inter•noise 96
25th Anniversary Congress – Liverpool

INCE: 63.2

CORRELATIONS BETWEEN SUBJECTIVE AND OBJECTIVE EVALUATIONS OF REFRIGERATOR NOISE

P May, P Davies & J S Bolton

Ray W Herrick Laboratories, School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907-1077, USA

1. INTRODUCTION

There is an increasing interest in developing models to predict human annoyance to appliance noise. One method of developing predictive models is to correlate human subjective responses with objective sound quality metrics. In this paper, the results of two such investigations are presented. A goal in both investigations was to determine whether a linear combination of well known sound quality metrics could be used to predict annoyance resulting from exposure to refrigerator noise. Another goal was to determine if either dB(A) or Zwicker's Unbiased Annoyance (UBA), used individually, were highly correlated with subjective evaluations of refrigerator noises. This ongoing research is supported by LG Electronics, Inc. of Korea.

In both studies, the objective measures (i.e., sound quality metrics) were determined by using a Brüel & Kjær Type 2144 Analyzer (with the Zwicker Loudness Option Type 7638) and the Brüel & Kjær ZWLOUD software. Loudness spectra were calculated every 2 milliseconds. The sound quality metrics included in the studies were: maximum loudness, mean loudness, rms loudness, standard deviation loudness, maximum sharpness, maximum peak amplitude, impulsiveness, N4, N10, N20, maximum fluctuation, maximum roughness and Zwicker's Unbiased Annoyance (UBA) [1]. Peak amplitude is a measure of the height of the largest peaks in the loudness spectra.

2. PREFERENCE TEST 1

The objectives of Test 1 were: 1) to measure the preferences of the subjects to 17 recordings of refrigerator noise in order to rank the signals in terms of annoyance, 2) to determine which objective measures were most strongly correlated with subject preferences, 3) to find if there are preference differences concerning refrigerator sounds between Korean and American subjects, and 4) to find if linear combinations of metrics could be used to predict subject preferences of the refrigerator sounds.

The stimuli were composed of 17, five second refrigerator recordings made from DAT recordings provided by the refrigerator manufacturer. Three refrigerators of varying capacities were considered. The recordings were made with the refrigerators in various operating configurations (e.g., steady state running condition shortly after a warm start, and steady state shortly after a cold start), as well as with different component combinations (e.g., such as fan off/compressor on, and compressor off/fan on). The signals in Test 1 were presented to the subjects at a higher sound pressure level than that at which they were recorded. Therefore, the signals as presented ranged from an averaged dB(A) value of 49 to 63.

The subjects consisted of 24 subjects ranging from 19 to 43 years in age. Twelve of the subjects were Koreans, temporarily residing in the United States, and 12 were North American born American residents. All participants were given a hearing threshold level test to establish that their hearing level fell within normal limits. The subjects were seated in an anechoic chamber and took a computerized test in which the refrigerator recordings were played to the subjects through headphones. The subjects heard the same signal through each headphone earpiece. For this test, a paired comparisons format was used in which each subject heard a different randomized sequence of pairs of signals. All possible pairings of the 17 signals were used. Each subject was asked to answer the question, "Which sound is more annoying: 1 or 2?"

The preference results were analyzed using the averaged responses of three groups: the Korean subjects, American subjects, and both groups combined. The preferences were compared using two methods. The first method was based on the fraction of times that each signal was chosen as most annoying out of the total number of times that the signal was presented. The second method was based on the Bradley-Terry-Luce (BTL) values, which show how much more annoying the subjects found one signal than another. The Koreans and Americans essentially chose the signals in the same rank order from most to least annoying, however, the BTL values revealed that the Koreans were more sensitive and reacted more strongly to the differences in the signals than did the Americans.

In developing a model for perceived annoyance for this set of refrigerator signals, regressions were performed using the subjects' preferences and the 13 sound quality metrics. All combinations of 1, 2 and 3 metrics were tested in the regressions. The R^2 statistic, which is a measure of the amount of variation in the dependent variable (perceived annoyance) that is explained by the sound quality metrics [3], was used to evaluate each of the regression combinations. Also, regressions were performed using subject preferences and the averaged dB(A) values for the signals. The R^2 values based on dB(A) levels were under 0.2, meaning that dB(A) alone could not be used to predict subjects' annoyance for this set of signals. Zwicker's UBA, when used alone, was found to give an R^2 value of only 0.42. As expected, the highest R^2 values were found for the models based on 3 metrics, with R^2 values being as high as 0.95. The 5 combinations of sound quality metrics yielding the highest R^2 values based on BTL values for all 24 subjects are shown in Table 1. The p values represent the significance level for each of the models. Shown in Figure 1 is a plot of the predicted versus actual BTL values using the best model from Table 1.

Table 1: Results of the multiple regressions for all 24 subjects using 3 variables (metrics) for Test 1. Shown are the combinations of 3 metrics yielding the highest R^2 values using the paired comparisons BTL values as the preference measures.

