Direct Comparison of Two Statistical Methods for Determination of Evoked-Potential Thresholds

By

Ted L. Langford
James H. Patterson, Jr.

Aircrew Protection Division

July 1994

94-31397

Approved for public release; distribution unlimited.

United States Army Aeromedical Research Laboratory
Fort Rucker, Alabama 36362-0577
Notice

Qualified requesters

Qualified requesters may obtain copies from the Defense Technical Information Center (DTIC), Cameron Station, Alexandria, Virginia 22314. Orders will be expedited if placed through the librarian or other person designated to request documents from DTIC.

Change of address

Organizations receiving reports from the U.S. Army Aeromedical Research Laboratory on automatic mailing lists should confirm correct address when corresponding about laboratory reports.

Animal use

In conducting the research described in this report, the investigators adhered to the Guide for Laboratory Animal Facilities and Care, as promulgated by the Committee on the Guide for Laboratory Animal Resources, National Academy of Sciences-National Research Council.

Disposition

Destroy this document when it is no longer needed. Do not return it to the originator.

Disclaimer

The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation. Citation of trade names in this report does not constitute an official Department of the Army endorsement or approval of the use of such commercial items.

Reviewed:

Kevin T. Mason
LTC, MC, MFS
Director, Aircrew Protection Division

Released for publication:

Roger W. Wiley, O.B., Ph.D.
Chairman, Scientific Review Committee

David H. Karney
Colonel, MC, SFS
Commanding
Several statistical procedures have been proposed as objective methods for determining evoked-potential thresholds. Data have been presented to support each of the methods, but there have not been direct comparisons using the same data. The goal of the present study was to evaluate correlation and variance ratio statistics using common data. A secondary goal was to evaluate the utility of a derived potential for determining thresholds. Chronic, bipolar electrodes were stereotaxically implanted in the inferior colliculi of six chinchillas. Evoked potentials were obtained at 0.25, 0.5, 1.0, 2.0, 4.0 and 8.0 kHz using 12-ms tone bursts and 12-ms tone bursts superimposed on 120-ms pedestal tones which were of the same frequency as the bursts, but lower in amplitude by 15 dB. Alternate responses were averaged in blocks of 200 to 4000 depending on the size of the response. Correlations were calculated for the pairs of averages. A response was deemed present if

(Continued on next page)
the correlation coefficient reached the 0.05 level of significance in 4000 or fewer averages. Thresholds were defined as the mean of the level at which the correlation was significant and a level 5 dB below that at which it was not. Variance ratios were calculated as described by Elberling and Don (1984) using the same data. Averaged tone burst and tone burst-plus pedestal data were differenced and the resulting waveforms subjected to the same statistical analyses described above. All analyses yielded thresholds which were essentially the same as those obtained using behavioral methods. When the difference between stimulus durations is taken into account, however, evoked-potential methods produced lower thresholds than behavioral methods.
Table of contents

List of figures ........................................... 1
Introduction ............................................. 3
Methods ..................................................... 4
Equipment and procedures .............................. 5
Results and discussion ................................. 7
Conclusions ............................................... 9
References ................................................. 10

List of figures

Figure Page

1. Behavioral thresholds compared with conventional tone-burst evoked thresholds of present study. Tone-burst evoked thresholds represent combined data from correlation and variance ratio methods ................. 8

2. Conventional tone-burst evoked thresholds compared with derived thresholds. Tone-burst evoked thresholds represent combined data from correlation and variance-ratio methods. Correlation and variance-ratio derived thresholds are plotted separately ................. 9
Introduction

The chinchilla currently is the animal of choice for use as a model to establish the effects of various parameters of noise and blast exposure on auditory structure and function.

The behavioral methods in use for measuring hearing in the animal model are based on conditioning techniques (Miller, 1970; Blakeslee, et al., 1978). These methods provide valid and reliable measures of auditory thresholds when properly used, but do have limitations. First, the data may be affected by differing animal motivation from session to session and by the training procedures, particularly when multiple experimenters contribute to a single data pool. Second, the large investment of time and effort for both the training and measurement phases of an experiment limits the rate at which animals can be processed and makes the loss of an animal from a project extremely costly to the overall effort. A rapid, objective method for measuring hearing in the animal model would eliminate these limitations of the behavioral procedures.

A rapid method of threshold assessment, based on the detection of electrical responses of the auditory system to sound stimuli, now is being used in some laboratories (Bancroft, et al., 1991). A chronic electrode is implanted in the auditory system, usually in the inferior colliculus, and time-locked responses evoked by brief, rapidly repeated tone bursts are recorded, then averaged to improve the signal-to-noise ratio. Several stimulus presentations are made per second and a response can be detected visually in the averaged waveform after several hundred presentations. The thresholds obtained in this manner compare well with those derived from behavioral methods when signals of the same duration are used for both measures (Henderson, et al., 1983). Visual detection of the response is not objective, however, so it is subject to problems such as criterion differences between experimenters and criterion shifts across time for individual experimenters.

