. Department of Transportation
Federal Aviation Administration
Office of Environment and Energy
Washington, D. C. 20591

13. Type of Report and Period Covered Final

14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The existing tone correction procedure in the Effective Perceived Noise Level (EPNL) calculation procedure required for aircraft certification under Part 36 of the Federal Aviation Regulations was compared with other tone correction procedures, including the SAE Aerospace Recommended Practice 1071 and a multitone procedure due to Kryter and Pearsons. Different amounts of tone correction ("level-weightings") and varying degrees of tone correction at different times in the flyover ("time-weightings") were also explored. Also studied was a measure of spectral fluctuation, developed by NASA and known as "spectral change." The research was limited to considering revisions within the framework of one-third octave, 0.5 second interval analysis, since such revisions can be quite easily implemented. The various tone correction noise metrics were tested against subjective judgments furnished by NASA of the noise from a range of propjet, turbojet, low and high bypass ratio turbofan, and supersonic commercial aircraft. It was found that a revision based on "spectral change" could, after further development, be a means to improve the accuracy of the EPNL metric. However, the success of the various other potential revisions depended on the characteristics of the data base tested. It was shown that research into improved metrics should be based on experimental plans which account for the correlations among the noise variables and the presence of any interactions. While the prospects for developing improved tone correction procedures are quite good, no change from the FAR Part 36 procedure was indicated without further research. A separate, psychoacoustical pilot experiment was also performed into the effects of pseudotones on judged noisiness. (Pseudotones are low frequency tones introduced into a measured spectrum by ground reflections near the microphone.) Subjects compared pairs of flyover sounds in which one sound measured at 1.2 m microphone height had a strong pseudotone, and the other sound from the same flyover (measured at 10 m height) had few or no pseudotones. They were not able, in any consistent manner, to detect a change in noisiness due to the pseudotones. This preliminary study therefore indicated no reason to change the FAR Part 36 procedure due to the subjective effects of pseudotones.

17. Key Words
Noise, Acoustics, Aircraft, Tones, Certification, Correction Procedures, Pseudotones, Aviation Regulations

18. Distribution Statement Unlimited

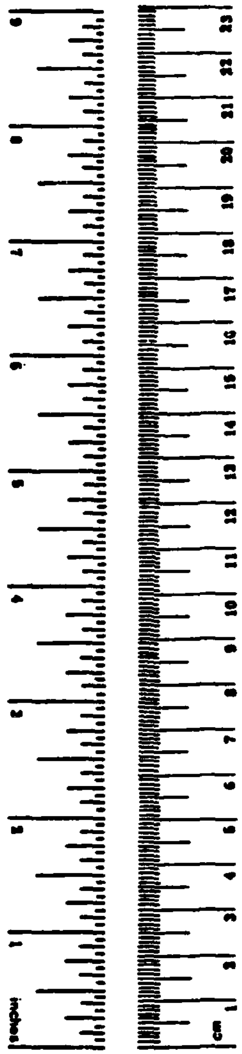
19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	73	

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

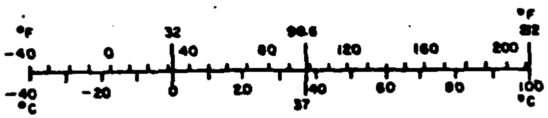
Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	*2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

*1 in ≡ 2.54 (exactly). For other exact conversions and more detailed tables, see NBS Misc. Publ. 286, Units of Weights and Measures, Price $2.25, SD Catalog No. C13.10:286.



Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



ABSTRACT

The existing tone correction procedure in the Effective Perceived Noise Level (EPNL) calculation procedure required for aircraft certification under Part 36 of the Federal Aviation Regulations was compared with other tone correction procedures, including the SAE Aerospace Recommended Practice 1071 and a multitone procedure due to Kryter and Pearsons. Different amounts of tone correction ("level-weightings") and varying degrees of tone correction at different times in the flyover ("time-weightings") were also explored. Also studied was a measure of spectral fluctuation, developed by NASA and known as "spectral change." The research was limited to considering revisions within the framework of one-third octave, 0.5 second interval analysis, since such revisions can be quite easily implemented. The various tone correction noise metrics were tested against subjective judgments furnished by NASA of the noise from a range of propjet, turbojet, low and high bypass ratio turbofan, and supersonic commercial aircraft. It was found that a revision based on "spectral change" could, after further development, be a means to improve the accuracy of the EPNL metric. However, the success of the various other potential revisions depended on the characteristics of the data base tested. It was shown that research into improved metrics should be based on experimental plans which account for the correlations among the noise variables and the presence of any interactions. While the prospects for developing improved tone correction procedures are quite good, no change from the FAR Part 36 procedure was indicated without further research.

A separate, psychoacoustical pilot experiment was also performed into the effects of pseudotones on judged noisiness. (Pseudotones are low frequency tones introduced into a measured spectrum by ground reflections near the microphone.) Subjects compared pairs of flyover sounds in which one sound measured at 1.2 m microphone height had a strong pseudotone, and the other sound from the same flyover (measured at 10 m height) had few or no pseudotones. They were not able, in any consistent manner, to detect a change in noisiness due to the pseudotones. This preliminary study therefore indicated no reason to change the FAR Part 36 procedure due to the subjective effects of pseudotones.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

The FAA's aircraft noise certification requirements described in FAR Part 36[1] are based upon the Effective Perceived Noise Level (EPNL) metric which contains a tone correction. The purpose of the tone correction is to make the metric account for the noisiness of pure tone components imbedded in the noise that are not adequately described by the primary building block in the calculation procedure: the Perceived Noise Level (PNL) term. The tone correction procedure is necessary because PNL was derived from experiments in which subjects judged the noisiness of bands of white noise,[2] in comparison with which tonal sounds (of the same SPL) have been judged to be more noisy.[3]

The tone correction procedure in FAR Part 36 took effect in 1969 and was based on experience mainly with subsonic turbojet and low bypass ratio turbofan airplanes. Since then, wide body airplanes with high bypass ratio turbofan engines have entered more widespread service, and the Concorde SST has introduced the very large, afterburning turbojet to certain airports. Many of the pure turbojet airplanes have also been retired. This changing tonal quality of the airport noise environment has dictated this reassessment of the FAR Part 36 tone correction procedure to ensure that it reflects the noisiness of the present day airplane fleet, both in community response and for aircraft certification.

Reassessment of the tone correction procedure in FAR Part 36 is also opportune for three further reasons.

First, the Society of Automotive Engineers published a recommended practice for computing the tone correction[4] which differs slightly from that used in FAR Part 36: there has been a continuing debate as to which of these methods is better and, indeed, whether an alternative procedure due to Kryter and Pearsons[5] may be superior to both. The Kryter-Pearsons procedure differs from the others in that it can correct for more than one tone in the spectrum (i.e., for the so-called multitones found in some turbofan spectra). The three procedures are described in detail later in this report.

Secondly, the comparative assessment of these tone correction procedures affords an opportunity to consider options to change them in other ways, for example by time-weighting the tone corrections through the flyover to reflect any changes in noisiness that depend on whether an aircraft is approaching an observer, is overhead, or is receding.

The third reason for reassessing the tone correction calculation derives from recognition of the presence of pseudotones in certain aircraft noise spectra. Pseudotones are not pure tones in the direct sound wave, but rather are manifested by peaks in the measured one-third octave band noise spectrum that are the result of interference between the direct wave and the ground-reflected wave.[6] These peaks, though they may not be pronounced or perceived as tones, can result in tone corrections in the EPNL calculation. There is therefore concern as to whether they exert an influence on measured EPNL values unjustified by their actual subjective effect.

This report, therefore, describes work* to reassess tone correction procedures in order to determine whether:

- The FAR Part 36 tone correction should be replaced by another procedure such as one in which multitones in the aircraft noise spectrum are considered, and/or one in which the tone correction is weighted according to when it occurs in the flyover. This work is described in Section 2.
- The presence of pseudotones is judged by people as important enough to warrant changes to the procedure. This work is described in Section 3.

Note that throughout this report, the word "tone" has been used as an abbreviation of the term "pure tone" (thus "tone correction," not "pure tone correction"), and that "multitone" has been considered synonymous with "complex tone." Pure tones and complex tones are defined in ANSI S3.20-1973 ("Psychoacoustical Terminology") as follows:

"<u>Pure Tone</u>:

(1) A pure tone is a sound wave, the instantaneous sound pressure of which is a simple sinusoidal function of the time.

(2) A pure tone is a sound sensation characterized by its singleness of pitch.

Note: Whether or not a subject hears a tone as pure or complex is dependent upon ability, experience, and listening attitude."

---

*Dr. Robert Rackl made a major contribution to this work by computing the values of most of the aircraft noise metrics under study here.

"<u>Complex Tone</u>:

(1) A complex tone is a sound wave containing simple sinusoidal components of different frequencies.

(2) A complex tone is a sound sensation characterized by more than one pitch."

2.0 TONE CORRECTION PROCEDURES ANALYSIS

2.1 Available Procedures

Available tone correction procedures are:

- o FAR Part 36,[1]
- o SAE ARP 1071,[4] and
- o Kryter-Pearsons.[5]

In Sections 2.1.1, 2.1.2 and 2.1.3, each of these procedures is described in turn, and the same example sound spectrum is used to calculate the tone-corrected Perceived Noise Level (PNLT) for each method, in order to highlight their differences. The calculation procedure to reach Effective Perceived Noise Level (EPNL) is also indicated to show how the tone correction influences the "final product."

2.1.1 FAR Part 36

The FAR Part 36 tone correction procedure, which has been adopted by ICAO,[7] is sometimes referred to as "the 10 step method." It is described in paragraph B36.5 of FAR Part 36[1] as a "correction for spectral irregularities." Ten steps are performed to detect and correct for tones in the one-third octave bands from 80 Hz to 10 kHz, as follows:

Step 1. Starting with the sound pressure level (SPL) in the 80 Hz one-third octave band (band number 3), calculate the changes in sound pressure level (or "slopes") in the remainder of the one-third octave bands as follows:

$$s(3,k) = \text{no value}$$

$$s(4,k) = SPL(4,k) - SPL(3,k)$$

$$\vdots$$

$$s(i,k) = SPL(i,k) - SPL\left[(i-1),k\right]$$

$$\vdots$$

$$s(24,k) = SPL(24,k) - SPL(23,k),$$

where k is a time index designating the 0.5 second* intervals of the flyover, and i is a one-third octave band number in the range 3 to 24, as allocated above. (The calculation therefore operates on a range of frequencies through the one-third octave band spectrum at intervals through the flyover when the Perceived Noise Level is within 10 dB of its maximum value.)

Step 2. Encircle the value of the slope, s (i,k), where the absolute value of the change in slope is greater than five; that is, where

$$\left|\Delta s\,(i,k)\right| = \left|s\,(i,k) - s\left[(i-1),\,k\right]\right| > 5$$

Step 3.

(a). If the encircled value of the slope s (i,k) is positive and algebraically greater than the slope $s\left[(i-1),\,k\right]$, encircle SPL (i,k).

(b). If the encircled value of the slope s (i,k) is zero or negative and the slope $s\left[(i-1),\,k\right]$ is positive, encircle $SPL\left[(i-1),\,k\right]$.

(c). For all other cases, no sound pressure level value is to be encircled.

Step 4. Omit all SPL(i,k) encircled in Step 3 and compute new sound pressure levels SPL' (i,k) as follows:

(a). For nonencircled sound pressure levels, let the new sound pressure levels equal the original sound pressure levels,

$$SPL'\,(i,k) = SPL\,(i,k)$$

(b). For encircled sound pressure levels in bands 1-23, let the new sound pressure level equal the arithmetic average of the preceding and following sound pressure levels,

$$SPL'\,(i,k) = (1/2)\left[SPL\left[(i-1),\,k\right] + SPL\left[(i+1),\,k\right]\right]$$

(c). If the sound pressure level in the highest frequency band (i = 24) is encircled, let the new sound pressure level in that band equal

*Under Appendix B, paragraph B36.9(c) of FAR Part 36, the intervals may be less than 0.5 second, but 0.5 second is an industry practice.

$$SPL'(24,k) = SPL(23,k) + s(23,k).$$

Step 5. Recompute new slopes s' (i,k), including one for an imaginary 25th band, as follows:

$$s'(3,k) = s'(4,k)$$

$$s'(4,k) = SPL'(4,k) - SPL'(3,k),k$$

.

.

.

$$s'(i,k) = SPL'(i,k) - SPL'\left[(i-1),k\right]$$

.

.

.

$$s'(24,k) = SPL'(24,k) - SPL'(23,k)$$

$$s'(25,k) = s'(24,k)$$

Step 6. For i from 3 to 23, compute the arithmetic average of the three adjacent slopes as follows:

$$\bar{s}(i,k) = (1/3)\left[s'(i,k) + s'\left[(i+1), k\right] + s'\left[(i+2), k\right]\right]$$

Step 7. Compute final adjusted one-third octave band sound pressure levels, SPL" (i,k) by beginning with band number 3 and proceeding to band number 24 as follows:

$$SPL''(3,k) = SPL(3,k)$$

$$SPL''(4,k) = SPL''(3,k) + \bar{s}(3,k)$$

.

.

.

$$SPL''(i,k) = SPL''\left[(i-1), k\right] + \bar{s}\left[(i-1), k\right]$$

.

.

.

$$SPL''(24,k) = SPL''(23,k) + \bar{s}(23,k)$$

Step 8. Calculate the differences, F (i,k), between the original and the adjusted sound pressure levels as follows:

$$F\,(i,k) = SPL\,(i,k) - SPL''\,(i,k)$$

and note only values greater than zero.

Step 9. For each of the 24 one-third octave bands, determine tone correction factors from the sound pressure level differences F (i,k) and Table I or Figure I.

Step 10. Designate the largest of the tone correction factors, determined in Step 9, as C (k).

PNL(k) is then calculated in the normal way for each kth interval - see, for example, Appendix B, paragraph B36.3 of Reference I.

Tone corrected perceived noise levels PNLT (k) are determined by adding the C(k) values to corresponding PNL (k) values, that is,

$$PNLT\,(k) = PNL\,(k) + C\,(k).$$

A narrow band analysis may be done if there is a suspicion that any irregularity in the spectrum is there as a result of tones from other than aircraft noise. A revised tone correction factor for that band is to be done if necessary. Pseudotones that are clearly not identified as related to engine noises are neglected if below 800 Hz. Once the value of PNLTM (Maximum Tone Corrected Perceived Noise Level) is calculated, the time intervals around the PNLTM interval are examined for possible band sharing of the tone by seeing whether the average of the four tone corrections, two immediately before and two immediately after the PNLTM interval, exceeds the tone correction at PNLTM. If so, this average tone correction is used to compute a new value for PNLTM.

EPNL is then calculated from

$$EPNL = 10 \log \left[\sum_{k=0}^{2d} \text{antilog} \left[PNLT(k)/10 \right] \right] - 13,$$

where d is the time interval to the nearest 1.0 second during which PNLT(k) is within a specified value (10 dB) of PNLTM.

Table 1

Tone Correction as a Function of Level Difference[1]

(This is a tabular representation of Figure 1)

Frequency f, Hz	Level Difference F, dB	Tone Correction C, dB
$50 \leq f < 500$	$0 \leq F < 20$	F/6
	$20 \leq F$	3-1/3
$500 \leq f \leq 5000$	$0 \leq F < 20$	F/3
	$20 \leq F$	6-2/3
$5000 < f \leq 10{,}000$	$0 \leq F < 20$	F/6
	$20 \leq F$	3-1/3

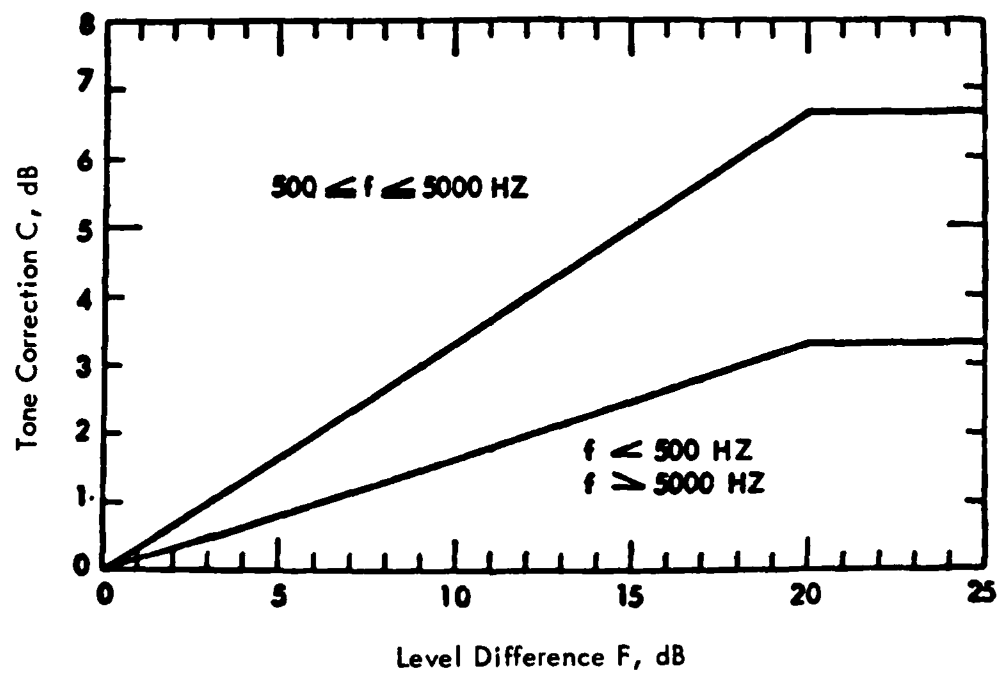


Figure 1. Tone Correction as a Function of Level Difference for the FAR Part 36 Procedure (from Reference 1)

The main features of the FAR Part 36 tone correction procedure which distinguish it from the SAE ARP 1071 or the Kryter-Pearsons procedures are as follows:

Differs from SAE ARP 1071 and Kryter-Pearsons	o It is a 10-step procedure involving the calculation of a "smoothed" spectrum, i.e., the supposed spectrum without tones, prior to a tone correction computation.
Differs from Kryter-Pearsons	o The tone correction is added to the PNL(k).
"	o Only the largest tone correction in the one-third octave band spectrum at each 0.5 second interval applies; the others are disregarded.
"	o The value of the tone correction is discontinuous with frequency (see Table 1 and Figure 1).[*]

The significance of these features can be gauged by comparison with the paragraph corresponding to the above in each of Sections 2.1.2 and 2.1.3 for the other two tone corrections.

