NASA Contractor Report 198252

Noise-Induced Sleep Disturbance in Residences Near Two Civil Airports

Sanford Fidell, Richard R. Howe, Barbara G. Tabachnick, Karl S. Pearsons, and
Matthew D. Sneddon
BBN Acoustic Technologies Division, Canoga Park, California

Contract NAS1-20101

December 1995

National Aeronautics and Space Administration
Langley Research Center
Hampton, Virginia 23681-0001

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The authors are grateful for the cooperation of test participants in the observations of in-home sleep disturbance reported here. Dr. Kevin P. Shepherd of the Structural Acoustics Branch of NASA Langley Research Center served as Technical Monitor throughout the planning and conduct of this study. John Ollerhead of National Air Traffic Services, Civil Aviation Authority, London, Professor Jim Horne of Loughborough University, UK, and Mark Rosekind of NASA Ames Research Center, provided valuable advice and assistance with equipment. The U.S. Air Force's Armstrong Laboratory at Wright-Patterson Air Force Base made available noise monitoring instruments, and the U.S. Army's Construction Engineering Research Laboratory in Champaign, Illinois made available additional instrumentation for recording behavioral awakening responses. David Barber, Elizabeth Fletcher, Linda Solar, and Suyeo Tomooka assisted with equipment assembly, installation, and servicing.

ABSTRACT

A large-scale field study of noise-induced sleep disturbance was conducted in the vicinities of Stapleton International Airport (DEN) and Denver International Airport (DIA) in anticipation of the closure of the former and opening of the latter. Both indoor and outdoor measurements of aircraft and other nighttime noises were made during four time periods. Measurements were made in 57 homes located as close as feasible to the runway ends of the two airports. Sleep disturbance was measured by several indices of behaviorally-confirmed awakening (button pushes upon awakening) and body movement (as measured with wrist-worn actimeters). A total of 2,717 subject-nights of observations was made over the course of the study.

Although average noise event levels measured outdoors decreased markedly at DEN after closure of the airport and increased slightly at DIA after its opening, indoor noise event levels varied much less in homes near both airports. No large differences were observed in noise-induced sleep disturbance at either airport. Indoor sound exposure levels of noise events were, however, closely related to and good predictors of actimetrically defined motility and arousal.

1 EXECUTIVE SUMMARY

This report describes a field study of the effects of nighttime noise exposure on the sleep of residents near two large civil airports. Observations of gross bodily movements (motility), behaviorally-confirmed awakenings, and self-reported awakenings were made in residences of test participants in neighborhoods as close as feasible to Stapleton International (DEN) and Denver International (DIA) airports, while noise levels produced by aircraft and other sources were monitored both outdoors and within sleeping quarters. The study period spanned the closing of DEN and the opening of DIA.

1.1 BACKGROUND

Sleep disturbance in airport neighborhoods remains a matter of considerable interest for both environmental assessment and regulatory purposes, since a fully satisfactory dosage-response relationship for predicting sleep disturbance from noise exposure in residential settings is not yet available (*cf.* FICON, 1992; Pearsons, Barber, Tabachnick, and Fidell, 1995). Two recent field studies of noise-induced sleep disturbance (Ollerhead, Jones, Cadoux, Woodley, Atkinson, Horne, Pankhurst, Reyner, Hume, Van, Watson, Diamond, Egger, Holmes, and McKean, 1992, and Fidell, Pearsons, Howe, Tabachnick, Silvati and Barber, 1995) have greatly increased the stock of information about noise-induced sleep interference in field settings. Although the studies of Ollerhead *et al.* and of Fidell *et al.* both measured behavioral indications of sleep disturbance, and although their findings are in reasonable agreement, they focused on different aspects of sleep disturbance and also differed in details of noise measurement. Ollerhead *et al.*, for example, considered the gross bodily movement ("motility") of test participants in their beds as an indication of sleep disturbance, while Fidell *et al.* measured behaviorally-confirmed awakenings. Ollerhead *et al.* measured noise levels produced outdoors by confirmed aircraft overflights, while Fidell *et al.* measured both outdoor and indoor noise exposure from all sources.

The primary goal of this study was to supplement the stock of field observations of aircraft noise-related sleep disturbance, and to document any changes in such disturbance associated with changes in aircraft operations. Another goal of the current study was to investigate whether motility and behavioral awakening measure the same kind of noise-induced sleep disturbance, and whether the two measures are equally sensitive to noises of indoor and/or outdoor origin. The study began in January of 1994 in anticipation of changes in aircraft noise exposure associated with the (then) imminent closure of DEN and the opening of the newly constructed DIA. Unanticipated delays in the opening of DIA required several modifications of the original test plan, eventually leading to four rounds of data collection.

Both behavioral awakening and motility measurements were made in the first round of data collection in residences near DEN for two weeks prior to its closure, along with outdoor measurements of aircraft noise and indoor measurements of household noise in test participants' sleeping quarters. Behavioral awakening was measured in a manner identical to that described by Fidell *et al.* (1994). Motility measurements were made with two types of actimeters. All of these measurements were originally planned to continue for at least two weeks after closure of DEN.

When postponement of the closing of DEN was announced in the midst of the initial data collection on 1 March 1994, it was decided to continue these measurements for an additional two weeks in any event. When a second opening date for DIA was announced for 15 May 1994, additional data were collected in a relatively quiet neighborhood to the north of DIA starting three weeks prior to the announced opening date. When another postponement of the opening of DIA was announced in the midst of this round of data collection, it was decided to continue these measurements for an additional two weeks as well.

The third round of data collection began approximately 10 days prior to the actual opening date for DIA, 28 February 1995. To the extent possible, participants who had contributed data in the second round of data collection served in this third round as well. Data collection continued for a total of five weeks.

The final round of data collection was started on 1-2 April 1995 in the areas near DEN in which observations had previously been made in the first round of data collection. Most of the same people who had contributed data earlier also participated in this final round of data collection.

1.2 RESULTS AND DISCUSSION

Table 1 summarizes the conditions under which 2,717 subject-nights of data were collected. Figures 1 through 4 summarize the distributions of indoor and outdoor nighttime noise events in each round of data collection. As expected, outdoor nighttime noise event levels decreased greatly after flight operations ceased at DEN. Outdoor nighttime noise event levels increased, although less dramatically, near test participants' homes after flight operations began at DIA. Indoor nighttime noise event levels as measured in sleeping quarters were much less affected by the changes in aircraft operations at both airports.

Table 1 Summary of data collection effort.

Site	Data Collection Round	Number of Homes	Number of Test Participants	Number of Subject-nights of Data Collection
DEN	Before closure (February/March 1994)	15	30	677
DIA	Before opening (April/May 1994)	14	29	712
DIA	Spanning opening (February/March 1995)	13	30	848
DEN	After closure (April 1995)	15	28	480
	TOTAL	57 (38 different homes)	117 (77 different people)	2717

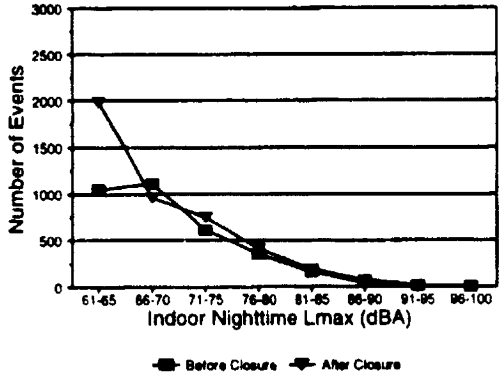


Figure 1 Distribution of noise events recorded inside test participants' sleeping quarters at DEN from 2200 to 0700 hours.

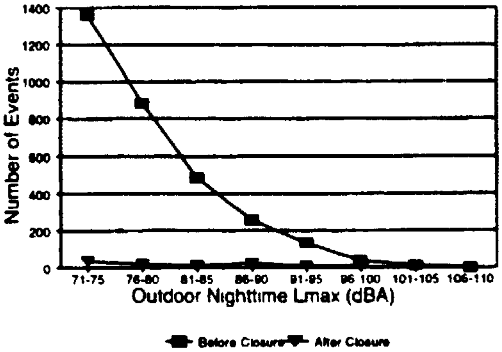


Figure 2 Distribution of outdoor noise events at DEN between 2200 and 0700 hours.

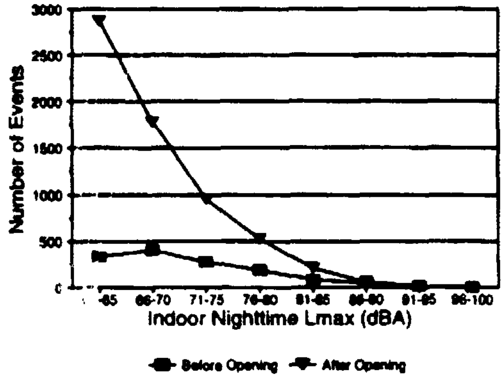


Figure 3 Distribution of noise events recorded inside test participants' sleeping quarters at DIA from 2200 to 0700 hours.

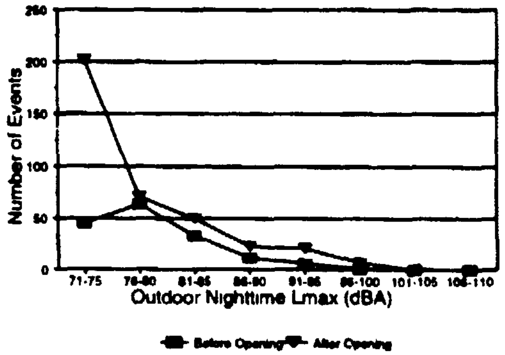


Figure 4 Distribution of outdoor noise events at DIA between 2200 and 0700 hours.

Figures 5 through 8 show dosage-response relationships developed for the various measures of sleep disturbance from the data of the present study. Several measures of sleep disturbance were reliably associated with indoor sound exposure levels of noise events. Motility was a more sensitive measure of sleep disturbance than awakening.

Figure 9 compares the average rate of behavioral awakening responses in the presence of aircraft noise from operating airports with the average rate of behavioral awakening responses in the absence of aircraft noise from operating airports for individual test subjects. Figure 10 compares motility in the presence of aircraft with motility in the absence of aircraft. The pattern of findings summarized in these figures indicates that neither awakenings nor motility were greatly affected by the changes in aircraft flight operations at the two airports.

Table 2 summarizes analyses performed on the collected data.

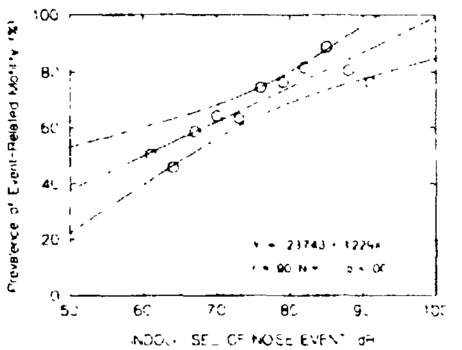


Figure 5 Prevalence of actimetric blips (defined by Ollerhead's criterion) at DEN before airport closure, aggregated over test participants in 3 dB increments of indoor SEL values of noise events. Curved lines bound the 95% confidence interval.

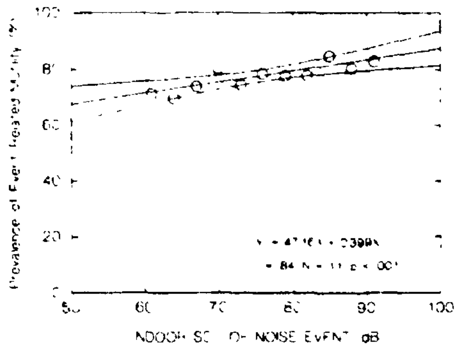


Figure 6 Prevalence of actimetric threshold crossings (defined by Cole's criterion) at DEN and DIA, aggregated over test participants in 3 dB increments of indoor SEL values of noise events. Curved lines bound the 95% confidence interval.

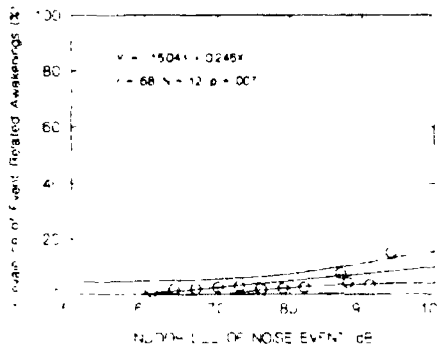


Figure 7 Prevalence of behavioral awakening responses at DEN and DIA aggregated over test participants in 3 dB increments of indoor noise measurements. Curved lines bound the 95% confidence interval.

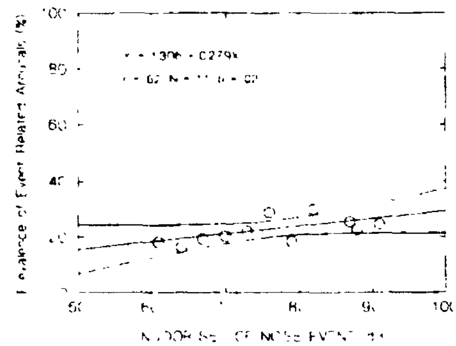


Figure 8 Prevalence of arousal responses by U.S. actimetric criterion at DEN and DIA aggregated over test participants in 3 dB increments of indoor noise measurements. Curved lines bound the 95% confidence interval.

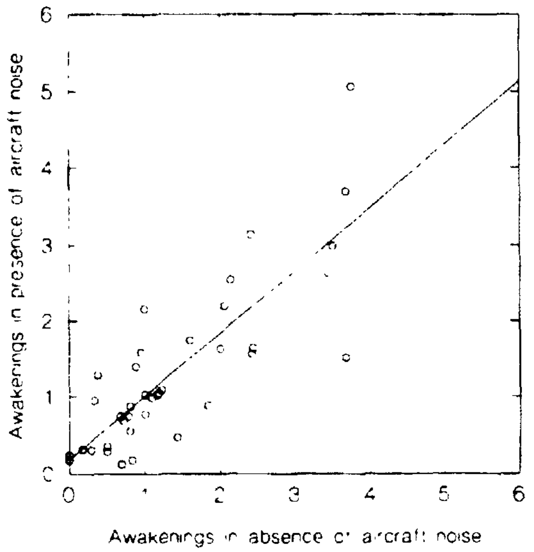


Figure 9 Average number of behavioral awakenings per night in the presence (ordinate) and absence (abscissa) of aircraft noise. Each data point represents responses of a single participant.

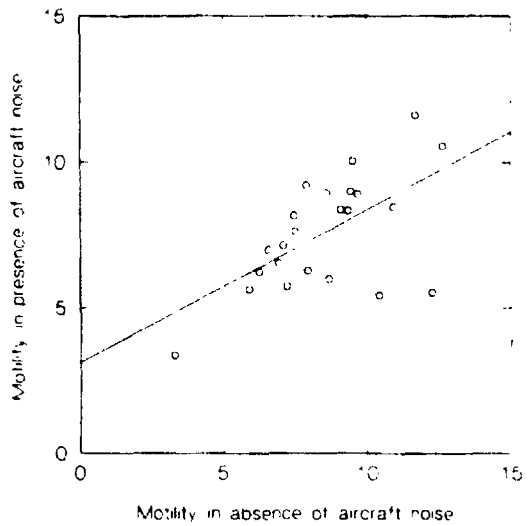


Figure 10 Average motility response in each 30-second epoch in the presence (ordinate) and absence (abscissa) of aircraft noise. Each data point represents responses of a single participant aggregated over the entire study.

1.3 CONCLUSIONS

Because no effort was made to rigorously define the complete population exposed to nighttime noise exposure, nor to obtain a representative sample of any wider population, conclusions drawn from the present study apply strictly only to test participants. To the extent that generalizations are made from the present findings, they should be restricted to the effects of noise on the sleep of long-term residents of neighborhoods without sudden, large changes in nighttime noise exposure.

The following are among the major findings of the present study:

1) The current findings closely resemble those of prior field studies of noise induced sleep disturbance.

2) Outdoor nighttime L_{eq} decreased about 12 dB on average at DEN upon closure of the airport, but increased only about 3 dB at DIA after opening of the airport.

Table 2 Guide to analyses performed in this study.

ANALYSIS	DATA SET	RESULTS SECTION
DESCRIPTIVE ANALYSES		
Indoor and outdoor noise environments	DEN before airport closure	4.2.1
Indoor and outdoor noise environments	DEN after airport closure	4.2.2
Indoor and outdoor noise environments	DIA before airport opening	4.2.3
Indoor and outdoor noise environments	DIA after airport opening	4.2.4
Behavioral awakening responses, motility, and self-reported awakenings	DEN before airport closure	4.3.1
Behavioral awakening responses, motility, and self-reported awakenings	DEN after airport closure	4.3.2
Behavioral awakening responses, motility, and self-reported awakenings	DIA before airport opening	4.3.3
Behavioral awakening responses, motility, and self-reported awakenings	DIA after airport opening	4.3.4
INFERENTIAL ANALYSES		
Dosage-response analysis	All data	4.4.1
Temporal adaptation of behavioral awakening responses at DIA	DIA before and after airport opening (1995)	4.4.1
Temporal adaptation of behavioral awakening responses and recalled awakenings	DIA before and after airport opening (1995)	4.4.2.1
Temporal adaptation of behavioral awakening responses and recalled awakenings	DIA one year before airport opening (1994)	4.4.2.1
Temporal adaptation of behavioral awakening responses and recalled awakenings	DEN before and after airport closure	4.4.2.2
Indoor L_{eq} before and after airport opening	DIA before and after opening (1995)	4.4.2.1
Outdoor L_{eq} before and after airport opening	DIA before and after airport opening (1995)	4.4.2.1
Analysis of variance on indoor and outdoor L_{eq}	DEN before and after airport closure (1995)	4.4.2.2
Multiway frequency analysis of awakenings and arousals as defined by three criteria	DEN before airport closure (1994)	4.4.3.1
Relationship between motility and behavioral awakening responses	DEN before airport closure (1994)	4.4.3.2
Relationship between initial sleep latency and time spent awake	DEN before airport closure (1994)	4.4.4
Relationship between behavioral awakening responses and recalled awakenings	All data	4.4.5
Prediction of Swiss-made actimeter measured motility from noise event levels and control variables	DEN before airport closure (1994)	4.4.6.1
Prediction of U.S.-made actimeter measured motility from noise event levels and control variables	All data from participants using AMI actimeters	4.4.6.2
Prediction of behavioral awakening responses from noise levels and control variables	All data	4.4.6.3
Prediction of U.S.-made actimetric arousals as defined by Cole's criterion	All data from participants using AMI actimeters	4.4.6.4
Attempted replication of Ollerhead's (1992) analysis	DEN before airport closure	Appendix F

3) Indoor nighttime L_{eq} varied little at either location with the transfer of flight operations from DEN to DIA.

4) The average number of behavioral awakening responses per night was 1.8 at DEN and 1.5 at DIA. The number of spontaneous behavioral awakening responses (unassociated with noise events) was 1.5 per night at DEN and 1.3 at DIA.

5) Statistically reliable relationships were observed between sound exposure levels of individual noise intrusions as measured inside sleeping quarters and several measures of sleep disturbance. These were:

   - SEL of individual noise intrusions accounted for about 81% of the variance in motility as measured by the Swiss-made actimeter. The linear relationship between the percentage of test participants exhibiting motility following a noise event was % motility = -23.74 + 1.23(SEL).

   - SEL of individual noise intrusions accounted for about 71% of the variance in motility as measured by the U.S.-made actimeter. The linear relationship between the percentage of test participants exhibiting motility following a noise event was % motility = 47.16 + 0.4(SEL).

   - SEL of individual noise intrusions accounted for about 45% of the variance in behavioral awakening responses. The linear relationship between the percentage of test participants exhibiting a behavioral awakening response following a noise event was % noise-induced awakening = -15.04 + 0.25(SEL).

   - SEL of individual noise intrusions accounted for about 38% of the variance in arousals as measured by the U.S.-made actimeter and defined and processed in accordance with the criteria of Cole *et al.* (1992). The linear relationship between the percentage of test participants exhibiting arousal following a noise event was % arousal = 1.31 + 0.28(SEL).

6) Indoor SEL accounted for less than one-third of the predictable variance in sleep disturbance in logistic regression models including other predictors.

7) Relationships among measures of sleep disturbance were reliable but weak to moderate:

- About 19% of variance was shared between motility as measured, processed, and defined by the two types of actimeter.

- About 1% to 5% of variance was shared among behaviorally-confirmed awakening and the two actimetric criteria for awakening.

- About 25% of variance was shared between behaviorally-confirmed and self-reported awakenings; participants recalled awakening slightly less than twice per night and pushed buttons to indicate awakenings about 1.6 times per night.

- About 4% of variance was shared between actimetrically-defined sleep latency and recalled time to fall asleep; recalled and actimetrically-defined sleep latency was about 17-18 minutes on average.

- About 25% of variance was shared between actimetrically-defined and recalled time spent awake; recalled time awake (about 12 minutes on average) was considerably shorter than actimetrically-defined (about 34 minutes on average).

2 INTRODUCTION

Sleep disturbance in airport neighborhoods remains a matter of considerable interest for both environmental assessment and regulatory purposes, since a reliable dosage-response relationship for predicting sleep disturbance from noise exposure in residential settings is not yet available (*cf.* FICON, 1992; Pearsons, Barber, Tabachnick, and Fidell, 1995). Two recent field studies of noise-induced sleep disturbance (Ollerhead, Jones, Cadoux, Woodley, Atkinson, Home, Pankhurst, Reyner, Hume, Van, Watson, Diamond, Egger, Holmes, and McKean, 1992, and Fidell, Pearsons, Howe, Tabachnick, Silvati and Barber, 1995) have nonetheless greatly increased the stock of information of this sort collected in field settings. Although the studies of Ollerhead *et al.* and of Fidell *et al.* both measured behavioral indications of sleep disturbance, and although their findings are in reasonable agreement, they focused on different aspects of sleep disturbance and also differed in details of noise measurement. Ollerhead *et al.*, for example, considered the gross bodily movement ("motility") of test participants in their beds as an indication of sleep disturbance, while Fidell *et al.* measured behaviorally-confirmed awakenings. Ollerhead *et al.* measured noise levels produced outdoors by confirmed aircraft overflights, while Fidell *et al.* measured both outdoor and indoor noise exposure from all sources.

Such differences among studies in measurements of noise exposure and sleep disturbance have created difficulties of interpretation and comparison. For example, Ollerhead *et al.* (1992) define "arousal" as "...the onset of sleep disturbance as measured by an actimeter..."; "awakening" as "... at least 15 seconds of 'wakefulness' or 10 seconds of 'movement time' in the EEG record"; and "(sleep) disturbance" as "... both awakenings and actimetrically-determined arousals...." Ollerhead *et al.* note that the latter term includes "events ... such as EEG-awakenings." Fidell *et al.* adopted an operational definition of sleep disturbance based on a behavioral confirmation (*i.e.*, "awake enough to push a bedside button"). The uncertain relationship between motility-based measures of sleep disturbance and behavioral awakening is a hindrance to development of a unified dosage-response relationship for noise-induced sleep disturbance.

2.1 OVERVIEW OF PRESENT STUDY

The present study began in January of 1994 in anticipation of changes in aircraft noise exposure associated with the (then) imminent closure of DEN and the opening of the newly-constructed DIA. Unanticipated delays in the opening of DIA required several modifications of the original test plan, eventually leading to four rounds of data collection. Both behavioral awakening and motility measurements were made in the first round of data collection in residences near DEN for two weeks prior to its closure, along with outdoor measurements of aircraft noise and indoor measurements of household noise in test participants' sleeping quarters. Behavioral awakening was measured in a manner identical to that described by Fidell *et al.* (1994). Motility measurements were made with two types of actimeters. All of these measurements were originally planned to continue for at least two weeks after closure of DEN.

When postponement of the closing of DEN was announced in the midst of the initial data collection on 1 March 1994, it was decided to continue these measurements for an additional two weeks in any event. When a second opening date for DIA was announced for 15 May 1994, additional data were collected in a relatively quiet neighborhood to the north of DIA starting three weeks prior to the announced opening date. When another postponement of the opening of DIA was announced in the midst of this round of data collection, it was decided to continue these measurements for an additional two weeks as well.

The third round of data collection began approximately 10 days prior to the actual opening date of DIA, 28 February 1995. To the extent possible, participants who had contributed data in the second round of data collection were contacted and served in this third round. Data collection continued for a total of five weeks in the vicinity of DIA.

The final round of data collection was started on 1-2 April 1995 in the areas near DEN in which observations had previously been made in the first round of data collection. Once again, the same people who had contributed data earlier also participated in this final round of data collection.

2.2 CLARIFICATION OF TERMS AND ANALYTIC APPROACHES

A glossary (*cf.* Chapter 8) defines statistical and acoustic terms and expands abbreviations. This section discusses major concepts in the design and analysis of this study.

2.2.1 Definition of Noise Events

Even though the noise exposure that causes sleep disturbance cannot be identified *a priori* in a field study because it is not under experimental control, a definition of a noise event is required for analysis of the relationship between noise exposure and sleep disturbance. A common definition of a noise event is a time series of noise levels that begins when a threshold level is exceeded for some period of time and continues until the level remains below the same or another threshold for another period of time. Threshold parameters may vary with the peculiarities of measurement sites, the noise sources of interest, and the tolerable false alarm rate for classifying noise events.

Ollerhead *et al.* (1992) adopted a 60 dB (outdoor) A-weighted threshold for defining aircraft noise events, and also required independent (non-acoustic) confirmation of the occurrence of an aircraft operation during the same time period. Fidell *et al.* (1994) found an (indoor) threshold level of 60 dB useful for defining an event in areas exposed to nighttime aircraft noise. They employed a 50 dB (indoor) threshold for control sites lacking nighttime aircraft noise, however.

Noise events were defined in the present study with respect to A-weighted threshold levels of 70 dB for outdoor noise and 60 dB for indoor noise. Noise levels had to exceed these threshold levels for at least 2 seconds, and could not dip 2 dB or more below these thresholds at any time.

2.2.2 Definition of Noise Epochs

An alternative approach to characterizing noise exposure that may be associated with sleep disturbance in an observational study is to construct a time series of contiguous noise measurements that can be analyzed in consecutive epochs of specified duration. These epochs may be examined for evidence of association between noise levels and sleep disturbance. Fidell *et al.* (1994) found 1-minute analysis epochs to be superior to 2- or 5-minute epochs for purposes of predicting behavioral awakening.

Epochs in the present study were 1 minute long for noise data gathered in data collection Rounds 1 and 2, 2 seconds long for noise data gathered during Rounds 3 and 4, and 30 seconds long for actimetric data.

2.2.3 Measures of Motility

Motility may be measured in a number of ways of varying cost and appropriateness for different purposes. According to Ollerhead *et al.*, gross body movements cease only during periods of deep sleep. Sleep disturbance may be measured by a wrist-worn recording device sensitive to arm and body movement (an "actimeter"). The efforts of Ollerhead *et al.* to relate motility to EEG activity were not fully successful in reliably distinguishing "arousals" (shifts from deeper to lighter sleep states) from "awakenings" (departure from an intuitively reasonable definition of sleep). Ollerhead *et al.* believed that about 40% of arousals inferred from highly processed motility data represented awakenings, but despite extensive analyses, were unable to distinguish arousals from awakenings on an episode-by-episode basis. Further, the number of awakenings predicted by their actimetric criteria after the 40% adjustment is still far in excess of the number of behaviorally-confirmed awakenings observed by Fidell *et al.* For reasons described in further detail in Section 3.4, motility was measured by two different instruments in the current study.

2.2.4 Definitions of Arousal

Arousal is defined in the current study with respect to actimeter type. The first definition of arousal is that of Ollerhead *et al.* (1992), based on measurements made by the Swiss-made actimeter during analysis epochs. Ollerhead's algorithm defined an arousal simply as any indication of actimetric activity after at least one epoch of no activity.

The second definition of arousal is based on data collected from the U.S.-made actimeter and proposed by Cole *et al.* (1992). Cole applied the following formula to the actigraph data collected in 30-second epochs:

$$D=0.0001(50A_{-4}+30A_{-3}+14A_{-2}+28A_{-1}+121A_{0}+8A_{+1}+50A_{+2})$$

where subscripts indicate activity during epochs preceding or succeeding the present epoch. When this expression is evaluated, Cole *et al.* considered any value greater than 1.0 to be an arousal.