	R^2	Metric 1	Metric 2	Metric 3	p
1	0.950	Std. Dev. Loudness	Max. Sharpness	N4	0.000
2	0.949	Mean Loudness	Std. Dev. Loudness	Max. Sharpness	0.000
3	0.949	RMS Loudness	Std. Dev. Loudness	Max. Sharpness	0.000
4	0.948	Max. Loudness	Std. Dev. Loudness	Max. Sharpness	0.000
5	0.947	Std. Dev. Loudness	Max. Sharpness	N10	0.000

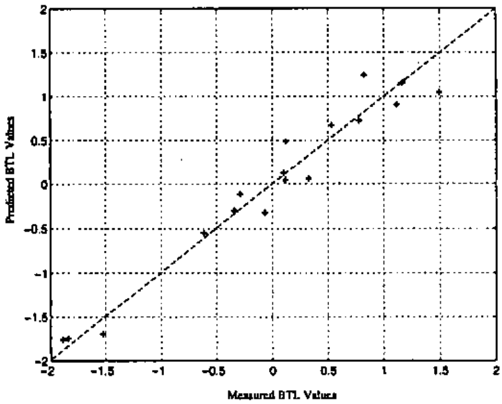


Figure 1: Plot of measured BTL values versus predicted BTL values for all 24 subjects in Test 1. The predicted values are from the preference model with 3 variables, yielding an R^2 of 0.950.

3. PREFERENCE TEST 2

In Test 2, 12 refrigerator recordings consisting of 8 steady state signals and 4 signals containing start up or shut down transients were used. None of the signals from Test 1 were used in this test. One objective was to study the effects of signal length on the perception of annoyance to refrigerator noise. Therefore, 5 second and 10 second versions of each of the 12 signals were used. Another goal was to compare the results obtained using two subjective testing designs, namely paired comparisons and a parametric testing method. The signals were presented to the subjects at the actual recorded levels which ranged from 17 dB(A) to 35 dB(A).

For this study, the subjects consisted of 24 North American born American residents between 21 and 44 years old; they were required to pass a hearing threshold level test. The subjects, tested individually, first took the paired comparisons test using the 12, five second refrigerator recordings. Immediately following the paired comparisons test, each subject took two versions of the parametric test design. One version of the parametric design used the 5 second signals while the other version used the 10 second signals. The testing order was alternated, so that some subjects took the test in the order: paired comparisons, parametric (using 5 sec. signals), parametric (using 10 sec.

signals), while other subjects took the test with the order of the two parametric tests reversed.

A Likert-style rating scale was used in the parametric tests; i.e., the subject rated each signal on a scale of 1 to 7 with a 1 being the least annoying and a 7 being the most annoying. Before taking each of the parametric tests, all 12 signals were played so that each subject could develop a range for the least and most annoying noises in the test.

As in Test 1, regression analysis was performed using the subject responses and the sound quality metrics. The results again indicated that averaged dB(A) sound pressure level was not significantly correlated with perceived annoyance, with R^2 values less than 0.1. When using UBA values, the R^2 values were less than 0.23. Multiple regressions using combinations of 1, 2 and 3 of the metrics yielded a maximum R^2 value of 0.85, which was for the parametric test using the 5 second signals as shown in Table 2. Preference models obtained using the three test designs (one paired comparisons test and two parametric tests) were very similar, indicating that the shorter signals and the parametric design could be used to develop models of human annoyance to refrigerator noise.

Table 2: Results of the multiple regressions for all 24 subjects using 3 variables (metrics) for Test 2. Shown are the combinations of 3 metrics yielding the highest R^2 values using the parametric tests (with 5 sec. signals) values as the preference measures.

	R^2	Metric 1	Metric 2	Metric 3	p
1	0.857	Max. Peak Amplitude	Max. Sharpness	Mean Loudness	0.001
2	0.828	Max. Peak Amplitude	Max. Sharpness	RMS Loudness	0.002
3	0.731	N20	Max. Peak Amplitude	Max. Sharpness	0.011
4	0.723	N4	Max. Peak Amplitude	Max. Sharpness	0.013
5	0.715	N10	Max. Peak Amplitude	Max. Sharpness	0.014

4. CONCLUSIONS

It was found that linear models using three sound quality metrics could be devised to predict the average annoyance of human subjects to refrigerator noise with R^2 values in excess of 0.80. For the sets of signals used in both studies, dB(A) was not significantly correlated with subjects' responses in predicting annoyance. Additionally, regressions using subject responses and Zwicker's UBA was found to yield R^2 values of only 0.2 to 0.45.

These results also indicate that parametric tests made using a Likert-type rating scale can be used to produce similar models to those obtained using the more time consuming paired comparisons test design. The more time efficient parametric tests will enable the use of a greater number of signals, allowing fuller use of the metric parameter space. This, in turn, will facilitate the development of a more generally applicable and reliable appliance noise annoyance model.