Example. Following the step-by-step instructions given above, for the example spectrum[+] (occurring at a certain 0.5 sec interval) shown in Table 2, the tone correction of 2.5 dB is calculated, and this amount is added to the PNL(k) of the spectrum as a whole. This PNL(k) can be calculated by reference to FAR Part 36[1] Appendix B, paragraph B36.3, and has the value 112.2 PNdB. The PNLT(k) therefore equals 114.7 PNdB.

---

[*] The discontinuity appears to derive from a 1961 paper by Little (Ref. 3) in which he stressed that the accuracy of his study could not justify presenting a more continuous relationship. It seems surprising that the intervening years have not led to revision of this relationship. The absence of this discontinuity in the Kryter-Pearsons procedure as used here may be one of its virtues.

[+] Note that this example spectrum is the same as that used in Sections 2.1.2 and 2.1.3 to illustrate the SAE ARP 1071 and Kryter-Pearsons procedures. It has been chosen to demonstrate the differences rather than similarities between these and the FAR Part 36 methods. The different tone corrections indicated for the FAR Part 36 and SAE ARP 1071 procedures might also be reduced for this example spectrum if a band-sharing correction were applied – see Section 2.1.2.

Table 2

Calculation Steps for the FAR Part 36 Tone Correction,[1] Illustrated by an Example Spectrum

Band i	Frequency, Hz	SPL, dB	s Step 1	|Δs| Step 2	SPL', dB Step 4	s' Step 5	$\bar{s}$ Step 6	SPL'', dB Step 7	F, dB Step 8	C, dB Step 9	C(k) Step 10
1	50										
2	63										
3	80	74			74	-8	-2.3	74	0		
4	100	66	- 8		66	-8	3.3	71.7	-		
5	125	74	8	16	75	9	6.7	75	-		
6	160	84	10	2	84	9	2.7	81.7	2.3	0.4	
7	200	86	2	8	86	2	-1.3	84.4	1.6	0.3	
8	250	87	1	1	83	-3	-1.3	83.1	3.9	0.7	
9	315	80	-7	8	80	-3	0.3	81.8	-		
10	400	84	4	11	82	2	1	82.1	1.9	0.3	
11	500	84	0	4	84	2	0	83.1	0.9	0.3	
12	630	83	- 1	1	83	-1	-0.3	83.1	-		
13	800	82	- 1	0	82	-1	0.7	82.8	-		
14	1000	83	1	2	83	1	2.7	83.5	-		
15	1250	85	2	1	85	2	2	86.2	-		
16	1600	95	10	8	90	5	-0.7	88.2	6.8	2.3	
17	2000	95	0	10	89	-1	-2.7	87.5	7.5	2.5	2.5
18	2500	83	-12	12	83	-6	-2.7	84.8	-		
19	3150	84	1	13	82	-1	-1.7	82.1	1.9	0.3	
20	4000	81	- 3	4	81	-1	-2.3	80.4	0.6	0.1	
21	5000	78	- 3	0	78	-3	-2.7	78.1	-		
22	6300	75	- 3	0	75	-3	-2.3	75.4	-		
23	8000	73	- 2	1	73	-2	-2	73.1	-		
24	10,000	77	4	6	71	-2	-	71.1	5.9	1.0	
						-2					

Note: See instructions for Step 3.

2.1.2 SAE ARP 1071

The SAE ARP 1071 tone correction procedure is sometimes referred to as "the 7 step method." It is described in Reference 4 and has been approved by the American National Standards Institute as ANSI S6.4-1973.[8] Seven steps are described to detect and correct for tones in the one-third octave bands from 80 Hz to 10 kHz. In the following summary, we have employed similar notation to that used in the FAR Part 36 description (Section 2.1.1) for the sake of uniformity. (Those familiar with SAE ARP 1071 may therefore prefer to consult their own texts.)

Step 1. Compute a spectrum slope:

$$s\left[(i+1), k\right] = SPL\left[(i+1),k\right] - SPL\,(i, k)$$

where

i is the one-third octave band number ranging from 19 (corresponding to 80 Hz) up to 39 (corresponding to 8000 Hz), and k is the time index at 0.5 sec time intervals. Note $s\left[(i+1),k\right] = 0$ for 50 and 63 Hz ($i < 19$).

Step 2. *Designate, e.g., encircle, values of* $s\left[(i+1),k\right]$ *where*

$$\left|\Delta s\left[(i+1),k\right]\right| = \left| s\left[(i+1),k\right] - s\ (i,k) \right| > 5.$$

Step 3. If the encircled $s\left[(i+1), k\right]$ is positive and algebraically greater than s (i,k), encircle $SPL\left[(i+1),k\right]$. If the encircled $s\left[(i+1),k\right]$ is zero or negative and algebraically less than s(i,k), encircle SPL(i,k).

Step 4

(a). For encircled values of SPL(i,k) located between adjacent and non-encircled values, $SPL\left[(i-1),k\right]$ and $SPL\left[(i+1),k\right]$, set

$$SPL'\,(i,k) = \left\{ SPL\left[(i+1),k\right] + SPL\left[(i-1),k\right] \right\} /2.$$

If the level in the highest band SPL (40,k) is encircled: set SPL' (40,k) = SPL (39,k) + s (39,k) if SPL (39,k) and SPL (38,k) are not encircled; set SPL' (40,k) = SPL (39,k) + $s_1/2$ if SPL (38,k) is encircled but SPL (37,k) is not; set SPL' (40,k) = SPL (39,k) + $S_2/3$ if SPL (37,k) and SPL (38,k) are encircled but SPL (36,k) is not. (Here s_1 and s_2 are the changes in SPL between bands 39 and 37, and between 39 and 36, respectively.)

(b). For two successive encircled values, SPL(i,k) and $SPL\left[(i+1),k\right]$, set $SPL'(i,k) = \left\{2\, SPL\left[(i-1),k\right] + SPL\left[(i+2),k\right]\right\} /3$

and $SPL'\left[(i+1),k\right] = \left\{SPL\left[(i-1),k\right] + 2\, SPL\left[(i+2),k\right]\right\} /3.$

If the levels in the two highest frequency bands are encircled:

set $SPL'(39,k) = SPL\ (38,k) + s\ (38,k)$

and $SPL'(40,k) = SPL\ (38,k) + 2\, s\ (38,k)$

if SPL (37,k) and SPL (38,k) are not encircled; or

set $SPL'(39,k) = SPL\ (38,k) + s_3/2$

and $SPL'(40,k) = SPL\ (38,k) + s_3$

if SPL (37,k) is encircled but SPL (36,k) is not; or

set $SPL'(39,k) = SPL\ (38,k) + s_4/3$

and $SPL'(40,k) = SPL\ (38,k) + 2\, s_4/3$

if SPL (36,k) and SPL (37,k) are encircled, but SPL (35,k) is not. (Here s_3 and s_4 are the changes in SPL between bands 38 and 36, and between 38 and 35, respectively.)

Step 5. For each encircled band level, determine

$$F\ (i,k) = SPL\ (i,k) - SPL'\ (i,k) > 0.$$

Where F values greater than 5 dB occur in adjacent bands, $F(i,k)$, $F\left[(i+1),k\right]$, and where $|s_5| < 5$ for 2 adjacent bands (here s_5 is the difference in SPL between the (i+2) and (i-1) bands), then define

$$F' = 10 \log \left[\text{antilog}\ (F(i,k)/10) + \text{antilog}\ (F(i+1,k)/10)\right].$$

Where one of two adjacent F values occurs in a band outside the frequency range 500-5000 Hz, the values shall be halved, and the F' value ascribed to the 500-5000 Hz range.

Step 6. Determine the tonal correction C by a procedure which is identical to that in Table I except that C=0 when Table I would assign it a value less than 1.

Step 7. The largest value of C in the kth time interval is designated C(k) and used to give PNLT(k) = PNL(k) + C(k). EPNL is then calculated as before.

An account is also taken of any tone corrections that are suspected to be incorrect: "When the tone corrections determined from one-third octave band spectra . . . are suspected to be incorrect, additional analyses may be made with filter bandwidths narrower than one-third octave bands. Results of the narrow band analysis may then be used to compute revised tone corrections . . . Because procedures for narrow band frequency analysis of time-varying noise spectra are not standardized and may be subject to possible errors, the use of such procedures should be noted, and the procedure described, when employed."

The main features of the SAE ARP 1071 tone correction procedure which distinguish it from the FAR Part 36 or the Kryter-Pearsons procedures are as follows:

Differs from FAR Part 36 and Kryter-Pearsons	o It is a seven-step procedure, involving the calculation of a "smoothed" spectrum, i.e., the supposed spectrum without tones, prior to a tone correction computation.
Differs from Kryter-Pearsons	o The tone correction is added to the PNL(k).
"	o Only the largest tone correction in the one-third octave band spectrum at each 0.5 second interval applies; the others are disregarded.
Differs from FAR Part 36 and Kryter-Pearsons	o Any tone correction less than 1 dB is made equal to zero.
Differs from Kryter-Pearsons	o The value of the tone correction is discontinuous with frequency (see Table 1 and Figure 1).

The SAE ARP 1071 procedure is superficially similar to FAR Part 36: as the paragraph above shows when compared with corresponding paragraphs for FAR Part 36 in Section 2.1.1 and for the Kryter-Pearsons method in Section 2.1.3, the SAE and FAR methods differ between one another far less than they each do with the Kryter-Pearsons method. However, there are detailed differences which can sometimes be important, for example:

- o The "smoothed" spectrum differs in SAE ARP 1071's omission of FAR Part 36's Steps 5, 6 and 7. When two adjacent one-third octave bands protrude above neighboring bands, the different smoothed spectra result in a higher tone correction being calculated for the SAE method. However this tendency is lessened by a band-sharing provision in Part 36, as explained further below.

- A FAR Part 36 provision allows pseudotones up to 800 Hz to be removed if they are diagnosed (e.g., if regarded under Appendix B, Paragraph B36.5(m) as not due to the aircraft).

A number of comparisons of these two methods were reported to an SAE A-21 Subcommittee on Aircraft Noise Metrics in the period February-April, 1975, to the effect that the SAE procedure, when applied to the same spectrum as the Part 36 procedure, produces, in relatively isolated cases, corrections which are higher than Part 36's by up to about 3 dB. However, the use of a "band-sharing" provision in Paragraph B36.5(n) of Part 36 tends to reduce the number of occasions that this occurs, by increasing the tone correction in the FAR Part 36 procedure whenever examination of the spectra at intervals before and after PNLTM suggests that tone suppression has occurred at PNLTM. In addition to this, tone correction differences of x dB are not necessarily translated into EPNL differences of x dB because the EPNL calculation considers a <u>succession</u> of time intervals, throughout which a given tone correction difference seldom persists. The FAA, in an information brief,[9] reported on 17 flyovers (spanning 707, 727, 737, 747, DC-8, DC-9, DC-10, L-1011 and Concorde aircraft), which had only two instances of significant EPNL difference, the maximum instance being 0.7 dB.

<u>Example</u>. Following the step-by-step instructions given above, for the same example spectrum as in Table 2, a tone correction of 4.7 dB can be calculated using the SAE ARP 1071 procedure, as shown in Table 3. This can be added to the same PNL(k) of 112.2 PNdB, to give a PNLT(k) of 116.9 PNdB.

Table 3

Calculation Steps for the SAE ARP 1071 Tone Correction,[4]
Illustrated by an Example Spectrum

Band i	Frequency, Hz	SPL, dB	s Step 1	|Δs| Step 2	SPL', dB Step 4	F, dB Step 5	F', dB Step 5	C, dB Step 6	C(k) Step 7
19	80	74							
20	100	66	- 8						
21	125	(74)	(8)	16	75	-		-	
22	160	84	10	2					
23	200	86	(2)	8					
24	250	(87)	1	1	83	4		0	
25	315	80	(-7)	8					
26	400	(84)	(4)	11	82	2		0	
27	500	84	0	4					
28	630	83	- 1	1					
29	800	82	- 1	0					
30	1000	83	1	2					
31	1250	85	2	1					
32	1600	(95)	(10)	8	84.3	10.7 }	14.0	4.7	4.7
33	2000	(95)	(0)	10	83.7	11.3 }			
34	2500	83	(-12)	12					
35	3150	(84)	(1)	13	82	2		0	
36	4000	81	- 3	4					
37	5000	78	- 3	0					
38	6300	75	- 3	0					
39	8000	73	- 2	1					
40	10,000	(77)	(4)	6	71	6		1	

Note: See instructions for Step 3.

2.1.3 Kryter-Pearsons

The Kryter-Pearsons tone correction procedure, so-named because it is taken from Reference 5 by these authors, has not been endorsed by a standards organization or government agency and, perhaps for this reason, no step-by-step calculation procedure has been enunciated for it. For the purposes of this research, Wyle has extracted and slightly modified (see later) one of a number of options for a tone correction procedure put forward in Reference 5. It is this specific "option" rather than all those put forward in Reference 5, that is termed the Kryter-Pearsons procedure and described on a step-by-step basis below. For simplicity, the same notation as in Section 2.1.1 is used.

Step 1. SPL(i,k) in each one-third octave band, i, between 80 Hz (i=3) and 10,000 Hz (i=24) is obtained for all k 0.5 sec intervals in the flyover. For each SPL(i,k), the quantity F(i,k) is calculated, where

$$F(i,k) = SPL(i,k) - \left\{ SPL\left[(i-1),k\right] + SPL\left[(i+1),k\right] \right\}/2 \gtrsim 3.$$

(Values of F less than or equal to 3 are considered zero.)

F (1,k) and F (24,k) are made equal to zero.

Step 2. Referring to Figure 2, a value C(i,k) is calculated for all i,k using interpolation to obtain C(i,k) for values of F between 3 and 25 for which curves are not given in Figure 2. Note in Figure 2 that the extrapolated curves below 500 Hz are due to Kryter and Pearsons, while those above 8000 Hz are a Wyle modification to avoid the discontinuities which are considered drawbacks of the other two procedures.

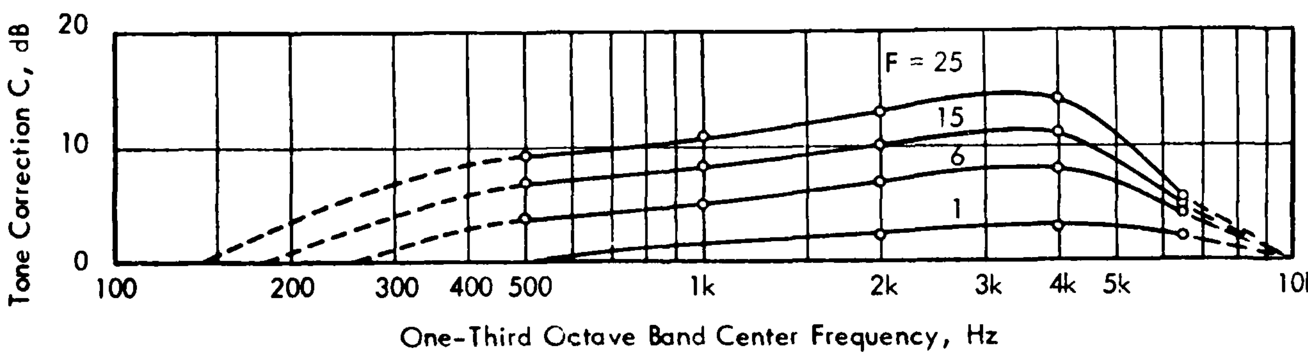


Figure 2. Tone Correction as a Function of Level Difference (adopted from Reference 5 and identified here as the Kryter-Pearsons procedure)

Step 3. Each value of C(i,k) is added to the corresponding value of SPL(i,k) to give tone-corrected spectra, SPL' (i,k). PNLT(k) is then calculated directly from the values of SPL'(i,k) in exactly the same way as PNL(i,k) is calculated from SPL(i,k).