2.2.5 Behaviorally-Confirmed Awakening

Behavioral indications of sleep disturbance other than motility are common in prior studies of noise-induced sleep disturbance. Perhaps the simplest of these is awakening confirmed by a button press (*cf.* Horonjeff *et al.*, 1982, and Fidell *et al.*, 1995). Although behaviorally-confirmed awakening does not provide fine detail about sleep state changes, disturbance so defined is relatively unambiguous, lends itself to straightforward interpretation, and can be cost-effectively measured with good temporal resolution in a large-scale field study. Behavioral confirmation of awakening was accomplished in the present study by means identical to those used by Fidell *et al.*

2.2.6 Analysis of Associations Between Noise Exposure and Sleep Disturbance

Three sorts of analyses may be undertaken of the association of noise exposure with behavioral indications of sleep disturbance:

- "noise event-based" (prospective) analyses,
- "awakening-based" (retrospective) analyses, and
- "entire night" analyses.

A **noise event-based** analysis seeks indications of sleep disturbance within some period (*e.g.*, one or five minutes) *after* the occurrence of a noise event exceeding a site-specific level and duration threshold. An **awakening-based** analysis attempts to associate sleep disturbance with noise measured in epochs (*e.g.*, of one or five minutes duration) *prior* to the occurrence of an awakening. **Entire night** analyses may be conducted on longer term, cumulative noise measures and both behavioral and self-report responses to a whole night's sleep.

2.3 ORGANIZATION OF THIS REPORT

Chapter 3 describes the procedural details of field data collection, including descriptions of the sites, test participants, instrumentation and instructions. Noise and actimetric measurements are analyzed in Chapter 4. Chapter 5 discusses the findings, while Chapter 6 presents conclusions.

A Glossary and a set of Appendices provide supporting detail. Appendix A contains details of recruitment of test participants. Appendix B contains a detailed description of data extraction procedures. Appendix C provides detailed figures showing the time-course of behavioral awakening responses for 22 participants at DIA. Appendix D summarizes the noise environments at DEN and DIA. Appendix E provides a summary of responses to the nighttime and morning questionnaires. Appendix F details the logistic regression analyses performed in this study. Appendix G describes efforts to replicate the analyses conducted by Ollerhead *et al.* (1992).

3 METHOD

This Chapter describes procedures used to select sites, to measure noise exposure, and to collect, reduce and analyze sleep disturbance data.

3.1 STUDY SITES AND DATA COLLECTION SCHEDULES

Observations of noise exposure and sleep disturbance were made in single-family detached homes near DEN and DIA during four rounds of data collection, as described below. Table 3 summarizes the circumstances of data collection.

Table 3 Summary of data collection conditions.

Round	Site	Data Collection Condition	Time
1	DEN	Before closure	(February/March 1994)
2	DIA	Before opening	(April/May 1994)
3	DIA	Spanning opening	(February/March 1995)
4	DEN	After closure	(April 1995)

3.1.1 First Round of Data Collection

The initial selection of data collection sites was based on the announced closing of DEN in March of 1994. Data collection began in two neighborhoods in the vicinity of DEN two weeks prior to the announced closing date. Data collection continued for an additional two weeks after the closing date had been postponed. Noise exposure and sleep disturbance were measured in 15 residences in the vicinity of DEN, mostly in the city of Aurora. Figure 11 shows the locations of test participants' homes in areas to the immediate south and east of DEN.

3.1.2 Second Round of Data Collection

A residential neighborhood as close as possible to DIA was selected for a second round of data collection after the next announcement of an opening date for DIA was made. Data collection began three weeks before the announced opening date (15 May 1994) and continued for an additional two weeks after the opening was once again postponed. resulting in a total of five weeks' data collection. Figure 12 shows the locations of fourteen test participants' homes near DIA.

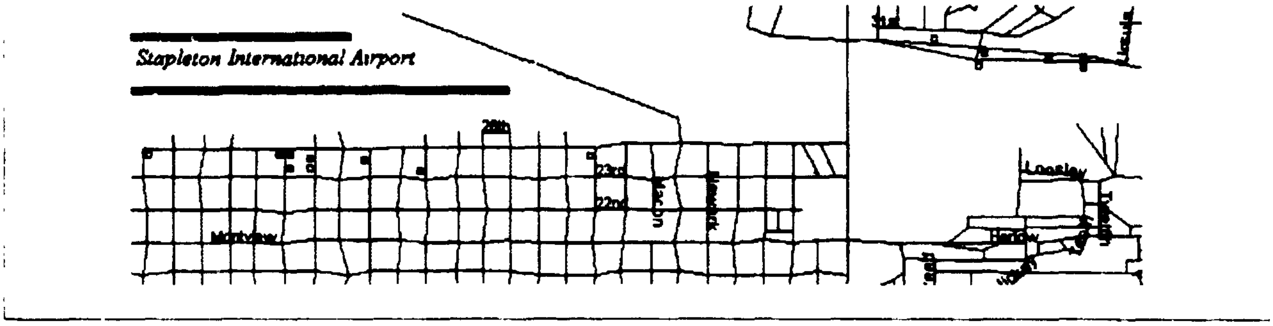


Figure 11 Map of study area near DEN. Squares indicate participants' homes. Open squares denote outdoor noise monitoring sites. Noise monitors were installed in all sleeping quarters.

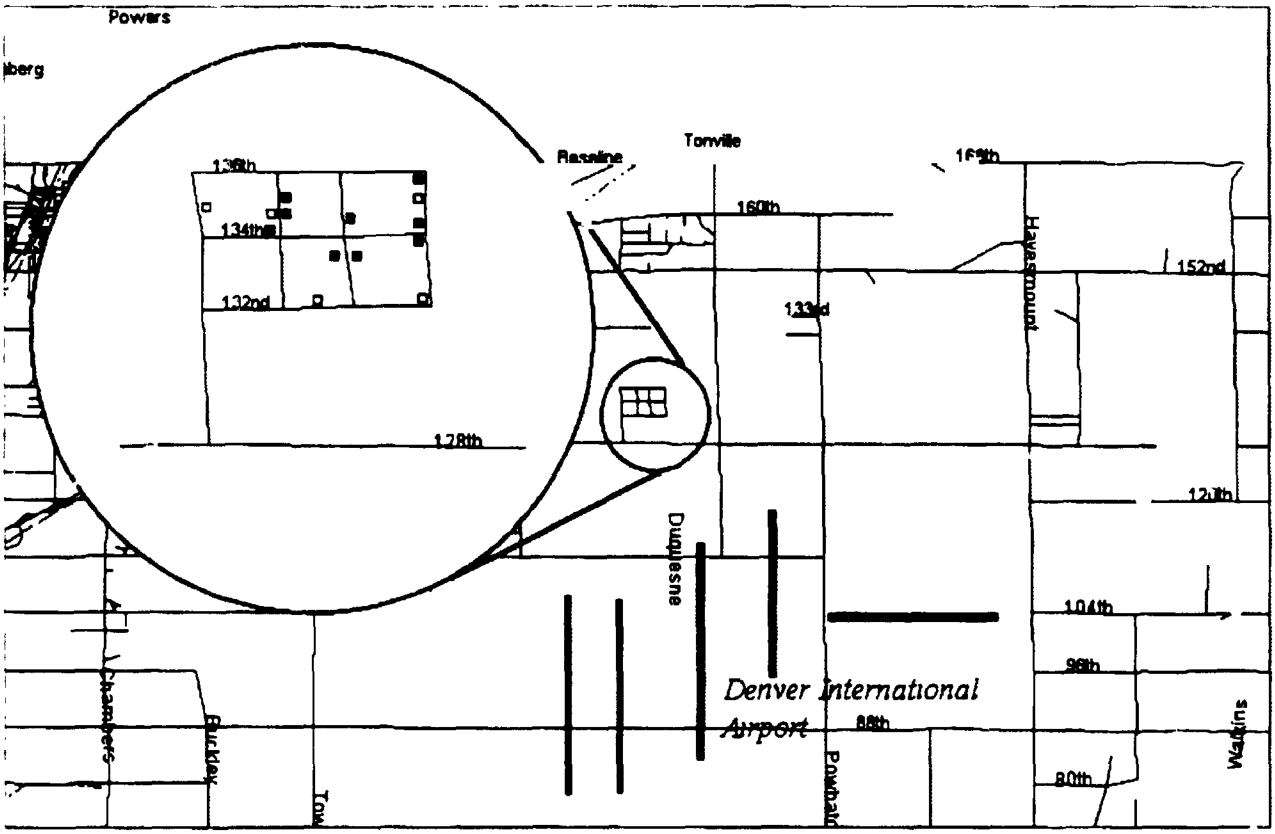


Figure 12 Map of study area near DIA. Squares indicate participants' homes. Open squares denote outdoor noise monitoring sites. Noise monitors were installed in all sleeping quarters.

3.1.3 Third Round of Data Collection

A third round of data collection was conducted in the same neighborhood used for the second round of data collection and started approximately 10 days prior to the final announced opening date of DIA, 28 February 1995. To the extent possible, the same people who took part in Round 2 were contacted and agreed to participate again. Several new participants also were chosen. Data were collected from a total of 13 residences and 30 test participants. DIA opened as scheduled and data collection continued for a total of five weeks.

3.1.4 Fourth Round of Data Collection

A fourth round of data collection commenced during the first week of April, 1995 at homes in two neighborhoods near DEN. Many of the participants selected for Round 1 also took part in this round. Several new participants also were chosen. Data were collected from a total of 15 residences and 28 participants for a period of three weeks.

3.2 TEST PARTICIPANTS

Test participants for the first two rounds of data collection were recruited through mailings to residences within address ranges determined by site visits. Address lists were assembled from direct observation of street addresses, reverse telephone directories, and information purchased from commercial re-sellers of public property records. The initial mailing included letters describing the study and a return form for those interested in participation. Follow-up of returned indications of interest was accomplished via telephone. The constraints of the spatial distribution of aircraft noise exposure, presumed self-selection biases of neighborhood residence, and relatively small numbers of eligible households and test participants precluded any efforts to obtain a random sample.

Test participants for Rounds 3 and 4 of data collection were selected from available participants from Rounds 1 and 2. Additional participants were recruited by telephone. An honorarium of $100-$150 was offered for up to five weeks participation in the study. Site visits were made to inspect residences to verify their noise exposure and overall suitability, to make an informal determination of potential test participants' hearing ability, to install equipment, to train test participants, and to schedule equipment maintenance visits. An instruction booklet was provided to prospective test participants prior to equipment installation. Use of the response recording instrumentation was explained and demonstrated at the time of initial installation, and reiterated during service visits. A toll-free telephone number was provided to encourage test participants to ask for clarification of procedures at all times. Appendix I contains further detail about recruitment procedures, instructions to test participants, and questionnaires.

Selection of households in the Denver area was made on the basis of the following overall requirements and preferences:

- approximately equal numbers of men and women;
- at least two people participating in each household;
- a range of ages, from young adult couples to the elderly;

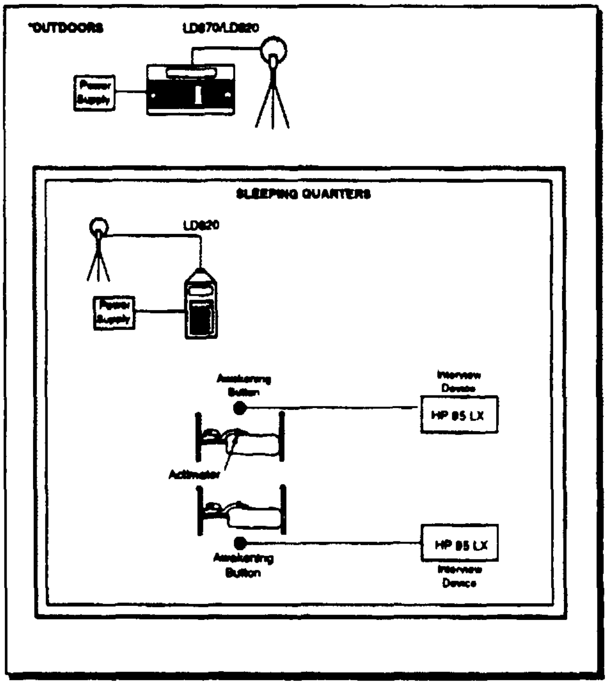
- neighborhood residence for at least 3 months;
- good general health;
- households in which occupants of shared sleeping quarters were likely to be present for all test nights; and
- households with differing ambient noise environments in sleeping quarters.

No formal determination of hearing acuity was made. Potential test participants who were observed to have difficulty using the telephone or understanding other spoken communication were not permitted to take part in the study.

3.3 NOISE MEASUREMENTS

3.3.1 Data Collection Rounds 1 and 2

Instrumentation was assembled to support automated data capture, and processing and analysis of large amounts of noise exposure information, as shown in Figure 13. This instrumentation preserved time synchronization among data streams for time series of A-weighted sound pressure measurements recorded indoors and outdoors, awakening responses from test participants' hand switches, and motility measurements from actimeters worn by test participants.



* Limited use

Figure 13 Schematic diagram of field instrumentation

Indoor noise measurements were made continuously with Larson-Davis 820 noise monitors for the four-week data collection period with microphones placed inside test participants' sleeping quarters. L_{eq} values were recorded every 60 seconds, as were 1-second time histories of noise events. Data captured by these monitors were downloaded approximately once per week, in conjunction with visits to test participants' homes for other purposes.

Outdoor noise measurements were made using five Larson-Davis noise monitors (models 820 and 870) in the vicinity of all test participants' residences using the same parameters used to collect indoor noise data.

3.3.2 Data Collection Rounds 3 and 4

The instrumentation used in Rounds 3 and 4 of data collection was identical to that used in previous rounds. However, the equipment was reprogrammed to record L_{eq} values every 2 seconds. The higher recording rate necessitated shorter downloading intervals. Several noise monitors that operated continuously were connected to phone lines and remotely downloaded every other day. The remaining noise monitors operated only from 2000 hours to 0800 hours to allow at least one week of operation between downloadings.

3.4 RESPONSE MEASUREMENTS

A palmtop computer (HP-95LX) was provided to each test participant to administer the evening and morning questionnaires. A pushbutton attached by a short cable to the computer served as the behavioral confirmation of awakening during the night. Participants were asked to push this button whenever they woke up for any reason during the night.

Nighttime motility was recorded via actimeters. All 30 test participants in Round 1 living near DEN were provided with the same Swiss-manufactured actimeters employed in the study of Ollerhead *et al.* (1992). In addition, six test participants wore actimeters manufactured by Ambulatory Monitoring, Inc. (AMI). The latter actimeters were also used to measure motility in the homes of six test participants near DIA during the second round of data collection. All test participants in Rounds 3 and 4 were provided with AMI actimeters to measure motility.

Table 4 compares characteristics of the two types of actimeters. In the modes used in this study, the AMI actimeters and those formerly used by Ollerhead *et al.* did not produce directly comparable outputs. The AMI model was capable of operating in two modes. The preferred operating mode for the AMI unit in sleep research (Personal Communication with M. Rosekind of NASA Ames Research Center, 1994) was the "zero crossing" mode, in which the actimeter summed the number of times that a threshold was exceeded during a measurement epoch (30 seconds). The actimeters used by Ollerhead *et al.* recorded only "time above threshold" for each epoch.

Table 4 Comparison of characteristics of two actimeters used in present study.

Characteristic	Gaehwiler (Swiss-made)	AMI (U.S.-made)
Size	51 x 37 x 21 mm	38 x 33 x 10 mm
Weight	68 grams	57 grams
Sensitivity	Fixed (0.1 g)	Adjustable (0.01 or 0.5 g)
Bandwidth	0.25 - 3 Hz	2 - 3 Hz, adjustable over the range of 0.16 - 10 Hz
Endurance	10 days (storage-limited)	10 days (battery-limited) - model used in data collection Rounds 1 and 2 10 days (storage-limited) - model used in data collection Rounds 3 and 4
Metrics of Motility	Time above threshold only	Number of zero-crossings within analysis epoch; Time above threshold

3.5 DATA REDUCTION PROCEDURES

Methods used to extract information from field records are described in this section. Appendix J contains a detailed description of software and procedures.

3.5.1 Quality Control Measures

All data collected in the field were carefully screened for quality control purposes. Data collected using the palmtop computers were checked for file-formatting errors and data falling outside permissible range values. A missing final awakening response that test participants were supposed to make before leaving bed in the morning was the most common fault. Any nights that contained uninterpretable and irreparable data were excluded from further analysis.

All actimetric data were likewise checked for acceptability. Any actimetry data files that contained suspect data or no substantive information were excluded from further consideration. Noise data collected from the Larson-Davis noise monitors were checked to ensure that the data appeared reasonable and contained no indication of equipment malfunctions. Any nights for which noise data were unavailable were noted, and no further analyses of these nights were conducted.

3.5.2 Data Processing

All data files were processed by BBN/Probe time series analysis software, as described in Appendix J. Each actimeter data file was converted to an appropriate format, noise files were converted either to ASCII files or into binary form, and all palmtop computer data files (containing sleep/wake times, interview data and button push data) were combined.

3.5.3 Definition of Aircraft Noise Events

Noise measurements in the current study were made both outside and inside participants' homes. During data collection Rounds 1 and 2, an outdoor noise event was considered to have occurred when the noise level exceeded 70 dB for at least two seconds. During data collection Rounds 3 and 4, an outdoor noise event was defined when the noise level exceeded 60 dB for at least two seconds.

During data collection Rounds 1 and 2, an indoor noise event was defined when the noise level exceeded 60 dB for at least two seconds. During data collection Rounds 3 and 4, an indoor noise event was defined when the noise level exceeded 50 dB for at least two seconds. No attempt was made to eliminate noise events from sources other than aircraft except in Round 3, when availability of information about the times of occurrence of aircraft operations permitted separate analyses of confirmed overflights.

3.5.4 Data Extraction Procedures

A summary plot of the sort shown in Figure 14 was automatically prepared and displayed during data reduction to evaluate indoor noise levels (1-minute L_{eq} values) and event levels, outdoor noise levels and event levels, unprocessed actimeter data, blips, and any behavioral awakenings. Options were presented for saving only indoor data, only outdoor data, or both. Once suitable data were selected, the set of variables noted above was written to a file for later combination with all other participants' data for inferential analyses.

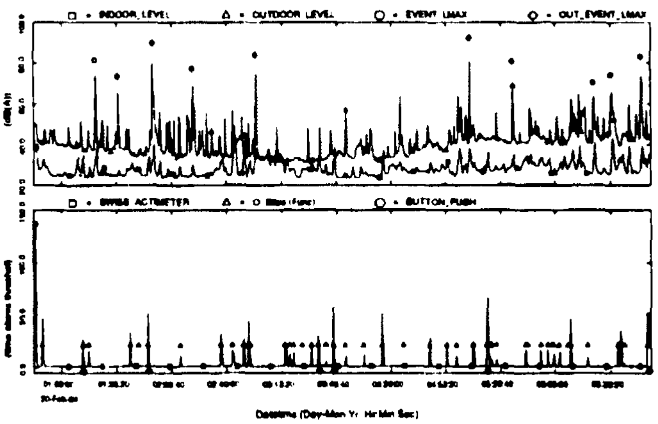


Figure 14 Example of display used to evaluate suitability of data for current analyses.

4 RESULTS

This chapter describes findings of analyses of acoustic measurements (indoor, outdoor, and aircraft only), motility measurements, behavioral awakenings, self-reports of sleep disturbance, and relationships among them.

4.1 OVERVIEW OF DATA COLLECTION AND ANALYSES

Table 5 summarizes the completed field data collection schedule at each site. Table 6 is a guide to the analyses performed on these data. Simple descriptive accounts of findings are presented in this section. More detailed inferential analyses begin in Section 4.4.

Table 5 Summary of data collection conditions.

Site	Data Collection Round	Number of Data Collection Sites	Number of Test Participants	Number of Subject-Nights of Data Collection	Actimeter Type
DEN	Before closure (February/March 1994)	15	30	677	30 Swiss-made 6 U.S.-made
DIA	Before opening (April/May 1994)	14	29	712	6 U.S.-made
DIA	Spanning opening (February/March 1995)	13	30	848	30 U.S.-made
DEN	After closure (April 1995)	15	28	480	28 U.S.-made
	TOTAL	**57** (37 different homes)	**117** (77 different people)	**2717**	

4.2 DESCRIPTION OF INDOOR AND OUTDOOR NOISE ENVIRONMENTS

Tables 20 through 23 in Appendix D summarize the indoor and outdoor noise environments at each site during each data collection period. Although considerable variability was observed in numbers of outdoor noise events during different time periods at the various sites, variability in indoor noise event levels was considerably smaller. Figures 15 and 16 illustrate the distribution of noise events recorded indoors and outdoors at DEN before and after closure of the airport. Figures 17 and 18 illustrate the distribution of noise events recorded indoors and outdoors at DIA before and after opening of the airport.

Table 6 Guide to analyses performed in this study.

ANALYSIS	DATA SET	RESULTS SECTION
DESCRIPTIVE ANALYSES		
Indoor and Outdoor Noise Environments	DEN before airport closure	4 2.1
Indoor and Outdoor Noise Environments	DEN after airport closure	4 2.2
Indoor and Outdoor Noise Environments	DIA before airport opening	4.2 3
Indoor and Outdoor Noise Environments	DIA after airport opening	4.2.4
Behavioral awakening responses, motility, and self-reported awakenings	DEN before airport closure	4.3 1
Behavioral awakening responses, motility, and self-reported awakenings	DEN after airport closure	4 3.2
Behavioral awakening responses, motility, and self-reported awakenings	DIA before airport opening	4 3 3
Behavioral awakening responses, motility, and self-reported awakenings	DIA after airport opening	4 3 4
INFERENTIAL ANALYSES		
Dosage-response analysis	All data	4 4.1
Temporal adaptation of behavioral awakening responses at DIA	DIA before and after airport opening (1995)	4.4.1
Temporal adaptation of behavioral awakening responses and recalled awakenings	DIA before and after airport opening (1995)	4 4.2.1
Temporal adaptation of behavioral awakening responses and recalled awakenings	DIA one year before airport opening (1994)	4 4.2 1
Temporal adaptation of behavioral awakening responses and recalled awakenings	DEN before and after airport closure	4 4 2 2
Indoor Leq before and after airport opening	DIA before and after opening (1995)	4 4 2 1
Outdoor Leq before and after airport opening	DIA before and after airport opening (1995)	4 4.2 1
Analysis of variance on indoor and outdoor Leq	DEN before and after airport closure (1995)	4.4 2 2
Multiway frequency analysis of awakenings and arousals as defined by three criteria	DEN before airport closure (1994)	4 4 3.1
Relationship between motility and behavioral awakening responses	DEN before airport closure (1994)	4 4 3 2
Relationship between initial sleep latency and time spent awake	DEN before airport closure (1994)	4 4 4
Relationship between behavioral awakening responses and recalled awakenings	All data	4 4 5
Prediction of Swiss-made actimeter measured motility from noise event levels and control variables	DEN before airport closure (1994)	4 4 6 1
Prediction of U S -made actimeter measured motility from noise event levels and control variables	All data from participants using AMI actimeters	4 4 6 2
Prediction of behavioral awakening responses from noise levels and control variables	All data	4 4.6 3
Prediction of U S -made actimetric arousals as defined by Cole's criterion	All data from participants using AMI actimeters	4 4 6 4
Attempted replication of Ollerhead's (1992) analysis	DEN before airport closure	Appendix F

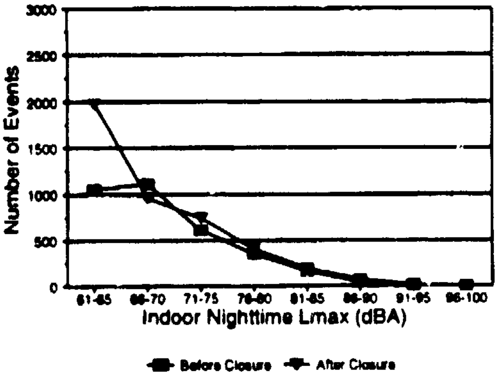


Figure 15 Distribution of noise events recorded inside test participants' sleeping quarters at DEN from 2200 to 0700 hours.

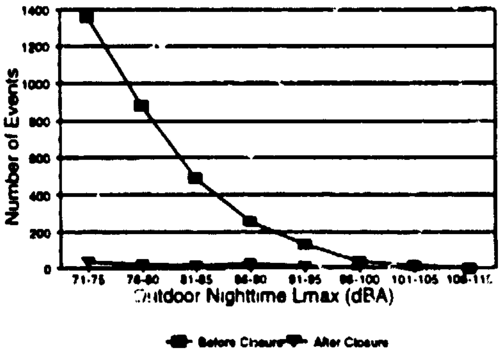


Figure 16 Distribution of outdoor noise events at DEN between 2200 and 0700 hours.

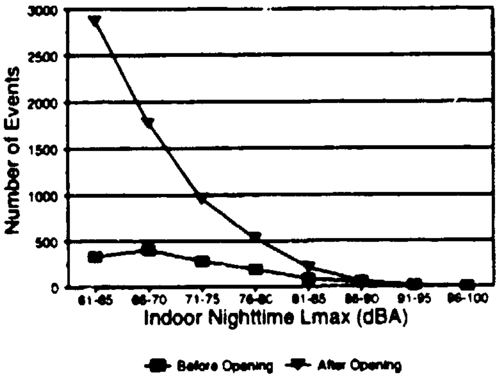


Figure 17 Distribution of noise events recorded inside test participants' sleeping quarters at DIA from 2200 to 0700 hours.

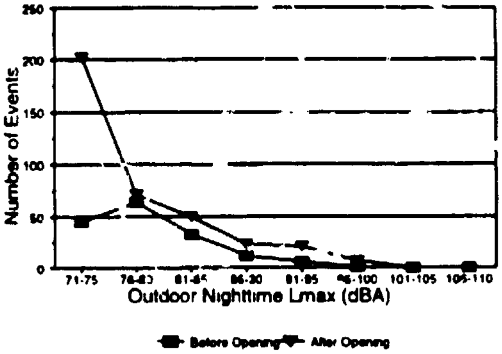


Figure 18 Distribution of outdoor noise events at DIA between 2200 and 0700 hours.

4.2.1 Noise Environment at DEN Before Closure of Airport

Table 20 in Appendix D summarizes the noise environment at DEN before closure of the airport. The number of noise events as defined in Section 2.2.1 recorded inside participants' sleeping quarters ranged from 130 to 7,500 at 15 different monitoring sites, with a total of 48,397. A total of 47,814 noise events were recorded outdoors, ranging from 3,316 to 12,635 at five different monitoring sites. Mean values of indoor noise event L_{max} ranged from 66.2 to 74.1 dB among the various test participants' homes. Outdoor L_{max} values ranged from 78.0 to 82.1 dB.

4.2.2 Noise Environment at DEN After Closure of Airport

Table 21 in Appendix D summarizes the noise environment at DEN after closure of the airport. The number of noise events as defined in Section 2.2.1 recorded inside participants' sleeping quarters ranged from 479 to 18,129 at 15 different monitoring sites, with a total of 72,701. A total of 20,826 noise events were recorded outdoors, ranging from 1,645 to 15,542 at three different monitoring sites. Mean values of indoor noise event L_{max} ranged from 57.0 to 71.9 dB among the various test participants' homes. Average outdoor L_{max} values ranged from 58.3 to 78.3 dB.

4.2.3 Noise Environment at DIA Before Opening of Airport

Table 22 in Appendix D summarizes the noise environment at DIA before opening of the airport. The number of noise events as defined in Section 2.2.1 recorded inside participants' sleeping quarters ranged from 58 to 2,461 at 14 different monitoring sites, with a total of 11,792. A total of 6,220 noise events were recorded outdoors, ranging from 669 to 2,851 at five different monitoring sites. Mean values of indoor noise event L_{max} ranged from 67.6 to 82.5 dB among the various test participants' homes. Average outdoor L_{max} values ranged from 75.0 to 84.2 dB.