5. REFERENCES

[1] E. Zwicker and H. Fastl, Psychoacoustics: Facts and Models (Springer Verlag, 1990)
[2] H. A. David, The Method of Paired Comparisons, 2nd edition (Charles Griffin & Co. LTD, 1988)
[3] J. O. Rawlings, Applied Regression Analysis: A Research Tool (Wadsworth & Brooks, 1988)

inter•noise 96
25th Anniversary Congress – Liverpool

INCE: 63.4

SOME POLICY AND REGULATORY IMPLICATIONS OF RECENT FINDINGS OF FIELD STUDIES ON NOISE-INDUCED SLEEP DISTURBANCE

S Fidell

BBN Acoustic Technologies, 21120 Vanowen Street, Canoga Park, CA 91303, USA

Who has not been awakened at one time or another by nighttime noise exposure? Who has not been annoyed when so awakened? Who could be churlish enough to imagine that his neighbor had not also been similarly afflicted? Why then has it proven so difficult to find a persuasive basis for regulatory policies for effectively limiting or otherwise discouraging nighttime noise exposure?

As is often the case, the difficulties are attributable in part to economic and technical factors, and in part to the complexity of the phenomenon of interest (sleep) itself. Although no society can afford to guarantee its members that their sleep will never be disturbed in any degree at all by noises of their fellow citizens' nighttime activities, opinions vary widely about both tolerable degrees of sleep disturbance and their predictably from physical measures of noise exposure alone.

Regulatory policies intended to provide noise-exposed populations with some degree of protection or relief from adverse effects of noise exposure are most readily accepted when they are accompanied by explicit statements of their elementary goals, clear definitions of the populations to be protected or relieved of noise exposure impacts, and a reasonably complete understanding of the nature of relief or protection to be conferred. To be credible, useful, and enforceable, regulatory decisions must also be supported by systematic analyses of empirical information, and must be expressed in appropriate and consistent units.

No noise regulations in the United States are systematically based on formal analyses of information about noise-induced sleep disturbance, in large part because present understanding of the ability of noise to disturb sleep satisfies so few of the above conditions. Consider, for example, the difficulty of defining the proper goal for limiting nighttime noise exposure. Is it reasonable (or affordable by society) to expect a regulation to prevent any noise-induced disturbance, of any nature, however slight, at any time, of anybody's sleep? If the goal is **not** to protect everyone at all times from any disturbance whatsoever, then should policy decisions focus on minimizing awakenings, motility, or shifts from one sleep stage to another, and what population proportion should receive how much benefit under what circumstances from pre-existing and novel nighttime noise sources?

Even assuming for the sake of argument that clear goals and definitions of population segments of interest could be articulated and agreed upon, understanding of noise-induced sleep disturbance is not sufficiently advanced to support the intensive analysis and close scrutiny needed to justify regulatory decisions. Some basic issues which remain unresolved in this area include the physiological purposes served by various forms of sleep, the consequences of chronic, intermittent, or minor disturbances of various sleep states, the degree to which residential populations self-select for tolerance to nighttime noise exposure, and whether sleep disturbance is more directly controlled by acoustic or nonacoustic (e.g., cognitive) factors.

A dosage-response relationship is perhaps the firmest basis for regulatory decisions. The meta-analysis of Pearsons, Barber, Tabachnick and Fidell (1995) of the noise-induced sleep disturbance literature has documented irreconcilably large differences in the findings of laboratory and field studies. In the judgment of Pearsons et al., the results of the two types of studies can not be usefully represented in a single dosage-response relationship. Since publication of this meta-analysis, three major field studies of noise-induced sleep disturbance have been conducted: those of Ollerhead, Jones, Cadoux, Woodley, Atkinson, Horne, Pankhurst, Reyner, Hume, Van, Watson, Diamond, Egger, Holmes, & McKean (1992); of Fidell, Pearsons, Tabachnick, Howe, Silvati, and Barber (1995); and of Fidell, Howe, Tabachnick, Pearsons, and Sneddon (1995). The findings of these three studies are in good agreement with the regression line that Pearsons et al. (1995) developed for the data of field studies.

The dosage-response relationship for noise-induced sleep disturbance developed by FICON (1992), intended to serve as a compromise between laboratory and field observations, does not closely represent either set of findings. Furthermore, the independent variable of this form of dosage-response relationship - indoor sound exposure level - is a single event metric that is difficult to estimate and awkward to manipulate within the conventional context of regulatory reliance on cumulative, long term integrated noise metrics. Worse yet, the dependent variable of the FICON relationship is a poorly-defined, all or nothing quantity derived from observations of the short term reactions of individuals to discrete noise intrusions. "Disturbance" in this sense is only loosely interpretable on a population basis or in real-world noise exposure circumstances.

Since the basic datum on which the relationship is built is the likelihood of disturbance (variously defined in different data sets) of self-selected individuals' sleep by discrete noise events, it is unclear how to use this information without making a great many unsupportable assumptions about the distribution of sensitivity to sleep disturbance of larger populations, and about the time course of adaptation to repeated noise intrusions. For example, how can the FICON curve be interpreted to yield estimates of the numbers of people whose sleep will be disturbed in varying degrees by multiple occurrences of novel and familiar noise events of different SEL values?

Although the recent field observations of more than 10,000 subject-nights of noise-induced sleep disturbance demonstrate that the sensitivity of long term residents of airport neighborhoods to the absolute sound exposure levels of individual noise intrusions is considerably lower than previously believed, it does not necessarily follow that nighttime "penalties" incorporated in daily average noise metrics should be modified. Since no meaningful evidentiary basis in real-world sleep disturbance data was available when nighttime noise penalties were first adopted (Fidell and Schultz, 1980), new evidence alone can hardly be considered a sufficient basis for revising them.