EPNL is then computed in the normal way.

Kryter and Pearsons[5] also describe octave and 1/10 octave band variants of this procedure, but they are not included here as a result of the need to consider (at this time) spectral time histories similar (i.e., one-third octave band spectra) to those considered for FAR Part 36 and SAE ARP 1071. This has the virtue of facilitating comparison of the three methods, but also has the more important significance of testing the Kryter-Pearsons method in the context of the data analysis system that is the industry norm for this type of measurement.

The main features of the Kryter-Pearsons tone correction procedure that distinguish it from the other two are as follows:

Differs from FAR Part 36 and SAE ARP 1071	o It is a simple, three-step procedure (in our formulation) involving no attempts at smoothing the spectrum or considering bandsharing of tones.
"	o The tone correction in each band is added to the SPL in that band.
"	o Multiple tone corrections can be produced and all, not just the largest, contribute to PNLT.
"	o Tone corrections are considered to be zero for F 3 dB.
"	o The value of the tone correction is assigned without discontinuities across the frequency range (see Figure 2).

Example. Following the step-by-step instructions for the modified Kryter-Pearsons method as defined above, for the same example spectrum shown in Table 2, the calculation shown in Table 4 (using Figure 2 also) produces a PNLT(k) of 116.2 PNdB.

Table 4

Calculation Steps for the Kryter-Pearsons Tone Correction,[5] Illustrated by an Example Spectrum

Band i	Frequency, Hz	SPL, dB	F, dB Step 1	C, dB Step 2	SPL', dB Step 3
3	80	74	-		74
4	100	66	-		66
5	125	74	-		74
6	160	84	4	0	84
7	200	86	-		86
8	250	87	4	0	87
9	315	80	-		80
10	400	84	-		84
11	500	84	-		84
12	630	83	-		83
13	800	82	-		82
14	1000	83	-		83
15	1250	85	-		85
16	1600	95	5	5	100
17	2000	95	6	6.5	101.5
18	2500	83	-		83
19	3150	84	-		84
20	4000	81	-		81
21	5000	78	-		78
22	6300	75	-		75
23	8000	73	-		73
24	10,000	77	-		77

2.2 Prospects for Revised Procedures

The appropriateness of any tone correction procedure depends for its success in satisfying the two requirements of:

- o Properly describing the noisiness of tones - a "perceptual" criterion, and
- o Being straightforwardly measured and analyzed within the framework of existing aircraft industry acoustical facilities - a "practical" criterion.

Given one-third octave analysis equipment and appropriate computer backup as already used in the industry, a tone correction that is also based on one-third octave analysis and the current 0.5 second interval analysis poses no difficulty satisfying the "practical" criterion. For this reason, all three tone correction procedures considered herein - FAR Part 36, SAE ARP 1071, and Kryter-Pearsons (as defined here) - can be considered "practical" procedures that can be (or are being) applied with a reasonable amount of effort.

Tone corrections requiring narrower band analyses or lesser intervals than 0.5 sec. might not, however, satisfy this criterion as they would dictate the use of different filters or impose the need for spectrum analysis by computer. Wyle is therefore reluctant to recommend them without overwhelming evidence that a revised one-third octave band, 0.5 sec. interval, tone correction procedure cannot be found that is significantly superior to the current FAR Part 36 procedure.

Section 2.2.1 of the report, however, contains a review of the inadequacy, from a "perceptual" standpoint, of tone corrections using one-third octave analysis and 0.5 sec. intervals in order to provide reference material for future use.

Section 2.2.2 contains proposals for improving tone corrections within the framework of current one-third octave band, 0.5 sec. interval, practice.

In Section 2.3, the various procedures are tested against an available subjective judgment data base, leading in Section 2.4 to recommendations.

2.2.1 Tone Perception as a Function of Frequency Bandwidth and Time

A major reason for using one-third octave analysis in acoustics is that the one-third octave bandwidths match to some degree the "filter bandwidths" of the human aural system, which indeed can be crudely modeled as a set of such filters spanning the audible frequency range.

The "filter bandwidths" that partially correspond with the one-third octave bandwidths are known as critical bands, a term defined in ANSI S3.20-1973 (where they are known as the "critical bands (for loudness)") as "that frequency band within which the loudness of continuously distributed sound of constant sound pressure level is independent of bandwidth." More simply, the loudness of a band of noise of constant SPL remains constant while its bandwidth is increased - until a certain bandwidth is reached, the critical bandwidth, after which its loudness increases. The process that seems to occur here is one in which the sound frequency components within this bandwidth "mask" one another to some extent, i.e., each reduces the loudness of the others. But sounds separated in frequency by more than a critical bandwidth exert no such influence, and their combined loudness is therefore greater.

The critical bandwidths increase with the frequency of the center frequency of the band under consideration, as shown in Figure 3. It is seen that the critical bandwidths correspond quite closely to the one-third octave bandwidths at frequencies above about 400 Hz. Since the ear is more sensitive to sound above this frequency, the one-third octave bands have served as a useful basis to analyze sounds when computing their loudness (and noisiness). They are used, for example, in calculating the PNL of aircraft flyovers. This usage doubtless made it inviting to base tone corrections on one-third octave band analysis also.

In tone perception, however, at least two other recognized psychoacoustical bandwidths, both considerably less wide than one-third octave bandwidths, may be equally or more relevant than the critical bandwidths. One of these is called the critical ratio, and the other the difference limen. Both are also shown in Figure 3.

The critical ratio is defined in ANSI S3.20-1973* as "that frequency of sound, being a portion of a continuous-spectrum noise covering a wide band, that contains sound power equal to that of a pure tone centered in the critical (ratio) and just audible in the presence of the wide band noise." The critical ratio is similar to the critical band - and different from the difference limen - in that it is a bandwidth within which one sound is masked by another. Being defined in terms of the audibility of a pure tone in a broadband background noise, however, the

*To avoid confusion with traditional practice, we have not adopted the practice used in ANSI S3.20-1973 of renaming the traditional term, "critical ratio," as "critical band (for masking)."

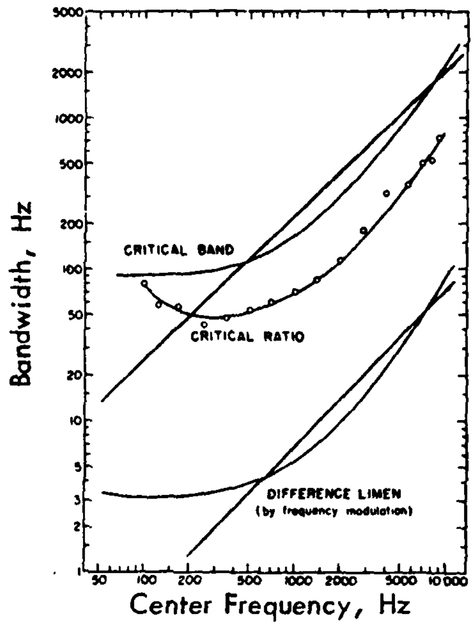


Figure 3. *Frequency Bandwidths Important in Tone Perception* (from Reference 10). Superimposed are the one-third octave bandwidths (upper straight line) and 1/100 octave bandwidths (lower straight line).

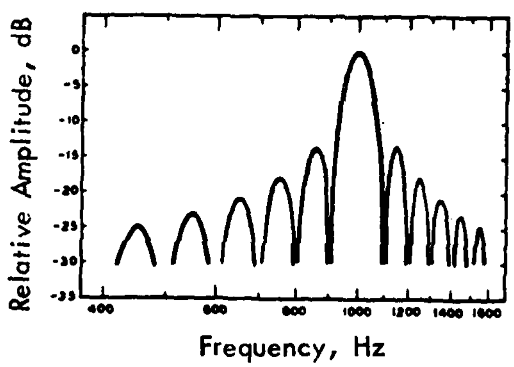


Figure 4. The Fourier Spectrum of a 1000 Hz Toneburst of 10 ms Duration (from Reference 20)

critical ratio seems to be a bandwidth at least as relevant to tone corrections as does the critical band. The critical ratio is shown in Figure 3, where it is seen to be about half as wide as the one-third octave bands.

The difference limen is defined in ANSI S3.2-1973 - where it is called the "differential threshold" - as the "minimum change in stimulus that can be correctly judged as different from a reference stimulus . . .". In the context of frequency, it can have a number of different meanings, all of which are related to the minimum frequency change required in a sound for that change to be detectable. For example, the difference limen shown in Figure 3 represents the just detectable frequency change in a sound signal whose frequency is being varied sinusoidally (i.e., modulated) four times per second.[10] This particular difference limen has a bandwidth of less than 1/100 of an octave, while - depending on the signal - other difference limens can also be obtained. The importance of these difference limens in the tone correction context is that, within bandwidths two orders of magnitude less than the one-third octave bandwidths, the ear still has the ability to detect the frequency shift in a pure tone. Since no one really knows what properties of a pure tone are annoying, it is not possible to exclude changing frequency content* within the one-third octave bandwidths as significant perceptually, even though the current practice of using one-third octave analysis would generally fail to reveal this.

Aside from reasonings in terms of established psychoacoustical concepts like the critical band, critical ratio and difference limen, other information is available which leads one to regard one-third octave band usage for tone correction procedures with suspicion.

Thurlow and Bernstein[12] showed that when people were presented with two simultaneous tones separated in frequency by about half the one-third octave bandwidths, they were able to detect the presence of two tones, or at least to detect a change in the overall quality of the resulting sound compared with a single tone. These two tones could have the same one-third octave band SPL as a single

---

*There is, of course, considerable evidence that changing sound level adds to annoyance, and some evidence that changing frequency (e.g., as sirens do) adds to detectability. In fact, changes in any stimulus characteristic stimulate the so-called orienting response (Ref. 11) and therefore add to the annoyance caused by any stimulus that, in the absence of this change, was already judged to be annoying.

tone but be perceived differently in terms of annoyance. This possibility was directly examined by Pearsons, et al.,[13] who found, for example, that five-tone complexes (i.e., a sound composed of five tones) were judged noisier than two-tone complexes when the frequency spacing between the highest and lowest frequency tones was as little as 1/10 octave. The difference averaged about 3 dB, which is significant when tone corrections generally used in aircraft noise are of about this order. The same reference also reports the extra noisiness occurring when 500 and 2000 Hz tones modulated at 5 Hz were presented: this seems to support the difference limen for frequency change as an important bandwidth for tone corrections since (as discussed above) the difference limen is the bandwidth at which tone frequency modulation of this type becomes detectable.

The preceding discussion on the appropriateness of one-third octave band analysis in formulating tone corrections should be supplemented by mentioning two other frequency-related matters which current tone corrections do not deal with. One is our perception of a tone in the presence of other noise, both tonal and broadband in nature, not necessarily close to that tone in frequency; the other is the relationship between tone perception and tonal duration.

A number of researchers[14, 15] have found that subjects may perceive more than two tones when two tones are presented, due either to beats within the ear's cochlea or outside it; the significance of beats within the cochlea is that the beat frequency tone(s) are not detected by the measurement microphone and therefore any required tone correction would need to be calculated from the two tones which generate the "beat" tone, a practice neither followed today, nor indeed possible within our current knowledge of the annoyance caused by the beats. Small, et al.,[16] found that when two tones are heard (separated in frequency by more than the critical bandwidths involved), the loudness of the higher frequency tone changes relative to its loudness when heard by itself. This too is not dealt with by current tone corrections. Finally, Webster, et al.,[17] and Schubert[18] found that a tone heard in the presence of white noise is perceived as of different frequency (i.e., it has a different pitch) than when heard alone. The change in pitch of the pure tone components in an aircraft noise spectrum that are due to changes in the broadband characteristics of the noise may add to annoyance in the same way that tone modulation does (see above). Bilger and Hirsh[19] showed that a tone of low frequency is audible only at a high level when presented in the presence of a much higher frequency band of intense noise. These two tone perception characteristics may well be relevant to aircraft tone corrections by virtue of aircraft tones occurring within a broadband spectrum of comparatively high SPL, but they are not dealt with by current methods.

The other facet of tone perception introduced for discussion above is the importance of tonal duration. When a short duration sinusoidal signal, i.e., a toneburst, is Fourier-analyzed, its spectrum is found to be composed not only of the sinusoidal frequency but also of frequencies each side of it. The significance of this is discussed by Garner[20] in the context of very short (10 ms) tonebursts - and Figure 4 decribes such a spectrum. Although aircraft noise sinusoids can last throughout the flyover, i.e., 10 seconds or longer, there is a considerable fluctuation in their level at a given receiver point, deriving perhaps from the high directivity of individual frequency components of the compressor noise as the airplane flies by (and the compressor axis changes in angle relative to the wavepath to a fixed observer).

These level fluctuations are easily evidenced in the 0.5 sec interval spectra as discussed later in this report in the context of a quantity known as "spectral change." It is possible moreover that these fluctuations last for considerably less than 0.5 sec and, though they are not short enough to sound like impulses, their duration may still be short enough for the Fourier spectrum of certain sinusoids to be sufficiently broadband to cause them to no longer sound tonal - while still appearing as "tonal" spikes in one-third octave analysis.

Other relevant aspects of tonal duration and tone perception are that tones and other sounds have reduced loudness when their duration is less than a certain time (known as the "integration time" of the ear). The masking process also takes time to be "set up," i.e., it has a "latency." The relevant times[2] associated with both these hearing characteristics range up to about 0.5 sec and they may therefore have some importance in aircraft tone perception, which is not catered to in current tone correction procedures.

All the preceding discussion in Section 2.2.1 throws doubt on the correspondence between tone perception, on the one hand, and current tone correction procedures on the other. To examine, more accurately, the true nature of tonal features in the airplane noise context, Wyle analyzed the difference between the tones indicated by one-third octave band analysis and those indicated by a narrower band analysis, as follows.

Narrow Band Analysis of Aircraft Sounds

Noise spectra from a B-727 aircraft were chosen since this aircraft frequently displays one dominant tone in its noise signature (at the fan blade passage frequency). Since the study of multiple tones was also an objective of this work, an A-300 Airbus was also chosen for analysis as representative of new technology high bypass ratio fanjet aircraft, whose noise signatures have "buzz-saw" tones at about 200 to 400 Hz and blade passage frequency tones at 2000 to 3000 Hz.

Analog tapes for the analysis were obtained from the Federal Aviation Administration, U. S. Department of Transportation.[21] These noise level measurements were taken at Dulles International Airport by members of the Transportation Systems Center (TSC) Noise Measurement and Assessment Laboratory. The flyovers simulated* both approach and departure conditons for the B-727 and the A-300 aircraft.

The equipment used by Wyle Laboratories in the frequency analysis of the analog tapes of these flyovers was:

Frequency Analyzer:	Nicolet Scientific Corporation 444A Mini-Ubiquitous FFT Computing Spectrum Analyzer
Tape Recorder:	Nagra IV-SJ Tape Recorder
X-Y Recorder:	Hewlett-Packard 7035B X-Y Recorder

Tapes were made of the beginning of the departure flyovers. The loops ended at the position where the Doppler shift of the dominant tone was noted subjectively. The tape loops for the departures of the B-727 and A-300 were then averaged by the analyzer's summing circuit over 64 spectra (true power averaging) for each of two frequency ranges, as set out in Table 5. This was sufficient to average out random signals and consequently clearly establish the presence of tones where they existed in the broadband noise. In the case of the low frequencies (up to 500 Hz), a spectrum measurement was made with the tape recorder running

*"Aircraft simulated a noise abatement departure procedure, with engine thrust reduced to a cutback EPR (Engine Pressure Ratio) prior to passing over the listening site. The B-727(s), in a clean configuration, at 210 knots and level at 2000 (ft) increased power to cutback EPR and initiated a climb when approximately 1 mile from the monitoring site. At 3000 (ft) MSL (Mean Sea Level) the aircraft reduced power, turned onto a race track pattern and descended to 2000 (ft) for the next simulated departure. The A-300 flew a similar pattern but initiated its climb at 2500 (ft) and climbed to 4000 (ft) MSL." (Ref. 21).

without tape to record noise generated by its mechanical and electrical components. This spectrum was then subtracted, on an energy basis, from the flyover spectrum. The spectra at frequencies above 500 Hz were not subjected to this background noise, so this procedure was not carried out in these ranges.

Table 5

Characteristics of the Frequency Analyses Performed on B-727 and A-300 Airplanes

Analysis No.	Frequency Range, Hz	Bandwidth, Hz	Memory Period of Analyzer at the Stated Bandwidth, sec	Sampling Duration for 64 Samples, sec
1	12.5 - 5000	12.5	0.02	5.1
2	25 - 10,000	25	0.04	2.6

The real-time analyzer has the capability of 400 line resolution for each spectrum, consequently giving the bandwidths described in Table 5. Table 5 also shows that the sampling duration was well above the 0.5 second (or less) which our hearing mechanism uses. The spectra obtained therefore indicate fewer tones than we may actually perceive, but even so are strikingly different from those obtained by conventional one-third octave analysis, as shown in Figures 5 and 6.