4.2.4 Noise Environment at DIA After Opening of Airport

Table 23 in Appendix D summarizes the noise environment at DIA after the opening of the airport. The number of noise events as defined in Section 2.2.1 recorded inside participants' sleeping quarters ranged from 359 to 10,308 at 13 different monitoring sites, with a total of 47,952. A total of 15,155 noise events were recorded outdoors, ranging from 3,484 to 5,885 at four different monitoring sites. Mean values of indoor noise event L_{max} ranged from 59.9 to 73.5 dB among various test participants' homes. Average outdoor L_{max} values ranged from 58.0 to 68.8 dB.

4.3 DESCRIPTION OF SLEEP DISTURBANCE OBSERVATIONS

Figure 19 compares the average rate of behavioral awakening responses in the presence of aircraft with the average rate of behavioral awakening responses in the absence of aircraft for individual test subjects. Figure 20 compares the average motility in the presence of aircraft with the average motility in the absence of aircraft. The pattern of findings summarized in these figures indicates that awakenings and motility were little affected by the changes in aircraft flight operations.

Table 7 summarizes the number of awakenings confirmed by button pushes averaged over the two sites. (Few of the responses were associated with noise events.)

4.3.1 Observations at DEN Before Airport Closure

Analyzable data were collected from 28 test participants living in 15 homes near DEN in February/March of 1994.[1] Twelve pairs of these participants shared sleeping quarters (and hence noise environments). Six hundred fifteen subject-nights of data were collected from these 28 test participants. A total of 1,234 behavioral awakening responses (button pushes) were logged during this round of data collection, for an average of 2 per night.

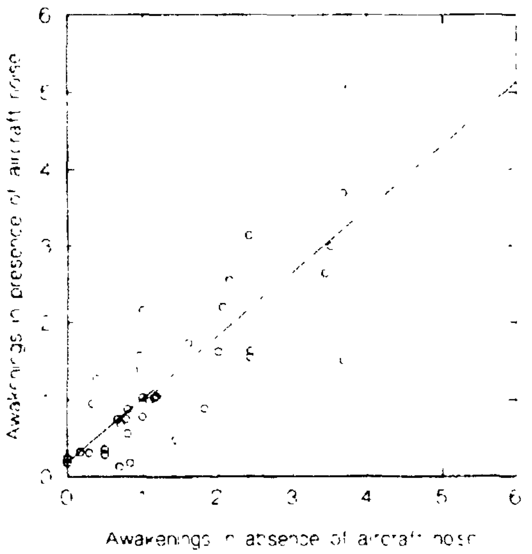


Figure 19 Average number of behavioral awakening responses per night in the presence (ordinate) and absence (abscissa) of aircraft noise. Each data point represents responses of a single participant.

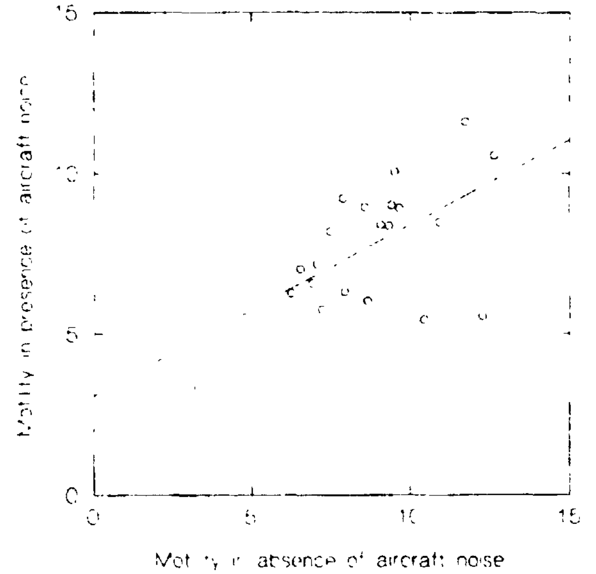


Figure 20 Average motility in each 30-second epoch in the presence (ordinate) and absence (abscissa) of aircraft noise. Each data point represents responses of a single participant aggregated over the entire study.

Figures 32 through 37 (located in Appendix M) summarize the responses to the nighttime and morning questionnaires. Participants' answers to questionnaire items in this data set indicated that they felt very or extremely tired during the day 18% of the time. Self-reports of number of awakenings averaged 2.3 times per night. Responses on 82% of the nights indicated that test participants fell asleep within 20 minutes of retiring. For 66% of the nights, participants recalled being awake less than 20 minutes during the night. About 52% of the responses from the morning questionnaire indicated that no aircraft were heard during the previous night. Reports of high annoyance due to nighttime noise were made on about 3% of the subject-nights.

---

[1] Numbers of participants may not sum to the totals seen in Table 5 because data that did not meet quality control standards were omitted from this and further analyses.

Table 7 Summary of behavioral awakening responses for all subject-nights at DEN and DIA.

	Variable		
	Average number of behaviorally-confirmed awakenings per night	Average number of spontaneous awakenings per night	Average number of noise-related awakenings per night
In presence of aircraft			
Mean	1.53	1.27	0.26
Standard deviation	1.99	1.80	0.59
Range	0 - 23	0 - 21	0 - 4
In absence of aircraft			
Mean	1.69	1.49	0.19
Standard deviation	2.12	1.98	0.56
Range	0 - 19	0 - 16	0 - 5
Averaged over all nights			
Mean	1.61	1.39	0.22
Standard deviation	2.06	1.90	0.57
Range	0 - 23	0 - 21	0 - 5

Individual 30-second actimeter epochs were analyzed only for this data set for purposes of comparison with the findings of Ollerhead *et al.* (1992). The great majority of these analysis epochs were unaccompanied by noise events. Figures 21 and 22 show the percent of noise event epochs in which actimetric "blips" (as defined by Ollerhead *et al.*, 1992) occurred during three time periods (0100-0130 hours, 0300-0330 hours, and 0500-0530 hours) throughout the night. These periods were chosen to facilitate direct comparisons with the findings of Ollerhead *et al.* Arousal rates related to outdoor noise events during these three time periods (Figure 21) ranged from 5% for noise events between 80 and 84 dB to 11% for noise events above 90 dB. Arousal rates related to indoor (Figure 22) noise events ranged from 17% for events between 65 and 69 dB to 31% for events between 70 and 74 dB. No clear trend is apparent in the relationship between noise levels and arousals.

4.3.2 Observations at DEN After Airport Closure

Analyzable data were collected from 28 participants living in 15 homes near DEN during this round of observations. Seventeen of these participants had also participated in data collection prior to closure of DEN. Eleven pairs of participants shared sleeping quarters. Seven hundred fifty-one behavioral awakening responses were logged during 457 subject-nights, for an average of 1.64 per night.

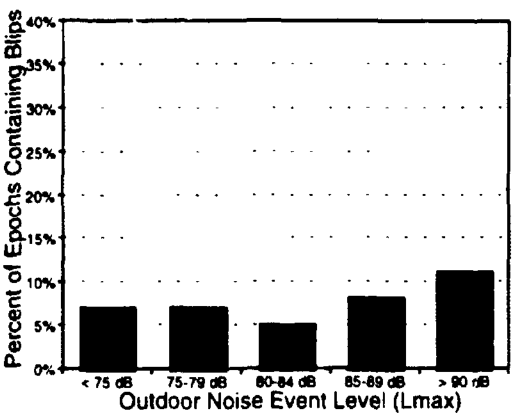


Figure 21 Arousal rate by outdoor noise events within noise level categories.

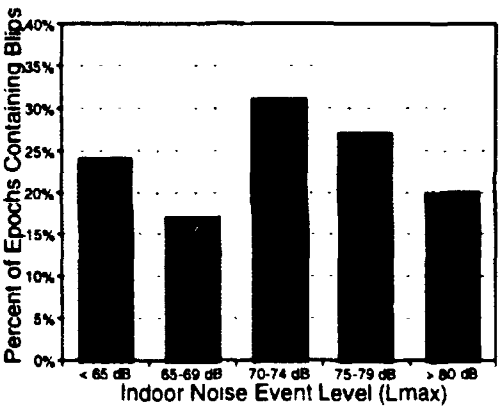


Figure 22 Arousal rate by indoor noise events within noise level categories.

Figures 32 through 37 (located in Appendix M) summarize the responses to the nighttime and morning questionnaires. Participants' answers to questionnaire items for the entire set of measurement nights indicated that they felt very or extremely tired during the day 19% of the time. Self-reports of number of awakenings averaged 1.8 times per night. Responses on 77% of the nights indicated that test participants fell asleep within 20 minutes of retiring. For 72% of the nights, participants recalled being awake less than 20 minutes during the night. Reports of high annoyance due to nighttime noise were made on about 2% of the subject-nights.

4.3.3 Observations at DIA Before Airport Opening

Of the 29 participants in 14 homes at DIA during the second round of data collection (April/May, 1994), thirteen pairs shared sleeping quarters and noise environments. Twenty-three of the participants, including nine pairs of bedmates, also contributed data at the start of the next round of data collection, which included several nights before airport opening. An additional 7 participants, including 2 pairs sharing sleeping quarters, participated only in the second round of data collection at DIA. Fifteen hundred and two behavioral awakening responses were logged during 880 subject-nights, for an average of 1.71 per night.

Figures 32 through 37 (located in Appendix M) summarize the responses to the nighttime and morning questionnaires. Data are presented separately for 1994 and 1995 collection rounds. Participants' answers to questionnaire items for the entire set of measurement nights indicated that they felt very or extremely tired during the day 26% of the time in 1994 and 22% of the time in 1995. Self-reports of number of awakenings averaged 1.8 times per night in 1994 and 2.5 times per night in 1995. Responses on 81% of the nights in 1994 and 82% of the nights in 1995 indicated that test participants fell asleep within 20 minutes of retiring. For 74% of the nights in 1994 and 78% of the nights in 1995, participants recalled being awake less than 20 minutes during the night. Reports of high annoyance due to nighttime noise were made on about 4.5% of the subject-nights in 1994 and about 1% of the subject nights in 1995.

4.3.4 Observations at DIA After Airport Opening

Thirty participants, including 11 pairs sharing sleeping quarters, participated in the third round of data collection at DIA in February/March of 1995. Six hundred forty-one behavioral awakening responses were logged in the 565 nights of data collection, for an average of 1.13 per night.

Figures 32 through 37 (located in Appendix E) summarize the responses to the nighttime and morning questionnaires. Participants' answers to questionnaire items for the entire set of measurement nights indicated that they felt very or extremely tired during the day 23% of the time. Self-reports of number of awakenings averaged 1.4 times per night. Responses on 84% of the nights indicated that test participants fell asleep within 20 minutes of retiring. For 83% of the nights, participants recalled being awake less than 20 minutes during the night. Reports of high annoyance due to nighttime noise were made on about 2% of the subject-nights.

4.4 INFERENTIAL ANALYSES

Pre-planned analyses were conducted on four data sets developed for each round of observations at each airport, as shown in Table 8:

1) whole nights (each case is a single subject-night of data);
2) behavioral awakening responses (each case is an individual button push);
3) subject-specific noise events (each case is a noise event occurring during the sleep times of a single test participant between 2200 and 0700 hours, defined by either indoor or outdoor criteria); and
4) subject-specific confirmed aircraft noise events (each case is an aircraft noise event occurring during the sleep times of individual test participants between 2200 and 0700 hours).

The second round of data collection at DIA spanned the transfer of flight operations from DEN to DIA. This round was partitioned for analytic purposes into data collected before operations began at DIA (combined with the first round of DIA data collection) and data collected after the start of flight operations at DIA.

Table 8 Description of data sets analyzed.

		Number of Cases			
		DEN		DIA	
DATA FILE	DEFINITION OF CASE	Before closure	After closure	Before opening	After opening
Whole night	Single subject-night of data	615	457	880	565
Behavioral awakening responses	Single behavioral awakening as defined by a button push	1234	751	1502	641
Subject-Specific Noise Events	Noise event exceeding fixed indoor or outdoor noise level thresholds occurring during individual test participant's sleep times between 2200 and 0700 hours	9660	8774	3625	9473
Monitored indoors		1664	3527	1318	2114
Monitored outdoors		7996	5247	2307	7359
Confirmed aircraft	Noise event created by aircraft producing a radar flight track occurring during individual test participant's sleep times between 2200 and 0700 hours	(no data)		390	7864

4.4.1 Dosage-Response Relationships

All dosage-response relationships were restricted to noise event data collected between 2200 and 0700 hours, since earlier time periods in the evening and later time periods in the morning contained too high a density of noise events for reliable association with individual responses. Dosage-response relationships were constructed for five indicators of sleep disturbance:

1) behavioral awakening responses (button pushes),
2) arousals defined by Ollerhead *et al.* (1992) criteria for the Swiss-made actimetric data,
3) arousals defined by Cole *et al.* (1992) criteria for the U.S.-made actimetric data,
4) motility as recorded by the Swiss-made actimeters, and
5) motility as recorded by U.S.-made actimeters.

Analyses related to confirmed aircraft noise events at DIA were possible only for the month prior to opening and the month following the opening of the airport.

The independent (predictor) variable for all dosage-response relationships was either indoor or outdoor SEL, quantized in 3-dB intervals. Data points reflect the proportion of noise events in each noise level interval that produced a response. Data were combined for all test participants and all data collection sessions for behavioral awakening and U.S.-made actimeter responses. Swiss-made actimeter recordings were available only at DEN for the data collection session before airport closure. Table 9 shows the definitions of awakening, arousal, and motility adopted for the various data collection devices.

Table 9 Definitions of awakening and motility adopted for various data collection devices.

INDICATION OF SLEEP DISTURBANCE	RECORDING DEVICE	CRITERION OF EFFECT
Awakening	Push button	Occurrence of response within five minutes of start of noise event
Arousal	Swiss-made actimeter	Identical to that of Ollerhead *et al.* (1992)
Arousal	U.S.-made actimeter	As defined by Cole *et al.* (1992), using base algorithm without iteration
Motility	Swiss-made actimeter	Any activity occurring in any of the ten 30 second epochs after the start of a noise event
Motility	U.S.-made actimeter	Any activity occurring in any of the ten 30 second epochs after the start of a noise event

One-sided analyses of significance of associations of sleep disturbance and noise events were tested at $\alpha = .025$. This seemingly lax criterion was adopted because of the relatively low power associated with the sample sizes generated by the 3-dB wide SEL categories--in the neighborhood of $N = 10$ to 14. At this level of significance, a correlation coefficient of about .60 is required for statistical reliability. Any 3-dB interval containing fewer than 10 noise events was excluded from analysis.

Correlations for the various dosage-response relationships are summarized in Table 10. Four of the dosage response relationships, all based on SEL of noise events measured indoors, were statistically reliable. The SEL value of indoor noise events successfully predicted (1) behavioral awakening responses, (2) motility as recorded by the Swiss-made actimeters, (3) motility as recorded by the U.S.-made actimeters, and (4) U.S.-made actimetric arousals as defined by Cole *et al.* (1992). None of the sleep disturbance measures varied reliably with SEL of noise events measured outdoors, nor did they vary reliably with SEL of confirmed aircraft noise events only.

Table 10 Summary of dosage-response correlations for events occurring between 2200 and 0700 hours. (Data aggregated over DEN and DIA for button push responses and U.S.-made actimeter. Data available only at DEN for Swiss-made actimeter).

Measure of Sleep Disturbance	Criterion for Sleep Disturbance	Number of Indoor Noise Events	Number of Outdoor Noise Events	Noise Measurement Type		
				Indoor Criterion	Outdoor Criterion	Outdoor Confirmed Aircraft
Motility	Swiss-made actimeter (time above threshold)	1519	6915	.90*	ns	nd
	U.S.-made actimeter (zero-crossings)	466	1535	.84*	ns	ns
Arousal	Ollerhead (Swiss-made actimeter)	1519	6915	ns	ns	nd
	Cole (U.S.-made actimeter)	466	1535	.62*	ns	ns
Awakening	Behavioral awakening response	2169	8572	.68*	ns	ns

*$p < .025$, one-sided test ns: not significantly different from a correlation of 0 nd: no data

Figure 23 shows that the probability of occurrence of at least one actimetric response recorded by a Swiss-made actimeter within five minutes of the start of a noise event was strongly related to indoor SEL, $r(9)=.90$, $p < .001$. The data set in which this relationship was observed was composed of noise events recorded for the participants at DEN before airport closure. The slope of the regression equation shown in Figure 23 is fairly shallow: each 1.0 dB increase in SEL raised the probability of an actimetric blip by about 1.23%. Polynomial regression revealed no significant higher order (quadratic or cubic) relationships.

Figure 24 shows that the probability of occurrence of an average number of zero crossings greater than 0 as measured by the U.S.-made actimeter also was reliably related to indoor SEL, $r(9) = .84$, $p < .025$. The data set in which this relationship was observed was based on six participants in the first rounds of data collection at DEN and DIA, and all participants for remaining data collection periods. Each 1 dB increase in SEL raised the probability of occurrence of a motility indication by about 0.4%.

Polynomial regression revealed no higher order relationships. The difference between correlations with SEL for the two actimetric criteria was not statistically reliable.

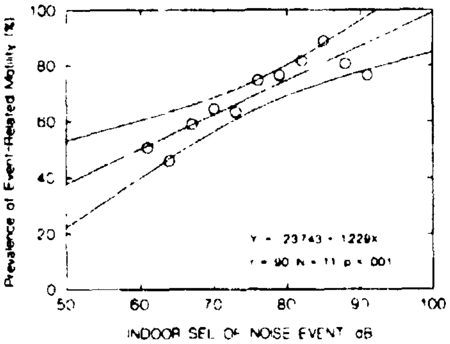


Figure 23 Prevalence of an actimetric response recorded by Swiss-made actimeters at DEN before airport closure, aggregated by test participants in 3 dB intervals of indoor SEL values of noise events. Curved lines bound the 95% confidence interval.

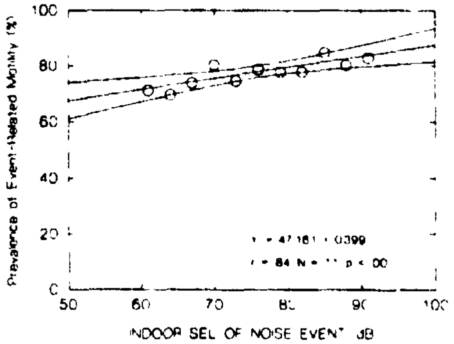


Figure 24 Prevalence of actimetric zero-crossings as recorded by the U.S.-made actimeter at DEN and DIA, aggregated by test participants in 3 dB intervals of indoor SEL values of noise events. Curved lines bound the 95% confidence interval.

Figure 25 shows that the indoor SEL of noise events predicted behavioral awakening responses moderately well, $r(10) = .68, p < .025$. The probability of awakening increased by about 0.25% with each 1 dB increase in SEL. Polynomial regression revealed no statistically reliable higher order relationships.

Arousals as scored by the Cole *et al.* (1992) actimetric criterion also were predicted reasonably well by the indoor SEL of noise events, $r(9) = .62, p < .025$, as shown in Figure 26. Motility measurements were collected from six participants in the first rounds of data collection at DEN and DIA, and from all participants during the remaining data collection periods. The probability of arousal increased about 0.28% with each 1 dB increase in SEL of indoor noise events. Polynomial regression showed no quadratic or cubic relationship. Neither the differences between correlations with indoor noise event SEL for the two indicators of awakening or arousal, nor the differences between correlations for awakening, arousal, and motility, were statistically reliable, $p > .05$.

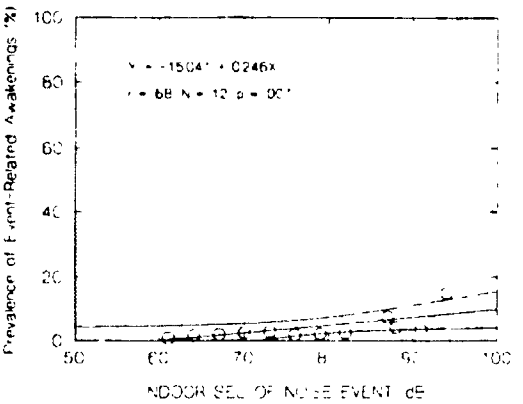


Figure 25 Prevalence of behavioral awakening responses at DEN and DIA aggregated by test participants in 3 dB intervals of indoor noise measurements. Curved lines bound the 95% confidence interval.

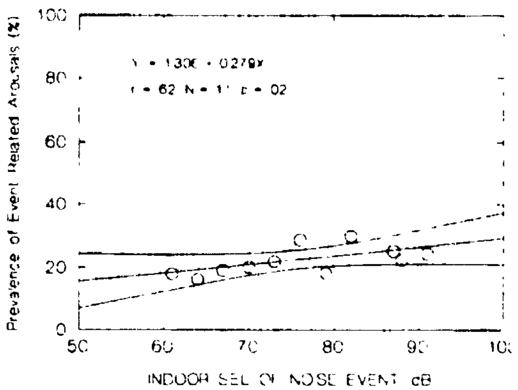


Figure 26 Prevalence of arousals defined by U.S. actimetric criterion (Cole, 1992) at DEN and DIA aggregated by test participants in 3 dB intervals of indoor noise measurements. Curved lines bound the 95% confidence interval.

4.4.2 Temporal Adaptation of Behaviorally-Defined and Self-Reported Sleep Latency

4.4.2.1 Effects Observed at DIA

Twenty-two test participants living near DIA provided analyzable data immediately before and after the start of flight operations at DIA in February, 1995. Analyses of temporal effects in this data set were based on whole night data, ignoring the first 3 nights of data collection as a period of familiarization with the in-home instrumentation. The entire data collection period was divided into five sequential intervals for purposes of this analysis:

1) the 6 or 7 nights of data collection (following the first three nights of data collection) prior to start of flight operations,
2) the first two nights after the start of flight operations,
3) the third through fifth nights after start of flight operations,
4) the sixth through eighth nights after start of flight operations, and
5) the remaining 4 to 19 nights following the eighth night.

A profile analysis of repeated measures of behavioral awakening responses and recalled time to fall asleep, after adjustment for total time slept as a covariate, showed a significant relationship with time period, multivariate $F(8, 164) = 2.92, p < .05$. Sequential interval accounted for 12% of the variability in the combination of sleep disruption measures. A stepdown analysis was performed on the two sleep disruption variables, in which the test for time to fall asleep was adjusted for awakenings by behavioral awakening responses as well as time slept, but the test for behavioral awakening responses was adjusted only for time slept. That is, behavioral awakening responses were assigned greater importance as a sleep disruption measure than was recalled latency to fall asleep. Each test was done with a probability of Type I (α) error set at .025.

The sequential data collection interval was related only to behavioral awakening responses, not recalled latency to fall asleep, after adjustment for sleep time, $F(4, 83) = 3.40, p < .025, \eta^2 = .14$. Figure 27 shows the average number of behavioral awakening responses as a function of time period, as well as the indoor and outdoor L_{eq} for those time periods. A planned contrast revealed no significant difference in average number of behavioral awakening responses before and after start of flight operations, $p > .025$. However, a planned trend analyses revealed that the apparent negative linear trend of Figure 27 was statistically significant, $F(1, 20) = 9.18, p < .025$.

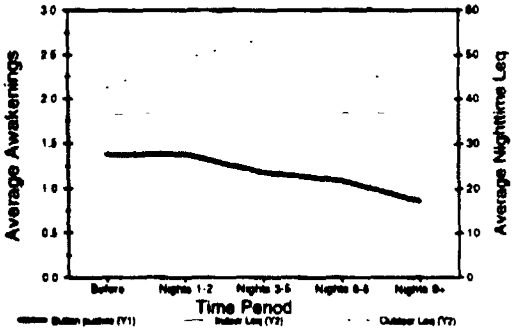


Figure 27 Behavioral awakening responses, indoor L_{eq}, and outdoor L_{eq} during intervals following start of flight operations at DIA.

A parallel analysis was performed on the data collected one year prior to start of flight operations at DIA, using data only from those who participated in both data collection rounds. A statistically reliable negative quadratic trend for behavioral awakening responses, after adjustment for sleep time, was also observed for these data, $F(1, 14) = 7.85, p < .025$. Figure 28 shows that the average number of behavioral awakening responses over the course of data collection in April and May, 1994 (during which time there were no flight operations) was similar to that over the course of data collection in February and March of 1995 (which included periods with and without flight operations). Note also the higher average number of button-push awakenings in 1994 than 1995.

Not all of the participants awakened less often over the course of data collection in 1995. Indeed, as seen in Figure 31 in Appendix K, a few (*e.g.*, participants 612, 620, and 625) exhibited a rise in the number of behavioral-awakening responses in the two nights following start of flight operations. However, these participants promptly returned to their previous level of awakenings or below.

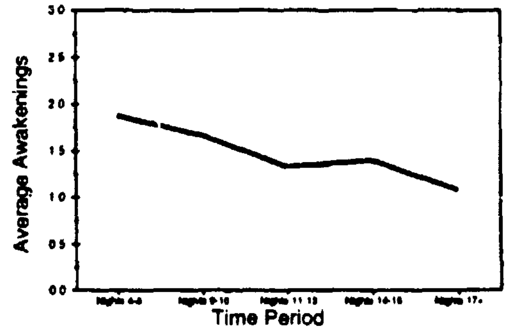


Figure 28 Behavioral awakening responses during intervals one year prior to start of flight operations at DIA.

The apparent quadratic trend of indoor L_{eq} over the 5 time periods in 1995 (*cf.* Figure 27) was statistically reliable, $F(1, 21) = 7.60, p < .025$. No reliable linear trend was found, however. Indoor L_{eq} before start of flight operations (mean = 36 dB) was significantly less than after start of flight operations (mean = 38 dB), $F(1, 21) = 7.48, p < .025$. A stronger relationship was found between time period and outdoor L_{eq}. The quadratic trend was statistically reliable, $F(1, 21) = 115.19$. Outdoor L_{eq} was significantly less before start of flight operations (mean = 43 dB) than after (mean = 48 dB), $F(1, 21) = 98.58, p < .025$.

A similar profile analysis of repeated measures was performed on motility as measured by the U.S.-made actimeter and time spent awake during the night. Data for this analysis were provided by only 14 of the participants. No relationship was found between time period and either of these two measures of sleep disturbance.

4.4.2.2 Effects Observed at DEN

Data were collected during non-adjacent time periods from participants in the flight path of DEN before and after cessation of flight operations: February and March, 1994 in the presence of flight operations, and April 1995 after flight operations had transferred to DIA. Sixteen participants provided usable data for both time periods.

A multivariate analysis of covariance of behavioral awakening responses awakenings and latency to fall asleep, after adjusting for time slept, showed no reliable difference between time periods, $p > .05$. Too few participants provided data to analyze motility or time awake during the night.

Analyses of variance revealed that outdoor, but not indoor, L_{eq} varied as a function of time period, $F(1, 15) = 134.16, p < .025$. Outdoor noise level dropped from an average of 58 dB for the 16 participants during flight operations in 1994 to 46 dB after cessation of operations in 1995.

4.4.3 Relationships Among Behavioral Awakenings and Motility

4.4.3.1 Awakening and Arousal

A multiway frequency analysis explored the relationships among three criteria for awakening and arousal: behavioral awakenings, motility using Ollerhead's (1992) criterion, and motility using Cole's (1992) criterion. This analysis was limited to the six participants at DEN who wore both Swiss- and U.S.-made actimeters, providing a total of 1,271 noise events. Noise events measured outdoors were chosen for analysis because of the greater number of outdoor events. Table 11 shows the distribution of noise events. The small cell sizes yielded expected frequencies that were too small to provide adequate power for analysis.

Table 11 Distribution of outdoor events producing awakenings or arousals by three criteria: behavioral awakening responses and Swiss and U.S.-made actimetrically-defined arousals.