Furthermore, ample grounds other than sleep disturbance may be found for imposing a nighttime penalty on noise exposure. In many residential neighborhoods, for example, the nighttime population exceeds the daytime population by a factor of 3 or more. This difference in relative sizes of the daytime and nighttime populations by itself suggests a rationale for a 5 dB penalty. Additionally, background noise levels in high population density (urban) areas generally drop by 3 dB or more during nighttime hours, increasing the apparent duration of many noise intrusions and making them correspondingly more audible.

Even assuming that clear goals, definitions, and purposes for nighttime noise regulation can be formulated, and that improved dosage-response relationships can be developed to support them, a set of difficult implementation issues remains. Nighttime noise penalties incorporated within daily average noise metrics are a clumsy means for discouraging noisy nighttime activities, which are often more heavily constrained by economic than by regulatory considerations in any event. It has long been known that incorporation of a time of day weighting factor into an energy summation does not yield benefits that scale directly with the size of the weighting factor. Galloway, Eldred and Simpson (1974), for example, demonstrated two decades ago that the difference on a nationwide basis between 24 hour L_{eq} and L_{dn} values is only 3.4 dB.

In short, newly available information from large scale field studies has made it clear that substantive progress in developing rational policies for dealing with noise-induced sleep disturbance must await better defined goals, more rigorous analyses, and a clearer sense of purpose than common prior practice in this area.

References

Federal Interagency Committee on Noise (FICON), (1992) Final Report: Airport Noise Assessment Methodologies and Metrics, Washington, D.C.

Fidell, S., Howe, R., Tabachnick, B., Pearsons, K., and Sneddon, M. (1995) Noise-Induced Sleep Disturbance in Residences near Two Civil Airports", NASA Contractor Report 198252, NASA Langley Research Center, Hampton, VA.

Fidell, S., Pearsons, K., Tabachnick, B., Howe, R., Silvati, L., and Barber, D. (1995) "Field study of noise-induced sleep disturbance", J. Acoust. Soc. Am, 98(2), Pt. 1, pp. 1025-1033.

Fidell, S., and Schultz, T.J. (1980) "A Critical Review of Time-of-Day Weighting Factors for Cumulative Measures of Community Noise Exposure" BBN Report No. 4216, Cambridge, MA.

Galloway, W. Eldred, K.., and Simpson, M., (1974) "Population Distribution of the United States as a Function of Outdoor Noise Level", Environmental Protection Agency, EPA-550/9-74-009-A, Washington, D.C.

Ollerhead, J. B., C. J. Jones, R. E. Cadoux, A. Woodley, B. J. Atkinson, J. A. Horne, F. Pankhurst, L. Reyner, K. I. Hume, F. Van, A. Watson, I. D. Diamond, P. Egger, D. Holmes, & J. McKean (1992). *Report of a Field Study of Aircraft Noise and Sleep Disturbance.* London: Department of Safety, Environment and Engineering.

Pearsons, K, Barber, D., Tabachnick, B., and Fidell, S. (1995), "Predicting noise-induced sleep disturbance", J. Acoust. Soc. Am., 97 (1), pp. 331-338.

inter•noise 96
25th Anniversary Congress – Liverpool

INCE: 63.4

NOISE AND SLEEP QUALITY IN TWO HOSPITALS IN THE CITY OF BELO HORIZONTE, BRAZIL

F Pimentel-Souza, J C Carvalho, A L Siqueira, P A S Alvares & C J S Rodrigues

Departamento de Fisiologia e Biofisica, ICB, POB 484, Universidade Federal de Minas Gerais, 31060-900-Belo Horizonte, Brazil

SUMMARY: Patients of two hospitals in Belo Horizonte (BH) were questionned on the eve of discharge to evaluate the sleep conditions related to noise. Patients had a worse perception of sleep at University Hospital, whose internal nocturnal Leq=53.7 dB(A) was greater than the one at Baleia Hospital, L_{eq}=45.5 dB(A). Noise was the only sleep-disturbing factor to be statistically significant. Cortisol levels were not statistically different.

The quality of sleep is frequently impaired by the excess of noise of the modern society and in overpopulated cities it needs to be better characterized, especially for patients in hospitals, because it is assumed to be important for patients recovery (1-7).

The external diurnal mean level, L_{eq}=69.5 dB(A), in City of Belo Horizonte, is considered to be excessive (8). To compare the patients perception of their own sleep, a questionnaire was applied during 1994 to male patients in two public hospitals: Baleia, a quiet one, and University, a noisy one. The 35 patients selected did not present any disease that would be expected to directly affect the sleep and were not users of somniferous or sedatives drugs. Only patients on the eve of discharge were evaluated, concerning the previous night sleep. The patients were between 17 and 60 years old, averaging 35.6. The reliability were evaluated by K index, the comparison of rates by the Fisher's exact and the Binomial tests (9).