Table 5 indicates that the bandwidths used were roughly equivalent to those of the difference limen (see Figure 3) at the mid-to-upper frequencies in each range.

Figure 5(a) (for the B-727) shows at least two tones in the narrow band spectrum, at about 2 kHz and 4 kHz - yet neither is incontrovertibly a tone in the one-third octave spectrum, and only one would be recognized by the FAR Part 36 and SAE ARP 1071 procedures. In the narrow band analysis of Figure 5(a), there is also some evidence of multiple tones in the 1-2 kHz region, which may to some extent have been lost in the slightly wider, narrow bandwidth analysis shown in Figure 6(a).

Figures 5(b) and 6(b) display A-300 spectra easily characterized as multiple tone spectra by the series of bunched spikes in the lower, narrow band curves. Again, these are more evident in the 12.5 Hz bandwidth analysis of Figure 5(b) than

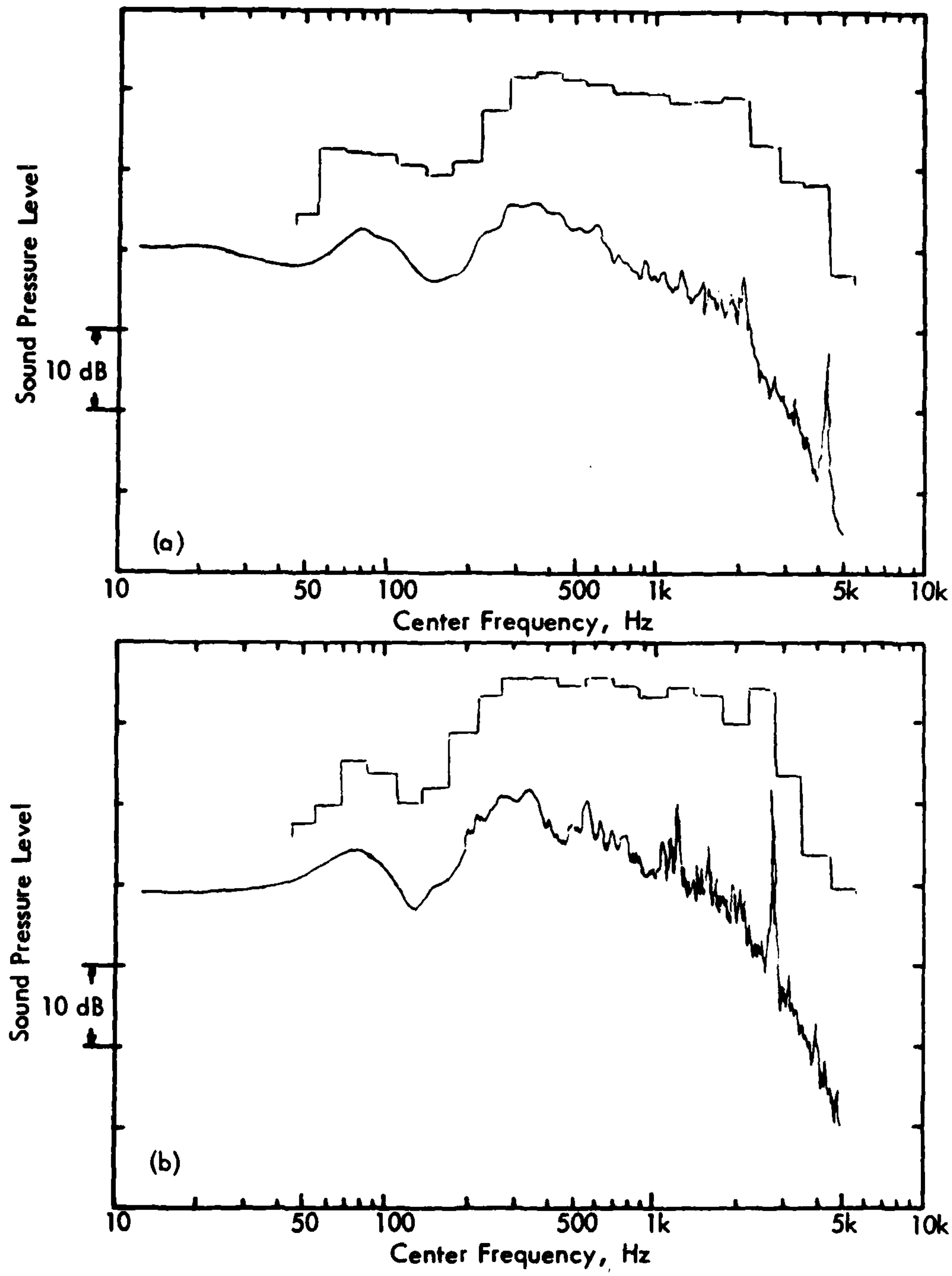


Figure 5. Narrow Band and One-Third Octave Analyses of Simulated Takeoffs by (a) B-727 (top) and (b) A-300 (bottom).

Analysis 1: 12.5 Hz bandwidths covering the 12.5 - 5000 Hz region (spiky curves); one-third octave analysis covering the same region (flat-top curves).

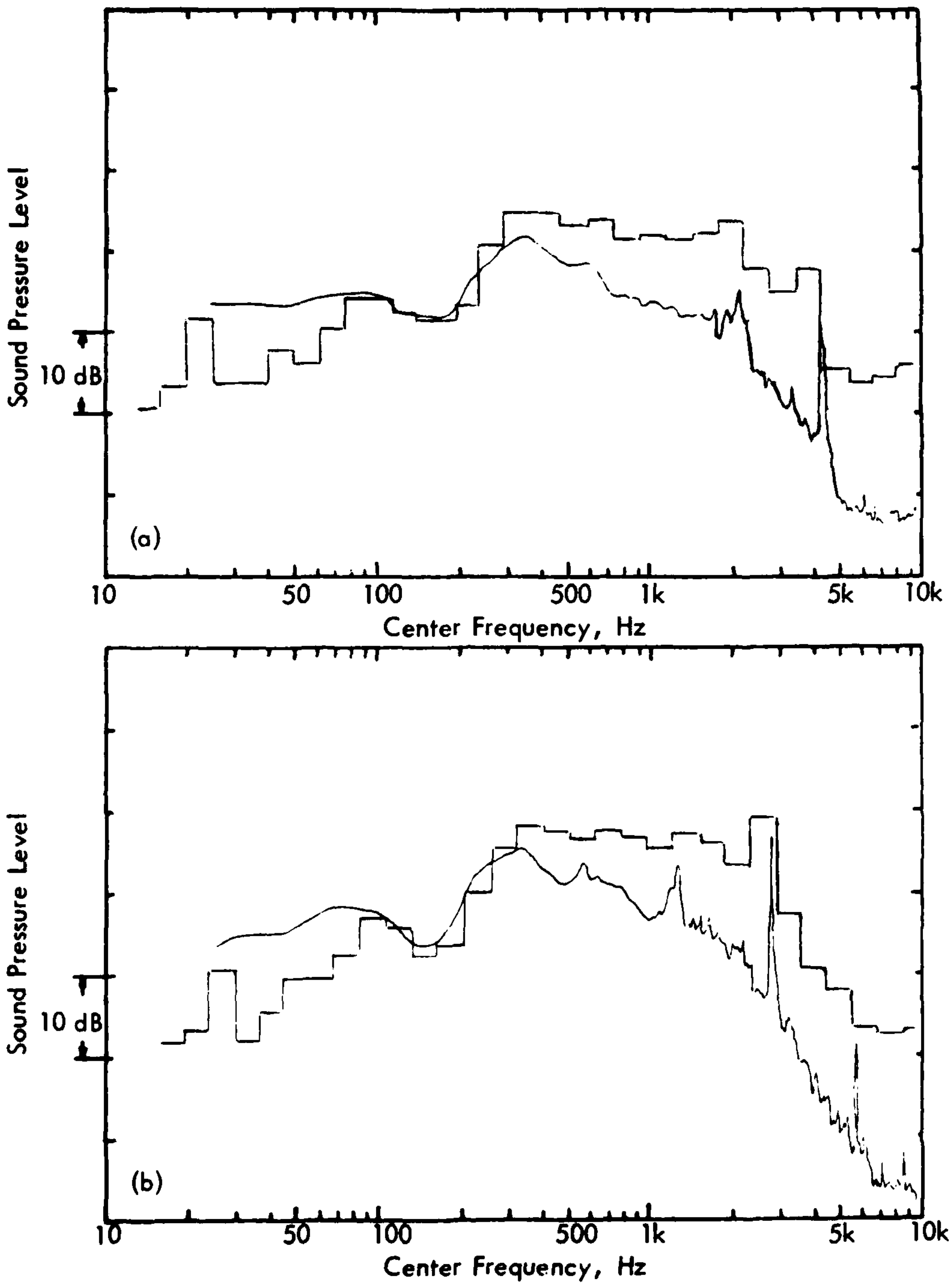


Figure 6. Narrow Band and One-Third Octave Analyses of Simulated Takeoffs by (a) B-727 (top) and (b) A-300 (bottom).

Analysis 2: 25 Hz bandwidths covering the 25 - 10,000 Hz region (spiky curves); one-third octave analysis covering the same region (flat-top curves).

in the 25 Hz bandwidth analysis of Figure 6(b). Nearly all the bunched spikes are lost in the one-third octave band analyses, where the only clearly established tones are at 2500 Hz and 25 Hz. In Figure 5(b), a narrow band spike at 1300 Hz stands nearly 10 dB above the immediately neighboring spectrum, yet the corresponding octave protrusion is only about 2 dB. In Figure 6(b), a narrow band spike at 5500 Hz protrudes by 13 dB, while the corresponding one-third octave protrusion is not visible at all. The bunched spikes between 1 and 2 kHz in the narrow band spectrum of Figure 5(b) provide clear examples of how more than one spike can exist within a one-third octave bandwidth yet be undetected by one-third octave analysis.

This section has indicated that there are many reasons for questioning the form of current tone correction procedures based on one-third octave analysis and 0.5 second time intervals. Alternate procedures, especially those using narrower frequency bandwidths (anywhere down to 1/100 octaves), might correct more adequately for tones in general and multiple tones in particular. The discussion and analysis in this section is intended to illustrate this for future reference. As stated at the outset in Section 2.2, existing industry practice is to use one-third octave analysis, and we have therefore concentrated on one-third octave analysis in the tone corrections analysis in the rest of this report. Section 2.2.2 describes revised one-third octave band tone correction procedures that are tested against one another in Section 2.3.

2.2.2 Comparison of Alternative One-Third Octave Band Tone Corrections

In addition to the FAR Part 36, SAE ARP 1071 and Kryter-Pearsons tone corrections described in Section 2.1, a number of revisions to these procedures, and a completely different one were examined. The various features which differed from those in FAR Part 36 were as follows:

- o Spectrum smoothing (see Section 2.1) differed in the case of the SAE ARP 1071* procedure (and therefore also in our revised variants of it).
- o Multitones (see Section 2.1) were considered by the Kryter-Pearsons* procedure (and therefore also in our revised variants of it).
- o Time-weighting, defined by weighting the amount of the tone correction to give it a different value through the flyover, was applied in various forms to the FAR Part 36, SAE ARP 1071 and Kryter-Pearsons procedures. This revision to those procedures was to test out a number of hypotheses as to when, during a flyover, tones might be most annoying. (These hypotheses are spelled out in Table 6.)
- o Level-weighting, defined by multiplying the amount of the tone correction by a constant independent of time, was applied in various degrees to all the procedures, for two reasons. One reason was to test out a prediction by Wells[22] that the Kryter-Pearsons tone correction is too small; the other reason was to nondimensionalize the time-weighting**. The level-weightings were designated by the variable, a, which consisted of the multiplier applied to the value of the tone correction; for the unaltered tone correction, a=1.

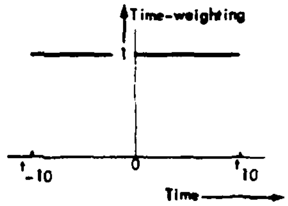
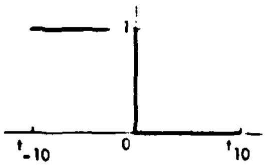
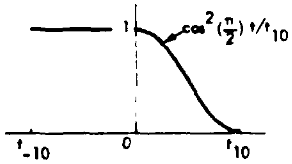
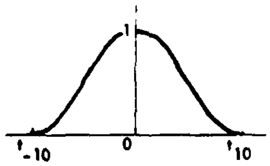
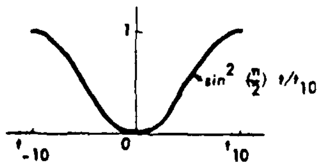
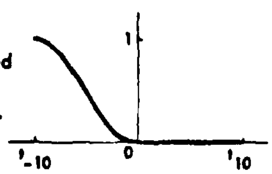
---

*A fuller explanation of the features of the SAE ARP 1071 and Kryter-Pearsons procedures which differ from FAR Part 36 is given in Section 2.1.

**The time-weighting reduced the magnitude of the tone correction at some points of the flyover while holding it the same at others - it never increased it. Therefore the question might arise as to whether any benefit due to time-weighting should be attributed to (1) shaping the tone correction in time, or (2) merely decreasing its amount. To ensure that the first of these was fully explored, the whole function was therefore multiplied by various level-weightings so that a whole range of "amounts" of tone correction were explored.

Table 6

The Hypotheses for and Descriptions of the Various Time-Weighting Functions Considered. Each function is given the code, b, shown in the right-hand column. The time at PNLTM is designated t=0; t_{-10} and t_{10} are the times at which PNLT is 10 dB lower than PNLTM, the times in all cases being established from the non time-weighted PNLT history. In the instance (b=02) when the time-weighting is discontinuous at t=0, the value at t=0 is the value defined by the time-weighting to the left, i.e., earlier.

HYPOTHESIS	TIME-WEIGHTING FUNCTION	CODE, b
1. The current time-weighting is appropriate.		00
2. The tones are annoying only on the approach.		02
3. The tones are annoying on the approach and become gradually less so thereafter.		03
4. The tones are much more annoying when the airplane noise is greatest than at other times.		33
5. The tones are annoying throughout the flyover but their effect at times other than t=0 is not adequately reflected in the energy summation of PNLTM (which leads to EPNL); they are therefore accentuated as shown.		55
6. The tones are annoying throughout the approach (but their effect is then not adequately conveyed by the energy summation of PNLTM which leads to EPNL); after t=0, the tones are not annoying.		52

- No tone correction. A level-weighting of zero (a=0) "cancelled" the tone correction. In this report, the ensuing EPNL is termed Integrated Perceived Noise Level (IPNL),* and was included as a datum against which to judge the performance of all the various methods.

- Averaged tone correction. All the various accepted procedures add a tone-correction to each PNL(k) before summing PNLT(k) over the relevant time history (the time for which PNLT(k) $\geq$ PNLTM - 10) to obtain EPNL. In this way, the considered PNLT(k) values contribute to EPNL in proportion to their sound power - a rational procedure which may, however, neglect much of the noisiness of the airplane's tones if the tone corrections at or near PNLTM happen to have small values. The revised procedure computes EPNL by adding to IPNL the average tone correction, $\bar{C}(k)$, across the relevant time history, which in this case is the time for which PNL(k) $\geq$ PNLM - 10. $\bar{C}(k)$ can be defined for each of the three baseline tone correction methods and is referred to here as $\bar{C}(k)_{FAA}$, $\bar{C}(k)_{SAE}$ or $\bar{C}(k)_{KP}$.

- Spectral fluctuation. An additional correction procedure, defined as Spectral Change in Reference 23, was also tested. Spectral Change, designated here as SC, is a variable describing the amount by which the SPL in each of the individual frequency bands varies with time over and above the variation of the overall A-weighted sound level. The period considered is the noisiest 3 seconds of the flyover. The computation of SC is defined in, and illustrated by, Table 7. Spectral Change is not a direct tone correction, but it may account for the noisiness of tones indirectly if tonal SPLs fluctuate with time more than the sound's broadband SPLs do. Shepherd[23] tested a number of coefficients, i.e., multipliers, for Spectral Change when adding this quantity to PNL and regressing the result with judged annoyance. One good value for the coefficient was 0.5.

*There are two forms of IPNL. The one used here is as defined above. The other form is the energy summation of PNL over the EPNL interval. The two forms differ by only a constant and the adequacy of each of them is treated equally by our analysis.