U.S. Actimeter Arousal	Swiss Actimeter Arousal	Button Push	
		Present	Absent
Yes	Yes	0	691
	No	9	122
No	Yes	2	39
	No	5	103

All of the two-way, but not the three-way, associations were statistically reliable at $\alpha = .0125$ using the test for marginal association.[2] The relationship between awakening by behavioral awakening responses and Ollerhead's criterion for arousal, $\chi^2(1) = 14.34$, $\phi^2 = .01$, indicated that 88% of the noise events producing behavioral awakening responses also triggered an arousal by the Ollerhead criterion, but 42% of the noise events without behavioral awakening responses also constituted Ollerhead-defined arousals. The relationship between behavioral awakening responses and the U.S.-made actimeter criterion for arousals, $\chi^2(1) = 10.52$, $\phi^2 = .01$, indicated that almost half (44%) of the noise events producing behavioral awakening responses also produced arousals by the U.S.-made actimeter criterion, but only 11% of the noise events without behavioral awakening responses produced enough response in the U.S.-made actimeter to be recorded as an arousal. (The strength of the relationship between behavioral awakening responses and arousals by Cole's criterion for the U.S.-made actimeter did not change in a reanalysis that added data collected in later rounds.)

The relationship between arousals as defined by Ollerhead's criterion and as defined by Cole's criterion, $\chi^2(1) = 62.62$, $\phi^2 = .05$, showed greater sensitivity of the Ollerhead *et al.* algorithm. Almost three-quarters (72%) of the U.S.-made actimeter-recorded arousals also produced an arousal by the U.S. actimetric criterion; however; 38% of the noise events that failed to produce an arousal by the U.S. actimetric criterion were followed by an Ollerhead-defined arousal.

4.4.3.2 Motility and Behavioral Awakening Responses

Relationships between behavioral awakening responses and motility were also explored through multiway frequency analysis of the 1,271 noise events recorded outdoors in which participants wore both

---

[2] The less conservative test was used because of the low power created by small expected frequencies in some cells.

kinds of actimeters. Table 12 shows the distribution of noise events. The cutoff criterion used for presence of U.S.-made actimeter motility was an average of 10 threshold crossings per 30-s interval in the 5-m period following an event.

Table 12 Distribution of outdoor events producing sleep disturbance by three criteria: behavioral awakening responses, and Swiss and U.S.-made actimetrically-defined motility.

U.S. Actimeter Motility	Swiss Actimeter Motility	Button Push	
		Present	Absent
Yes	Yes	10	220
	No	0	18
No	Yes	6	375
	No	0	642

Motility recorded by both the Swiss- and U.S.-made actimeters reliably predicted awakenings. The relationship between behavioral awakening responses and at least one Swiss actimetric threshold crossing, $\chi^2(1) = 23.66$, $\phi^2 = .02$, indicated that while all of the noise events that produced a button push also were followed by a response recorded by the Swiss-made actimeter, almost half of the noise events (47%) that did not trigger a button push were also followed by an actimetric response. The relationship between button push awakenings and motility recorded by the U.S.-made actimeter, $\chi^2(1) = 14.36$, $\phi^2 = .01$, indicated that 63% of the behavioral awakening responses were accompanied by U.S.-made actimeter-recorded motility, and 19% of the noise events without a button push also were followed by movement according to the U.S.-made actimeter.

A stronger association was noted between movement as recorded by the two actimeters, $\chi^2(1) = 280.15$, $\phi^2 = .19$. Most of the noise events (93%) triggering sufficient zero-crossings on the U.S.-made actimeter also produced a threshold crossing on the Swiss device, whereas 37% of the noise events that failed to elicit a response by the U.S.-made actimeter produced a response on the Swiss instrument.

4.4.4 Initial Sleep Latency and Time Spent Awake

The data permitted identification of two sets of awakening: those for which a noise event occurred in the preceding 5 minutes and those that were not preceded by a noise event. Latency was defined as the time between the behavioral awakening response and sleep as determined by Ollerhead's criterion (*i.e.*, 7 minutes with no actimetric response) for the Swiss-made actimeter. There were 962 behavioral awakening responses for which latency was recorded, collected at DEN before airport closure. Logistic regression was employed because of the large discrepancy between the number of behavioral awakening responses preceded by a noise event (38 indoors and 95 outdoors) and the number not preceded by a noise event (924 indoors and 867 outdoors). Latencies did not differ between behavioral awakening responses that were and were not preceded by a noise event, $p > .05$, whether the noise event was defined by indoor or outdoor criteria. That is, time to fall asleep after awakening was no greater for a noise event-induced awakening than for any other awakening.

The relationship between time to fall asleep upon retiring and recalled time to fall asleep was assessed using the all-night data set. Mean time to fall asleep upon retiring, as determined by the Ollerhead criterion for the Swiss-made actimeter, was about 17 minutes (SD = 13.5 minutes), after deletion of two outlying cases with initial latencies in excess of 2.5 hours. Mean recalled latency was 1.8 (SD = 0.97) on a scale in which 1 represents less than 10 minutes and 2 represents 10 to 20 minutes. Both variables were positively skewed. After applying a log transform to both latency measures, a small but statistically significant relationship was found between recalled and Ollerhead-derived initial latencies, $r(540) = .19, p < .001$.

The relationship between total time awake after a button push and recalled time awake was also assessed using the all-night data. Average time awake during the night, as measured by latency between a button push and return to sleep by Ollerhead's criterion, was 34.08 minutes (SD = 35.93 minutes). Mean recalled time awake was 2.2 on a scale in which 1 reflects less than 10 minutes and 2 represents 10 to 20 minutes. This indicates a strong discrepancy between recalled time awake and that estimated by Ollerhead's criterion. Both measures were positively skewed. After applying a log transform to measured latency and a square root transform to recalled time awake, a moderate relationship was found between the two measures of time spent awake, $r(379) = .51, p < .001$.

No reliable linear relationships were found between time awake during the night (log transformed) and cumulated L_{eq} for the night, as measured either indoors or outdoors, $p > .05$. Polynomial regression analysis revealed no statistically significant quadratic or cubic trends, either, $p > .025$.

4.4.5 Behavioral and Recalled Awakenings

The relationship between behavioral awakening responses and the number of awakenings recalled the following morning was assessed using the data for all nights over the four rounds of data collection. Average number of recalled awakenings was 1.95 (SD = 1.23); average number of behaviorally-confirmed awakenings was 1.61 (SD = 2.06). After applying a logarithmic transform to compensate for positive skewness in both measures, a moderate relationship was found between them, $r(2515) = .51, p < .001$.

4.4.6 Predicting Sleep Disturbance from Noise Level and Control Variables

Direct logistic regression analyses were employed to predict sleep disturbance following indoor noise events from the levels of the noise events, ambient noise levels, personal characteristics of respondents, time-related characteristics, and rating of tiredness the previous evening. Logistic regression is an appropriate analytic tool when the predicted variable represents the probability of an outcome (in this case whether sleep is disturbed) and predictor variables are a mixture of discrete and continuous measures.

Noise events used were those occurring between 2200 and 0700 hours; each event constituted a case for analysis. The four measures of sleep disturbance were those showing statistically significant dosage-response relationships with noise measured indoors: Swiss-made actimeter-recorded motility, U.S.-made actimeter-recorded motility, behavioral awakening responses, and arousals by the U.S. actimetric criterion. For all analyses, Type I error rate was controlled by setting $\alpha = .005$ for each predictor. Contribution of each predictor variable was assessed after controlling for all other predictor variables in direct logistic regression.

Predictors included two sound level measures: SEL of noise events as measured indoors and L_{eq} of ambient level in sleeping quarters. Personal characteristics included gender, the linear effect of years of age, the quadratic effect of age (in which younger and older participants were combined and compared with participants 35-49 years of age), and spontaneous (non-event related) numbers of awakenings for the night in which the event occurred. This latter measure was poorly distributed, so a transform of it was used in analysis, in which the inverse was taken of spontaneous number of awakenings + 1, and then the measure was reflected (*i.e.*, the analyzed measure was 1 minus the inverse) to mimic the direction of the original measure.

Time-related characteristics were time since retiring in 15-m intervals, duration of residence in months, and study duration as indicated by number of nights in the study when the event occurred. A final predictor was a rating of tiredness during the previous day, on a scale of 1-5 in which 1 indicated not at all tired and 5 indicated extremely tired.

Table 13 summarizes the results of the logistic regression analyses. Noise events considered in each analysis were those for which data were available for all 10 predictors and the sleep disturbance measure of interest.

Table 13 Summary of logistic regression analyses of four indicators of sleep disturbance by indoor SEL of individual events and additional predictors.

Characteristic	Measure of Sleep Disturbance			
	Motility (Swiss)	Motility (U.S.)	Behavioral Awakening	Arousal (Cole *et al.*, 1992)
Number of events with disturbance/total events	857/1337	3921/5104	184/7685	1060/5104
Significant predictors (each adjusted for all others)	SEL, Age (quadratic), Gender	All except SEL	SEL, Ambient, Night, Age (linear)	Spontaneous awakenings, Gender, Age (linear and quadratic), Night, Ambient, Tiredness,
Full model (10 predictors)				
Variance accounted for	8%	11%	7%	4%
Prediction success	59%	69%	95%	68%
d′	0.69	0.73	0.71	0.63
SEL alone				
Variance accounted for	3%	< 1%	2%	< 1%
Prediction success	56%	64%	95%	67%
d′	0.62	0.53	0.55	0.53
Average Indoor SEL that did/did not disturb sleep	74 dB/69 dB	65 dB/64 dB	69 dB/66 dB	66 dB/64 dB

4.4.6.1 Prediction of Motility as Recorded by Swiss-made Actimeter

Data for these analyses were provided by the six participants at DEN prior to airport closure who wore Swiss-made actimeters, and who were exposed to a total of 1,337 noise events between 2200 hours and 0700 hours.

The model with all 10 predictor variables worked significantly better than a chance model, $\chi^2(10) = 15.62$, $p < .001$. Three of the 10 predictors were reliably associated (at $\alpha = .005$) with presence of motility (at least one actimetric blip) within 5 minutes of a noise event recorded indoors. Table 24 in Appendix N shows coefficients and odds ratios for each predictor variable, as well as significance tests for adding each predictor to a model containing all other predictors.

For each 1 dB increase in the SEL of an indoor noise event, the probability of motility increased by 7%. Average SELs for indoor noise events that did and did not awaken participants were 73 dB and 69 dB, respectively. Participants in the 35 to 49 year age range were 2.5 times less likely to register an actimetric blip in response to a noise event than those who were younger or older. Men were about 50% more likely to move in the presence of a noise event.

Despite the strong confidence in the ability to predict the presence of at least one actimetric blip statistically, the size of the relationship between motility and the set of predictors was small; McFadden's $\rho^2 = .08$. Prediction success was only 59%, as compared with 50% correct prediction by chance. Prediction success by SEL alone was 56%, with $\rho^2 = .03$, although that model was reliably better than a chance model $\chi^2(1) = 56.59, p < .001$. Inclusion of residence, to account for individual differences in noise sensitivity, raised prediction success to 61% and McFadden's ρ^2 to .13. (This was the only analysis in which residence could be included as a set of dummy-coded predictors; the remaining analyses were based on all four rounds of data collection and contained too many residences for stable analysis.)

Note that the prediction equations accounted for little variance in sleep disturbance relative to that provided by dosage-response analyses, despite strong statistical power. This is due to the great variability in analyses based on individual noise events experienced by individual participants (in logistic regression analyses) as compared with aggregated data (in dosage-response analyses).

The performance of the logistic regression model in predicting the presence of movement may be summarized by a receiver operating characteristic (ROC) curve. An ROC curve plots the probability of a correct decision -- a "hit" -- against the probability of an incorrect decision -- a "false alarm" -- to show the entire range of performance (ratios of hits to false alarms) that a decision maker (a statistical prediction model, in this case) can exhibit. The area under the ROC curve, d', is a measure of the detectibility of movement by the model. The ROC curve for the performance of SEL alone as a predictor had a $d' = 0.62$, whereas the ROC curve for the performance of a model based on all of the predictor variables had a $d' = 0.69$.

Additional multiple logistic regression analyses of Swiss-recorded motility data appear in Appendix O. These analyses were based on the probability of an actimetric blip in a 30-s epoch, whether or not that epoch included a noise event. Separate analyses were performed on indoor and outdoor noise measurements. Outdoor noise level was unrelated to actimetric blips in analysis epochs.

Inclusion of epochs in which noise events did not occur changed results using indoor SEL somewhat. Age effects were similar, with a stronger tendency for greater responsiveness of younger and older participants to noise events than to SEL of epochs in general. Individual differences were prominent in both analyses. Gender effects were found for analyses of noise events, but not epochs. However, tiredness was related to actimetric blips in the epoch, but not to event data.

4.4.6.2 Prediction of Motility as Recorded by the U.S.-made Actimeter

Data for this analysis were provided by the 6 participants who wore U.S.-made actimeters in the first round of data collection at DEN, and 56 participants in the remaining rounds of data collection at both sites. A total of 5,104 noise events occurring between 2200 and 700 hours was included in this analysis.

The model based on the set of 10 variables predicted motility as recorded by the U.S.-made actimeter better than a chance model, $\chi^2(10) = 605.31$, $p < .001$. All of the predictors except noise level significantly added to prediction after adjustment for all other variables. Table 25 in Appendix N shows that the number of spontaneous awakenings (after inverting and reflecting to compensate for severe skewness) was negatively associated with motility. This means that the greater the rate of spontaneous awakenings, the less the likelihood of motility in the presence of a noise event. Men were almost twice as likely as women to move in the presence of a noise event. Age showed both linear and quadratic relationships with motility. The direction of these relationships indicates that older participants (50 years old and above) were 4% more likely to move than younger participants (less than 35 years old), but participants between those ages were almost 40% less likely to move than the average of the extreme groups. That is, the difference in motility between younger and middle age participants was less than the difference between the middle age and older participants.

The probability of movement grew about 6% with each 15 minutes since retiring and decreased less than 1% with each month of residence. Motility increased 2% with each night in the study. Tiredness the previous day decreased motility about 15% for each unit on a scale of 1 (not at all tired) to 5 (extremely tired). Each dB of ambient level decreased motility by about 2%.

The full 10-predictor model accounted for 11% of the variability in sleep disturbance, and correctly predicted the motility outcome of 69% of the noise events. A model that included only SEL accounted for less than 1% of the variance in motility, although it was a significant improvement over a chance model, $\chi^2(1) = 8.48$, $p = .004$. The model based on SEL alone correctly predicted the outcome for 64% of the noise events.

The model based on SEL alone, with $d' = 0.53$, was less successful in detecting the presence of movement than was the model based on all 10 predictors, $d' = 0.73$.

4.4.6.3 Prediction of Behavioral Awakening Responses

Data for this analysis were provided by all participants, responding to a total of 7,685 noise events occurring between 2200 and 0700 hours.

The 10 variables reliably predicted awakening as recorded by behavioral awakening responses, with a model produced by those variables better than a chance model, $\chi^2(10) = 124.07$, $p < .001$. Four variables significantly added to the remaining variables in prediction of awakening: ambient level, age, study duration, and SEL, as seen in Table 26 of Appendix N. With each dB decrease in ambient level, the probability of awakening increased by about 6%. Probability of awakening increased about 4% with each year of age, and decreased by about 4% with each subsequent night in the study. Each dB of SEL of an event increased its probability of awakening participants by about 4%. Average SELs for noise events that did and did not awaken participants were 69 dB and 66 dB, respectively.

The prediction success rate of 95% for the full model reflects the extreme rarity of noise events that elicited behavioral awakening responses: 184 out of 7.685. Using McFadden's ρ^2 criterion, the model accounted for 7% of the variance in awakening. A model based solely on SEL of noise events also predicted awakening better than a chance model, $\chi^2(1) = 25.64$, $p < .001$. Prediction success on the basis of SEL alone also was .95, with McFadden's $\rho^2 = .02$.

The ROC analysis of awakening as signaled by behavioral awakening responses indicated that the ROC curve for the performance of SEL alone as a predictor had a $d' = 0.55$, whereas the model with all 10 predictors had a $d' = 0.71$.

4.4.6.4 Prediction of Arousal by U.S.-made Actimeter Criterion

Data for this analysis were provided by the 6 participants who wore U.S.-made actimeters in the first two rounds of data collection at DEN and DIA, and 56 participants in the second two rounds of data collection at both sites. A total of 5,104 noise events, occurring between 2200 and 0700 hours, was analyzed.

The 10 variables reliably predicted arousal as determined by the U.S. actimetric criterion (Cole *et al.*, 1992), with a model produced by those variables better than a chance model, $\chi^2(10) = 188.3$[illegible], $p < .001$. Seven variables significantly added to the remaining variables in prediction of arousal: ambient level, age (linear and quadratic components), gender, spontaneous awakenings, study duration, and tiredness as seen in Table 27 of Appendix N. With each dB decrease in ambient level, the probability of arousal increased by about 2%. Probability of arousal was greater for older (50 years old or more) than younger (34 years old or less) participants, with participants in the middle closer to younger than older participants in arousal in response to noise events.

Probability of arousal as determined by Cole's (1992) criterion increased by about 2% with each subsequent night in the study. Men were about 30% more likely to be aroused in the presence of a noise event than women. Each unit of rating on the tiredness scale decreased the likelihood of arousal by about 12%. A dB increase in ambient level decreased the probability of arousal in the presence of a noise event by about 2%.

The prediction success rate for the full model was 68%. Using McFadden's ρ^2 criterion, the model accounted for 4% of the variance in arousal. A model based solely on SEL of noise events was also better than chance, $\chi^2(1) = 16.00$, $p < .001$. Prediction success on the basis of SEL alone was 67%, with McFadden's $\rho^2 < .01$.

The ROC analysis of arousal indicated that the ROC curve for the performance of SEL alone as a predictor had a $d' = 0.53$, whereas the model with all 10 predictors had a $d' = 0.63$.

4.4.7 Attempted Replication of Ollerhead's Analysis

Several additional analyses were designed to replicate the findings of Ollerhead *et al.* (1992) with data collected from 28 participants wearing the Swiss-made actimeters at DEN prior to airport closure. Results of these analyses are discussed in Appendix O.

5 DISCUSSION

5.1 COMPARISON OF PRESENT AND EARLIER FINDINGS

Table 14 compares the major characteristics of several recent field studies of the influence of aircraft noise on sleep disturbance. Appendix O contains a detailed comparison of the present findings with those of Ollerhead *et al.* (1992).

Figure 29 plots the data from the current dosage-response relationship between SEL and behavioral awakenings along with data from the six field studies reviewed by Pearsons *et al.* (1995), the data from Ollerhead *et al.* (1992) and the data from Fidell *et al.* (1995). The current findings are highly consistent with those of prior findings, such that inclusion of current data has little effect on the prior dosage-response relationship. The relationship is quite stable, but accounts for only about a third of the variance in the data set. Each 10 dB increase in SEL raises the prevalence of awakening by only about 1.5%.

The dosage-response relationship shows much greater variability at higher than lower noise levels. For example, the range of prevalence of awakening at 60 dB is from 0 to about 2%. The range at 100 dB is from 0 to over 15%, since even high level noise events sometimes fail to awaken test participants.

5.2 OUTDOOR SEL OF IDENTIFIED AIRCRAFT EVENTS AS PREDICTOR OF SLEEP DISTURBANCE

Outdoor noise event levels, whether defined by a level threshold or as confirmed aircraft flyovers, were not reliably related to participants' motility rates as measured by two actimeters, and to their awakening or arousal as measured by three criteria. Outdoor noise event levels cannot therefore be viewed as the principal cause of sleep disturbance.

5.3 ABILITY OF INDOOR SEL TO PREDICT SLEEP DISTURBANCE

Reliable dosage-response relationships were found between indoor SEL of noise events and motility as measured by the Swiss- and U.S.-made actimeters. A reliable relationship was also observed between indoor SEL of noise events and awakening as indicated by button pushes or arousal as determined by Cole's (1992) algorithm applied to the U.S.-made actimeter data, but not with arousal as determined by Ollerhead's (1992) algorithm applied to the Swiss-made actimeter data. No reliable difference was observed in the strength of the relationships between any of the indicators of sleep disturbance and indoor SEL. Estimates of sensitivity of sleep to a single dB increase SEL ranged from about 0.25% to about 1.23%.

Table 14 Comparison of design features of the current study with Ollerhead *et al.* (1992) and Fidell *et al.* (1995).

DESIGN FEATURE	OLLERHEAD *et al.*, 1992	FIDELL *et al.*, 1995	CURRENT STUDY
Venue	Eight neighborhoods near four British Airports	Neighborhoods and individual sites near Castle AFB, LAX, and control areas	Neighborhoods near two large civil airports
Independent Variables	Outdoor confirmed aircraft sound exposure levels	Indoor and outdoor noise event levels	Indoor and outdoor noise event levels; confirmed aircraft sound exposure levels
Dependent Variables	Motility within 30 seconds of a confirmed aircraft noise event, as measured by actimetric time-above-threshold criterion; some EEG	Behaviorally-confirmed awakening within varying time periods after occurrence of noise event	Motility as measured by actimetric time-above-threshold and number of zero crossings within analysis epoch; behaviorally-confirmed awakenings within five minutes of noise event
Principal Inferential Analysis	Multiple logistic regression with *post hoc* definition of sleep disturbance sensitivity categories	Multiple logistic regression;	Multiple logistic regression;
Predictor Variables Considered	Outdoor aircraft noise event levels; age, gender, duration of residence time of night, individual sensitivity, miscellaneous additional factors	Indoor and outdoor noise event, "whole-night" and ambient levels; age, gender, duration of residence, time since retiring, time of night, duration of participation in study; self-rated tiredness; miscellaneous additional factors	Indoor and outdoor noise event, aircraft noise, "whole-night" and ambient levels; age, gender, duration of residence, time since retiring, time of night, duration of participation in study; self-rated tiredness, miscellaneous additional factors
Subject-Nights of Observations	5742	1857	2717

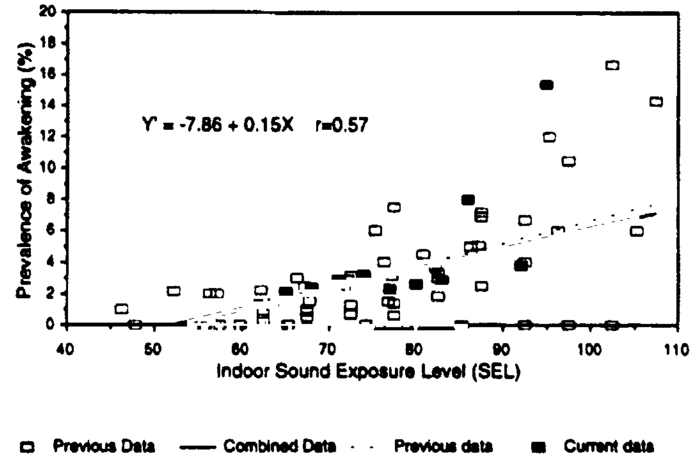


Figure 29 Composite of data from current study with findings of prior sleep disturbance field studies.

The algorithm used by Ollerhead (1992) was a rather lax one, counting the onset of *any* period of motility, regardless of duration, as a sleep disturbance. Cole's (1992) algorithm, as applied to the data gathered from U.S.-made actimeters, was more stringent in defining sleep disturbance, taking into account the duration of motility. Given that nighttime noise intrusions are relatively rare events, the less stringent criterion of arousal could lead to an overestimate of the number of sleep disturbances experienced during the night, thus paradoxically reducing the likelihood of finding a reliable dosage-response relationship.

5.4 ROLE OF OTHER PREDICTORS OF SLEEP DISTURBANCE

Multiple logistic regression analysis revealed several predictors of sleep disturbance (as measured by motility indicated by Swiss- and U.S.-made actimeters, by behavioral awakening responses, and by arousals defined by the U.S.-made actimetric criterion) independent of SEL. Ambient noise was negatively associated with sleep disturbance in the presence of an indoor noise event for all indicators except motility indicated by the Swiss-made actimeter. Each 1 dB increase in ambient L_{eq} decreased the effect of noise events on sleep disturbance by 2-6%.

All measures of sleep disturbance except behavioral awakening responses were gender-related. Men were 25-100% more likely to respond to a noise event than were women. Spontaneous awakenings were related to both U.S. actimetric measures of sleep disturbance: motility and arousal. Participants who spontaneously awoke more often during the night were less likely to awaken or be aroused in response to noise events.

The relationship between age and sleep disturbance was complex, and differed among the sleep disturbance measures. Behavioral awakening was linearly related to sleep disturbance, with about a 4% increase in response rate with each year of age. Motility as measured by the Swiss actimeter bore only a quadratic relationship with age, with older and younger respondents more likely to move than those between 35 and 49 years of age. Both linear and quadratic relationships between age and sleep disturbance were found for motility and arousal as determined by the U.S. actimeter. Older participants were more responsive to noise than younger, but the difference accelerated between middle-age and older participants.

Time since retiring predicted U.S. actimetric motility and behavioral awakening, with increases in sleep disturbance ranging from 1-6% for each 15 minutes since retiring. Duration of residence was related only to U.S. actimetric motility, but the increase in motility with each month of residence was less than 1%. Study duration was related to all measures of sleep disturbance except Swiss actimetric motility, but the nature of the relationship differed. Number of nights in the study increased the probability of motility and arousal as measured by the U.S. actimeter by about 2%, but decreased the probability of behavioral awakening by about 4%.

SEL failed to reliably predict motility or arousal as indicated by the U.S. actimeter after adjustment for the other 10 predictors. This indicated some relationship between SEL and one or a combination of other predictors. By itself, indoor SEL reliability predicted all four measures of sleep disturbance analyzed.

In all cases, the use of control variables more than doubled the predictability of sleep disturbance over that afforded by noise level alone.

5.5 EFFECTS OF CHANGES IN FLIGHT OPERATIONS AT TWO AIRPORTS

Neither the beginning nor ending of flight operations at DIA and DEN, respectively, reliably affected the number of behavioral awakenings, recalled time to fall asleep, time spent awake during the night, or motility. A general decrease in behavioral awakening responses at DIA during the third round of data collection was not shared by all participants. Further, this apparent habituation to data collection procedures was also observed at DIA over the weeks of data collection conducted one year prior to start of flight operations.

Change in indoor L_{eq} with the transfer of operations was minor (2 dB) at DEN, and not statistically significant at DIA. Outdoor noise levels increased by about 5 dB at DIA and decreased by about 12 dB at DEN following the closure of DEN and opening of DIA. The minor effects of flight operations on indoor noise level apparently were insufficient to disturb sleep.

The minimal changes in indoor noise levels may be related to season of the year and associated temperature control, *i.e.*, windows may be open more frequently or air conditioners may be turned on.

5.6 RELATIONSHIPS AMONG INDICATORS OF SLEEP DISTURBANCE

Weak but reliable relationships were observed among three indicators of awakening and arousal: button pushes, Ollerhead's criterion for the Swiss actimeter, and Cole's criterion for the U.S. actimeter. The two actimetric criteria for arousal were more highly related (about 5% of variance shared) than either of them were with behavioral awakening (about 1% of variance shared). However, dosage-response relationships suggest the possibility of a superiority of behavioral indication of awakening.

The strongest association found was between Swiss and U.S. actimetric measures of motility, with about 19% of variance shared. Small, reliable, relationships were also found between each of the measures of motility and behavioral awakening, with about 1-2% of variance shared. Dosage-response relationship suggest the possibility of the superiority of the Ollerhead criterion for motility as the best indication of sleep disturbance, but further research is needed to confirm this.

5.7 COMPARISON OF CURRENT DOSAGE-RESPONSE RELATIONSHIP AND LOGISTIC ANALYSES WITH THOSE OF FIDELL *et al.* (1995)

The current dosage-response relationship between indoor SEL and behavioral awakening closely resembled that of Fidell *et al.* (1995). Variance in awakening accounted for by SEL of noise events varied from about 30-45%. A single dB increase in SEL appeared to produce about a 0.2% increase in the probability of awakening.

The average number of spontaneous behaviorally-confirmed awakenings per night was somewhat lower than found by Fidell *et al.*: about 1.4 per night in the current study as compared with slightly over 2 per night found previously. The average number of awakenings per night associated with noise events, however, was comparable with the prior study at about 0.22 per night.

5.7.1 Logistic Regression Analyses of Behavioral Awakening Responses

Table 15 summarizes a comparison of the multiple logistic analyses of the current study and those of Fidell *et al.* (1995). The findings of the two analyses were generally consistent with respect to noise level of events as well as ambient noise, although the magnitude of the effect of noise level varied somewhat. The current study showed a 3% increase in awakening with each 1 dB increase in noise level, while Fidell, *et al.* reported a 6% increase in awakening. Findings with respect to age were in the opposite direction, however. Current data indicated an increase in responsiveness with age, whereas Fidell *et al.* reported a decrease in responsiveness with age.