The three modalities of insomnia are elevated in both hospitals. Nocturnal awakening affected 91.4% of the patients, with mean of 2.6 times, surprisingly uniformely ranged from 17 up 60 years old, against about 1.0 in three middle size cities in Europe. Difficulties inducing sleep, 51.4% of patients, is much greater than the 32% in Sao Paulo City, one of the greatest rates in the world. In Europe this values are about 7.5%. Early morning awakening, 31.4% of patients, against 30.8% ins Sao Paulo and 5.5% in Europe (10,11). Medical care disturbed the sleep of 43.0% of the patients, organic diseases for 29.0%, psychological problems for 17.0% and parasomnias for 28.0% per night. However, only in terms of environmental conditions, 66.0% of the patients reported disturbs, in which 45.7% were affect by noise, 31.4% were affected by illumination, 22.9% were affected by temperature, 20.0% were affected by the presence or attention to another people and 11% were affected by lack of adaptation to the hospital setting. These percentages are considerably greater than those of citizens of Sao Paulo City, which are 14% due to external factors and 9.5% due only to noise (10).

To assess the noise influence, it was studied separately the two hospitals. Patients who desired longer sleep duration at University was 61.5%, statistically different from the 27.3% at Baleia (Figure 1), which is better than the 44% at Sao Paulo. There was no statistical difference at home. At University the difference between hospital and home (23.1%) was statistically significant and at Baleia not, 27.3% and 33.3% respectively. The slightly better sleep rate for last hospital than at home indicates possibility of controlling sleep quality during hospitalization.

For all 24 questions the results indicated better sleep at Baleia patients, except "sleep perturbation due to temperature". This item and "sleep pertubation due to illumination" are the only ones in which the internal hospital noise could not be one of the intervening cause. However, only in four items the comparison of rates between the two groups were statistically different (9) (Figures 1,2). At University 30.8% of the patients reported having tired on awakening and was significantly greater than the 4.6% of the patients, at Baleia. At University the percentage of patients, who woke up tired, is greater than for the dwellers of British airports' neighbourhoods, where sound levels are considered disturbing (12).

FIG.1: NEED OF SLEEP in two groups of pacients at home and at hospitals: UNIVERSITY (large) and BALEIA (light)

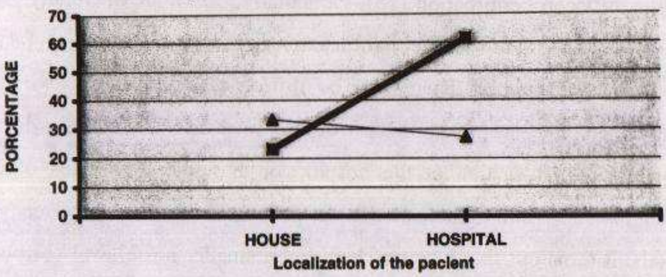
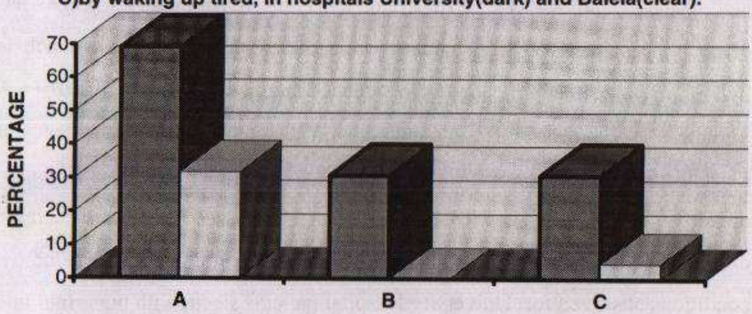


FIG.2: SLEEP DISTURBS,A)by noise in general,B)by traffic noise, C)by waking up tired, in hospitals University(dark) and Baleia(clear).



For all other evaluated possible causes for sleeping perturbation there was no difference between the two hospitals, except noise. At University (69.2%) this factor was significantly greater than at Baleia (31.8%). At home the difference was not significant (Figure 2). At home the majority of patients of both hospitals, who are poor and usually reside in the outskirks of city, seem already affected by sleep disturb induced by noise. Since the noise level increase toward downtown, the sleep pertubation by noise must still be higher, reaching those levels shown at University and corroborating the fact that BH is considered to be a noisy city (8). The rates of sleeping disorder induced by noise at University were greater than the one found in citizens of Oklahoma, USA, and in France (4), and they correspond to those found between Zaragoza's citizens, Spain, where noise is the first cause of claims of sleep perturbation (13). Particularly, 30.8% at University's patients pointed out traffic noise as the significantly greater sleep-disturbing factor than 0% at Baleia's (Figure 2). There was no difference between the two hospitals groups concerning sleep perturbation due to medical care, organic disease or psychological perturbation at the hospital or at home. The rates of diseases sleeping disorder seem to indicate that the patients are in the way of recovering at hospitals. It turns out that a better sleep and a smaller noise level corroborate to immunological recovery (2,7), whose better scores were found at Baleia.