Table 7

Computation of Spectral Change.[23] (Usage employed herein considered one-third octave bands and an arithmetic, rather than logarithmic, averaging of line 7 to give "Spectral Change")

		Octave Band Center Frequency, Hz								
		63	125	250	500	1k	2k	4k	8k	dB(A)
(1)	Spectrum 1-½ secs. before maximum†	78	80	81	72	81	75	72	70	83.46
(2)	Spectrum at maximum	82	84	85	83	79	78	75	70	85.53
(3)	Spectrum 1-½ secs. after maximum	84	85	88	80	74	73	70	68	83.27
(4)	Change between times (1) and (2)**	4	4	4	11	2	3	3	0	
(5)	Change between times (2) and (3)**	2	1	3	3	5	5	5	2	
(6)	Total Change ((4) + (5))	6	5	7	14	7	8	8	2	
(7)	Change exceeding dB(A) level change.*	1.67	0.67	2.67	9.67	2.67	3.67	3.67	0.00	

Spectral Change = (1.67 + 0.67 + 2.67 + 9.67 + 2.67 + 3.67 + 3.67 + 0)/8 = 3.09

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

†The time interval to which "maximum" refers is that at which the A-weighted sound level is a maximum. This generally, but not always, corresponds with the interval at which PNL = PNLM or PNLT = PNLTM.

**Absolute value.

2.3 Test of Existing and Revised Procedures

2.3.1 Data Base

A number of researchers have reviewed the literature to try to assess the superiority of various tone correction procedures over one another, and over no correction at all. These include Kryter,[2, 24] Ollerhead,[25] and Scharf and Hellman.[26] In general, the results of this type of review are rankings of the various metrics in order of their success in correlating with judged noisiness. A common feature of this type of review, however, is that these rankings often differ from study to study and that such differences are only partially diagnosed by references to the various experimental procedures and data bases employed.

As indicated in Section 2.3.3, the authors believe that research into the superiority of one metric over another should include deeper studies into the reasons why these experiments sometimes produce conflicting results. In this study, we have therefore preferred to investigate one data base in-depth rather than review several data bases cursorily. As shown in subsequent sections, this in-depth study has been successful in demonstrating how conflicting results can emerge from such studies depending, for example, on the data taken into the analysis. The results therefore offer better perspective conclusions as to the superiority of one metric over another, and indicate pitfalls to avoid in future such analyses.

The data base chosen for in-depth analysis was one by Powell,[27] which was derived from a carefully conducted psychoacoustical experiment, using a comparatively large number of subjects (96) and 120 flyover noises from six different airplanes on landing and takeoff. The airplane types spanned the range of interest (with the exception of helicopters) and involved a propjet (CV-640), a subsonic turbojet (DC-8TJ), two low bypass fanjets (DC-8TF and B-737), a high bypass fanjet (B-747), and a supersonic turbojet (Concorde). Their spectra are shown in Figure 7. Powell's data were kindly furnished to Wyle for this study, and various additional metrics were then computed and compared for their ability to predict the observed subjective judgments.

2.3.2 Type of Analysis

The type of analysis used was as follows. For each of the aircraft flyover signals, Powell[27] calculated the EPNL values for the signal levels which produce equal noisiness judgments by his subjects. The result is shown in Figure 8. The

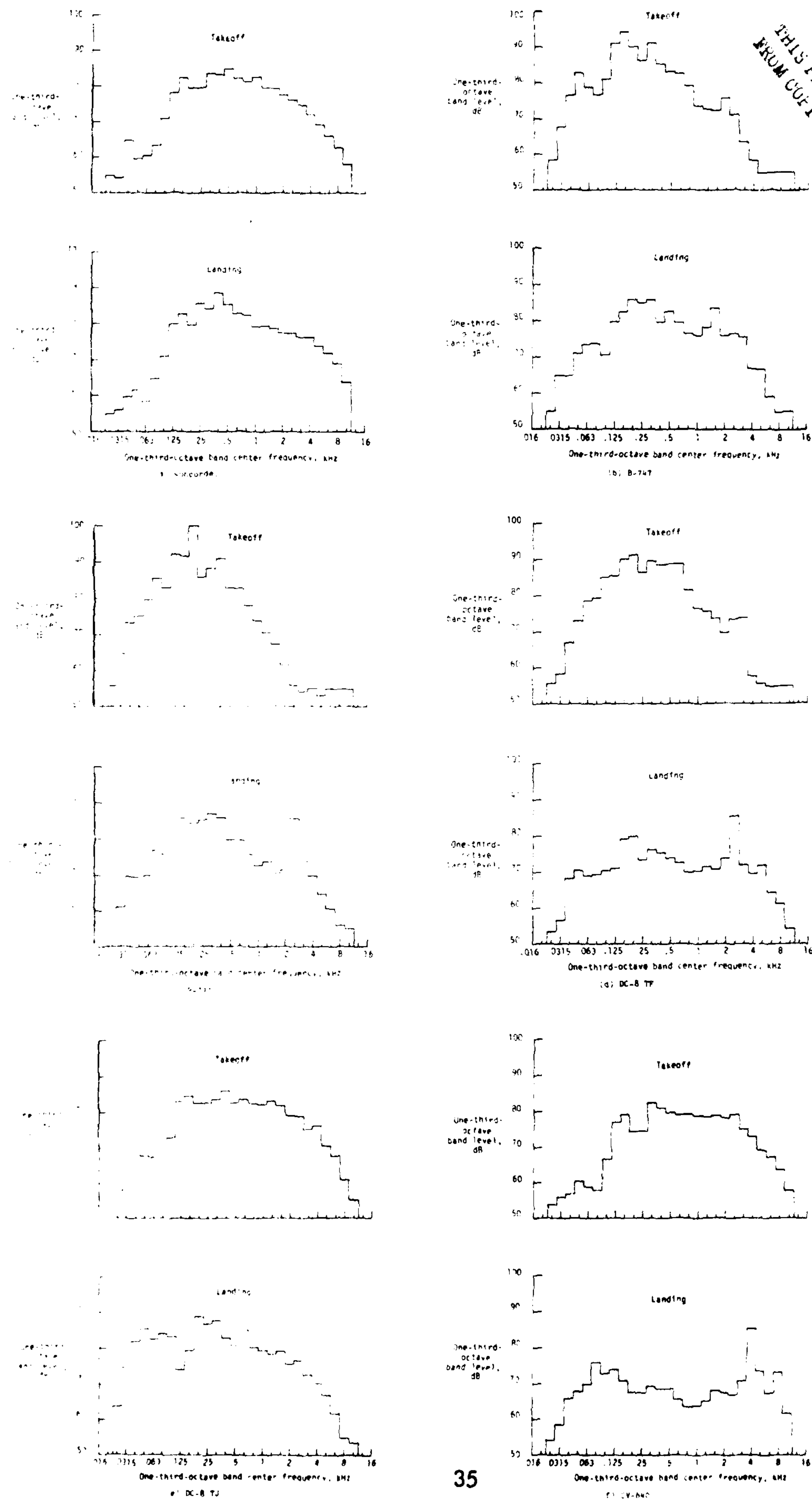


Figure 7. Spectra for the Airplane Flyovers in Powell's Data (Reference 27)

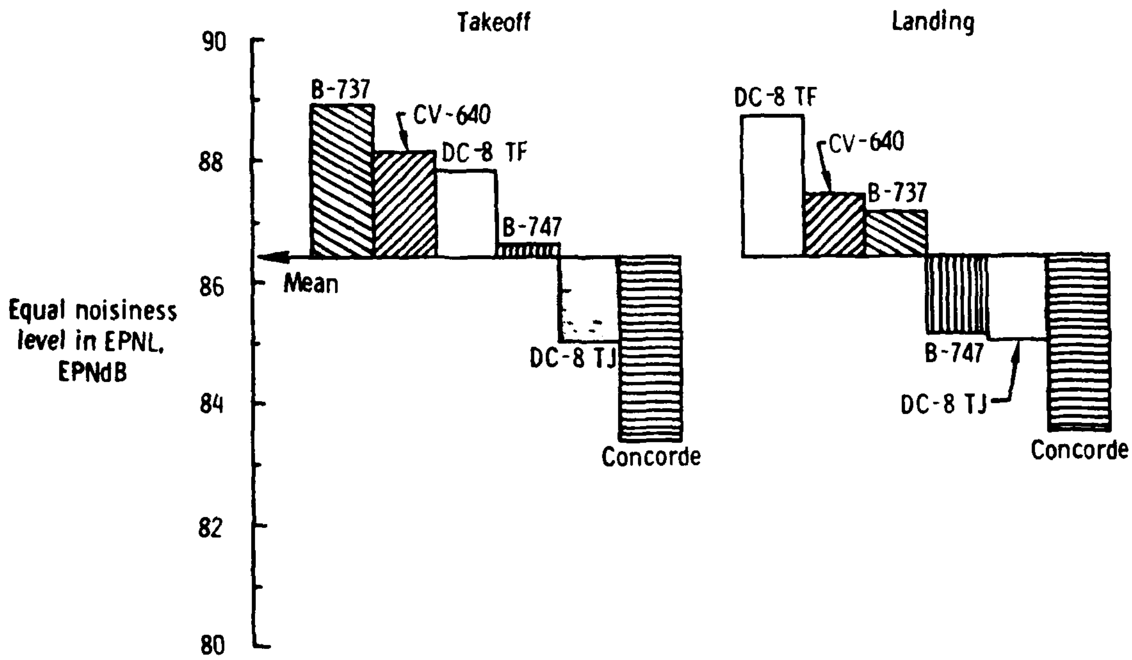


Figure 8. Equal Noisiness Level in EPNL for Different Airplane/Flyover Categories (from Powell, Reference 27)

success of EPNL can be measured by the standard deviation for the various aircraft signals in Figure 8 about their mean value: perfect success would imply zero standard deviation since then equal EPNL values would give equal noisiness. The standard deviation in Figure 8 for both the takeoff and landing data is about 2 dB.

To investigate the virtues of noise metrics other than EPNL, standard deviations were therefore calculated for them also, in all cases considering the airplane flyover levels judged to be equally noisy. The superiority or inferiority of these metrics was gauged by the degree to which these standard deviations were smaller or larger than one another.

Although any difference in these standard deviations can be regarded as some evidence for the superiority of one metric over another, a statistical test was also performed on some of them as described in Section 2.3.4.

In addition to standard deviation, the "range" of each metric is also indicated in some of the results. Its purpose was to expose any instances in which metric was successful in collapsing the data for all spectra except perhaps for a single, but important, aircraft type. Although the standard deviation might then be very small, a large "range" for the metric could imply its failure to deal fairly with a particular type of noise, a type of power plant, an airplane category, or the airport communities most exposed to that airplane category.

Section 2.3.3 also describes an important facet of the analysis method concerning the separation of data.

2.3.3 Separation of Data

Powell's data from Reference 27 were studied in two parallel analyses. One considered all 12 of his aircraft/flyover categories, i.e., six aircraft types times two flyover conditions (landing and takeoff). The other considered only seven aircraft/flyover combinations, which were chosen using the following rationale.

Experiments in which any dependent variable (e.g., subjective judgments) is influenced by more than one independent variable (e.g., several noise stimulus characteristics) require special care in experimental procedure and analysis. The reason for this is that, if the independent variables are correlated with one another, as is very often the case, an analysis which takes no account of those intercorrelations can produce misleading results. To illustrate this, consider a hypothetical study in which judged noisiness (dependent variable) is to be predicted from two variables: airplane type (jet versus prop), and aircraft height. Aircraft type and aircraft height are the independent variables. The experimental procedure involves presenting tapes of takeoff noise by a jet and a propeller plane as measured at a given point near an airport, but the experimental procedure and subsequent analysis fail to take account of the fact that the jet's takeoff is steeper than the prop-plane's, i.e., the independent variables, aircraft type and aircraft height, are correlated with each other. As a result of this failure, the analysis results in judged noisiness being found to be greatest not only with the jet aircraft but with the greater height that the jet has attained over the microphone. (A consequent absurd conclusion that considered all of the aircraft signatures would be that the lower the aircraft, the less intense the sound.)

This hypothetical example highlights an experimental problem in which the absurdity of the result is obvious enough that (1) the experiment is very unlikely to have been structured in the described manner, or (2) the analysis could have obviated the potential problem in a number of ways, including analyzing the effect of height for the two aircraft types separately.

In most experiments, the intercorrelations of the independent variables are not as strong as in the example. However, they might *still be substantial. In* addition, the result being researched is usually less obvious in advance, and so the misleading result may go undetected.

An additional problem in certain data is that there may be an interaction between the dependent variables; for example, one aircraft type's flyover may be judged less noisy with increasing height and the other aircraft type's flyover may be judged more noisy with increasing height (say, because of a slower flyby). In this case, the experiment, if not properly analyzed, might yield an effect on noisiness by height that depended on the numbers of each airplane type that were featured in the experiment. Therefore, if these numbers were not proportional to those in the aircraft fleet in general, the experimental result would be erroneous.

These two data characteristics are termed *multicollinearity* (for the intercorrelations of the independent variables) *and interactions* (for the differing effects of one variable depending on the value of another). A fuller explanation of multicollinearity is given in Reference 28, and an example of an experiment designed to avoid it is Reference 29. Note that such experiments invariably require either synthetic stimuli or very carefully selected real stimuli. An easy-to-understand explanation of interactions is given in Reference 30. Note also that both these effects are now being recognized by some aircraft noise researchers; for example, synthetic stimuli and factorial experimental designs[30] are being used by NASA;[31] the latter reference also points out that "current tone-correction procedures did not adequately account for the effects of the interaction of tonal content with sound pressure level." This suggests that different tone corrections are therefore indicated for different levels of noise, i.e., EPNL should not just be a linear function, f, of PNL and tone correction like f(PNL, tone correction) but also should recognize their interrelationship as a function like f(PNL, tone correction, PNL x tone correction).

In the context of our extended analysis of Powell's data,[27] the above considerations led to a study of whether the noise variables might be correlated with one another. Figure 9 presents a plot of the tone correction quantity, TC, against the spectral descriptor, IPNL-SEL, for Powell's data. Here TC is a simple tone correction based on the maximum amount by which any one-third octave band SPL, at the time interval corresponding to PNLM, stands above the average SPL of the two adjacent bands.* The spectral descriptor, IPNL-SEL, where SEL is the Sound Exposure Level (an integration of the A-weighted level between the 10 dB-down points), is analogous to the difference between the D-weighted and A-weighted sound levels, since PNL and therefore IPNL have similar origins to dB(D). The relationship between TC and IPNL-SEL is therefore the relationship between a tone correction on the one hand and the spectrum shape** for the aircraft type on the other. As shown in Figure 9, a correlation (r = 0.60) exists between these quantities in Powell's data. Neglect of this correlation could lead to selecting a tone correction that is not so much a tone correction as a method of reducing whatever variance is introduced by any inadequacy in PNL, for example. A lower correlation (r = 0.54) is present, however, for the flyovers shown circled, as discussed further below.

---

*i.e., $TC = \left\{ SPL(i,k) - \frac{1}{2}\left[(SPL(i-1,k) + SPL(i+1,k)\right]\right\}_{max}$ for k corresponding with PNLM.

**A more obvious, geometrically-derived relationship was sought but not found in the spectra shown in Figure 7.

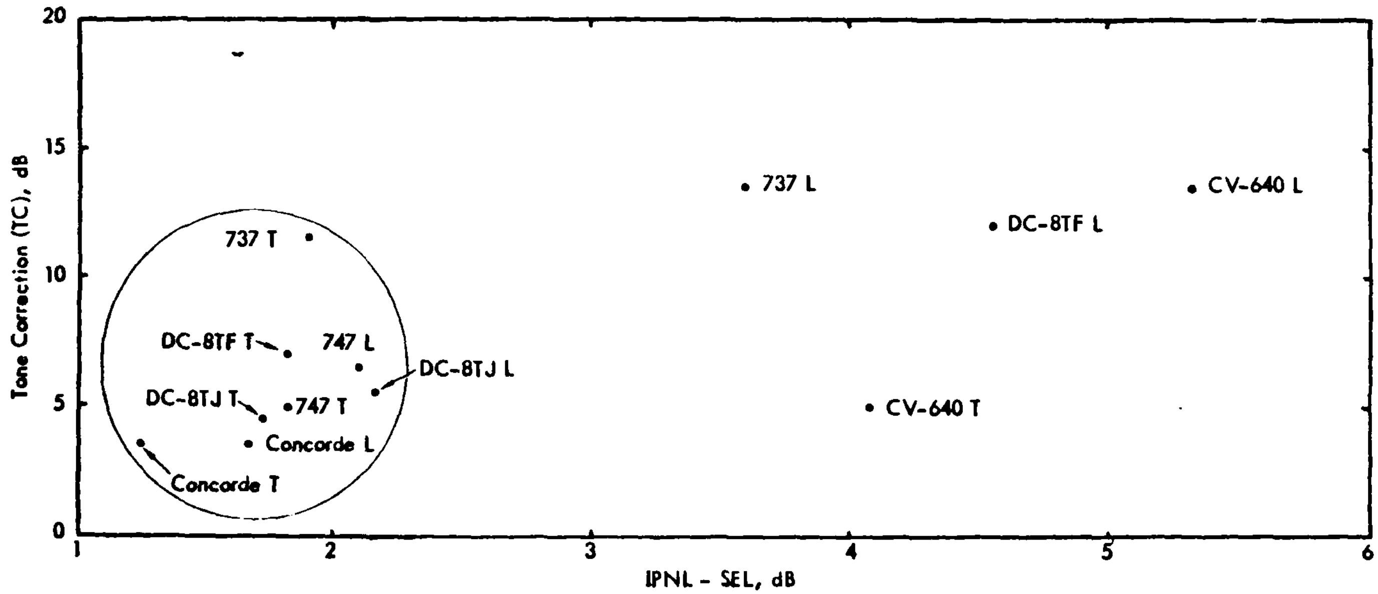


Figure 9. Correlation Between a Simple Tone Correction and the Shape of the Spectrum in Powell's Data (Reference 27). See Section 2.3.3 for definition of TC.