5.7.2 Event-Detection Analyses

The indoor noise event data for the six participants wearing Swiss actimeters were divided into 30-s epochs between the hours of 2200 and 0700. These indoor data were modeled as an event-detection process, as described by Fidell *et al.* In this analysis, an arousal (Swiss actimeter blip) was considered to be a consequence of a decision that a change had occurred in the short-term noise environment. This decision-making process is characterized by the ratio of "hits" (assertions that a signal is present when it is truly present) to "false alarms" (assertions that a signal is present when it is fact absent) that can be achieved (Green and Swets, 1966).

Table 15 Comparison of current behavioral awakening analysis results, using indoor noise event data, with behavioral awakening findings reported by Fidell *et al.* (1994)

PREDICTOR	FINDINGS OF CURRENT ANALYSIS	FINDINGS OF FIDELL *ET AL.*	COMMENTS
Indoor Noise level	Positive linear effect	Positive linear effect	Consistency of findings is noteworthy despite different environments
Ambient level	Negative linear effect	Negative linear effect	
Time of night	No effect	Positive linear effect	Defined as time since retiring
Nights in study	Negative linear effect	No effect	
Age	Positive linear effect, no quadratic effect	Negative linear effect	Quadratic effect not tested by Fidell *et al*
Gender	No effect	No effect	
Duration of residence	No effect	Positive linear effect	
Tiredness	No effect	Positive linear effect	
Performance of logistic model as predictor of awakening to noise events	$d'=0.71$	$d'=0.79$	

Epochs containing noise events as well as actimeter blips may be viewed as hits, while epochs containing actimeter blips but no noise events can be viewed as false alarms. The standard index of sensitivity is a scalar quantity known as d'. When d' is zero, a detector has no information about the presence or absence of a signal and thus is completely insensitive to it. When $d' = 4$, a detector can make essentially perfect decisions about the presence or absence of a signal.

The gross hit rate (as defined above) in the present data set was about 24%, while the gross false alarm rate was about 6%. Assuming equivalent Gaussian distributions of numbers of epochs with and without noise events, the value of the sensitivity index, d', which corresponds to this ratio of hits and false alarms is .88. The detection performance of test participants in the study of Fidell *et al.* was at a level of .23. Thus, the current data show motility to be about 6 dB more sensitive to noise than behaviorally-confirmed awakenings.

5.8 IMPLICATIONS FOR FURTHER STUDY

The current study attempted to take advantage of changes in flight operations occurring with the opening of DIA and closure of DEN. However, these changes turned out to have little effect on outdoor noise exposure near participants' homes, and virtually no effect on noise exposure in participants' sleeping quarters. Therefore, there was no opportunity to observe the effect of large changes in noise environments on sleep. Future research should concentrate on a venue in which large numbers of nighttime aircraft flights are either initiated or ended during the data collection period. Greater changes in exposure than were available during the current study will be necessary to resolve issues of adaptation to changes in nighttime noise exposure.

The current study suggests that motility may be better predicted by noise exposure than are behaviorally-confirmed awakenings. Dosage-response relationships with motility appear to be stronger (although not reliably so) and there was evidence of habituation to the instrumentation required for behavioral confirmation of awakenings. Future research, therefore, should include actimeters for all participants.

The current study seems to resolve the dosage-response relationship between indoor noise exposure and behavioral awakening responses, at least at lower levels of exposure. Instrumentation for behaviorally-confirmed awakenings is nevertheless recommended in future research because of the greater interpretability of button pushes than movement as indicators of sleep disturbance. Additionally, the use of behaviorally-confirmed awakening responses at a site with higher indoor noise exposure would help to confirm the linear nature of the relationship and to provide greater opportunity for a nonlinear relationship to emerge.

6 CONCLUSIONS

Because no effort was made to rigorously define the complete population exposed to nighttime noise exposure, nor to obtain a representative sample of any wider population, conclusions drawn from the present study apply strictly only to test participants. To the extent that generalizations are made from the present findings, they should be restricted to the effects of noise on the sleep of long-term residents of neighborhoods without sudden, large changes in nighttime noise exposure.

The following are among the major findings of the present study:

1) The current findings closely resemble those of prior field studies of noise-induced sleep disturbance.

2) Outdoor nighttime L_{eq} at test sites near DEN decreased by about 12 dB on average upon closure of the airport, but increased by only about 3 dB at test sites near DIA after opening of the airport.

3) Indoor nighttime L_{eq} varied little at either location with changes in flight operations from DEN to DIA.

4) The average number of behavioral awakening responses per night was 1.8 at DEN and 1.5 at DIA. Spontaneous behavioral awakening responses (those unassociated with noise events) accounted for 1.5 awakenings per night at DEN and 1.3 awakenings per night at DIA.

5) Statistically reliable relationships were observed between sound exposure levels of individual noise intrusions as measured inside sleeping quarters within five minutes of their occurrence and several measures of sleep disturbance. These were:

   - SEL of individual noise intrusions accounted for about 81% of the variance in motility as measured by the Swiss-made actimeter. The linear relationship between the percentage of test participants exhibiting motility following a noise event was:

$$\% \text{ motility} = -23.74 + 1.23(\text{SEL})$$

   - SEL of individual noise intrusions accounted for about 71% of the variance in motility as measured by the U.S.-made actimeter. The linear relationship between the percentage of test participants exhibiting motility following a noise event was:

$$\% \text{ motility} = 47.16 + 0.4(\text{SEL})$$

- SEL of individual noise intrusions accounted for about 45% of the variance in behavioral awakening responses. The linear relationship between the percentage of test participants exhibiting a behavioral awakening response following a noise event was:

$$\% \text{ noise-induced awakening} = -15.04 + 0.25(\text{SEL})$$

- SEL of individual noise intrusions accounted for about 38% of the variance in arousals as measured by the U.S.-made actimeter and defined and processed in accordance with the criteria of Cole *et al.* (1992). The linear relationship between the percentage of test participants exhibiting arousal following a noise event was:

$$\% \text{ arousal} = 1.31 + 0.28(\text{SEL})$$

6) Indoor SEL accounted for somewhat less than one-third of the predictable variance in sleep disturbance in logistic regression models that included other predictor variables.

7) Relationships among measures of sleep disturbance that were reliable, but weak to moderate in magnitude, included the following:

   - About 19% of variance was shared between motility as measured, processed, and defined by the two types of actimeter.

   - About 1% to 5% of variance was shared among behaviorally-confirmed awakening and the two actimetric criteria for awakening.

   - About 25% of variance was shared between behaviorally-confirmed and self-reported awakenings; participants recalled awakening slightly less than twice per night and pushed buttons to indicate awakenings about 1.6 times per night.

   - About 4% of variance was shared between actimetrically-defined sleep latency and recalled time to fall asleep; on average recalled and actimetrically-defined sleep latency was about 17-18 minutes.

   - About 25% of variance was shared between actimetrically-defined and recalled time spent awake; recalled time awake (about 12 minutes on average) was considerably shorter than actimetrically-defined (about 34 minutes on average).

7 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),* 461-469.

Fidell, S., K. Pearsons, R. Howe, B. Tabachnick, L. Silvati, & D. S. Barber (1995). "Field study of noise-induced sleep disturbance." *Journal of the Acoustical Society of America, 98(2),* 1025-1033.

Horonjeff, R. D., S. Fidell, S. R. Teffeteller, & D. M. Green (1982). "Behavioral awakening as functions of duration and detectability of noise intrusions in the home." *Journal of Sound and Vibration, 84,* 327-336.

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 .S., D. S. Barber, B. G. Tabachnick, & S. Fidell (1995). "Predicting noise-induced sleep disturbance." *Journal of the Acoustical Society of America, 97(1),* 331-338.

8 GLOSSARY AND ABBREVIATIONS

Terms in this Glossary are defined in the sense in which they are used in the body of this report, not necessarily in their broadest sense.

α: The probability of making a Type I error (*q.v.*).

AL_{max}: Abbreviation for maximum A-level (*q.v.*).

Annoyance: A general adverse attitude toward noise exposure.

Analysis of variance: Analysis of the relationship between one or more discrete independent variables and a single continuous dependent variable.

ANOVA: Abbreviation for analysis of variance (*q.v.*).

A-weighted sound level: A single number index of a broadband sound that has been subjected to the A-weighting network (*q.v.*)

A-weighting network: A frequency-equalizing function intended to approximate the sensitivity of the human hearing to sounds of moderate sound pressure level.

β: The symbol for a standardized regression coefficient, indicating the change in standardized units in a criterion variable with a standard deviation change in a predictor variable. In multiple regression (*q.v.*), change is evaluated with all other predictor variables held constant.

B: The symbol for an unstandardized regression coefficient, indicating the change in a criterion variable predicted from a one-unit change in a predictor variable. In multiple regression (*q.v.*), change is evaluated with all other predictor variables held constant.

Between-subjects analysis: ANOVA in which each case provides data for only one level of a discrete independent variable, such as gender.

Bivariate regression: Statistical technique for assessing the prediction of a continuous dependent variable from a single continuous independent variable, and the linear correlation between the variables.

BMDPLR: Commercial statistical software for logistic regression analysis.

Confidence interval: The range of population values of a statistic (*e.g.*, a mean or regression line) that is reasonable within some probability level.

Confounding: A potential cause (the confound or confounder) of a response has not been controlled and, therefore, cannot be isolated from the presumed causal agent (noise exposure).

Covariate: Variable for which statistical adjustment or control has been made.

C-weighting network: A frequency-equalizing function intended to approximate the sensitivity of the human hearing to sounds of high sound pressure level. Essentially limits the bandwidth to include only unweighted 1/3 octave band levels from 31.5 to 8000 Hz.

d': Abbreviation and symbol for the scalar index of signal detectability

Day Average Sound Level: Time-average sound level between 0700 and 2200 hours. Unit, decibel (dB); abbreviation, DL; symbol, L_d.

NOTE:

Day average sound level in decibels is related to the corresponding day sound exposure level, L_{Ed}, according to:

$$L_d = L_{Ed} - 10 \log (54000/1)$$

where 54,000 is the number of seconds in a 15-hour day.

Day-Night Average Sound Level: Twenty-four hour average sound level for a given day, after addition of 10 decibels to levels from 0000 to 0700 hours and from 2200 (10 p.m.) to 2400 hours. Unit, decibel (dB); abbreviation, DNL; symbol, L_{dn}.

NOTES:

1. Day-night average sound level in decibels is related to the corresponding day-night sound exposure level, L_{Edn}, according to:

$$L_{dn} = L_{Edn} - 10 \log (86400/1)$$

where 86,400 is the number of seconds in a 24-hour day.

2. A-frequency weighting is understood, unless another frequency weighting is specified explicitly.

dB: Abbreviation for decibel (*q.v.*).

decibel: Unit measure of sound pressure level and other kinds of levels. It is expressed mathematically as the product of 10 times the logarithm to the base 10 of the ratio of a quantity of interest to a reference quantity.

Dependent variable: The response variable (effect) in a statistical analysis.

Direct logistic regression: Logistic regression in which a set of variables are evaluated simultaneously for their influence in the assessment of the probability of an outcome.

DNL: Abbreviation for Day-Night Average Sound Level (*q.v.*).

Dosage-response relationship: A plot (and analysis) showing a response (*e.g.*, prevalence of disease or awakening) to a dose of noise exposure; also known as dosage-effect relationship.

η^2: In analysis of variance (*q.v.*), the proportion of variance in the dependent variable associated with the independent variable.

Effective Perceived Noise Level: The perceived noise level of a single event that has been modified for the additional annoyance caused by duration and tones.

EGRET: Commercial statistical software package for logistic regression analysis.

EPNL: Abbreviation for Effective Perceived Noise Level (*q.v.*).

Equivalent Noise Level: The sound level typical of the sound levels at a certain place during a stated time period. The time-average sound level in decibels is the level of the mean-square A-weighted sound pressure during the stated time period, with reference to the standard sound pressure of 20 micropascals.

Hosmer-Lemeshow χ^2: An inferential goodness-of-fit test to assess how far a logistic regression model (*q.v.*) departs from observed data.

Independent variable: A presumed causal (or predictor) variable in a statistical analysis.

L_{10}: The level of noise that is exceeded 10 percent of the time.

L_{50}: The level of noise that is exceeded 50 percent of the time.

L_{eq}: Abbreviation for equivalent noise level (*q.v.*).

Logistic regression: A statistical technique for assessing the probability of an outcome from a set of other variables, also known as multiple logistic regression.

Maximum A-level: The maximum A-weighted sound level in a given time period.

Maximum Sound Level; Maximum Frequency-weighted Sound Pressure Level: Greatest fast (125-ms) A-weighted sound level within a stated time interval. Alternatively, slow (1000 ms) time-weighting and C-frequency-weighting may be specified. Unit, decibel (dB); abbreviation, MXFA; symbol, L_{AFmx} (or C and S).

McFadden's ρ^2: In logistic regression, the proportion of variance in the outcome predictable from one or more predictor variables.

Multicollinearity: Extremely high relationships among variables, preventing stable statistical analysis.

Multivariate ANOVA: Analysis of the relationship between one or more discrete independent variable and multiple continuous dependent variables.

Multiway frequency analysis: Analysis of relationships among discrete independent variables.

Night Average Sound Level: Time-average sound level between 0000 and 0700 hours and 2200 and 2400 hours. Unit, decibel (dB); abbreviation, NL; symbol, L_n.

NOTE:

Night average sound level in decibels is related to the corresponding night sound exposure level, L_{En}, according to:

$$L_n = L_{En} - 10 \log(32400/1)$$

where 32,400 is the number of seconds in a 9-hour night.

Odds ratio: In logistic regression, the change in odds of an outcome with a one-unit increase in a predictor variable.

One-sided test: Inferential test in which differences only in one direction between two populations are evaluated or relationships between variables in only direction (positive or negative) are evaluated..

ϕ^2: The size of the relationship between discrete variables on a scale of 0 (no relationship) to 1.00 (perfect relationship).

PNL: Abbreviation for perceived noise level (*q.v.*)

Perceived Noise Level: A single number index obtained by a computational procedure that combines the 24 one-third octave frequency band sound pressure levels in bands centered from 50 to 10,000 Hz to obtain a single level. The number computed by this calculation procedure gives an approximation to the perceived noise level as judged by subjective experiment on a fundamental psychoacoustic basis. Perceived noise level is numerically equal to the sound pressure level of a reference sound that is judged by listeners to have the same perceived nosiness as the given sound. Perceived noise level is generally computed for each 0.5-second time interval during an aircraft flyover.

Polynomial regression: Bivariate regression (*q.v.*) in which relationships more complex than linear are evaluated.

Power: Sensitivity of a statistical analysis to finding a true difference among populations or relationship among variables, defined as $1 - \beta$.

Planned contrast: A pre-planned analysis in which component comparisons within a complex ANOVA are analyzed; for example, the difference between two levels (*e.g.*, time periods) of an independent variables, ignoring all other levels.

Profile analysis of repeated measures: A form of multivariate ANOVA (*q.v.*) in which cases provide data for all levels of a discrete independent variable, such as time period.

r: Index of bivariate linear correlation, the relationship between two continuous variables.

R^2: Symbol for squared multiple correlation, the variance in the criterion variable that is predictable from the set of predictor variables in multiple regression (*q.v.*).

Random effects model: An ANOVA model in which levels of the discrete independent variable (such as participants) are selected randomly.

Receiver operating characteristic (ROC) curve: A plot of the sensitivity of a receiver showing the proportion of hits (decision that an event has occurred when it has in fact occurred) as a function of false alarms (decision that an event has occurred when it has not in fact occurred). The area between the ROC curve and the major diagonal is a measure of d' (*q.v.*).

SD: Abbreviation for standard deviation.

SEL: Abbreviation of sound exposure level (*q.v.*).

Sound Exposure Level: Time integral of squared, instantaneous frequency-weighted sound pressure over a stated time interval or event. Unit: pascal-squared second; symbol, E.

NOTES:

1. If frequency weighting is not specified, A-frequency weighting is understood. If other than A-frequency weighting is used, such as C-frequency weighting, an appropriate subscript should be added to the symbol; *e.g.*, E_C.

2. Duration of integration is implicitly included in the time integral and need not be reported explicitly. For the sound exposure measured over a specified time interval such as one hour, a 15-hour day, or a 9-hour night, the duration should be indicated by the abbreviation or letter symbol, for example one-hour sound exposure (1HSE or E_{1h}) for a particular hour; day sound exposure (DSE or E_d) from 0700 to 2200 hours; and night sound exposure (NSE or E_n) from 0000 to 0700 hours plus from 2200 to 2400 hours.

3. Day-night sound exposure (DNSE or E_{dn}) for a 24-hour day is the sum of the day sound exposure and ten times the night sound exposure.

4. Unless otherwise stated, the normal unit for sound exposure is the pascal-squared second.

Sound Level; Weighted Sound Pressure Level: Ten times the logarithm to the base ten of the ratio of A-weighted squared sound pressure to the squared reference sound pressure of 20 μPa, the squared sound pressure being obtained with fast (F) (125-ms) exponentially weighted time-averaging. Alternatively, slow (S) (1000-ms) exponentially weighted time-averaging may be specified; also C-frequency weighting. Unit, decibel (dB); symbol L_A, L_C.

NOTES:

1. In symbols, A-weighted sound level $L_{A\tau}(t)$ at running time t is:

$$L_{A\tau}(t) = 10 \log \left\{ \left[(1/\tau) \int_{-\infty}^{t} p_A^2(\xi) e^{-(t-\xi)/\tau} d\xi \right] / p_0^2 \right\}$$

where τ is the exponential time constant in seconds, ξ is a dummy variable of integration, $p_A^2(\xi)$ is the squared, instantaneous, time-varying, A-weighted sound pressure in pascals, and p_0 is the reference sound pressure of 20 μPa. Division by time constant τ yields the running time average of the exponential-time-weighted, squared sound-pressure signal. Initiation of the running time average from some time in the past is indicated by $-\infty$ for the beginning of the integral.

2. ANSI S1.4-1983, *American National Standard Specification for Sound Level Meters*, gives standard frequency weightings A and C and standard exponential time weightings fast (F) and slow (S).

Sound Pressure; Effective Sound Pressure: Root-mean-square instantaneous sound pressure at a point, during a given time interval. Unit, pascal (Pa).

NOTE:

In the case of periodic sound pressures, the interval is an integral number of periods or an interval that is long compared to a period. In the case of non-periodic sound pressures, the interval should be long enough to make the measured sound pressure essentially independent of small changes in the duration of the interval.

Sound pressure level: A measure of sound taken as ten times the common logarithm of the square of the ratio of sound pressure to the reference sound pressure of 20 micropascals. The frequency bandwidth must be identified.

Statistical adjustment: Holding adjusted variables constant in order to reveal the unique effect of other variables. See covariate.

Stepdown analysis: Supplemental analysis to multivariate ANOVA in which multiple dependent variables are evaluated in a pre-set priority order; each dependent variable, in turn, is evaluated after statistical adjustment for higher priority dependent variables.

Subject-night: The amount of data collected from one subject for one night.

Two-sided test: Inferential test in which differences in either direction between two populations are evaluated or relationships between variables in only direction (positive or negative) are evaluated.

Type I error: Declaring populations different when in fact they are not different, or relationships among variables to exist when they do not.

Type II error: Failure to declare populations different when in fact they are different, or failing to find relationships among variables when in fact they exist.

Within-subjects analysis: ANOVA in which cases provide data for all levels of a discrete independent variable, such as time period, also known as repeated measures ANOVA.

APPENDIX A RECRUITING PROCEDURES AND INSTRUCTIONS TO TEST PARTICIPANTS

A.1 LETTERS OF SOLICITATION OF TEST PARTICIPATION

The initial mailing to prospective test participants near DEN included a letter of explanation on NASA letterhead, a letter on BBN letterhead, and a return form with a stamped, pre-addressed envelope.

The wording of the NASA letter was as follows:

> *"The Acoustics Division of NASA's Langley Research Center is conducting a field study of the effects of noise on sleep as part of our national Advanced Subsonic Technology Noise Reduction Program. The findings of this study are expected to help NASA in evaluating aircraft noise effects on people, and for more general environmental analysis purposes. One of the sites at which this research will be conducted is in Denver. If you are interested in taking part in this important study, NASA would greatly appreciate your cooperation.*
>
> *"The attached letter explains what your job would be and how to learn about this study in more detail. Please feel free to contact me at XXX-XXX-XXXX if you would like additional information about NASA's Acoustic Research Program. Thank you in advance for your interest."*

The wording of the BBN letter was as follows:

> *"BBN Systems and Technologies is conducting a scientific study during February and March of sleep disturbance in your neighborhood. As described in the attached letter, this study is being conducted for the National Aeronautics and Space Administration (NASA).*
>
> *"Everyone who takes part in this study will push a button if they wake up at night. Some people will also be asked to wear a wristwatch-like device at night. We would like to tell you more about this study, and to find out if you and/or other members of your family might be interested in taking part. Each person who participates will be paid $150 at the end of the 4 week study period.*
>
> *"If you would like to learn more about this study, please fill out and mail the attached form in the stamped envelope. Returning the form does not obligate you to take part in the study. BBN will contact people who return the form to explain details of the study and to answer questions*

about it. If you wish to speak to someone about the study, please call BBN's toll free number (X-XXX-XXX-XXXX) at your convenience."

Prospective test participants were asked to provide information useful for contacting them and assessing their suitability for participation on the returned form.

A.2 INSTRUCTIONS TO TEST PARTICIPANTS

Test participants were sent an instruction booklet following a telephone interview during which (1) they were informed about the study and their roles as test participants, (2) their willingness and suitability as test participants were determined, and (3) an initial equipment installation appointment was set up. Follow-up telephone calls were made to answer any additional questions test participants had upon examining the instruction booklet.

The contents of the instruction booklet are reproduced on the following pages.

YOUR JOB IN THE SLEEP STUDY

This booklet explains what you are expected to do in the sleep study.

You have three things to do every day:

1. Answer the Nighttime Questionnaire on the small computer before you go to bed for the night.
2. Push the red button if you wake up for any reason during the night.
3. Answer the Morning Questionnaire on the small computer when you get out of bed in the morning.

What to Do Just Before You Go to Bed at Night:

1. Make sure that the two black cables are firmly plugged into the small computer, and that the computer is plugged into a wall outlet. Also, check to see that the noise monitoring equipment is plugged into an outlet.
2. Make sure that the red button you push when you wake up is within easy reach of your bed.
3. Open the hinged top of the small computer by lifting the lid from the front. If the screen is blank, press the "ON" button in the upper right-hand corner of the keyboard.
4. Press the "F10" key (toward the right of the top row of buttons) to start the nighttime questionnaire.
5. Answer the question by picking the number which best describes your answer, then press the "ENTER" key. Turn to page ** for an explanation of the nighttime question. If you make a mistake in your answer, you can correct it by pressing the "ESC" key in the upper left corner of the keyboard and answering the question again.
6. After you have answered the question and you are ready to go to sleep, press "F10" again to set the computer for the night. You may leave the computer lid open or closed as you like. Do not turn the computer off.

Note: You should answer the question only once each night as you are about to go to bed, not each time you wake up during the night.

What to Do During the Night:

Press the red button once, right away, each time you wake up for any reason at all during the night. If you stay awake for a while after pushing the button, do not press the button again.

If you forget to press the button when you wake up during the night, and you then stay awake for more than five minutes, do not press the button.

If you stay awake for a while after you wake up during the night and you can't remember (or are not sure) if you pushed the button when you first woke up, do not press the button again.

REMEMBER: Press the red button ONCE, as soon as you wake up, each time you wake up for any reason at all.

What to Do When You Wake up in the Morning:

1. As soon as you wake up in the morning press the red button once.
2. If you closed the lid of the small computer the previous evening, open it by lifting the lid from the front.
3. Press the "F10" key. You will then be asked to estimate how many times you woke up during the night. Press the number on the keypad in the lower right corner of the keyboard and press "ENTER."
4. Answer each of the following questions by picking the number which best describes your answer, then pressing the "ENTER" key. Turn to page ** for an explanation of each of the morning questions. If you make a mistake in your answer, you can correct it by pressing the "ESC" key in the upper left corner of the keyboard and answering the question again.

   If you forget to answer the morning questionnaire when you get up, then answer the questions as soon as you remember.

What to Do If You Take a Nap During the Day:

You don't have to do anything with the equipment if you take a nap during the day. There is no need for any interview or to push the button before or after napping.

If You Have Other Questions:

If you have any questions about the study procedures or experience any difficulty in operating the computer, please call

X-XXX-XXX-XXXX

How to Answer the Nighttime Question:

There is only one question to answer before you go to bed at night:

How tired did you feel today?

Please pick the phrase that best describes how you felt throughout the entire day (not just at the time you are answering the question). Your choices are:

1. Not at all tired
2. Slightly tired
3. Moderately tired
4. Very tired
5. Extremely tired

Press the number on the keypad in the lower right-hand corner of the keyboard corresponding to your choice.

How to Answer the Morning Questions:

How many times did you wake up last night?

Please estimate the number of times you woke up last night. Type in the number and press "ENTER."

How well did you sleep last night?

Please tell us how well you slept last night. Your choices are:

1. Not at all well
2. Fairly well
3. Moderately well
4. Very well
5. Extremely well

How long did it take you to fall asleep?

Please estimate how long it took you to fall asleep when you first went to bed last night. Your choices are:

1. Less than 10 minutes
2. 10 - 20 minutes
3. 20 - 30 minutes
4. 30 - 60 minutes
5. more than an hour

How much were you awake last night?

Please estimate the total amount of time you were awake during the night after you first went to sleep. For example, if you woke up twice during the night and were awake for approximately five minutes each time, then you were awake for a total of about 10 minutes. In this case, you should answer "10-20 minutes." If you did not wake up at all during the night, or if you fell back to sleep quickly after awakening, answer "Less than 10 minutes." Your choices are:

1. Less than 10 minutes
2. 10 - 20 minutes
3. 20 - 30 minutes

4. 30 - 60 minutes
5. more than an hour

How annoyed were you by noise last night?

If you heard any noise during the night (whether you were awakened by it or you were already awake), how annoyed were you by the noise? Your choices are:

1. Did not hear any noise last night
2. Not at all annoyed by noise last night
3. Slightly annoyed by noise last night
4. Moderately annoyed by noise last night
5. Very annoyed by noise last night
6. Extremely annoyed by noise last night

APPENDIX B DATA EXTRACTION PROCEDURES

B.1 GENERAL APPROACH

The field data were extensively preprocessed before performing any statistical analyses. This preprocessing was necessary because:

- The statistical analysis software could not directly cope with the raw data formats;
- Most statistical analyses were conducted on derivative or computed measures rather than raw data. Many of these computed measures were developed from searches through the field data files for very specific or complicated combinations of events; and
- Considerable quality control checking of the input data was required, a task for which the statistical analysis software was ill-suited.

The total quantity of raw data was also a formidable consideration. Approximately 120 megabytes were collected in the first two rounds of data collection alone. Preprocessing this permitted extraction of only those events and data relevant to the inferential analyses.

BBN/Probe interactive data analysis software package was used to perform the data extraction and preprocessing. BBN/Probe is a time-series data analysis software package designed for very large and complex data sets. Automated data extraction and processing of the sleep study data were performed using command files containing BBN/Probe commands put together in an analysis "script." These command files were used to:

1) Deal with the different input data formats, opening the data files and representing the different measurement variables in proper and consistent units, all with strict regard to time synchrony;
2) Perform all of the event searches and computation of the various test measures (*e.g.*, computation of actimetrically-predicted awakening);
3) Perform quality control screening by checking for unreasonable or missing input data; and
4) Provide the required output data, formatted exactly as required by the statistical analysis package.