Sleep perturbations due to noise have been reported in subjective (6,10,12) and objective studies (3,14,15,16), found a decreasing III, IV and V stage's lengths of deep sleep in EEG. Sleep quality was reduced since 45 dB(A), which is the maximum value recommended at residences and at hospitals in Brazil (1,3,17). The mean levels of noise recommended by WHO and the Brazilian Agency are of L_{eq}=35 and 40 dB(A) respectively (17,18). The quiet hospital, Baleia, with nocturnal internal L_{eq}=45.5 dB(A) and L_{max}=50.6 dB(A), monitored each minute during 500 minutes, is not at satisfactory standards at night. However, the overall conditions observed for University Hospital patients sleep, with nocturnal internal L_{eq}=53.7 dB(A) and L_{max}=59.1 dB(A), is worse due to its noisy environment. The nocturnal sound L_{eq} of the two hospitals were significantly different (t-Student test, $P<0.05$), but urinary cortisol were not statistically different, indicating that sound stress has not attained exhausting levels.

References: 1.A Chapon et al. Inst Recherche et Transports, France, (1972). 2.RW Cantrell. **The Laryngoscope**, suppl 1, LXXXIV(10/2),1-55, (1974). 3.MG Terzano et al. **Eletroenceph. clin. Neurophysiol.**, 76, 29-38, (1990). 4.J Mouret J and M Vallet. Min Santé et Action Hum, France, (1992). 5.M Soyer and JL Pichon. **Proc 6th Int Cong**, vol 2:621-624, France, (1993). 6.Ohrstrom E. **Proc 6th Int Cong**, vol 3:331-338, France, (1993). 7.K Altena and DGM Beersma. **Proc 6th Int Cong**, vol 2:575-578, France, (1993). 8.F Pimentel-Souza and PAS Alvares. **Proc 6th Int Cong**, vol 2:441-444, (1993) 9.F Pimentel-Souza et al. **Braz J Med Biol Res**, 29 (1996, in press). 10.S Braz. Doctoral thesis, Escola Paulista de Medicina, São Paulo, (1988). 11.C Janson et al. **Sleep**, 18(7):589-597, (1995). 12.J Ollerhead and I Diamond. **Proc 6th Int Cong**, vol 3:373-376, France, (1993). 13.PAS Sanches and JC Celme. **Proc 6th Int Cong**, vol.2, 323-326, (1993). 14.J Friedman and G Globus. NASA, USA, (1974). 15.M Vallet et al. Inst Recherche et Transports, France, (1975) 16.M Vallet et al. Inst Recherche et Transports, France, (1977). 17.ABNT. **NB-95.** ABNT, Brazil, (1966). 18.WHO. **Noise.** WHO, Geneva, (1980).

inter•noise 96
25th Anniversary Congress – Liverpool

INCE: 63.4

RECENT FIELD STUDIES IN THE UNITED STATES INVOLVING THE DISTURBANCE OF SLEEP FROM AIRCRAFT NOISE

K Pearsons

BBN Acoustic Technologies Division, Canoga Park, USA

A recent review (Pearsons, Barber, Tabachnick, and Fidell, 1995) of studies of noise-induced sleep disturbance indicated a large difference between behavioral awakening associated with noise for studies conducted in the laboratory and studies conducted in field (home) environments. Figure 1 shows the results of that review in terms of the prevalence of awakenings as a function of Indoor Sound Exposure Level (SEL). Since that review, a major field study conducted in the United Kingdom (Ollerhead, Jones, Cadoux, Woodley, Atkinson, Horne, Pankhurst, Reyner, Hume, Van, Watson, Diamond, Egger, Holmes, and McKean, 1992) has provided results that are in agreement with the field studies included in the review.

The great disparity between sleep disturbance findings under laboratory and field environments suggested the need for further research. However, as an interim measure, the results of the review by Pearsons et al. were reinterpreted in a relationship suggested by FICON (1992), as shown in Figure 2. Since more research was still necessary for environmental assessment purposes, studies supported by USAF (Fidell, Pearsons, Tabachnick, Howe, Silvati, and Barber, 1995a) and later by NASA (Fidell, Howe, Tabachnick, Pearsons, and Sneddon, 1995b) were conducted to provide more information on sleep disturbance due to noise in the home environment. This paper summarizes these research efforts.

USAF STUDY

Residential areas near a military training base (Castle Air Force Base), near a commercial international airport (Los Angeles International Airport) and in locations lacking appreciable nighttime aircraft noise were selected as test sites. The first two sites were designated as "high" and "medium" exposure sites, respectively, while the other locations were considered control sites. The number of subjects in the high, medium and control exposure sites were 27, 35 and 23 people respectively.

Noise events were measured in the test participants' bedrooms, between the hours of 10 pm and 8 am. A noise event was defined as a sequence of noise levels that began when an A-weighted threshold level was exceeded for at least 2 seconds and ended when the noise level dropped more than 2 dB below the threshold.

Sleep disturbance was recorded on a palmtop computer at beside. The test participant indicated the time of retiring, the time of first attempting to sleep and the time of finally waking up in the morning by pressing various response buttons on the computer. The computer queried the test participant in the evening about tiredness and possible alcohol and drug consumption. Morning queries included judgments of overall sleep quality, feelings of tiredness, recall of number and duration of nighttime awakenings and time to fall asleep. A small, hand-held push-button attached by cable to the computer was used to indicate when he/she woke up for whatever reason. The computer time-stamped the button push and stored the result in a file covering the time that the test participant attempted to sleep.