When it is not possible to obtain nonmulticollinear data when analyzing the usefulness of each component term in a metric, gross comparisons of the success of one metric compared with another in describing human response can still be made using multicollinear data. The validity of these comparisons depends on obtaining sample data that is representative of real-life data. In the context of the present study, the rankings of the various tone-corrected EPNLs according to their success in describing annoyance should be the same for the sample data as in real-life, if performed on a set of aircraft flyovers representative of the occurrence of these aircraft flyovers in the real world.

In our reanalysis of Powell's data, we have analyzed two sets of data in parallel - one set being all the applicable data in Reference 27, and the other a subset of these data in which the multicollinearity has been slightly reduced. Because, however, substantial multicollinearity remains in the reduced data base, and is likely to exert a substantial influence on the results, the reduction in the data base was accomplished by taking some cognizance of the prevalence of the various data base aircraft types in the aircraft fleet as a whole.

Examination of the circled data in Figure 9 shows that Concorde flyovers, in particular, are represented far beyond their incidence in real life. Therefore one Concorde flyover, the landing one, was removed. The resulting correlation coefficient between TC and IPNL-SEL for the remaining seven flyovers in the circle was 0.53, compared with the 0.60 correlation coefficient for all Powell's data, thus conferring a very slight further reduction in multicollinearity.

The seven flyovers in the reduced data base were, therefore, as follows:

747	Landing
DC-8TJ	Landing
737	Takeoff
747	Takeoff
DC-8TF	Takeoff
DC-8TJ	Takeoff
Concorde	Takeoff

Multicollinearity may, of course, still be present in the form of correlations involving other attributes of the noise. One such attribute is flyover duration, although duration has been shown[31] to be well accounted for by the EPNL metric. The effects of multicollinearity have not been investigated exhaustively in this study; they should ideally be taken account of when first developing metrics.

2.3.4 Results

The metrics evaluated in this analysis can be summarized as follows:

- EPNL (FAR Part 36) abbreviated to EPNL (FAA)
  EPNL (SAE ARP 1071) abbreviated to EPNL (SAE)
  EPNL (Kryter-Pearsons) abbreviated to EPNL (KP)

  These metrics are described in Section 2.1. Each was investigated for all 30 combinations of level-weighting, a=1, 0.5, 1.5, 2, 3, and time-weighting, b=00, 02, 03, 33, 55 and 52, as described in Section 2.2.2. a=1, b=00 are baseline values for these quantities.

- IPNL and SEL

  IPNL is described in Section 2.2.2. SEL is the Sound Exposure Level, an integration of the A-weighted sound level over the 10 dB-down interval. Neither IPNL nor SEL contains a tone correction; therefore neither can have a level-weighting or time-weighting. Note that each of the EPNL variants, when a=0, is the same as IPNL.

- IPNL + $\bar{C}(k)_{FAA}$, IPNL + $\bar{C}(k)_{SAE}$ and IPNL + $\bar{C}(k)_{KP}$

  These metrics are described in Section 2.2.2. $\bar{C}(k)$ is an arithmetic average of the FAA, SAE or KP tone correction values in each 0.5 second interval between the 10 dB-down points. The tone corrections may be level-weighted and time-weighted as before.

- SC

  SC is the quantity Spectral Change as defined in Section 2.2.2. Its usefulness was investigated (with the coefficient 0.5) when added to a limited selection of the other metrics.

A selection of results for the reduced and unreduced* data bases obtained from Reference 27, including the individual flyover sound level values of some representative metrics, is given in Table 8. Table 9 is a summary table which only presents the standard deviations, but does so for all the metrics. Figures 10 and 11 contain representative plots of the standard deviations for two of the metrics to show graphically the effect of changing level-weighting, a, and time-weighting, b.

*The nature and significance of the reduced and unreduced data sets are discussed in Section 2.3.3.

Table 8

Sound Levels, Standard Deviations and Ranges, in decibels, for Some Representative Metrics Analyzed Here from Powell's Data[27]

		Metric						
		EPNL (FAA)	EPNL (SAE)	EPNL (KP)	IPNL	SEL	EPNL (FAA)	EPNL (SAE)
Flyover		a=1 b=00	a=1 b=00	a=1 b=00	no a no b	no a no b	a=2 b=00	a=2 b=00
Reduced Data Base								
747	L	85.2	85.1	86.3	83.8	81.7	86.6	86.5
DC-8TJ	L	85.1	85.0	85.4	83.9	81.8	86.3	86.2
737	T	88.9	89.2	88.3	87.9	86.0	90.1	90.9
747	T	86.7	86.9	86.5	85.0	83.2	88.4	88.9
DC-8TF	T	88.2	88.5	87.8	86.7	84.9	89.7	90.5
DC-8TJ	T	85.0	85.1	85.3	83.8	82.1	86.3	86.5
Concorde	T	83.4	83.8	83.7	82.5	81.3	84.5	85.7
Std. Dev.		1.95	2.02	1.57	1.89	1.81	2.04	2.18
Range		5.5	5.4	4.6	5.4	4.7	5.6	5.3
Unreduced Data Base								
737	L	87.1	87.3	89.1	84.8	81.2	90.1	90.6
747	L	85.2	85.1	86.3	83.8	81.7	86.6	86.5
DC-8TF	L	88.8	89.3	91.4	85.3	80.7	92.3	93.3
DC-8TJ	L	85.1	85.0	85.4	83.9	81.8	86.3	86.2
Concorde	L	83.5	83.4	83.8	82.3	80.7	84.6	84.7
CV-640	L	87.5	88.1	90.8	83.8	78.5	91.4	92.4
737	T	88.9	89.2	88.3	87.9	86.0	90.1	90.9
747	T	86.7	86.9	86.5	85.0	83.2	88.4	88.9
DC-8TF	T	88.2	88.5	87.8	86.7	84.9	89.7	90.5
DC-8TJ	T	85.0	85.1	85.3	83.8	82.1	86.3	86.5
Concorde	T	83.4	83.8	83.7	82.5	81.3	84.5	85.7
CV-640	T	88.0	88.6	87.2	86.6	82.5	89.6	90.9
Std. Dev.		1.95	2.12	2.48	1.70	1.97	2.62	2.90
Range		5.5	5.9	7.7	5.5	7.5	7.9	8.7

L = landing, T = takeoff, a = level-weighting, b = time-weighting code
The metrics are defined in Sections 2.1 and 2.2.2 (a = 1, b = 00 are baseline values)

Table 8 (Continued)

Flyover		EPNL (KP) a = 2 b = 00	EPNL (FAA) a = 2 b = 55	EPNL (SAE) a = 2 b = 55	EPNL (KP) a = 2 b = 55	IPNL + $\bar{C}(k)_{FAA}$ a = 1 b = 00	IPNL + $\bar{C}(k)_{SAE}$ a = 1 b = 00	IPNL + $\bar{C}(k)_{KP}$ a = 1 b = 00
Reduced Data Base								
747	L	88.9	85.0	85.1	86.2	85.2	85.2	86.3
DC-8TJ	L	87.0	84.8	84.8	85.1	85.1	85.0	85.5
737	T	88.9	88.8	89.0	88.3	88.9	89.1	88.3
747	T	88.1	86.6	86.7	86.2	86.7	86.9	86.5
DC-8TF	T	89.0	87.9	88.1	87.4	88.1	88.2	87.6
DC-8TJ	T	86.8	84.9	85.0	85.1	85.1	85.2	85.3
Concorde	T	85.1	83.6	84.1	83.4	83.6	83.8	83.7
Std. Dev.		1.45	1.88	1.85	1.60	1.88	1.92	1.50
Range		3.8	5.2	4.8	4.8	5.3	5.3	4.5
Unreduced Data Base								
737	L	94.3	87.8	88.2	90.7	87.4	87.7	89.5
747	L	88.9	85.0	85.1	86.2	85.2	85.2	86.3
DC-8TF	L	97.6	88.2	88.9	91.5	88.6	89.2	91.1
DC-8TJ	L	87.0	84.8	84.8	85.1	85.1	85.0	85.5
Concorde	L	85.4	83.5	83.3	83.9	83.6	83.6	84.1
CV-640	L	97.9	87.5	88.2	91.8	87.3	88.1	90.5
737	T	88.9	88.8	89.0	88.3	88.9	89.1	88.3
747	T	88.1	86.6	86.7	86.2	86.7	86.9	86.5
DC-8TF	T	89.0	87.9	88.1	87.4	88.1	88.2	87.6
DC-8TJ	T	86.8	84.9	85.0	85.1	85.1	85.2	85.3
Concorde	T	85.1	83.6	84.1	83.4	83.6	83.8	83.7
CV-640	T	87.6	88.3	89.1	86.9	88.1	88.8	87.1
Std. Dev.		4.42	1.92	2.13	2.84	1.89	2.07	2.38
Range		12.8	5.2	5.8	8.3	5.3	5.5	7.4

L = landing, T = takeoff, a = level-weighting, b = time-weighting code
The metrics are defined in Sections 2.1 and 2.2.2 (a = 1, b = 00 are baseline values)

Table 8 (Continued)

Flyover		IPNL + $\bar{C}(k)_{FAA}$ a = 2 b = 00	IPNL + $\bar{C}(k)_{SAE}$ a = 2 b = 00	IPNL + $\bar{C}(k)_{KP}$ a = 2 b = 00	IPNL + $\bar{C}(k)_{FAA}$ a = 2 b = 55	IPNL + $\bar{C}(k)_{SAE}$ a = 2 b = 55	IPNL + $\bar{C}(k)_{KP}$ a = 2 b = 55	IPNL + ½ SC no a no b	
747	L	86.7	86.7	88.8	85.5	85.6	86.8	84.4	Reduced Data Base
DC-8TJ	L	86.3	86.1	87.1	85.2	85.3	85.7	85.5	
737	T	89.9	90.3	88.7	88.9	89.0	88.3	88.1	
747	T	88.4	88.9	88.1	86.9	87.1	86.5	86.1	
DC-8TF	T	89.4	89.7	88.6	86.4	88.2	87.5	87.5	
DC-8TJ	T	86.5	86.5	87.0	85.5	85.5	85.6	85.0	
Concorde	T	84.7	85.1	85.0	83.9	84.1	83.8	85.2	
Std. Dev.		1.85	2.01	1.37	1.58	1.74	1.48	1.35	
Range		5.2	5.3	3.7	5.0	4.9	4.5	3.7	
737	L	90.0	90.6	94.4	88.1	88.4	90.7	85.5	Unreduced Data Base
747	L	86.7	86.7	88.8	85.5	85.6	86.8	84.4	
DC-8TF	L	92.0	93.2	97.7	88.8	89.5	91.6	85.5	
DC-8TJ	L	86.3	86.1	87.1	85.2	85.3	85.7	85.6	
Concorde	L	84.7	84.8	86.1	83.9	84.2	85.1	86.8	
CV-640	L	90.9	92.6	97.6	88.4	88.9	92.7	84.5	
737	T	89.9	90.3	88.7	88.9	89.0	88.3	88.1	
747	T	88.4	88.9	88.1	86.9	87.1	86.5	86.1	
DC-8TF	T	89.4	89.7	88.6	86.4	88.2	87.5	87.5	
DC-8TJ	T	86.5	86.5	87.0	85.5	85.5	85.6	85.0	
Concorde	T	84.7	85.1	85.0	83.9	84.1	83.8	85.2	
CV-640	T	89.6	90.9	87.6	88.9	89.6	87.1	87.8	
Std. Dev.		2.38	2.88	4.34	1.89	2.07	2.74	1.27	
Range		7.3	8.4	12.7	5.0	5.5	9.0	3.7	

L = landing, T = takeoff, a = level-weighting, b = time-weighting code
The metrics are defined in Sections 2.1 and 2.2.2 (a = 1, b = 00 are baseline values)

Table 8 (Continued)

Flyover		EPNL (FAA) + ½ SC	EPNL (SAE) + ½ SC	EPNL (KP) + ½ SC	EPNL (KP) + ½ SC	IPNL + $\bar{C}(k)_{KP}$ + ½ SC	
		a = 1, b = 00	a = 1, b = 00	a = 1, b = 00	a = 2, b = 00	a = 2, b = 00	
747	L	85.8	85.7	86.9	89.5	89.5	Reduced Data Base
DC-8TJ	L	86.8	86.7	87.1	88.7	88.8	
737	T	89.1	89.5	88.5	89.1	88.9	
747	T	88.9	88.0	87.6	89.2	89.2	
DC-8TF	T	88.9	89.2	88.5	89.7	89.3	
DC-8TJ	T	86.2	86.3	86.4	87.9	88.2	
Concorde	T	86.1	86.5	86.4	87.8	87.7	
Std. Dev.		1.51	1.49	0.91	0.75	0.65	
Range		3.3	3.7	2.2	1.9	1.8	
737	L	87.8	88.0	89.8	95.0	95.1	Unreduced Data Base
747	L	85.8	85.7	86.9	89.5	89.5	
DC-8TF	L	89.0	89.5	91.6	97.8	97.9	
DC-8TJ	L	86.8	86.7	87.1	88.7	88.8	
Concorde	L	88.0	87.9	88.3	89.9	90.6	
CV-640	L	88.2	88.8	91.4	98.6	98.2	
737	T	89.1	89.5	88.5	89.1	88.9	
747	T	88.9	88.0	87.6	89.2	89.2	
DC-8TF	T	88.9	89.2	88.5	89.7	89.3	
DC-8TJ	T	86.2	86.3	86.4	87.9	88.2	
Concorde	T	86.1	86.5	86.4	87.8	87.7	
CV-640	T	89.2	89.8	88.4	88.8	88.8	
Std. Dev.		1.29	1.41	1.76	3.83	3.80	
Range		3.4	4.0	5.2	10.8	10.6	

L = landing, T = takeoff, a = level-weighting, b = time-weighting code
The metrics are defined in Sections 2.1 and 2.2.2. (a = 1, b = 00 are baseline values)

Table 9

Standard Deviations of All Metrics Evaluated Here from Powell's Data[27] (The asterisks and daggers can be interpreted with the assistance of Table 10 subject to the cautions described in Section 2.3.4. Each square refers to the reduced data base (top left), and the unreduced data base (bottom right.))