B.2 INPUT DATA FILES

The raw field measurement data were stored in up to seven different data files per test subject, each with its own data formats. These raw files were downloaded from the field measurement equipment to a laptop PC, and transferred to disk on a DEC VAX computer. The raw data files consisted of:

1) "SWISS" ACTIMETRY DATA - Data file containing 30-s samples from the Gaehwiler actimeter;
2) "U.S." ACTIMETRY DATA - Data file containing 30-s samples from the AMI (U.S.-made) actimeter;
3) INDOOR NOISE LEVEL DATA - Data files produced by the LD820 noise monitors, containing a continuous A-weighted noise level record, sampled every 60 s in data collection Rounds 1 and 2 and sampled every 2 s in data collection Rounds 3 and 4;
4) INDOOR NOISE EVENT DATA - Data files produced by the LD820 noise monitors, containing noise event records of the duration, L_{eq}, L_{max}, and SEL for each noise event above a preset threshold;
5) OUTDOOR NOISE LEVEL DATA - Data files produced by the LD820 and LD870 noise monitors, containing a continuous A-weighted noise level record, sampled every 60 s in data collection Rounds 1 and 2 and sampled every 2 s in data collection Rounds 3 and 4;
6) OUTDOOR NOISE EVENT DATA - Data files produced by the LD820 and LD870 noise monitors, containing noise event records of the duration, L_{eq}, L_{max}, and SEL for each noise event above a preset threshold; and
7) BUTTON-PUSH DATA - File containing the time tags for each button push recorded by the HP-95 palmtop computers.

B.3 OUTPUT DATA FILES

The data extraction procedures created five output files per subject-night:

1) A formatted output file containing all of the test variables needed for the "whole-night"-related statistical analyses;
2) A formatted output file containing all of the test variables needed for the "button-push"-related statistical analyses;
3) A formatted output file containing all of the test variables needed for the "noise event"-related statistical analyses;
4) A level-*vs*-time plot showing the indoor and outdoor noise levels, noise events, button-push events, and the actimetry levels for the entire night. This was used as a visual quality control check of each night's data; and
5) An ASCII text summary of the principal variables from files 1-3. This was also intended to be used as a data quality control check.

The output data were packed in three separate output files to simplify the analyst's task in importing and organizing the datasets for subsequent inferential analysis.

B.4 DATA EXTRACTION COMMAND FILES

Data extraction was performed by a suite of eleven BBN/Probe command files. This suite consisted of a master control procedure and ten subordinate procedures that performed specific data extraction tasks. The master control procedure (named <subj>_AUTO.PRB) handled all "once-per_subject" initialization tasks, made sure that the noise measurement and button-push data files for the subject were opened, and invoked the main data extraction procedure for each night for which there were valid data.

Figure 30 illustrates the hierarchy for these command procedures:

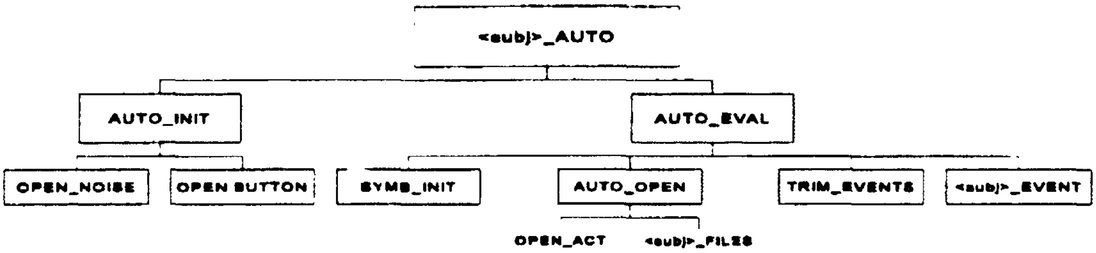


Figure 30 Hierarchy of BBN/Probe command procedures used to reduce and extract data.

The command procedure AUTO_INIT handled the initialization tasks for a given test subject. The command procedure AUTO_EVAL took care of extracting, night by night, all of the data required for the various statistical analyses. AUTO_EVAL wrote these data to ASCII text files in a format compatible with the statistical analysis software packages.

The suite of BBN/Probe command files and a brief description of their tasks are shown in the Table 16.

Table 16 Description of BBN/Probe command procedures.

PROCEDURE FILE NAME	FUNCTION
<subj>_AUTO.PRB	The main control procedure. This is a template file, customized for each test subject to include a list of nights for which data were available.
AUTO_INIT.PRB	This procedure performs all of the once-per-subject initialization tasks, such as opening the noise and button-push data files.
OPEN_NOISE.PRB	This procedure opens the various noise (level and event) data files.
OPEN_BUTTON.PRB	Opens the HP "button-push" data files.
AUTO_EVAL.PRB	This is the main data extraction procedure. It computes all of the variables needed for later analysis, opens the output data files, and writes these data to disk.
SYMB_INIT.PRB	Initializes global symbols used by AUTO_EVAL.
AUTO_OPEN.PRB	Determines which, if any, actimetry data files are to be opened.
OPEN_ACT.PRB	Procedure to open the actimetry data files.
TRIM_EVENTS.PRB	Procedure to delete extraneous Probe event definitions (*e.g.*, those defined from earlier nights).
<subj>_FILES.PRB	Contains catalog of relevant data files for a given test subject.
<subj>_EVENT.PRB	Contains Probe event definitions for a given test subject (e.g., retirement- and wake times).

The main data extraction procedure, AUTO_EVAL, was invoked once for each subject-night where field measurement data were available. Not all types of data were universally available (*e.g.*, no nearby outdoor noise measurements or missing data). AUTO_EVAL compiled as much information as possible, and substituted "missing data" codes where necessary. Tables 17 through 19 summarize the test parameters and how they were computed by AUTO_EVAL. The information in Table 18 was computed and stored for each button push during the night. The information in Table 19 was computed and stored for each noise event (indoor or outdoor) during the night.

Table 17 Test variables used for the "whole-night" statistical analysis.

PARAMETER	DESCRIPTION
night_number*	Night ID number, counting from beginning of the study
group_code*	Subject group ID code
site_code*	Test site ID code
subject_id*	Test subject ID code
age*	Subject age
sex_code*	Subject sex
residence*	Subject's time in residence
tiredness*	Subject's 'tiredness' at retirement
rec_latency*	Subject's recalled initial sleep latency
rec_awake*	Subject's recalled time awake
num_pushes	Total number of button pushes for that night
us_avg_mot	Average actimetric motility level that night (US)
sws_avg_mot	Average actimetric motility level that night
(Swiss) latency_code	Initial sleep latency, determined by Ollerhead 'blip' method
indoor_leq	The indoor Leq for the entire night
outdoor_leq	The outdoor Leq for the entire night
sleep_code	Subject's sleep duration, in integer minutes
int_count	Number of intdoor noise events for the night
ext_count	Number of outdoor noise events for the night
num_O_blips	Number of arousals as defined by Ollerhead *et al.* (1992)
us_awakenings	Number of arousals as defined by Cole *et al.* (1992)
int_avg_lev	Average indoor noise event Lmax
ext_avg_lev	Average outdoor noise event Lmax
int_bins(1:6)	Distribution of indoor noise events
ext_bins(1:7)	Distribution of outdoor noise events
oh_tot_time	Total time awake as defined by Ollerhead *et al.* (1992)
us_tot_time	Total time awake as defined by Cole *et al.* (1992)
BP_OH_code	Behavioral awakening match with Ollerhead-defined arousal

Table 18 Test variables used for the "button-push" statistical analysis.

PARAMETER	DESCRIPTION
night_number	Night ID number, counting from beginning of study
group_code	Group ID code
site_code	Site ID code
subject_id	Subject ID code
age	Subject age
sex_code	Subject gender
residence	Length of residence
tiredness	Tiredness previous dat
rec_latency	Recalled latency to fall asleep
rec_awake	Recalled number of times awake
latency_code	Calculated sleep latency based on Ollerhead *et al.* (1992)
push_no	Button push number, counting from 1 at start of night
push_code	
retire_code	
bp_in_leq5	Indoor Leq in previous five minutes
bp_out_leq5	Outdoor Leq in previous five minutes
bp_AMI_predict	Behavioral awakening response match to Cole *et al.* arousal algorithm
bp_OH_code	Behavioral awakening response match to Ollerhead *et al* arousal algorithm
num_in_events	Number of indoor noise events
num_out_events	Number of outdoor noise events
max_int_sel	Maximum indoor noise event SEL in previous five minutes
mean_int_sel	Mean indoor noise event SEL in previous five minutes
last_int_sel	Most recent indoor noise event SEL in previous five minutes
max_ext_sel	Maximum outdoor noise event SEL in previous five minutes
mean_ext_sel	Mean outdoor noise event SEL in previous five minutes
last_ext_sel	Most recent outdoor noise event SEL in previous five minutes

Table 19 Test variables used for the "noise-event" statistical analysis.

PARAMETER	DESCRIPTION
night_number	Night number, counting from start of study
group_code	Group ID code
site_code	Site ID code
subject_id	Subject ID code
age	Age
sex_code	Gender code
residence	Length of residence
tiredness	Tiredness previous day
rec_latency	Recalled latency to fall asleep
rec_awake	Recalled number of times awake
latency_code	Calculated sleep latency
retire_code	
ev_type_code	Event type code: 1=indoor 2=outdoor
event	
ev_time_code	Event time
match_code	Match to indoor/outdoor event
ev_lmax	Event Lmax
ev_sel	Event SEL
ev_ambient	Ambient Leq prior to event
ev_us_awk	Arousals as predicted by Cole *et al.* (1992) in following five minutes
ev_us_mot	Average U.S.-actimeter recorded motility in following five minutes
ev_sws_awk	Arousals as predicted by Ollerhead *et al.* (1992) in following five minutes
ev_sws_mot	Average Swiss-actimeter recorded motility in following five minutes

In addition, AUTO_EVAL produced a nightly data summary plus a time plot showing indoor and outdoor noise data, noise events, actimetry levels, and button-push events. These were manually examined for anomalies in the data that were undetected by the checks in AUTO_EVAL.

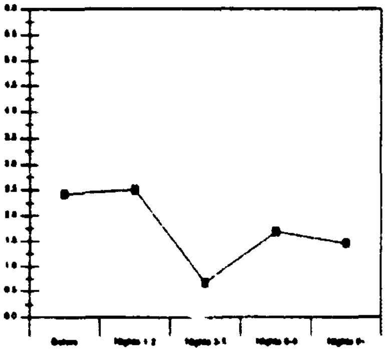
The final step in the process was to copy all of the output files to diskettes to be read by the statistical analysis software.

APPENDIX C BEHAVIORAL AWAKENING RESPONSES ON SUCCESSIVE NIGHTS AT DIA

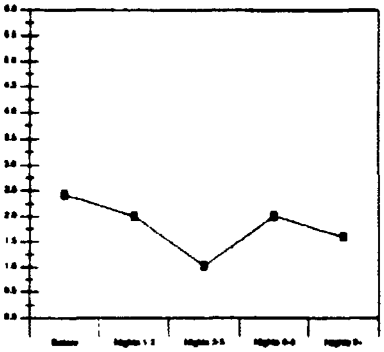
Figure 31 plots behavioral awakening responses (button pushes) for the 22 residents who participated in data collection at DIA during the period around the start of flight operations. Nights are grouped into five periods:

- nights before the closure of DEN (after deleting the first 3 nights of data collection);
- the first two nights after start of flight operations to DIA;
- nights 3-5 after the opening of DIA;
- nights 6-8 after the opening of DIA; and
- the remaining nights of data collection.

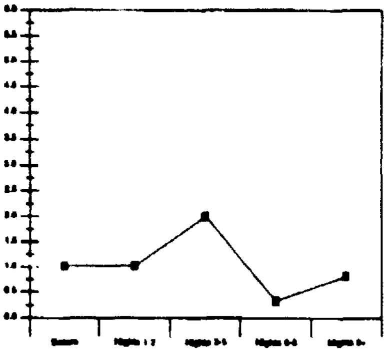
Participant 601



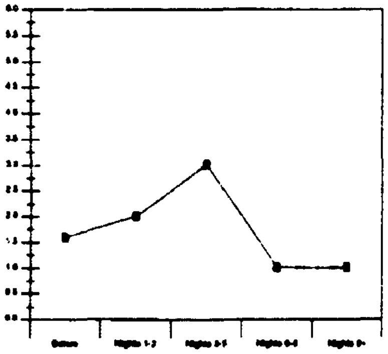
Participant 602



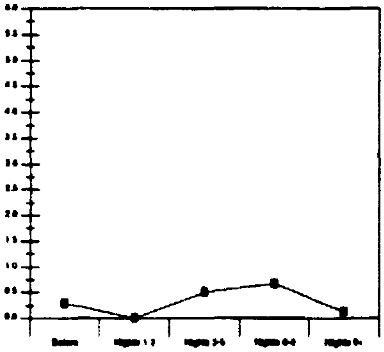
Participant 603



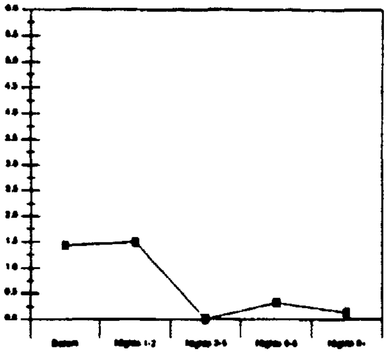
Participant 604



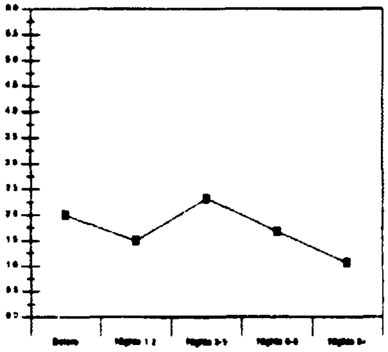
Participant 609



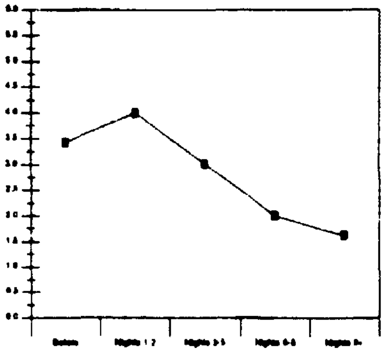
Participant 610



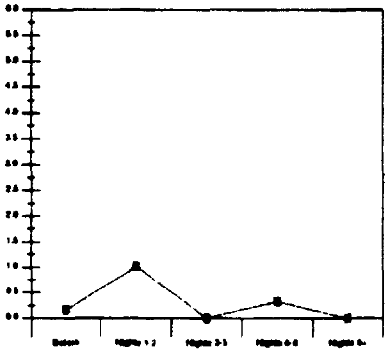
Participant 611



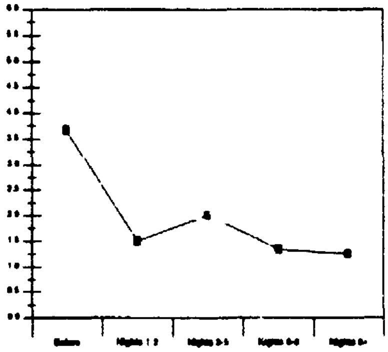
Participant 612



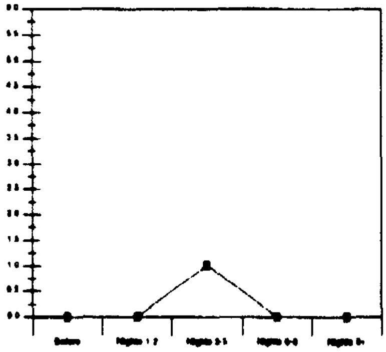
Participant 613



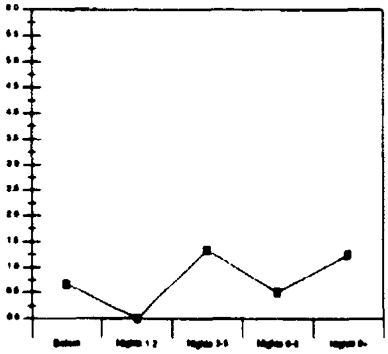
Participant 614



Participant 615



Participant 616



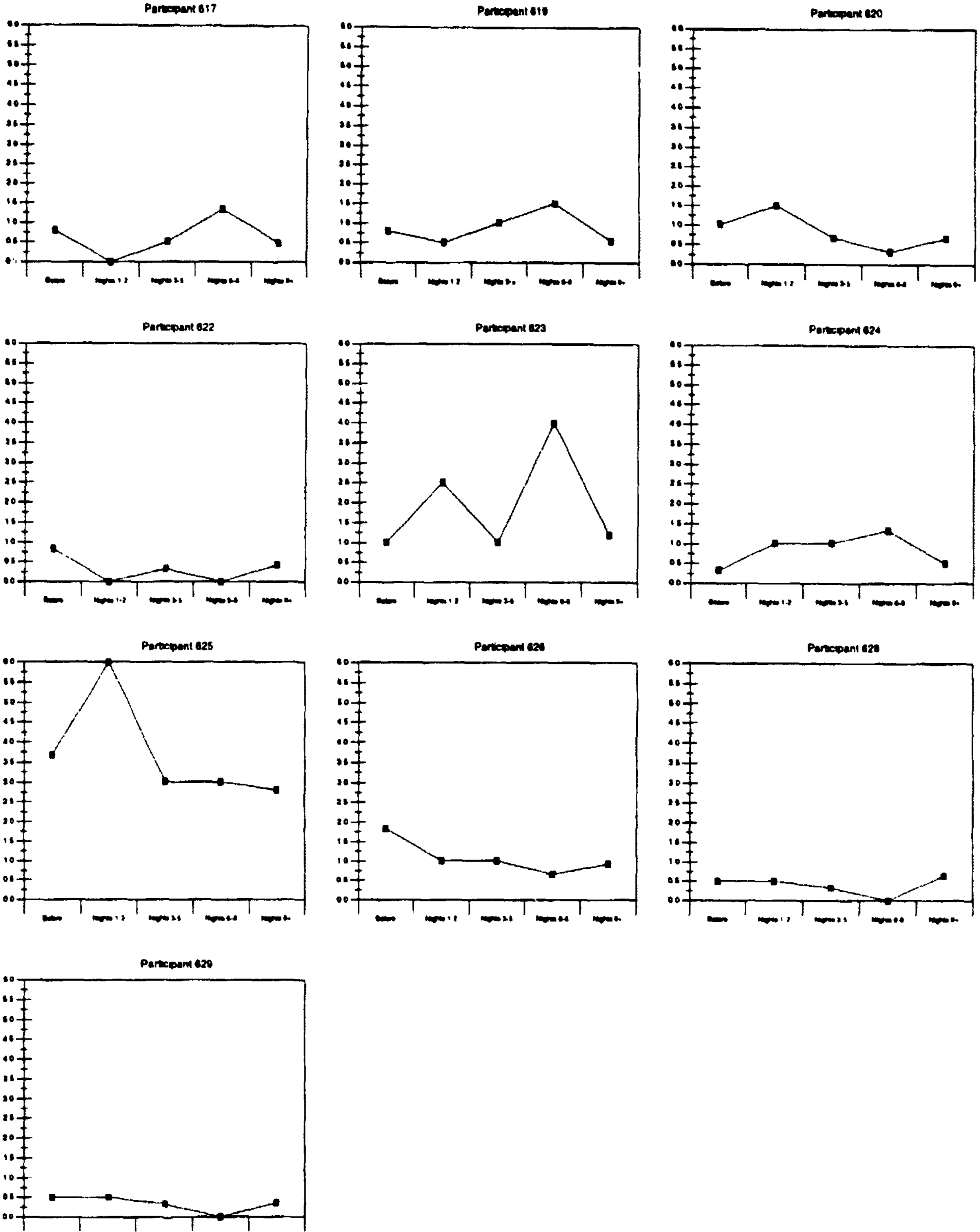


Figure 31 Time course of behavioral awakening responses for 22 individual participants at DIA just before and after start of flight operations at DIA.

APPENDIX D SUMMARY OF NOISE ENVIRONMENTS

D.1 SUMMARY OF NOISE ENVIRONMENTS

Tables 20 through 23 summarize the noise event data analyzed in the current study. A complete discussion of these analyses is located in Section 4

Table 20 Summary of noise measurements at test participants' homes near DEN before airport closure.

Site	Total Number of Noise Events		Number of Noise Events Between 2200 and 0700 hours		Average Event A-weighted L_{max} (dB)	
	Inside	Outside	Inside	Outside	Inside	Outside
A	2103	- -	170	- -	66.9	- -
B	3593	12635	86	453	74.1	82.1
C	162	- -	7	- -	73.2	- -
D	5950	- -	2451	- -	67.4	- -
E	1029	10691	82	892	71.0	78.0
F	3383	- -	310	- -	71.9	- -
G	6977	11651	229	1206	69.2	78.1
H	2849	3316	67	256	72.9	79.1
I	1857	- -	170	- -	73.5	- -
J	1632	- -	273	- -	70.0	- -
K	5340	- -	260	- -	73.0	- -
L	7500	- -	810	- -	68.1	- -
M	4520	- -	202	- -	68.1	- -
N	1371	- -	732	- -	66.2	- -
O	130	9521	49	365	73.7	80.0
TOTALS	**48397**	**47814**	**5898**	**3712**		

Table 21 Summary of noise measurements at test participants' homes near DEN after airport closure.

Site	Total Number of Noise Events		Number of Noise Events Between 2200 and 0700 hours		Average Event A-weighted L_{max} (dB)	
	Inside	Outside	Inside	Outside	Inside	Outside
A	10553	3639	738	568	65.9	78.3
B	1609	--	587	--	57.0	--
C	2519	--	1759	--	61.3	--
D	767	--	411	--	60.0	--
E	495	15542	326	2052	71.9	58.3
F	18129	1645	1425	162	59.9	67.3
G	1109	--	886	--	58.5	--
H	4833	--	1661	--	61.2	--
I	8128	--	432	--	58.6	--
J	4955	--	2542	--	58.2	--
K	479	--	333	--	61.4	--
L	1737	--	744	--	59.4	--
M	1428	--	615	--	65.3	--
N	2388	--	856	--	63.7	--
O	7897	--	4201	--	66.8	--
P	5042	--	2940	--	59.7	--
Q	633	--	339	--	62.7	--
TOTALS	**72701**	**20826**	**20795**	**2782**		

Table 22 Summary of noise measurements at test participants' homes near DIA before airport opening.

Site	Total Number of Noise Events		Number of Noise Events Between 2200 and 0700 hours		Average Event A-weighted L_{max} (dB)	
	Inside	Outside	Inside	Outside	Inside	Outside
A	176	- -	7	- -	72.0	- -
B	1087	- -	56	- -	71.4	- -
C	2461	- -	408	- -	70.6	- -
D	58	1180	28	424	72.4	84.2
E	702	- -	99	- -	73.8	- -
F	1353	- -	211	- -	69.0	- -
G	413	- -	13	- -	82.5	- -
H	376	748	106	29	71.8	78.9
I	251	- -	26	- -	77.5	- -
J	208	2851	18	55	75.0	75.0
K	820	669	102	36	67.6	79.0
L	178	772	14	41	70.3	77.0
M	840	- -	35	- -	73.3	- -
N	449	- -	13	- -	73.8	- -
O	477	- -	3	- -	69.0	- -
P	1943	- -	229	- -	72.1	- -
TOTALS	**11792**	**6220**	**1368**	**585**		

Table 23 Summary of noise measurements at test participants' homes near DIA after airport opening.

Site	Total Number of Noise Events		Number of Noise Events Between 2200 and 0700 hours		Average Event A-weighted L_{max} (dB)	
	Inside	Outside	Inside	Outside	Inside	Outside
A	1049	- -	848	- -	59.9	- -
B	887	- -	278	- -	73.5	- -
C	4746	- -	2496	- -	61.0	- -
D	3059	5786	408	201	63.0	68.8
E	2640	- -	800	- -	68.3	- -
F	5742	- -	2056	- -	60.7	- -
G	873	- -	504	- -	61.2	- -
H	1515	5885	933	2847	61.2	58.0
I	647	- -	355	- -	66.8	- -
J	3665	- -	754	- -	63.5	- -
K	10308	- -	1244	- -	62.8	- -
L	359	- -	164	- -	59.9	- -
M	2689	- -	804	- -	64.0	- -
N	895	- -	230	- -	60.3	- -
O	2073	- -	463	- -	64.2	- -
P	5902	- -	3208	- -	65.0	- -
Q*	- -	3484	- -	2167	- -	66.0
R	903	- -	328	- -	67.7	- -
TOTALS	47952	15155	15873	5215		

* Site Q was equipped only with an outdoor noise monitor.

APPENDIX E SUMMARY OF INTERVIEW DATA

E.1 SUMMARY OF NIGHTTIME INTERVIEWS

Figure 32 describes responses to the nighttime interview question based on data from 2,717 subject-nights. Responses are described separately for five rounds of data collection.

E.2 SUMMARY OF MORNING INTERVIEWS

Figures 33 through 37 describe the results of the morning interview questionnaire based on data from 2,717 subject-nights. Each figure describes responses separately for five rounds of data collection.