The data set analyzed included a total of 1887 subject nights from the three sites, which was made up of 632 from the high exposure site (Castle AFB), 783 from the medium exposure site (LAX) and 472 from the control (non-aircraft) site. The average number of awakenings per night at each of the three sites ranged from 2.2 to 2.5. Several noise metrics, including Maximum A-level, Sound Exposure Level (SEL), time-weighted average A-level (TAVA) and A-weighted signal-to-noise ratio (S/N) were used to describe noise events possibly responsible for awakenings. In addition, different amounts of time (1, 2 or 5 minutes) were included following an event to determine which provided the best dose response relationship for sleep disturbance. Correlations between the probability of awakenings and noise event metrics with various amounts of post event intervals were calculated. The highest correlation was found for SEL, within 5 minutes prior to an awakening. A linear dose-response relationship is shown in Figure 3 for the three sites included in this study.

Other, higher order regressions were attempted with the data shown in Figure 3. Although no significant improvement in correlation resulted, factoring in the results of the evening and morning questionnaires did improve overall correlation. The averaged results shown in Figure 3 did, however, agree with the averaged line for field studies shown in Figure 1.

NASA STUDY

The NASA study, as initially envisioned, differed from the USAF study in two ways. First, the NASA study was designed to test the effect of noise on sleep for a sudden offset (and onset) of aircraft flyover noise. Second, an additional method was included of assessing sleep disturbance using actigraphs (or actimeters). Actigraphs provide a measure of motility (gross body movement) that has been reported to be comparable to EEG measures of waking on a whole-night basis. The actigraph is a small, self-contained recording accelerometer worn on the wrist. Two types were used in different portions of the studies: a Swiss-made actimeter that records the time above a fixed threshold for every 30 second epoch, and a U.S.-made actigraph that records the number of times that a threshold is exceeded during every 30 second epoch.

The study was conducted at two sites in the greater Denver, Colorado area. One site was near the original Stapleton Airport (DEN) while the other was near the new Denver International Airport (DIA). The sites were chosen because of the sudden change in aircraft noise exposure that was expected when DEN closed and DIA opened. The transition was to take place in less than a day so that people who were exposed to noise near DEN would not hear aircraft noise the day after DEN closure, while the people located near DIA who had not been exposed to aircraft noise would suddenly hear aircraft noise. The closure of DEN was initially scheduled for 1 March 1994. Noise and sleep disturbance instrumentation was set up at DEN for two weeks prior to the scheduled closing. However, closing was rescheduled to 15 May 1994. In spite of the change, data were collected for two more weeks, for a total of four weeks. Three weeks prior to the new changeover date, data collection began in a relatively quiet neighborhood northwest of DIA. When another postponement of the opening of DIA was announced in the midst of this data collection, it was decided to continue these measurements for an additional two weeks.

Data collection for the third effort started 10 days before the actual DIA opening date of 28 February 1995. Data collection continued for 3 weeks at a location again northwest of DIA. The final round of data collection started a week after the conclusion of the third round and was located near DEN. This round lasted nearly 3 weeks.

Data were analyzed for a total of 2717 subject nights at the DEN and DIA sites. The average number of awakenings per night was 1.6, somewhat lower than that observed in the USAF study. Outdoor event noise levels changed at both locations after the opening of the new airport. Indoor nighttime noise event levels as measured in sleeping quarters were not as dramatic as expected at the two airports. This lack of a dramatic change in the noise environment in the sleeping quarters also manifested itself in a lack of observable changes in sleep disturbance. Dose response relationships were still possible using techniques described in the USAF study for behavioral awakening, as shown in Figure 4. In addition, actimeter data in various forms provided similar results as shown in Figure 5-7. Figure 5 shows the prevalence of actimetric blips defined by Ollerhead's criterion, resulting from use of the Swiss-made actimeter. Figure 6 shows the prevalence of actimetric threshold crossings (raw data) using the U.S.-made actigraph. Finally, Figure 7 shows the prevalence of arousal responses defined by Cole's criterion (Cole, Kripke, Gruen, Mullaney, and Gillin, 1992) using the U.S.-made actigraph. Although the meaning of these non-behavioral awakening responses is unclear, they do provide better correlation with event SEL's than does behavioral awakening, as shown in Figures 5 and 6, and comparable correlation with Event SEL's to behavioral awakening, as shown in Figure 8.

CONCLUSIONS AND RECOMMENDATIONS

Figure 8 presents all of the results for behavioral awakening collected in these two studies, along with the average field data results represented by Figure 1. The data are remarkably consistent. This is not to say that there is not a lot of variation in the responses across individuals, but the mean prevalence of awakenings does appear quite stable over the range of SEL's shown. The difference between the field and laboratory results has still not been resolved. Additional tests are now being planned to further investigate the effect of habituation.

All of the recent studies on sleep have been conducted in locations where people are accustomed to the noise, a possible exception being the recent study at DIA. However, no criteria currently exist for relatively infrequent events to which people have not yet or may never become accustomed.