EPNL (FAA)

Reduced / Unreduced		a: 1	0.5	1.5	2	3
b	00	1.95 / 1.95	1.91 / 1.77	1.98 / 2.26	2.04 / 2.62	2.16 / 3.50†
	02	1.87 / 1.92	1.88 / 1.75	1.88 / 2.26	1.88 / 2.70	1.97 / 3.87††
	03	1.95 / 1.94	1.91 / 1.77	1.99 / 2.27	2.02 / 2.66	2.10 / 3.69††
	33	1.95 / 1.85	1.91 / 1.72	1.99 / 2.05	2.03 / 2.37	1.96 / 3.28†
	55	1.90 / 1.74	1.88 / 1.70	1.87 / 1.81	1.88 / 1.92	1.97 / 2.38
	52	1.90 / 1.75	1.90 / 1.71	1.88 / 1.81	1.89 / 1.92	1.92 / 2.38

EPNL (SAE)

Reduced / Unreduced		a: 1	0.5	1.5	2	3
b	00	2.02 / 2.12	1.94 / 1.83	2.09 / 2.49	2.18 / 2.90	2.40 / 3.76††
	02	1.90 / 1.98	1.88 / 1.76	1.87 / 2.30	1.74 / 2.72	1.71 / 3.78††
	03	2.04 / 2.09	1.94 / 1.81	2.10 / 2.43	2.14 / 2.84	2.35 / 3.73††
	33	2.06 / 1.97	1.95 / 1.87	2.12 / 2.23	2.11 / 2.56	2.41 / 3.51†
	55	1.87 / 1.81	1.88 / 1.73	1.88 / 1.95	1.85 / 2.13	1.80 / 2.68
	52	1.89 / 1.80	1.88 / 1.74	1.90 / 1.92	1.89 / 2.05	1.81 / 2.52

EPNL (KP)

Reduced / Unreduced		a: 1	0.5	1.5	2	3
b	00	1.57 / 2.48	1.71 / 1.82	1.49 / 3.42†	1.45 / 4.42†††	1.61 / 6.65†††
	02	1.70 / 2.40	1.78 / 1.79	1.63 / 3.41†	1.61 / 2.72	1.78 / 7.55†††
	03	1.64 / 2.43	1.75 / 1.61	1.56 / 3.35†	1.52 / 4.43†††	1.68 / 6.73†††
	33	1.68 / 2.12	1.78 / 1.73	1.60 / 2.87	1.51 / 3.87††	1.31 / 6.16†††
	55	1.73 / 1.82	1.81 / 1.67	1.68 / 2.22	1.60 / 2.84	1.56 / 4.58†††
	52	1.81 / 1.79	1.85 / 1.67	1.80 / 2.12	1.75 / 2.70	1.75 / 4.46†††

Table 9 (Continued)

IPNL + $\overline{C}(k)_{FAA}$

Each cell: Reduced / Unreduced

b	a = 1	a = 0.5	a = 1.5	a = 2	a = 3
00	1.88 / 1.89	1.88 / 1.74	1.87 / 2.11	1.85 / 2.38	1.89 / $3.44^{\dagger}$
02	1.81 / 1.88	1.85 / 1.73	1.78 / 2.22	1.79 / 2.71	1.76 / $3.65^{\dagger\dagger}$
03	1.88 / 1.92	1.88 / 1.74	1.86 / 2.20	1.92 / 2.60	2.00 / $3.61^{\dagger}$
33	1.91 / 1.83	1.90 / 1.72	1.92 / 2.14	1.97 / 2.60	2.41 / $3.74^{\dagger\dagger}$
55	1.70 / 1.68	1.87 / 1.71	1.63 / 1.76	1.58 / 1.89	1.55 / 2.31
52	1.85 / 1.75	1.87 / 1.70	1.81 / 1.83	1.78 / 1.97	1.76 / 2.45

IPNL + $\overline{C}(k)_{SAE}$

Each cell: Reduced / Unreduced

b	a = 1	a = 0.5	a = 1.5	a = 2	a = 3
00	1.92 / 2.07	1.90 / 1.81	1.96 / 2.43	2.01 / 2.88	2.20 / $3.87^{\dagger\dagger}$
02	1.83 / 1.97	1.86 / 1.77	1.85 / 2.40	2.04 / 2.96	2.73 / $3.88^{\dagger\dagger}$
03	1.92 / 2.07	1.90 / 1.81	1.99 / 2.44	2.19 / 2.91	2.48 / $3.98^{\dagger\dagger}$
33	1.96 / 1.93	1.93 / 1.76	2.05 / 2.28	2.33 / 2.75	2.75 / $3.91^{\dagger\dagger}$
55	1.67 / 1.74	1.85 / 1.72	1.61 / 1.88	1.74 / 2.07	2.10 / 2.75
52	1.85 / 1.81	1.86 / 1.73	1.85 / 1.94	1.82 / 2.08	1.79 / 2.61

IPNL + $\overline{C}(k)_{KP}$

Each cell: Reduced / Unreduced

b	a = 1	a = 0.5	a = 1.5	a = 2	a = 3
00	1.50 / 2.38	1.68 / 1.78	1.39 / $3.28^{\dagger}$	1.37 / $4.34^{\dagger\dagger\dagger}$	1.61 / $6.37^{\dagger\dagger\dagger}$
02	1.64 / 2.49	1.76 / 1.76	1.57 / $3.44^{\dagger}$	1.59 / $4.37^{\dagger\dagger\dagger}$	2.07 / $7.17^{\dagger\dagger\dagger}$
03	1.59 / 2.40	1.73 / 1.76	1.50 / $3.47^{\dagger}$	1.52 / $4.46^{\dagger\dagger\dagger}$	1.74 / $6.53^{\dagger\dagger\dagger}$
33	1.68 / 2.16	1.79 / 1.71	1.57 / 3.21	1.52 / $4.21^{\dagger\dagger\dagger}$	1.80 / $6.60^{\dagger\dagger\dagger}$
55	1.64 / 1.86	1.75 / 1.65	1.54 / 2.25	1.48 / 2.74	2.00 / $4.40^{\dagger\dagger\dagger}$
52	1.77 / 1.81	1.82 / 1.67	1.73 / 2.14	1.67 / 2.76	1.63 / $6.75^{\dagger\dagger\dagger}$

Table 9 (Continued)

	IPNL no a no b	SEL no a no b	IPNL + 1/2 SC no a no b	EPNL (FAA) + 1/2 SC a = 1 b = 00	EPNL (SAE) + 1/2 SC a = 1 b = 00	EPNL (KP) + 1/2 SC a = 1 b = 00	EPNL (KP) + 1/2 SC a = 2 b = 00	IPNL + $\bar{C}(k)_{KP}$ + 1/2 SC a = 2 b = 00
Reduced	1.87	1.81	1.35	1.51	1.49	0.91	0.75**	0.65***
Unreduced	1.70	1.97	1.27	1.29	1.41	1.76	3.83††	3.80††

Table 10 describes the meaning of the asterisks and daggers in Table 9.

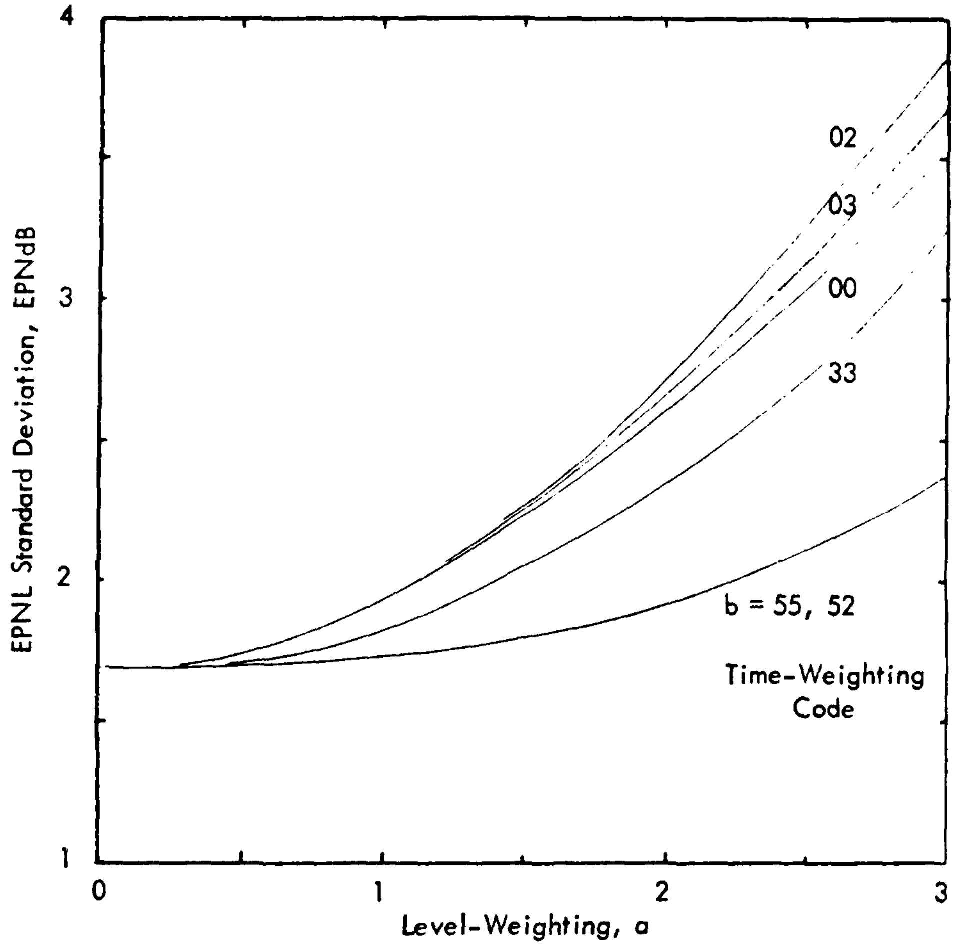


Figure 10. Standard Deviations of EPNL (FAR Part 36) for All Powell's Data (Ref. 27) as a Function of "Level-Weighting," a, and "Time-Weighting," b. a and b are defined in Section 2.2.2.

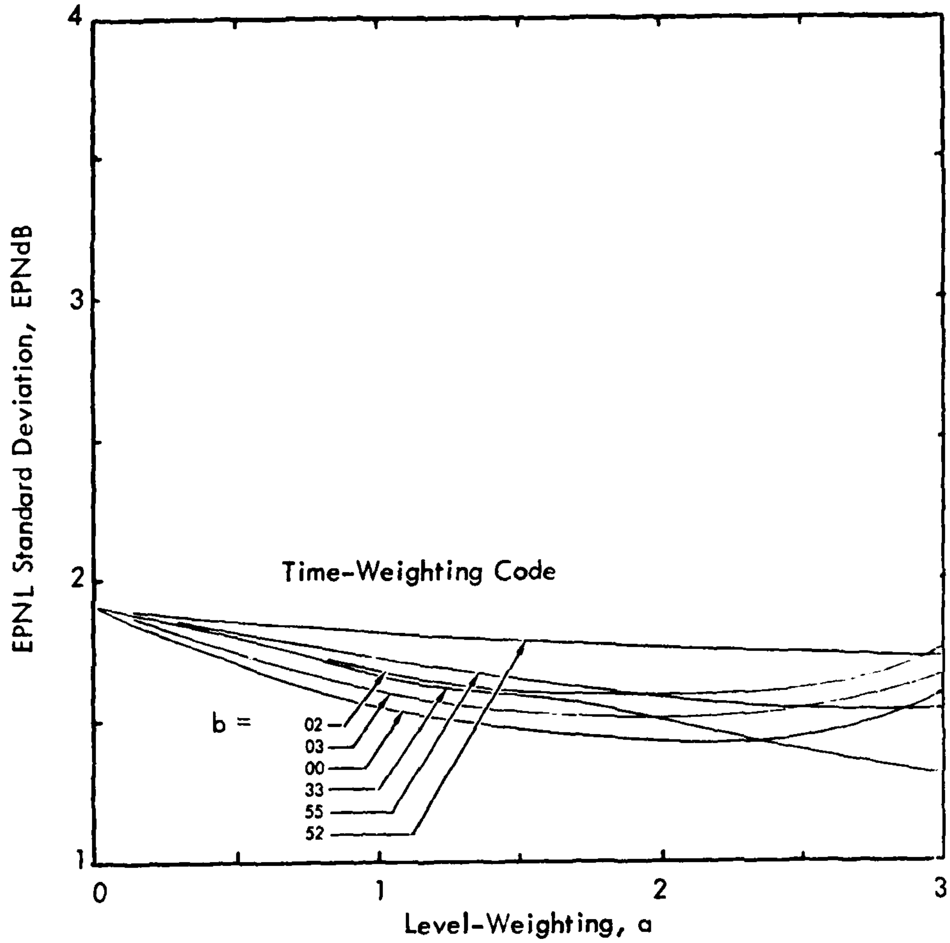


Figure 11. Standard Deviations of EPNL (Kryter-Pearsons) for a Reduced Set of Powell's Data (Ref. 27) as a Function of "Level-Weighting," a, and "Time-Weighting," b. a and b are defined in Section 2.2.2.

A statistical test to gauge the significance of the differences between the standard deviations is Snedecor's F test.[30] This was used (after applying Bessel's correction[30]) to see whether the metrics with standard deviations that were (1) lower, or (2) higher than those observed for EPNL (FAR Part 36) were truly (1) superior, or (2) inferior.

Using the F-test and the tables of F values in Reference 32, the quoted sample standard deviations have to take the values shown in Table 10 to be judged significantly different from the EPNL (FAR Part 36) standard deviations at the stated significance levels. Note, however, that there are approximately 400 standard deviations listed in Table 9, and it is therefore reasonable to expect that in any "fishing trip" in which <u>all</u> of them are compared with the values given in Table 10, some of the differences that are indicated to be significant are, in fact, due to chance. This is especially true at the $p = 0.050$ level, and probably also true at the $p = 0.025$ and $p = 0.010$ levels. The proper application of Table 10 is to permit a significance test to be conducted on any <u>few</u> metrics chosen to be of special interest <u>before</u> knowing their standard deviations.

An exception to this "proper application" rule is possible, however, when several metrics which share similar features to the one indicated to be significant are themselves indicated to be significant. In these instances only, a "fishing trip" through Table 9 may be justified. For example, a metric which is <u>indicated</u> to be superior to others when it has a level-weighting of $a = 1.5$ might be <u>accepted</u> as truly superior, if variants of the same metric with $a = 1$ and $a = 2$ are also indicated to be superior. The asterisks (*) and daggers (†) shown in Table 10 and applied to Table 9 must therefore be interpreted with the cautions indicated above.

Results for a selection of the more successful metrics are given in Table 11.

Table 10

Values Which a Standard Deviation Must Be Less Than (or Exceed) for Them to Be Judged Significantly Lower (or Higher) than the Standard Deviation Found for EPNL (FAR Part 36). (Use of this table is subject to the cautions described in Section 2.3.4.)

Significance Level, p		0.010	0.025	0.050
Reduced Data Base	Metric significantly worse than EPNL (FAR Part 36) if Std. Dev. exceeds . . .	5.661[†††]	4.692[††]	4.024[†]
	Metric significantly better than EPNL (FAR Part 36) if Std. Dev. is less than . . .	0.668[***]	0.806[**]	0.940[*]
Unreduced Data Base	Metric significantly worse than EPNL (FAR Part 36) if Std. Dev. exceeds . . .	4.121[†††]	3.636[††]	3.273[†]
	Metric significantly better than EPNL (FAR Part 36) if Std. Dev. is less than . . .	0.922[***]	1.045[**]	1.161[*]

Table 11

Standard Deviations of Selected Noise Metrics Analyzed Here from Powell's Data (Ref. 27) For each of the metrics in the top half of the table, the values of level-weighting, a, and time-weighting, b, are those which give the lowest standard deviation in Table 9. (The metrics in the lower half of the table were investigated for only the values of a and b shown.)

Metric	Reduced Data Base			Unreduced Data Base		
	σ	a	b	σ	a	b
EPNL (FAA)	1.87	{1, 1.5}	{02, 55}	1.70	0.5	55
EPNL (SAE)	1.71	3	02	1.73	0.5	55
EPNL (KP)	1.31	3	33	1.67	0.5	{55, 52}
IPNL + $\bar{C}(k)_{FAA}$	1.55	3	55	1.68	1	55
IPNL + $\bar{C}(k)_{SAE}$	1.61	1.5	55	1.72	0.5	55
IPNL + $\bar{C}(k)_{KP}$	1.37	2	00	1.65	0.5	55
IPNL	1.87	-	-	1.70	-	-
IPNL + ½SC	1.35	-	-	1.27	-	-
EPNL (FAA) + ½SC	1.51	1	00	1.29	1	00
EPNL (SAE) + ½SC	1.49	1	00	1.41	1	00
EPNL (KP) + ½SC	0.91	1	00	1.76	1	00
EPNL (KP) + ½SC	0.75	2	00	3.83	2	00
IPNL + $\bar{C}(k)_{KP}$ + ½SC	0.65	2	00	3.80	2	00

2.4 <u>Conclusions on Tone Correction Procedures</u>

The results in Section 2.3.4 allow the following conclusions to be drawn about the superiority of the various metrics when compared with EPNL (FAA).

For the reduced data base

- There was a consistent tendency for variants of EPNL(KP) to be superior. EPNL(KP) for a=1, b=00 had a standard deviation 19 percent lower than EPNL(FAA) for the same a and b. The best variant of EPNL(KP) - a=3, b=33 - had a standard deviation 30 percent lower than that of the best variant of EPNL(FAA).

- There was a tendency for many of the metrics to be superior with a level-weighting of 2 to 3 in the instances when the time-weightings were applied. The time-weighting most consistently successful was b=55.

- There was little or no benefit in time-averaging the tone corrections and adding them to IPNL.

- There is a substantial advantage to using Spectral Change, which is greatest when applied to the Kryter-Pearsons tone corrections, particularly EPNL (KP) + ½SC for a = 2, and IPNL + $\bar{C}(k)_{KP}$ + ½SC for a = 2.

- The minimum standard deviation of 0.65 occurred for this latter metric, and was 67 percent lower than the standard deviation for the currently used EPNL(FAA) metric.

For the unreduced data base

- EPNL(FAA), EPNL(SAE), and EPNL(KP) were not much different from one another. (EPNL(KP) may appear fractionally better for the values of a and b giving the lowest standard deviation in Table 11, but appears inferior to the other two when all are compared for a=1, b=00.)

- IPNL and SEL were no worse then EPNL (FAA).

- Level-weightings of about 0.5 may have improved the metrics slightly, while level-weightings greater than about a = 2 made them worse.

- A consistently successful time-weighting was b=55.
- There was little or no benefit in time-averaging the tone corrections and adding them to IPNL.
- There was an advantage to using Spectral Change in most instances when the metric to which it was applied had little or no level-weighting. (When the metric was level-weighted, SC made its standard deviation worse.)
- The minimum standard deviation of 1.27 occurred for IPNL + ½SC and was 35 percent lower than the standard deviation for the currently used EPNL(FAA) metric.

A metric which appeared consistently better for both data sets was IPNL + ½SC. The addition of ½SC to all three EPNL metrics was also beneficial for c=1 and b=00.

Few of the other metrics that seemed superior in one set of data were also superior in the other. While EPNL(KP) seemed superior in the reduced data for most values of a and b, its superiority is arguable, at best, in the unreduced data.

A general conclusion applicable to time-weighting is that it may well be useful to increase the tone correction early in the flyover.

No general conclusion appears possible with regard to level-weighting, except perhaps that level-weightings exceeding a value of about 3 are clearly excessive.