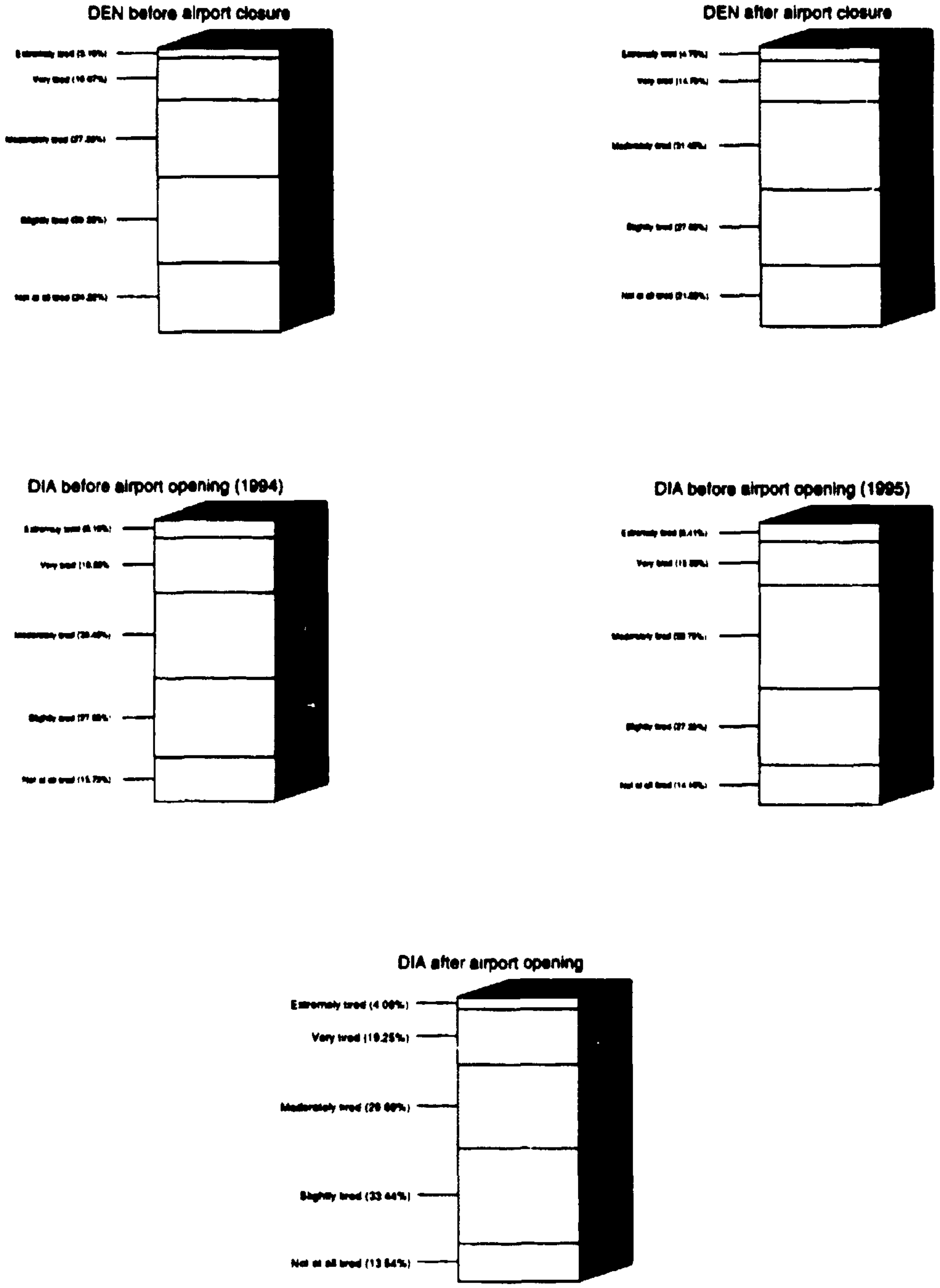
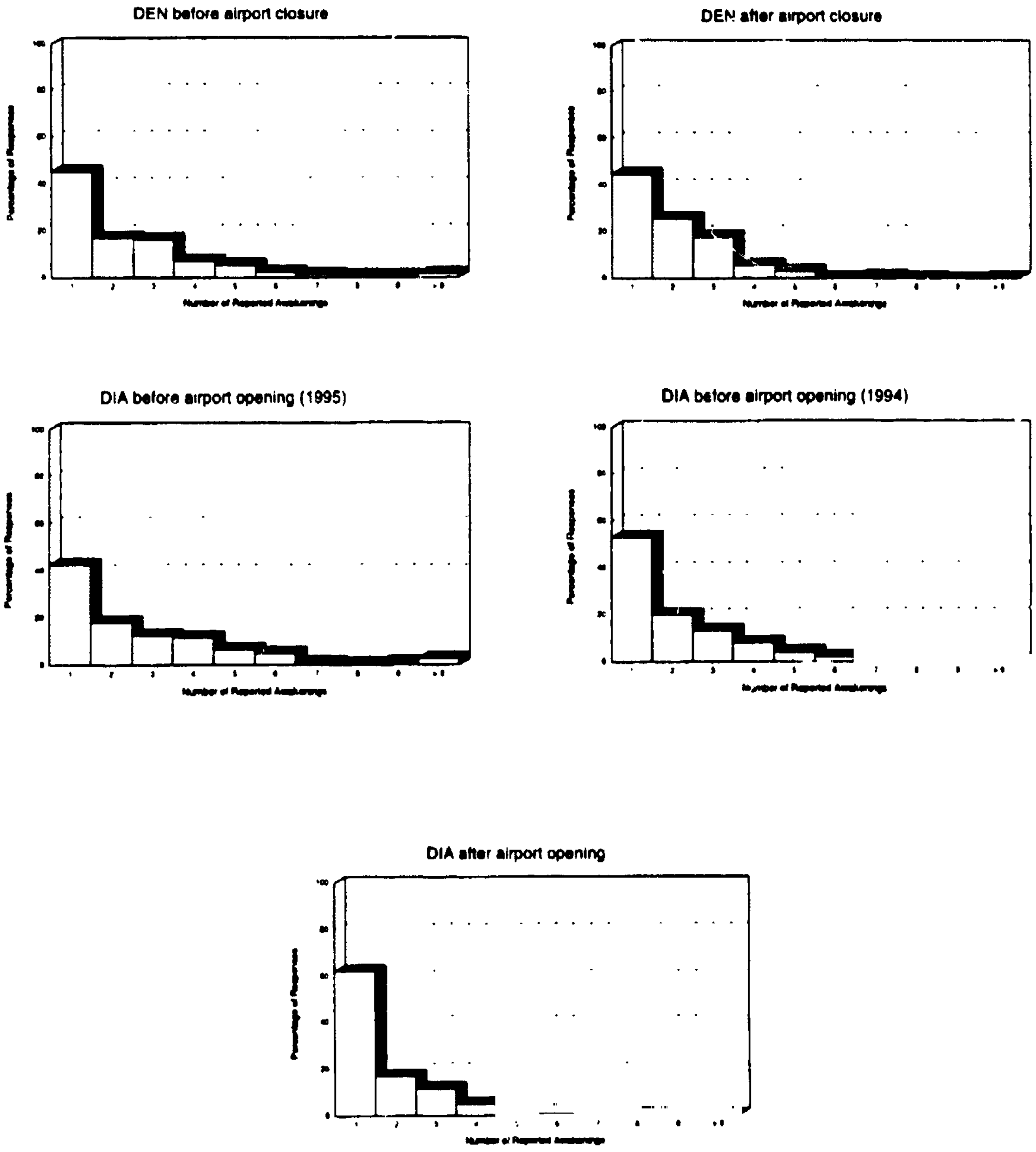


Figure 32 Summary of responses to: "How tired did you feel today?"


DEN before airport closure

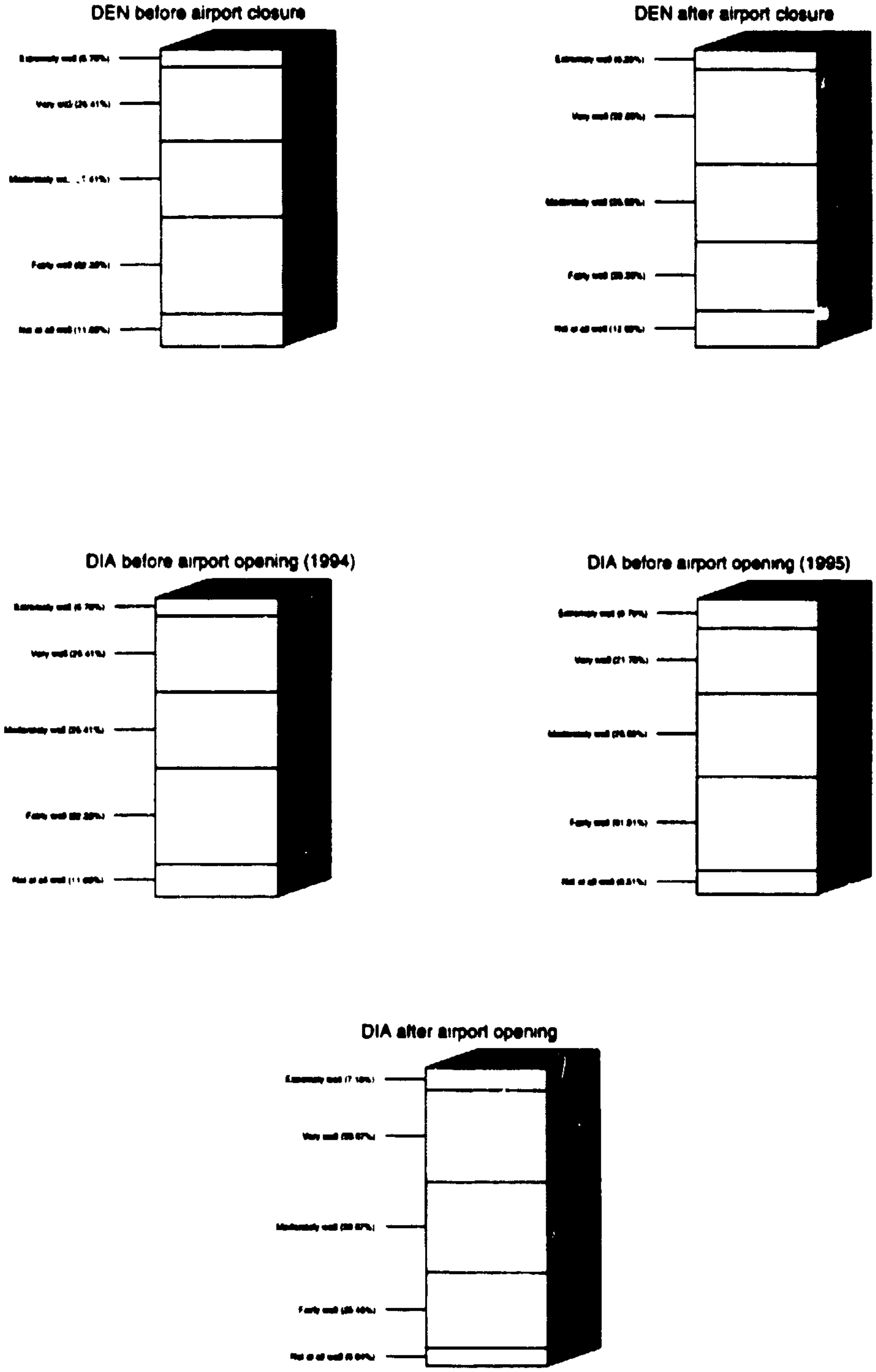
DEN after airport closure

DIA before airport opening (1995)

DIA before airport opening (1994)

DIA after airport opening

Figure 33 Summary of responses to: "How many times did you wake up last night?"


DEN before airport closure

DEN after airport closure

DIA before airport opening (1994)

DIA before airport opening (1995)

DIA after airport opening

Figure 34 Summary of responses to: "How well did you sleep last night?"

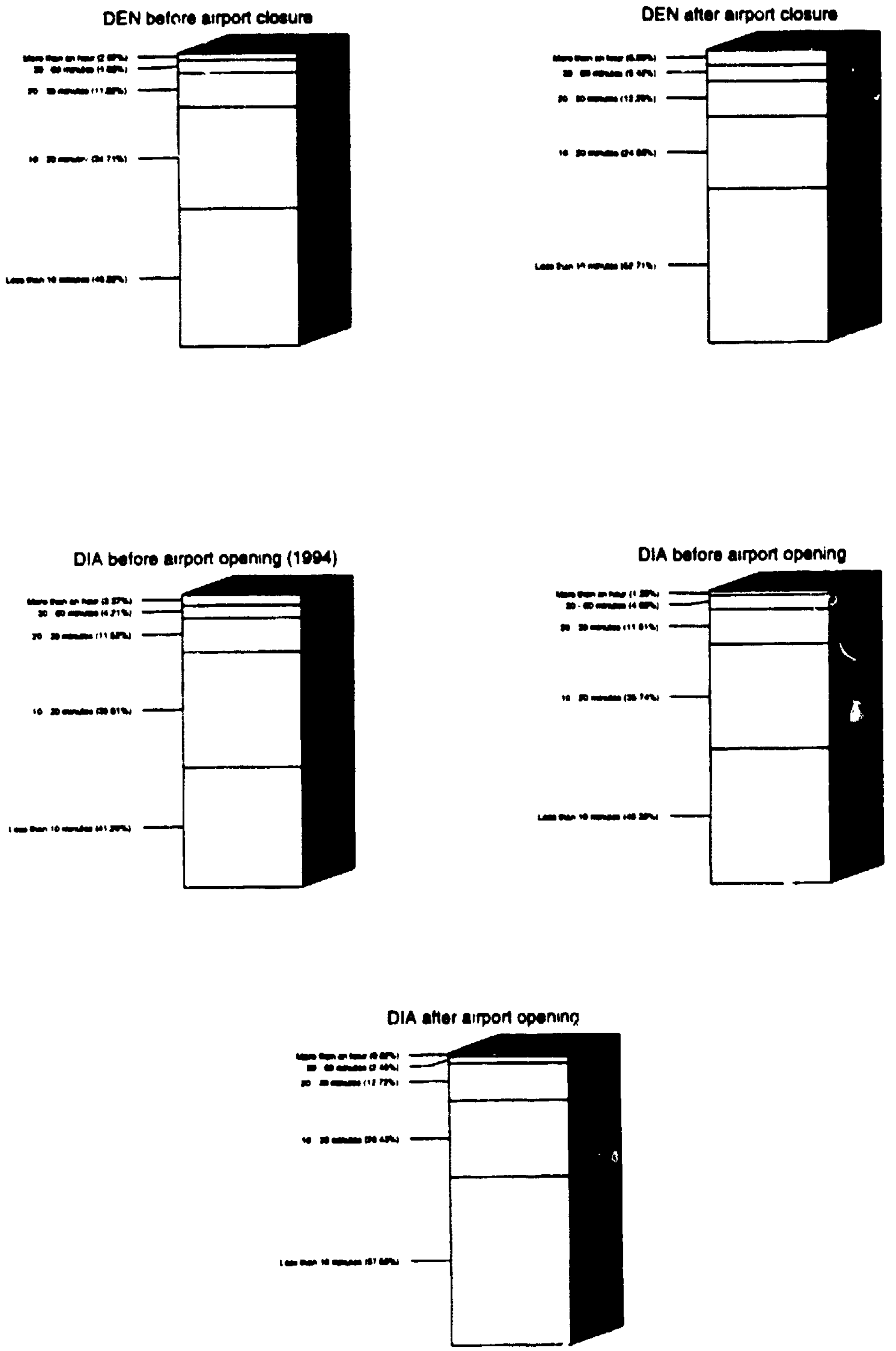
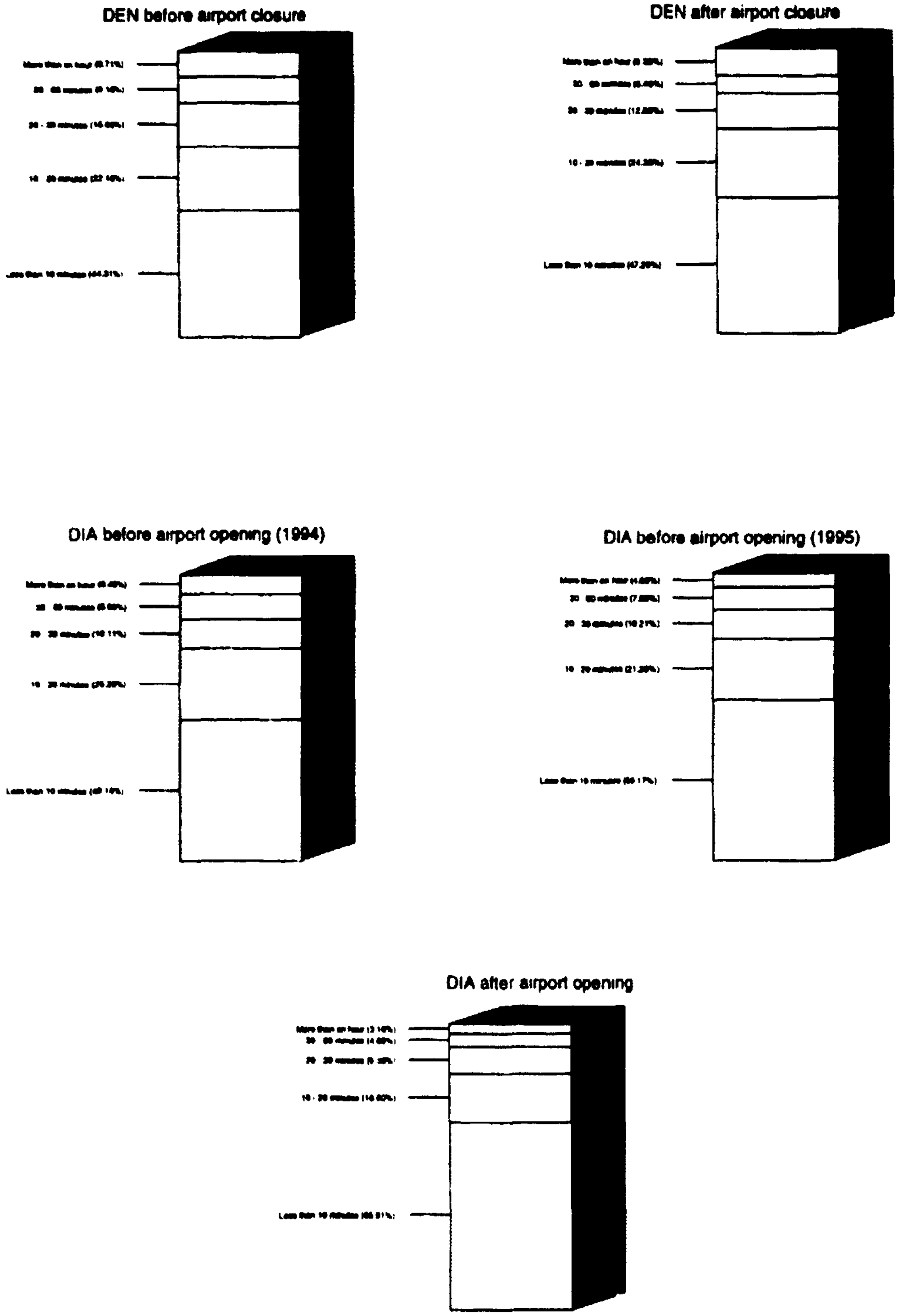


Figure 35 Summary of responses to: "How long did it take you to fall asleep?"



DEN before airport closure

DEN after airport closure

DIA before airport opening (1994)

DIA before airport opening (1995)

DIA after airport opening

Figure 36 Summary of responses to: "How much were you awake last night?"

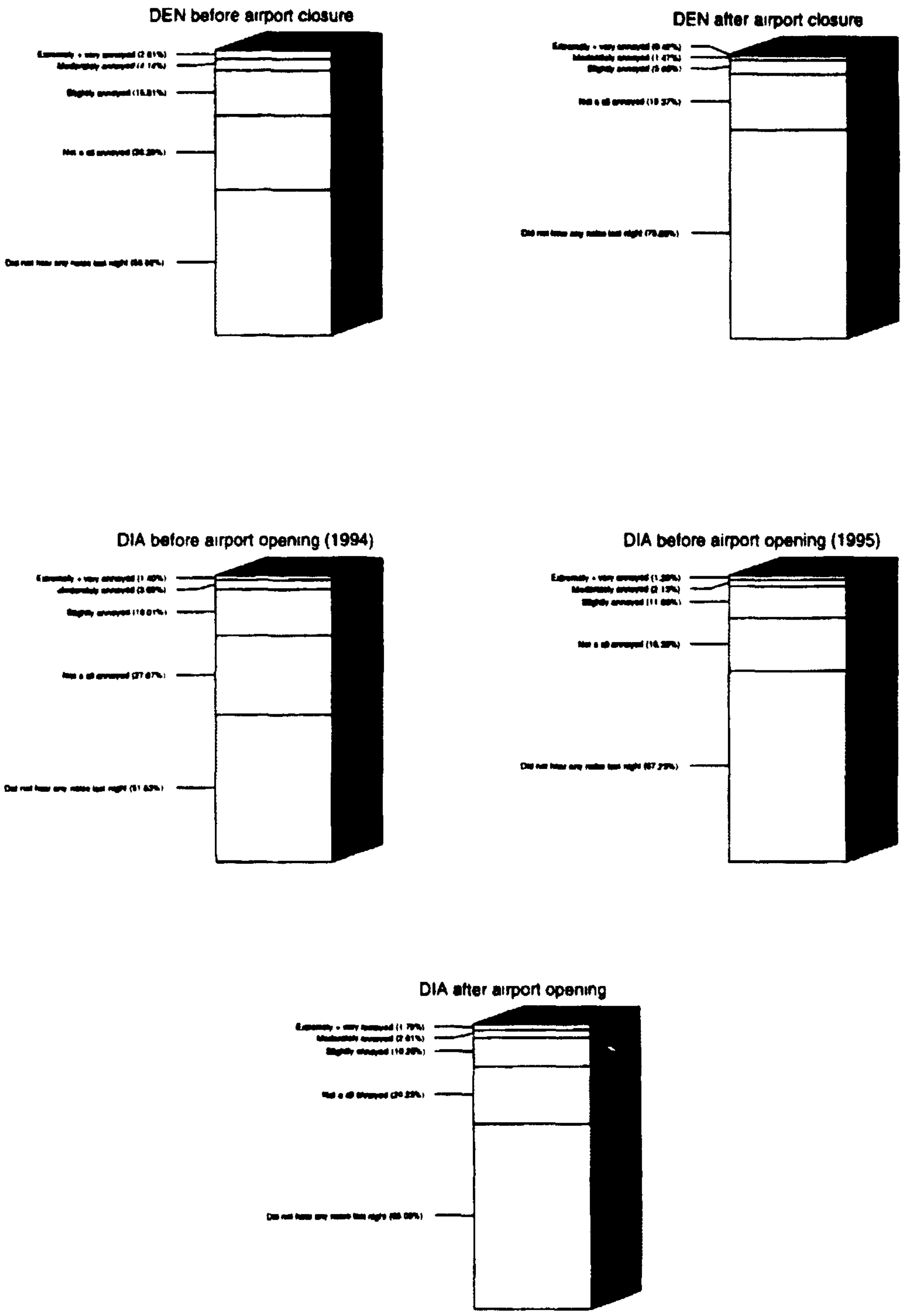


Figure 37 Summary of responses to: "How annoyed were you by noise last night?"

APPENDIX F RESULTS OF LOGISTIC REGRESSION ANALYSES

Tables 24 through 27 provide details of the logistic regression analyses of sleep disturbance by noise events discussed in Section 4.4.6. Each sleep disturbance measure is presented in a separate table. Each table shows the regression coefficients for each of 10 predictors, the odds ratios, 95% confidence interval for the odds ratios, and the contribution of each predictor to the model.

Regression coefficients (B) are of limited value in this nonlinear analysis, but are useful as indications of the direction of the relationship to each predictor with the probability of sleep disturbance. Positive coefficients indicate that an increase in the value of the predictor results in an increase in the probability of disturbance. The relative magnitudes of the coefficients do not indicate the relative strength of unique contribution to prediction of each variable, because variables are measured on different scales.

The odds ratio is a more easily interpreted transformation of the regression coefficient, defined as e^B. An odds ratio that is greater than one indicates not only that the likelihood of sleep disturbance increases with increasing magnitude of the predictor, but also the extent of increase in likelihood. For example, an odds ratio of 2 indicates that as the predictor increases by one unit, the odds in favor of disturbance doubles. Regression coefficients and their associated odds ratios are estimated values, subject to sampling error. The 95% confidence limits bound the likely range of values (odds ratios) given the sample data. A variable significantly enhances prediction of awakening if the confidence limits on its associated odds ratio do not include 1.0.

The last columns in Tables 24 to 27 present the results of a series of analyses of each sleep disturbance measure in which models are formed with each predictor separately removed from the full model containing all predictors. The performance of the reduced model for each predictor is then compared with the full model. A statistically significant result indicates that the model without a given predictor does a significantly poorer job of predicting sleep disturbance than one that includes that predictor, and thus serves as a test of the significance of prediction for that variable. This latter test is preferable to tests of variables based on confidence limits for odds ratios.

Table 24 Prediction of at least one blip measured by Swiss-made actimeter following within five minutes of noise events recorded indoors between 2200 and 0700 hours.

Variable (unit)	*B*	Odds ratio per unit	95% Confidence Interval for Odds Ratio		*F* to Remove df=1, 1326
			Lower	Upper	
Personal Characteristics					
Number of spontaneous awakenings (inverted and reflected)	0.157	1.17	0.78	1.75	0.60
Gender	0.413	1.51	1.44	2.01	8.17*
Age (linear, years)	0	0.99	0.98	1.02	0.02
Age (quadratic, 35-49 vs. others)	0.921	2.51	1.93	3.27	47.98*
Time related Characteristics					
Time since retiring (in 15 minute increments)	0.008	1.01	1.00	1.02	1.89
Duration of residence (months)	0	1.00	1.00	1.00	0.23
Number of nights in study	0.008	1.01	0.99	1.03	0.87
Pre-sleep characteristics					
Tiredness on retiring (scale of 1 to 5)	0.138	1.15	1.03	1.28	6.20
Noise characteristics					
Ambient level (dB)	-0.014	0.99	0.97	1.00	3.72
SEL (dB)	0.064	1.07	1.05	1.09	45.33*

$^*p < .005$

Table 25 Prediction of motility recorded by U.S.-made actimeter following within five minutes of noise events recorded indoors between 2200 and 0700 hours.

Variable (unit)	*B*	Odds ratio per unit	95% Confidence Interval for Odds Ratio		*F* to Remove
			Lower	Upper	df=1, 5093
Personal Characteristics					
Number of spontaneous awakenings (inverted and reflected)	-0.575	0.56	0.44	0.73	18.87*
Gender	0.680	1.92	1.70	2.29	76.56*
Age (linear, years)	0.041	1.04	1.03	1.05	54.32*
Age (quadratic, 35-49 vs. others)	-0.486	0.62	0.51	0.74	25.24*
Time-related Characteristics					
Time since retiring (in 15 minute increments)	0.061	1.06	1.06	1.07	288.71*
Duration of residence (months)	-0.002	1.00	1.00	1.00	10.15*
Number of nights in study	0.017	1.02	1.01	1.03	12.39*
Pre-sleep characteristics					
Tiredness on retiring (scale of 1 to 5)	-0.168	0.85	0.79	0.90	26.22*
Noise characteristics					
Ambient level (dB)	-0.022	0.98	0.97	0.99	34.69*
SEL (dB)	0.010	1.01	1.00	1.02	4.48

*$p < .005$.

Table 26 Prediction of an awakening by button push following within five minutes of noise events recorded indoors between 2200 and 0700 hours.

Variable (unit)	***B***	**Odds ratio per unit**	**95% Confidence Interval for Odds Ratio**		***F* to Remove df=1, 7674**
			Lower	**Upper**	
Personal Characteristics					
Number of spontaneous awakenings (inverted and reflected)	0.513	1.67	1.01	2.76	4.17
Gender	0.222	1.25	0.92	1.70	2.07
Age (linear, years)	0.037	1.04	1.01	1.06	9.63*
Age (quadratic, 35-49 vs. others)	-0.177	0.84	0.58	1.21	0.92
Time-related Characteristics					
Time since retiring (in 15 minute increments)	0.011	1.01	1.00	1.03	2.08
Duration of residence (months)	-0.002	1.00	1.00	1.00	3.28
Number of nights in study	-0.038	0.96	0.94	0.98	15.18*
Pre-sleep characteristics					
Tiredness on retiring (scale of 1 to 5)	0.574	1.06	0.93	1.21	0.73
Noise characteristics					
Ambient level (dB)	-0.058	0.94	0.92	0.96	54.17*
SEL (dB)	0.053	1.05	1.04	1.07	37.40*

*$p < .005$.

Table 27 Prediction of an awakening by U.S. actimetric criterion following within five minutes of noise events recorded indoors between 2200 and 0700 hours.

Variable (unit)	*B*	Odds ratio per unit	95% Confidence Interval for Odds Ratio		*F* to Remove
			Lower	Upper	df=1, 5093
Personal Characteristics					
Number of spontaneous awakenings (inverted and reflected)	-0.591	0.55	0.43	0.71	22.94*
Gender	0.261	1.30	1.12	1.50	12.43*
Age (linear, years)	0.026	1.03	1.02	1.04	24.61*
Age (quadratic, 35-49 vs. others)	-0.278	0.76	0.64	0.90	10.13*
Time-related Characteristics					
Time since retiring (in 15 minute increments)	0.004	1.00	1.00	1.01	1.48
Duration of residence (months)	<0.001	1.00	1.00	1.00	0.10
Number of nights in study	0.019	1.02	1.01	1.03	16.34*
Pre-sleep characteristics					
Tiredness on retiring (scale of 1 to 5)	-0.129	0.88	0.82	0.94	14.50*
Noise characteristics					
Ambient level (dB)	-0.021	0.98	0.97	0.99	35.79*
SEL (dB)	0.009	1.01	1.00	1.02	4.16

*$p < .005$.

APPENDIX G REPLICATION OF OLLERHEAD'S INFERENTIAL ANALYSIS

This appendix describes analyses of relationships between nighttime noise exposure and motility using the Swiss actimeter described in Section 3.4. The information analyzed was collected in the vicinity of DEN prior to airport closure.

The analysis methods described in this appendix closely resemble those of Ollerhead *et al.* (1992). The primary analysis thus is of the relationship between outdoor measurements and actimetric "blips." The unit of analysis was of epochs, as defined in Section 2.2.2. Epochs in the present study were 1 minute long for noise data and 30 seconds long for actimetric data.

G.1 METHOD

Methods used to extract information specifically for replication of Ollerhead's analysis are described in this section.

G.1.1 Duplication of Actimetric Analysis Algorithms

Ollerhead *et al.* predicted sleep onset by means of what they termed a "14,10" algorithm. Each night's data was searched to locate the first period of 14 consecutive non-movement ("0") epochs. Test participants were assumed to have been asleep 10 epochs into that period. Since analysis epochs were 30 seconds long, this definition was tantamount to declaring sleep onset five minutes into the first seven-minute long movement-free period of the night.

Ollerhead *et al.* defined "arousal" by searching actimeter records from the assumed onset of sleep for epochs during which a non-zero value of motility followed at least one epoch of no activity. Ollerhead *et al.* termed the epoch in which movement began a "blip." Each blip was considered an arousal.