Acknowledgment: The studies described in this presentation were supported by:

Armstrong Laboratory, Occupational & Environmental Health Directorate
Bioenvironmental Engineering Division, Human Systems Center
Air Force Material Command

National Aeronautics and Space Administration
Langley Research Center

REFERENCES

Cole, R. J., D. F. Kripke, W. Gruen, D. J. Mullaney, & J. C. Gillin (1992). "Automatic Sleep/Wake Identification from Wrist Activity," *Sleep*, *15*(5), pp. 461-469.

Federal Interagency Committee on Noise (FICON) (1992). Final Report: Airport Noise Assessment Methodologies and Metrics, Washington, D.C.

Fidell, S., K. Pearsons, B. Tabachnick, R. Howe, L. Silvati, & D. Barber (1995a). "Field Study of Noise-Induced Sleep Disturbance," *J. Acoust. Soc. Am.*, *98*(2:1), pp. 1025-1033.

Fidell, S., R. Howe, B. Tabachnick, K. Pearsons, & M. Sneddon (1995b). "Noise-Induced Sleep Disturbance in Residences Near Two Civil Airports," NASA Contractor Report 198252, NASA Langley Research Center, Hampton, VA.

Ollerhead, J. B., C. J. Jones, R. E. Cadoux, A. Woodley, B. J. Atkinson, J. A. Horne, F. Pankhurst, L. Reyner, K. I. Hume, F. Van, A. Watson, I. D. Diamond, P. Egger, D. Holmes, & J. McKean (1992). "Report of a Field Study of Aircraft Noise and Sleep Disturbance," London: Department of Safety, Environment and Engineering.

Pearsons, K., D. Barber, B. Tabachnick, & S. Fidell (1995). "Predicting Noise-Induced Sleep Disturbance," *J. Acoust. Soc. Am.*, *97*(1), pp. 331-338.

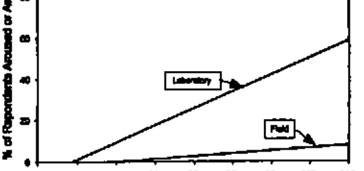


Figure 1. Comparison of Laboratory and Field studies

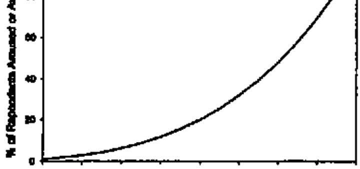


Figure 2. Dosage-Response Curve Results Adopted by FICON

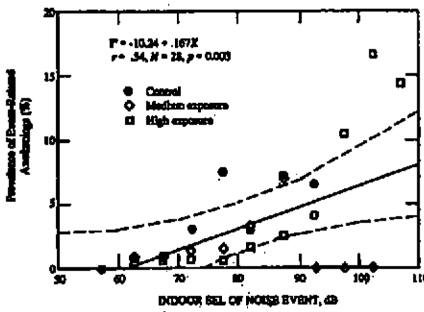


Figure 3. Prevalence of Behavioral Awakening Reponses [USAF]

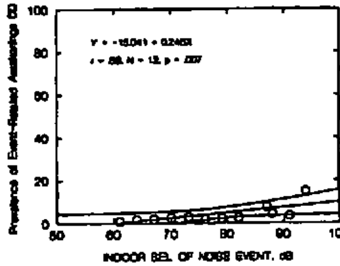


Figure 4. Prevalence of Behavioral Awakening Reponses [NASA]

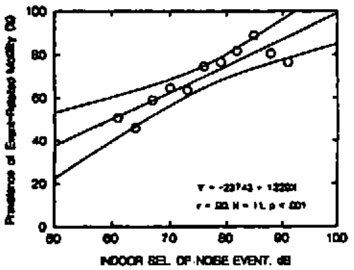


Figure 5. Prevalence of Actimetric Blips (defined by Ollerhead's criterion) [NASA]

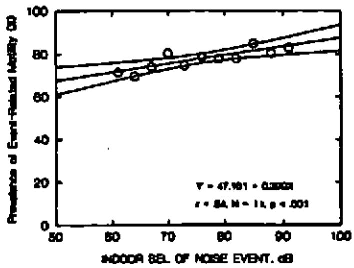


Figure 6. Prevalence of actimetric Threshold Crossings (raw data) [NASA]

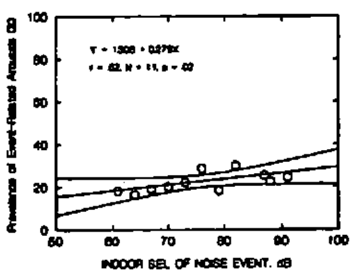


Figure 7. Prevalence of Arousal Responses (defined by Coles criterion) [NASA]

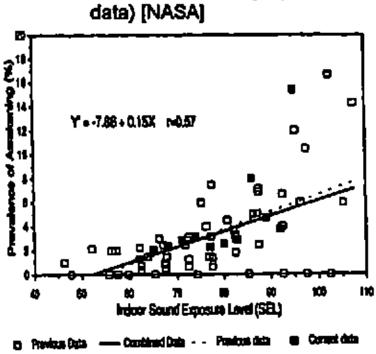


Figure 8. Composite of Data from NASA Study with Findings of Prior Studies