The results from the unreduced data base favor metrics like IPNL and IPNL + ½SC, but cannot be regarded as an indication that tone corrections are superfluous. The authors' data show a correlation between the tone corrections and Spectral Change, so that the reason for the superiority of IPNL + ½SC over IPNL is partly due to the influence of SC in accounting for the noisiness of the tones.

The results indicate that SC may be worth developing. Its present formulation operates only on 3 seconds of the flyover and has no frequency-weighting.

Despite the difficulty in drawing firm conclusions about the superiority of one tone correction over the other, there is conclusive evidence in the psychoacoustical literature that tone corrections can help describe the noisiness of tonal sounds, and ample evidence in our analysis that improved tone corrections may well be possible if developed from appropriate experimental plans (e.g., factorial designs), which take account of multicollinearity and the interactions between the noise variables. A clear illustration of the significance of multicollinearity is provided here by the disparate results yielded from a single experiment when multicollinearity is significant in one analysis and reduced in another.

There were indications from this research that improvements to tone correction procedures are possible within the context of one-third octave band, 0.5 second interval analysis, which is the existing industry standard. However, a literature survey and frequency analysis (Section 2.2.1) showed that tone corrections might be further improved if based on narrow band analysis.

Notwithstanding this potential for improvement, none of the tone correction procedures actually tested in this study proved to be consistently and significantly superior to the one in FAR Part 36.

3.0 PSEUDOTONE PERCEPTION

3.1 Experimental Procedure

3.1.1 Problem Definition

This section of the report addresses the question of whether a flyover noise having pseudotones,[6] for which the EPNL may consequently be slightly different from that for the same flyover noise without pseudotones, was indeed perceived to be more or less noisy. Behind this question is concern as to whether the FAR Part 36 measurement procedure should be changed, because the specified microphone height of 1.2 m is one at which pseudotones occur. Wyle has advised FAA[6] that a 10 m microphone height provides a way to reduce the pseudotone measurement anomaly. However, the necessity for making this change must also be judged in its perceived noisiness context.

The study was structured so that subjects judged the relative noisiness of pairs of flyover sounds - each pair being composed of (1) a flyover sound measured at 1.2 m microphone height, and (2) the sound of the same flyover measured at 10 m microphone height.

3.1.2 Noise Stimuli

Five pairs of aircraft flyover sounds were presented through Telephonics TDH-39 headphones, and were tape recordings on Nagra IV SJ recorders of the Los Angeles International Airport measurements described in Reference 6. The data in that reference were reviewed to pick out the flyovers for which a pseudotone was strongly evident at 1.2 m microphone height and not evident (or barely evident) at 10 m height. (These flyovers were generally those in which the ground surface near the microphone was acoustically hard, e.g., asphalt.)

The criterion for diagnosing a pseudotone is based on the frequency shift with time of a low frequency dip in the spectrum, as evident on the spectral time history plots in Reference 6 and as described in that report.

Flyovers exhibiting strong pseudotones at 1.2 m and minimal or no pseudotones at 10 m were selected so as to maximize any subjective pseudotone effect, on the reasoning that it should be researched first, on a pilot experiment basis, where it was most obvious. A follow-up experiment might then be required if a pseudotone effect were proved, to investigate the effect for a normal flyover "population." It turned out that only the pilot experiment was required.

The order of the flyover sound pairs (pseudotones vs nonpseudotone) were alternated so that order effects in the subjects' judgments could be controlled. The noise stimuli - and the associated airplane types and noise levels - are described in Table 12.

Table 12

Flyover Pairs in Order of Presentation With Their EPNL Values and Subjective Scores

Noise Pair No.	Aircraft and Flyover Procedure		Pseudotone	EPNL, EPNdB	Subjective Score	Subjective Score Excess Over Mean
1	747	Takeoff	Yes No	111.0 111.3	-6	-0.5
2	727	Landing	No Yes	104.2 104.5	-8.5	-3
3	727	Landing	Yes No	104.5 106.7	-3	+2.5
4	707	Landing	No Yes	112.1 112.9	-3	+2.5
5	747	Takeoff	Yes No	114.7 113.0	-7	-1.5

3.1.3 Subjects

Subjects were 10 Wyle employees, eight male and two female, aged 22 to 53 years. All subjects claimed no hearing impairment, but in this pilot experiment, no audiometric screening was undertaken. In any case, subjects acted as their own controls, thus obviating some of the need for ensuring minimum hearing standards. The subjects were volunteers and had a general idea of the object of the experiment (see also the Instructions in Section 3.1.4), but were uninformed about the experimental structure.

3.1.4 Instructions

Subjects received the instructions shown in Table 13, based for uniformity on somewhat similar instructions in References 27 and 31. They were also given the Questionnaire in Table 14. Note that the same tape was presented twice, since the 5-minute time for one presentation was short enough to permit a repeat presentation (thereby reducing experimental variance). The subjects were not told that the second tape was a repeat of the first, and the airplane sounds were sufficiently "characterless" that no subject appeared to detect the repetition. The first questionnaire was collected after the first tape was played, to prevent any tendency to copy it.

Subjects were assured that their names and individual results would be kept confidential.

3.2 Results

A scoring system was established as follows.

If the first sound in a pair was judged the more annoying, the noise pair was scored +1. If the second sound was judged the more annoying, the pair was scored -1. If the two sounds were judged equally annoying, the pair was scored zero. The overall score for each pair was then established by (1) averaging each subject's first tape and second tape scores in order to determine that subject's overall opinion of that pair, and (2) adding the scores so obtained from all the subjects.

In this way, a score of +10 would indicate total agreement that the first sound was considered the more annoying; -10 would indicate total agreement that the second sound was considered the more annoying; and zero would indicate that there were as many opinions that the first sound was more annoying as that the second sound was more annoying.

The above scoring system results are shown in the "subjective score" column of Table 12. There was a strong tendency for the second sound in each pair to be judged the more annoying, perhaps because it occurred closest in time to the subjects' judgments. This was removed in the column labeled "subjective score excess over mean" by subtracting the mean (-5.5) of the subjective scores from the scores themselves. The right-hand column is therefore intended to represent the subjective scores normalized for order effects. Table 15 presents the results shown in Table 12 in a form suitable for drawing possible conclusions.

Table 13

Instructions to Subjects in Pseudotone Pilot Experiment

INSTRUCTIONS

THE EXPERIMENT IN WHICH YOU ARE PARTICIPATING IS TO HELP US UNDERSTAND THE CHARACTERISTICS OF AIRCRAFT SOUNDS WHICH CAN CAUSE ANNOYANCE IN AIRPORT COMMUNITIES. WE WOULD LIKE YOU TO JUDGE HOW ANNOYING SOME OF THESE AIRCRAFT SOUNDS ARE. BY ANNOYING WE MEAN - UNWANTED, OBJECTIONABLE, DISTURBING, OR UNPLEASANT.

AIRCRAFT SOUNDS WILL BE PRESENTED IN PAIRS. WE WANT YOU TO TELL US WHICH SOUND IN EACH PAIR IS THE MORE ANNOYING. A SCORING SHEET WILL BE PROVIDED. CHECK THE APPROPRIATE BOX TO INDICATE WHICH SOUND IS THE MORE UNWANTED, OBJECTIONABLE, DISTURBING, OR UNPLEASANT.

IF YOU JUDGE THE TWO SOUNDS TO BE EQUALLY ANNOYING, CHECK BOTH BOXES.

THERE ARE FIVE PAIRS OF SOUNDS IN THE FIRST TAPE. AFTER THIS TAPE IS COMPLETE, PLEASE HAND IN YOUR SCORE SHEET. ANOTHER TAPE WILL THEN BE PRESENTED TO YOU, AND A NEW SCORE SHEET WILL BE USED.

IF YOU HAVE ANY QUESTIONS, PLEASE ASK THEM NOW. AFTER THE TEST BEGINS, IT SHOULD PROCEED WITHOUT INTERRUPTION. THANK YOU FOR YOUR COOPERATION.

Table 14

Questionnaire in Pseudotone Pilot Experiment

Subject Number________ Age______ Sex________ Date____________

Tape Number ________

Pair 1	First sound more annoying	☐
	Second sound more annoying	☐
Pair 2	First sound more annoying	☐
	Second sound more annoying	☐
Pair 3	First sound more annoying	☐
	Second sound more annoying	☐
Pair 4	First sound more annoying	☐
	Second sound more annoying	☐
Pair 5	First sound more annoying	☐
	Second sound more annoying	☐

Table 15

Correspondence Between EPNL and Noisiness Differences in Each Noise Pair

1	2	3	4	5	6
Noise Pair No.	Pseudotone Judged Noisier?	Order Effects Removed, Pseudotone Judged Noisier?	EPNL Higher for Pseudotone?	Correspondence-Columns 2 & 4?	Correspondence-Columns 3 & 4?
1	No	No	No	Yes	Yes
2	Yes	Yes	Yes	Yes	Yes
3	No	Yes	No	Yes	No
4	Yes	No	Yes	Yes	No
5	No	No	Yes	No	No

The most relevant columns in Table 15 are those in which the order effects are removed: columns 3 and 6. Column 3 indicates that the pseudotone flyover was judged to be the noisier about as often as it was not (2 out of 5 occasions). Column 6 indicates that the EPNL difference between the sounds in each pair correctly predicted the subjective difference on 2 out of 5 occasions.

3.3 Conclusions on Pseudotone Effects

The results in Section 3.2 showed that (1) the subjects did not consistently agree that the pseudotone flyovers were either less or more annoying than the others, and (2) the EPNL metric did not display any consistent tendency to disagree with the subjects' opinions.

Since the flyover pairs were selected on the basis of a strong pseudotone difference, the results can probably be taken to indicate that in a normal flyover "population" any pseudotone effect is, at the very least, as small as the other unaccounted for effects which govern subjective judgments. In the laboratory setting, such other unaccounted for effects are minimized and the above conclusion is therefore even more valid for real life.

The results were, however, derived from a pilot experiment using only a few subjects. The strength of the experiment in its use of noise pairs from the same flyover reproduced at the same level as measured in the field, was also a weakness:

an alternative experimental plan could have held the EPNL value constant in each noise pair and compared the subjective judgments (using a continuous scale) with a "geometric" descriptor of the pseudotone taken from the graphical spectral time histories in Reference 6. These changes might have increased the sensitivity of the experiment to detect pseudotone effects on subjective judgments. However, the experiment as performed had a high face-validity, i.e., it strongly and obviously tested the likelihood that a change in microphone height would change the noisiness of a flyover or improve the ability of the EPNL procedure to predict noisiness.

The present, preliminary study indicates that the pseudotone effects at the 1.2 m microphone height are likely to have little or no subjective consequence. Thus, an aircraft noise certification procedure may either be based on a measurement technique which effectively eliminates them from the data by the use of a 10 m microphone (see Volume I[5] for a more detailed discussion on this issue), or the FAR Part 36 procedure may remain unchanged.

REFERENCES

1. Federal Aviation Administration, "Noise Standards: Aircraft Type and Airworthiness Certification," Federal Aviation Regulations, Volume III, Part 36, Washington, D. C., 1969.

2. Kryter, K. D., "Loudness of Annoyance-Value of Bands of Noise," in Oralism and Autism: Transactions of the 30th Annual Meeting of the National Forum on Deafness and Speech Pathology, St. Louis, MO, 1948, 26-28.

3. Little, J. W., "Human Response to Jet Engine Noises," Noise Control 7(3), 11-13, 1961.

4. Society of Automotive Engineers, "Definitions and Procedures for Computing the Effective Perceived Noise Level for Flyover Aircraft Noise," Aerospace Recommended Practice (ARP) 1071, revised edition, Warrendale, PA, 1973.

5. Kryter, K. D. and Pearsons, K. S., "Judged Noisiness of a Band of Random Noise Containing an Audible Pure Tone," J. Acoust. Soc. Am. 38, 106-112, 1965.

6. Rackl, R., "Correction Procedures for Aircraft Noise Data - Volume I: Pseudotones," Report No. FAA-EE-80-1, Vol. I, prepared for the U. S. Department of Transportation, Federal Aviation Administration, 1979.

7. International Civil Aviation Organization, "Aircraft Noise, Annex 16 to the Convention on International Civil Aviation," International Standards and Recommended Practices, third edition, Montreal, Quebec, 1978.

8. American National Standards Institute, "Definitions and Procedures for Computing the Effective Perceived Noise Level for Flyover Aircraft Noise," ANSI S6.4-1973, New York, NY, 1973.

9. Office of Noise Abatement, U. S. Department of Transportation, "Comparison of Procedures for Calculation of Effective Perceived Noise Level for Typical Aircraft Flyovers," Information Brief, Washington, D. C., 1975.

10. Zwicker, E., Flottorp, G. and Stevens, S. S., "Critical Band Width in Loudness Summation," J. Acoust. Soc. Am. 29, 548-557, 1957.

11. Sokolov, E. N., "The Measurement of Sensitivity and Reactivity of a Conditioned Reflex or in Connection with the Interaction of Orienting and Defense Reactions," American Institute of Biological Sciences, New York, NY, 1965.

12. Thurlow, W. R. and Bernstein, S., "Simultaneous Two-Tone Pitch Discrimination," J. Acoust. Soc. Am. 29, 515-519, 1957.

13. Pearsons, K. S., Bishop, D. E. and Horonjeff, R. D., "Judged Noisiness of Modulated and Multiple Tones in Broadband Noise," J. Acoust. Soc. Am. 45, 742-750, 1969.

14. Wegel, R. L. and Lane, C. E., "The Auditory Masking of One Pure Tone by Another and its Probable Relation to the Dynamics of the Inner Ear," Phys. Rev. 21, 266-285, 1924.

15. Ehmer, R. H., "Masking Patterns of Tones," J. Acoust. Soc. Am. 31, 1115-1120, 1959.

16. Small, A. M., and Thurlow, W. R., "Loudness Relations in Two-Component Tones," J. Acoust. Soc. Am. 26, 381-388, 1954.

17. Webster, J. C., Miller, P. H., Thompson, P. O. and Davenport, E. W., "The Masking and Pitch Shifts of Pure Tones Near Abrupt Changes in a Thermal Noise Spectrum," J. Acoust. Soc. Am. 24, 147-152, 1952.

18. Schubert, E. D., "The Effect of a Thermal Masking Noise on the Pitch of a Pure Tone," J. Acoust. Soc. Am. 22, 497-499, 1950.

19. Bilger, R. C. and Hirsch, I. J., "Masking of Tones by Bands of Noise," J. Acoust. Soc. Am. 28, 623-630, 1956.

20. Garner, W. R., "The Effect of Frequency Spectrum on Temporal Integration of Energy in the Ear," J. Acoust. Soc. Am. 19, 808-815, 1947.

21. Federal Aviation Administration, U. S. Department of Transportation, "Results of Aircraft Noise Flight Demonstration, July 17, 1978, Dulles International Airport," Washington, D. C., 1978.

22. Wells, R. J., "A New Method for Computing the Annoyance of Steady State Noise Versus Perceived Noise Level and Other Subjective Measures," Paper presented at the 77th Meeting of the Acoustical Society of America, April 1969. (Also published by General Electric R&D Center, Schenectady, New York.)

23. Shepherd, K. P., "The Effect of the Duration of Jet Aircraft Flyover Sounds on Judged Annoyance," NASA CR 159132, National Aeronautics and Space Administration, Langley Research Center, Hampton, VA, 1979.

24. Kryter, K. D., "A Note on the Quantity Effective Perceived Noisiness and Units of Perceived Noise Level," J. Sound Vib. 25, 383-393, 1972.

25. Ollerhead, J. B., "An Evaluation of Methods for Scaling Aircraft Noise Perception," NASA CR 1883, National Aeronautics and Space Administration, Langley Research Center, Hampton, VA, 1971.

26. Scharf, B. and Hellman, R., "Comparison of Various Methods for Predicting the Loudness and Acceptability of Noise. Part II: Effects of Spectral Pattern and Tonal Components," U. S. Environmental Protection Agency Report 550/9-79-102, November 1979.

27. Powell, C. A., "Judgments of Relative Noisiness of a Supersonic Transport and Several Commercial-Service Aircraft," NASA TN D-8434, National Aeronautics and Space Administration, Langley Research Center, Hampton, VA, 1977.

28. Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K. and Bent, D. H., Statistical Package for the Social Sciences, McGraw-Hill, New York, second edition, 1974.

29. May, D. N., "Highway Noise Barrier Perceived Benefit," Report 79-AC-05, Research & Development Division, Ministry of Transportation & Communications, Downsview, Ontario, 1979.

30. Moroney, M. J., Facts from Figures, Penguin Books, Harmondsworth, Middlesex, England, 1962.

31. McCurdy, D. A. and Powell, C. A., "Effects of Duration and Other Noise Characteristics on the Annoyance Caused by Aircraft-Flyover Noise," NASA TP 1386, National Aeronautics and Space Administration, Langley Research Center, Hampton, VA, 1979.

32. Hays, W. L., Statistics for the Social Sciences, Holt, Rinehart and Winston, New York, 1973.