Ollerhead *et al.* used a custom program, ACCORD, to detect sleep onset and arousal and to format the data for further processing. BBN/Probe, a commercial time-series analysis software package, was used in the current study to duplicate these processes. Comparisons of the outputs of BBN/Probe procedures and ACCORD revealed no differences in results. Figure 38 is a sample plot of the output from BBN/Probe procedures that demonstrates the replication of Ollerhead's blip classification algorithm.

G.1.2 Definition of Aircraft Noise Events

Ollerhead *et al.* defined an Aircraft Noise Event (ANE) as any event that exceeded an A-weighted level of 60 dB and simultaneously triggered three noise monitors, placed outside in a triangular pattern in the study area, within the time frame of an aircraft flying over the area. These ANEs were then matched against air traffic control logs for independent confirmation of the occurrence of a known overflight. Ollerhead's ANEs were thus limited to noise events exceeding 60 dBA that were independently confirmed as aircraft flyovers.

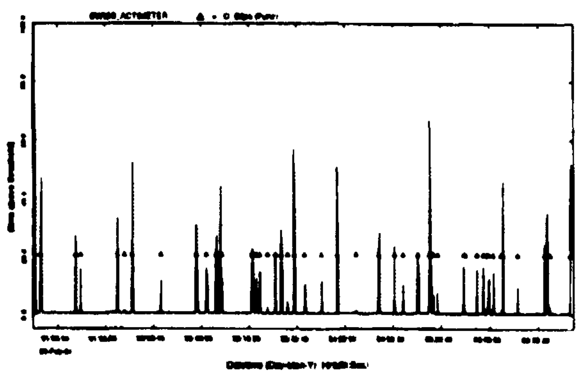


Figure 38 Demonstration of replication of blip classification algorithm.

Noise measurements in the current study were made both outside and inside participants' homes. An outdoor noise event occurred when the noise level exceeded 70 dB for at least two seconds. An indoor noise event occurred when the noise level exceeded 60 dB for at least two seconds. No attempt was made to eliminate noise events from sources other than aircraft. Thus, any noise event of whatever origin was an ANE in Ollerhead's terminology.

G.1.3 Data Epochs

Each actimetric data epoch in the Ollerhead *et al.* study was described by a set of summary statistics. Each epoch in the current study was similarly described by the following variables: date, time, subject, age, gender, length of residence, arousal (blip or no blip), L_{max} of event (if present), SEL of event (if present), self-rated tiredness during previous day, number of recalled awakenings next morning, annoyance due to noise, cumulative length of time spent awake during the night, and occurrence of a behavioral awakening response (button push).

G.1.4 Data Extraction Procedures

The data reduction procedures provided a quality control check for conformity with the constraints of the analysis. Following the procedure of Ollerhead *et al.*, each participant's nightly data were extracted for three time periods: between 1:00 and 1:30 AM, between 3:00 and 3:30 AM, and between 5:00 and 5:30 AM. Any night in which sleep started after 1:00 AM or ended prior to 5:30 AM was excluded from analysis.

A summary plot, shown in Figure 39, was automatically prepared and displayed during data reduction to evaluate indoor noise levels (1-minute L_{eq} values) and event levels, outdoor noise levels and event levels, unprocessed actimeter data, blips, and any behavioral awakening responses. Options were presented for saving only indoor data, only outdoor data, or both. Once suitable data were selected, the set of variables noted above was written to a file for later combination with all other participants' data for inferential analyses.

Separate data sets (one for indoor noise measurements and one for outdoor noise measurements) were produced to facilitate direct comparisons with the (outdoor only) noise measurements of Ollerhead *et al.*

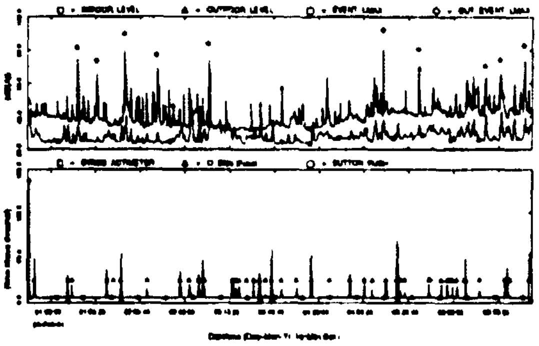


Figure 39 Example of display used to evaluate suitability of data for current analyses.

G.2 INFERENTIAL ANALYSES

The inferential analyses reported in this section are modeled as closely as practical on the multiple logistic regression analysis conducted by Ollerhead *et al.* (1992). Ollerhead's terminology and definitions of responses are preserved to the greatest extent possible. An exact replication of Ollerhead's statistical analyses was not possible for reasons described below. The general goal of these analyses was to identify variables that singly or in combination were strongly predictive of motility (that is, the occurrence of actimetric blips during analysis epochs).

Because field studies of sleep disturbance are observational rather than experimental in nature, no strictly causal relationship may be inferred between (say) noise exposure and sleep disturbance. However, multiple logistic regression analysis may yield a prediction equation -- a statistical "model" -- that can be used to quantify the individual and joint worth of noise and other variables as predictors of actimetric blips. If the level of a noise event, either by itself or in combination with a small number of other variables, proves to be a very reliable predictor of motility, then noise may be considered as a factor, if not a cause, in sleep disturbance.

G.2.1 Analysis Strategy

Ollerhead *et al.* (1992) identified a set of five variables from which they constructed a multiple logistic regression model to predict the occurrence of actimetric blips in analysis epochs. Several attempts are described below to undertake a similar multiple logistic regression analysis to determine the predictive utility of this set of five variables (plus three more found to be useful by Fidell *et al.* [1995]

in predicting behavioral awakening) in the present data set. These eight predictor variables were the sound pressure level of the noise event, the test participant's gender and age, the time of night, the sequential night of participation in the study, duration of residence in home, a self-rating of tiredness the previous night, and the test participant's susceptibility to sleep disturbance ("sensitivity").

Arguing that a within-subjects analysis constitutes a more conservative test of the predictability of actimetric blips, Ollerhead *et al.* adopted a random effects model for their multiple logistic analysis.[3] Differences between the size and nature of the current data set and that of Ollerhead *et al.* precluded an exact replication of the analysis performed by Ollerhead *et al.*, however. The present data set is not as large as that of Ollerhead *et al.*, and contains disproportionately fewer epochs with high level noise events. Furthermore, since Ollerhead *et al.* made no indoor noise measurements, separate data sets had to be constructed for indoor and outdoor noise measurements. All test participants with complete data for at least one night contributed data to this analysis. These data sets are summarized in Table 28.

Table 28 Summary of data sets for logistic regression analysis.

	Outdoor Noise Measurements	Indoor Noise Measurements
Noise Event Threshold	70 dB	60 dB
Number of Test Participants Contributing Data	27	28
Number of Noise Measurement Epochs in 3 time periods	53,673	68,904

G.2.2 First Approach to Replicating Analysis of Ollerhead *et al.*

An initial attempt to perform a logistic regression analysis of outdoor noise measurement data was made with EGRET, the statistical analysis program employed by Ollerhead *et al.* Ollerhead *et al.* selected EGRET in part because it allowed random effects models, and because they wished to consider individual differences in susceptibility to sleep disturbance as a potential predictor variable rather than simply as another source of error variance. EGRET was unable to provide a solution for the present data set, however. It failed to converge on a set of parameter estimates after more than 300 iterations.

---

[3] The random effects model was adopted as an approximation to a full within-subjects analysis. A within-subjects statistical analysis has fewer degrees of freedom for error and hence is more conservative than a between-subjects analysis unless individual differences are strong. A full within-subjects analysis may not be possible in a large data set composed of correlated observations and closely-related predictor variables, however, because of multicollinearity problems.

EGRET could not find a solution because of the wide range in numbers of epochs associated with each combination of predictor variables. For example, there were no epochs for analysis of some combinations of age, sensitivity, gender, time of night, and noise levels. Thousands of epochs were associated with other combinations of predictor variables, however.

G.2.3 Second Approach to Replicating Analysis of Ollerhead *et al.*

A different multiple logistic regression program, BMDPLR, was selected as an alternate means for performing multiple logistic regression analyses. BMDPLR does not permit random effects modeling directly, but can be used to predict actimetric blips from all combinations of predictor variables. The problem encountered in replicating Ollerhead's analysis through BMDPLR was in characterizing individual sensitivity to sleep disturbance as a predictor of actimetric activity.

Rather than treating sensitivity to sleep disturbance as a form of individual difference, Ollerhead *et al.* created categories of sensitivity to sleep disturbance for use as a predictor variable. In an effort to treat individual differences in the present study, they were coded at first as a categorical variable with as many categories as test participants. BMDPLR codes these categories as dummy variables, permitting as many variables as the number of test participants minus one.

A full logistic regression solution could not be obtained by approximating Ollerhead's sensitivity variable in this manner, however, because of multicollinearity produced by extremely strong associations between gender, age, and cases (individual participants). By including only some of the cases (participants) in the predictive model, individual differences were found to be strongly predictive of the occurrence of actimetric blips.

G.2.4 Third Approach to Replicating Analysis of Ollerhead *et al.*

A third approach modeled individual differences by coding household as a categorical variable. This was not an unreasonable strategy because no household had participants of the same age and gender. Associations among household, age, and gender were once again too strong to permit modeling of all households. Individual differences were found to be a good predictor of motility even with only some of the households modeled.

G.2.5 Fourth Approach to Replicating Analysis of Ollerhead *et al*

Sensitivity to sleep disturbance was finally approximated as a random effect (to partially account for nonindependence of observations of the same participant) by intentionally adding a component of random error, normally distributed with zero mean and unit standard deviation, to the estimated sensitivity. Test participants were assigned to a sensitivity category on the basis of their average number of blips across epochs, separately for indoor and outdoor noise measurements.

Table 29 shows cutoff criteria for defining high and low sensitivity categories. Remaining participants were assigned to the "intermediate" sensitivity category. Separate categorization was done for data sets based on indoor and outdoor noise measurements. Category scores of 2 (for low sensitivity), 6 (for intermediate sensitivity), and 10 (for high sensitivity) were assigned.

Table 29 Criteria for defining categories of sensitivity to sleep disturbance

CATEGORY	DESCRIPTION OF CATEGORY	NOISE EPOCHS
High Sensitivity	Overall blip response rate	> 9%
	Standard deviations from mean response rate	2
	Number of participants[*]	2
Low Sensitivity	Overall blip response rate	< 3.7%
	Standard deviations from mean response rate	1.4
	Number of participants[*]	1
Intermediate Sensitivity	All test participants (25) not in either of the above categories	

[*] The unit of analysis is not the individual, but the epoch. Each test participant may have contributed thousands of epochs to the analysis.

After addition of random error, cutoffs for the sensitivity categories were set at 4 and 8. Thus, the sensitivity category shifted when the random error component was greater than two standard deviations from the mean.

This coding scheme provided reasonable estimates of effects of sensitivity to sleep disturbance, but may have overestimated the statistical reliability of sensitivity as a predictor variable. The other coding schemes incorporating individual differences (another way of defining sensitivity), however, supported the strong effect of sensitivity. The lack of a true random effects model of sensitivity was also unlikely to have distorted the tests of other predictors. Those that were found to be unreliable would be even less likely to be statistically significant in a random effects model, and those that were found to be statistically significant reached a low enough probability level that confidence in the findings was warranted despite any underestimation of standard errors that may have occurred.[4]

Table 30 summarizes the treatment of continuous and categorical predictor variables. Category boundaries were chosen to the extent possible to be consistent with those selected by Ollerhead *et al.*

Table 30 Treatment of predictor variables for multiple logistic regression

VARIABLE	TYPE	CATEGORY BOUNDARIES AND UNITS
AGE	CATEGORICAL	20-34 years 35-49 50 or more
GENDER	CATEGORICAL	Male, female
TIME OF NIGHT	CATEGORICAL	0100-0130 hours 0300-0330 0500-0530
OUTDOOR NOISE LEVELS	CATEGORICAL	<75 dB 75-79 dB 80 dB or greater
INDOOR NOISE LEVELS	CATEGORICAL	< 65 dB 65 dB or greater
SENSITIVITY TO SLEEP DISTURBANCE	CATEGORICAL	Low, intermediate, and high.
DURATION OF RESIDENCE	CONTINUOUS	Months
SEQUENTIAL NIGHT IN STUDY	CONTINUOUS	Nights
SELF-RATED TIREDNESS	CONTINUOUS	Numeric value corresponding to rating scale categories

G.3 RESULTS OF MULTIVARIATE LOGISTIC REGRESSION

The logistic regression model predicts the probability of an actimetric blip in a given epoch. Epochs are characterized by values on each of the predictor variables. For example, a given epoch may be one in which noise was at a high level, occurring in the final time period of the 15th night in the study, for a woman of intermediate sensitivity in the lowest age category, who was moderately tired upon retiring, and who had been in her current residence for 5 years. A predicted response rate of 5% means that in an

---

[4] No attempt was made to replicate Ollerhead's Wilkinson-Diamond analysis, since the current data set did not permit estimation of the sensitivity variable in quiet periods.

epoch described by the these characteristics, the logistic regression model predicts a .05 probability of observing an actimetric blip.

G.3.1 Predictions for Outdoor Noise Measurement Data Set

The occurrence of blips in analysis epochs was well predicted by four variables: individual susceptibility to sleep disturbance, age, self-reported tiredness, and sequential night of data collection. No improvement in prediction was gained by including outdoor noise level, time of night, gender, and duration of residence among the predictor variables.

The predictive model based only on the four former variables fit the data well (Hosmer-Lemeshow $\chi^2(3) = 3.78$, $p = .29$). No significant difference was observed between this model and the one produced by all eight predictors, $\chi^2(5) = 4.24$, $p > .05$. Parsimony thus recommends the four-variable prediction model. All four predictors were strongly associated with the probability of an actimetric response, $p < .001$.

Table 31 shows average predicted percentages of actimetric blips for high-, intermediate-, and low-sensitivity participants in the three age groups, for an average tiredness level and night in study. Test participants in the intermediate age group reliably ($p < .001$) produced actimetric blips at a lower rate than test participants in older and younger age groups. Among test participants of intermediate sensitivity, an older or younger participant was about 20-25% more likely to produce a blip than one of intermediate age. The difference between younger and older participants was not statistically reliable, $p > .05$.

Table 31 Predicted percent of actimetric blips as a function of age group and sensitivity

Age	Sensitivity		
	High	Intermediate	Low
20-34	8.36%	5.84%	4.68%
35-49	6.78	4.69	3.76
≥ 50	8.19	5.69	4.58

The reliable difference in actimetric blip rate due to sensitivity reflects the way that the categories were defined.[5] For each of the age groups, participants with the highest sensitivity, as defined for this analysis, are about 80% more likely to produce actimetric blips than those with the lowest sensitivity.

A 6% increase in predicted actimetric blips was observed for each single category increase in tiredness rating (*e.g.*, from slightly to moderately tired). An 11% decrease in predicted actimetric blips was observed for each succeeding night of participation in the study.

G.3.2 Predictions for Indoor Noise Measurement Data Set

A predictive model of motility measurements based on two categories of indoor noise event levels and the additional seven predictors showed that gender, night in study, and time period failed to add significantly to prediction of responses. A model based on fewer predictor variables (sensitivity, indoor noise level, age, months in residence, and tiredness) also fit the data well, Hosmer-Lemeshow $\chi^2(8) = 5.31$, $p = .72$. There was no significant difference between the models, $\chi^2(4) = 3.26$, $p > .05$. Age, sensitivity, tiredness, and indoor noise level were all highly related to the likelihood of actimetric blips, $p < .001$, as was duration of residence, $p < .002$.

Table 32 shows the average predicted rate of actimetric blips for all combinations of age, sensitivity, and indoor noise level for which data were available, averaged over tiredness rating and duration of residence. High noise levels (A-level ≥ 65 dB) were rare in the sample of indoor measurements, so that some of the estimates may be unstable. Within epochs in which noise levels were lower than 65 dB, effects of age and sensitivity are similar to those for outdoor noise measurements. The quadratic effect of age, although reliable ($p < .001$), is weaker for indoor noise measurements. For example, a younger participant of intermediate sensitivity is about 15% more likely to awaken than a participant between 35 and 50 years of age. At the same time, the division of test participants into sensitivity categories is a bit stronger. Test participants of high sensitivity (as defined for this analysis) were about twice as likely to awaken as test participants of lowest sensitivity.

---

[5] Definitions of sensitivity categories are unavoidably arbitrary. Attempts to replicate Ollerhead's assignment criteria failed because none of the current participants was two standard deviations below the mean in overall response rate. The "effect" of sensitivity--differences in blip rates for the different sensitivity groups--depends completely on the category definitions.

Table 32 Percent of actimetric blips predicted by logistic model as a function of age group and sensitivity. Value in parentheses is number of epochs in category.

L_{max} (dB)	AGE (years)	ASSIGNED SENSITIVITY CATEGORY		
		High	Intermediate	Low
< 65	20-34	8.74% (893 epochs)	5.83% (31931 epochs)	4.43% (765 epochs)
	35-49	7.56 (418)	5.02 (16702)	3.80 (3406)
	≥ 50	8.18 (3136)	5.44 (11318)	4.13 (263)
≥ 65*	20-34	(0)	24.54 (30)	19.57 (1)
	35-49	(0)	21.73 (28)	17.21 (3)
	≥ 50	31.88 (6)	23.33 (4)	(0)

* Note that percentages in high noise categories are based on small numbers of epochs.

G.4 SUMMARY OF FINDINGS

Multiple logistic regression analysis yielded predictive models for both outdoor and indoor measurements of noise that show clear effects of age and of individual differences (coded in the manner of Ollerhead *et al.* as sensitivity to sleep disturbance). Participants in the middle age range (35-50 years) displayed less motility during the night than participants who were younger or older. Strong individual differences in response rates occured at all times of the night. Gender and time of night appeared unrelated to responses in both noise measurement locations.

Outdoor noise event levels appeared to be unrelated to participants' motility rates, while indoor noise levels were strongly related to them.

G.5 ROLE OF INDOOR NOISE EVENT LEVEL IN PREDICTION OF MOTILITY

The multiple logistic regression demonstrated a reliable effect of indoor noise level as a predictor of motility. Although the magnitude of the effect is difficult to evaluate in the present data set (because of the scarcity of high-level noise events), motility rates of participants exposed to high noise levels seem to have increased dramatically. Participants of intermediate sensitivity under the age of 50 may be more than 4 times more likely to respond when the A-weighted noise level is 65 dB or greater than at lower noise levels. As shown in Table 32, comparisons among older participants and those of higher and lower sensitivity are based on too few epochs to be interpretable. The effect produced by categorizing sensitivity to sleep disturbance on the basis of indoor noise data was slightly stronger than that for outdoor noise data, with the most sensitive participants about twice as likely to respond as the least sensitive ones.

G.6 COMPARISON OF FINDINGS WITH THOSE OF OLLERHEAD *et al.*

Table 33 summarizes findings from the current study and from that of Ollerhead (1992). The failure to replicate the outdoor noise effect found by Ollerhead *et al.* is probably due to differences in how those measurements were made, how events were defined, and the greater number and level of noise events in Ollerhead's data. Since Ollerhead *et al.* analyzed data from only their two noisiest sites in their logistic regression analysis, their data set included measurements for a noisier environment than that of the current study. In addition, Ollerhead's definition of a noise event included: (1) only aircraft overflights confirmed via control tower logs, and (2) only events that simultaneously triggered three noise monitors. Outdoor noise measurements in the current study were based on measurements collected by the noise monitor located nearest to each test participant's home.

Table 33 Comparison of current findings (outdoor noise measurements only) with those of Ollerhead *et al.* (1992).

PREDICTOR	FINDINGS OF CURRENT STUDY	FINDINGS OF OLLERHEAD *ET AL.*	COMMENTS
Outdoor Noise level	No effect	Positive linear effect	Probably due to different noise definitions and environments
Sensitivity to sleep disturbance	Positive linear effect	Positive linear effect	Essentially a recoding of individual differences
Time of night	No effect	Positive linear effect	Lack of effect may be due to differences in noise environments
Nights in study	Negative linear effect	No effect	No obvious methodological basis for difference in findings
Age	Quadratic effect: young and old more responsive	Negative linear effect	May be related to age distributions in study participants
Gender	No effect	Males more responsive	No obvious methodological basis for differences in findings
Duration of residence	No effect	Not evaluated	
Self-reported tiredness	Positive linear effect	Not evaluated	

Effects of sensitivity to sleep disturbance are guaranteed by the treatment in the original and replicated analyses, despite differences in definitions of sensitivity. While Ollerhead *et al.* failed to find a reliable relationship between responses and the sequential night in the study, the current study found a negative relationship: the longer the participants were in the study, the less motility they exhibited. On the other hand, the current study failed to replicate Ollerhead's finding of greater responsiveness over the course of a night.

Gender effects were found in both studies, but differed. Ollerhead found less responsiveness with increasing age while the current study found older and younger participants to be more responsive than those between the ages of 35-49. The current study failed to replicate Ollerhead's finding of greater responsiveness of male participants.

G.7 COMPARISON OF FINDINGS WITH THOSE OF FIDELL *et al.*

G.7.1 Logistic Regression Analyses

Table 34 compares results of the current study with those of Fidell *et al.* (1995). The logistic regression analysis performed by Fidell *et al.* was confined to epochs in which noise events occurred, and the dependent variable was an awakening defined by a button push rather than an actimeter blip. Further, the analysis by Fidell and colleagues was based solely on indoor noise measurements, which were incorporated in logistic regression analysis as a continuous predictor variable. Their analysis showed an increase in rate of awakening of about 6% with each decibel increase in SEL as compared with an estimated quadrupling of response rate for high *vs.* low level epochs in the current study.

Rating of tiredness showed a stronger effect in the study of Fidell *et al.*, who reported a 26% increase in response rate with each unit increase in tiredness rating, compared with a 7% increase found in the current study with indoor noise measurements. Sequential nights of study participation was not found to be related to response rate in either study with regard to indoor noise measurements.

The effect of duration of residence was similar in the two studies; a small positive effect was found in which the increase in response rate was less than 1% per month of residence. Participants in the current study had all resided in their homes for at least one year, so that effects of habituation to the noise environment could not be usefully estimated.

Table 34 Comparison of current actimeter analysis results, using indoor noise data, with behavioral awakening findings reported by Fidell *et al.* (1995).

PREDICTOR	FINDINGS OF CURRENT ANALYSIS	FINDINGS OF FIDELL *ET AL.* (1995)	COMMENTS
Indoor Noise level	Positive linear effect	Positive linear effect	Consistency of findings is noteworthy despite different noise definitions and environments
Sensitivity/individual differences	Positive linear effect	Individual differences effect	Different statistical treatment of these variables
Time of night	No effect	Positive linear effect	Defined as time since retiring by Fidell *et al.*
Nights in study	No effect	No effect	
Age	Quadratic effect: young and old more responsive	Negative linear effect	Treated as continuous variable by Fidell *et al.*
Gender	No effect	No effect	
Duration of residence	Positive linear effect	Positive linear effect	
Tiredness	Positive linear effect	Positive linear effect	
Performance of test participants as detectors of noise events while sleeping	$d' = .88$	$d' = .23$	Effect stronger in motility date by about 6 dB

Fidell *et al.* tested only the linear effect of age, and found a small negative trend. While the current study failed to replicate the linear trend, a fairly strong quadratic effect was evident.

G.7.2 Event-Detection Analyses

The indoor data of the current study were modeled as an event-detection process, as described by Fidell *et al.* In this analysis, an arousal (actimeter blip) was considered to be a consequence of a decision that a change had occurred in the short-term noise environment. This decision-making process is characterized by the ratio of "hits" (assertions that a signal is present when it is truly present) to "false alarms" (assertions that a signal is present when it is fact absent) that can be achieved (Green and Swets, 1966).

Epochs containing noise events as well as actimeter blips may be viewed as hits, while epochs containing actimeter blips but no noise events can be viewed as false alarms. The standard index of sensitivity is a scalar quantity known as d′. When d′ is zero, a detector has no information about the presence or absence of a signal and thus is completely insensitive to it. When d′ = 4, a detector can make essentially perfect decisions about the presence or absence of a signal.

The gross hit rate (as defined above) in the present data set was about 24%, while the gross false alarm rate was 5.6%. Assuming equivalent Gaussian distributions of numbers of epochs with and without noise events, the value of the sensitivity index, d′, which corresponds to this ratio of hits and false alarms is .88. The detection performance ot test participants in the study of Fidell *et al.* was at a level of .23. Thus, the current data show motility to be about 6 dB more sensitive to noise than behaviorally-confirmed awakenings.

G.7.3 Comparisons with Major Logistic Regression Analyses

The purpose of the logistic regression analysis described in this appendix was to replicate the analyses of Ollerhead *et al.* Analyses of motility and awakening described in the body of this report have a wider focus. One difference between analyses described here and the major analyses is the statistical treatment of individual differences. The current definitions of sensitivity to sleep disturbance are not predictively useful since they are arbitrarily based on *post hoc* evaluation of response data. Age, gender, night in study, *etc.* are known before analyzing the response data for a sample of participants. It is not possible to categorize test participants' sensitivity to sleep disturbance without independently measuring it prior to data collection. Similarly, the modeling of within-subjects effects (individual differences) is not predictively useful in the absence of external verification of them.

One of the major logistic regression analyses, based on responses measured by the Swiss actimeter, evaluated models with and without individual differences (within-subjects effects). Although individual differences have been found to be strong in all sleep research to date, it is of academic interest only to evaluate the gain in predictability offered by including individual differences in a model. Major interpretations in the body of the report are focused on the roles of "knowable" predictor variables.

The logistic regression analysis in this appendix was further limited by the treatment of some predictors as categorical variables, as per Ollerhead *et al.* This limitation may have diluted the effects of variables for which relationships with motility would be expected to be solely linear, such as noise level and time of night. On the other hand, the current analysis surprisingly found a quadratic effect of age, a finding that would have been obscured had age been modeled as a single continuous variable (capable of producing only a linear effect). Analyses in the body of the report therefore modeled quadratic effects as well as linear effects of age.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
December 1995

3. REPORT TYPE AND DATES COVERED
Contractor Report

4. TITLE AND SUBTITLE
Noise-Induced Sleep Disturbance in Residences Near Two Civil Airports

5. FUNDING NUMBERS
C NAS1-20101
WU 538-03-15-01

6. AUTHOR(S)
Sanford Fidell, Richard R. Howe, Barbara G. Tabachnick, Karl S. Pearsons, and Matthew D. Sneddon

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
BBN Acoustic Technologies Division
21120 Vanowen Street
Canoga Park, CA 91303-2853

8. PERFORMING ORGANIZATION REPORT NUMBER
BBN Report No. 8096

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Langley Research Center
Hampton, VA 23681-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NASA CR-198252

11. SUPPLEMENTARY NOTES
Langley Technical Monitor: Kevin P. Shepherd
Final Report - Task 1

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified - Unlimited

Subject Category 71

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A large-scale field study of noise-induced sleep disturbance was conducted in the vicinities of Stapleton International Airport (DEN) and Denver International Airport (DIA) in anticipation of the closure of the former and opening of the latter. Both indoor and outdoor measurements of aircraft and other nighttime noises were made during four time periods. Measurements were made in 57 homes located as close as feasible to the runway ends of the two airports. Sleep disturbance was measured by several indices of behaviorally confirmed awakening (button pushes upon awakening) and body movement (as measured with wrist-worn actimeters). A total of 2717 subject-nights of observations were made over the course of the study.

Although average noise event levels measured outdoors decreased markedly at DEN after closure of the airport and increased slightly at DIA after its opening, indoor noise event levels varied much less in homes near both airports. No large differences were observed in noise-induced sleep disturbance at either airport. Indoor sound exposure levels of noise events were, however, closely related to and good predictors of actimetrically defined motility and arousal.

14. SUBJECT TERMS
Noise-induced sleep disturbance, aircraft noise, behaviorially-confirmed awakening, actimeter, actigraph, noise measurements, motility measurements, nighttime noise exposure

15. NUMBER OF PAGES
135

16. PRICE CODE
A07

17. SECURITY CLASSIFICATION OF REPORT
Unclassified

18. SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

19. SECURITY CLASSIFICATION OF ABSTRACT

20. LIMITATION OF ABSTRACT