Several objective methods of detecting averaged electrical responses have been proposed for use in human evoked-response audiometry. Two of the methods outlined below have shown the most promise in terms of objectivity and reliability:

A statistical method based on correlation has been used by Arnold (1985), Salvi, et al. (1987), and Weber and Fletcher (1980). If time-locked responses to stimuli are present, the correlation between two averaged samples of those responses will be greater than that of the background activity. The background noise at the recording site is assumed to be random in the absence of a stimulus, so averaged samples of the noise taken at different times will not be correlated with one another. To
determine objectively the presence of responses, samples of activity time-locked to the stimulus are taken, averaged, and the correlation calculated between the two sets of averages. If the obtained correlation is greater significantly than that of the background, or zero, a response is present.

A second statistical method is based on evaluation of the ratio of the variance of the background noise to that of the response evoked by the stimulus (Elberling and Don, 1984; Don, Elberling and Waring, 1984). The variance for the background noise is estimated by sampling a single point of the response waveform across individual stimulus presentations. The variance of the response is calculated across points of the averaged recorded waveform. The ratio of the two variances follows the F-distribution and can be used to determine the presence of a response with any desired probability of error.

The third method, originally based on the visual detection of a derived response of the cochlear potential (Pantev and Pantev, 1982; Pantev et al., 1985), has been adapted by Berlin and his coworkers (Berlin et al., 1991; Hood et al., 1991) for use in evoked response work involving the central auditory system. A tone burst, some 20 dB above threshold, repeatedly is presented together with a continuous tone of the same frequency, but at a level near the expected threshold. The tone burst then is presented without the continuous tone. The responses to each of the stimulus configurations are averaged separately, and then the two sets of averaged data are differenced. The result is the response to the continuous tone. Lower thresholds have been obtained with this method than with more conventional methods. In addition, as the continuous tone has a narrower bandwidth than the tone burst, the derived response is more frequency-specific than can be obtained with conventional methods.

Each of these methods appears to have merit when compared with conventional, visual methods of threshold estimation, but they never have been directly compared with one another. The present experiment was designed to compare the two objective statistical methods with one another using both conventional tone-burst data and derived data from the differencing method of Berlin and his coworkers. The results will be compared with thresholds obtained using behavioral methods.

Methods

Six 1-to-2-year old male chinchillas bred and raised in the U.S. Army Aeromedical Research Laboratory (USAARL) colony were used as subjects. The animals were anesthetized by intramuscular injection of a combination of Ketamine (40 mg/kg) and Xylazine (2-5 mg/kg). The anesthetized animals were prepared for aseptic
surgery by plucking the fur from the top of the head, scrubbing
the plucked area, and swabbing it with Betadine. A lubricant was
applied to the eyes to prevent drying. The animal then was
placed in a stereotaxic instrument and draped. An incision was
made near the midline, the skin retracted, and the exposed skull
cleared of periosteum by scraping. The exposed portion of the
skull then was coated with a solution of silver nitrate followed
by a cyanoacrylate adhesive to provide a stable base for the
cement used in fixing the electrode in place. The skull was
marked according to previously determined stereotaxic coordinates
and a dental burr was used to make a 2 mm opening. After cutting
the dura, a concentric, bipolar electrode was lowered slowly into
the left inferior colliculus while electrical activity evoked by
a broadband click was monitored by an oscilloscope and an audio
system. At the proper depth, as determined by the best response
as well as by the stereotaxic coordinates, the electrode was
cemented in place. The incision was closed and a topical
antibiotic applied to the area.

Data collection was not initiated until a minimum of 2 weeks
was allowed for recovery from the procedure.

The animal was restrained for the electrophysiological
measurements in an apparatus designed for behavioral testing
(Blakeslee, et al., 1978). The apparatus fixed the animal’s head
with respect to movement in the horizontal and coronal planes,
but did not prevent pinna movement or rotation of the head about
the interaural axis.

Equipment and procedures

Stimuli for the experiment were computer-generated and
controlled. An inverse-FFT routine was used to generate the
digital representations of the acoustic waveforms with
experimenter-specified frequencies, durations, phases, and
amplitude envelopes. The stimulus waveforms were calculated
prior to the experimental sessions and stored in memory. During
experimental sessions, they were output via 16-bit D-to-A
converters and filtered appropriately for the sampling rate.
After conversion and low-pass filtering, the stimuli were
amplified, then led to an electrically-shielded speaker located
inside a double-walled, sound-insulated booth. The speaker was
placed 1 meter in front of the animal. The sound field in the
position occupied by the animal’s head was calibrated at each
frequency to be used prior to data collection. The calibration
results were used by the software to set the stimulus levels
desired for experimental conditions by means of a programmable
attenuator.
The stimuli were 12 ms tone bursts at 250, 500, 1000, 2000, 4000, and 8000 Hz. They had rise and decay durations of 6 ms. Previous evoked-potential work using these stimuli (Langford, Mozo and Patterson, 1989) showed them to have acceptable signal-to-noise ratios and bandwidths. The tone bursts were alternated with tone bursts temporally centered on a 120 ms tone of the same frequency, but at a level 15 dB below that of the burst. Since the time relation between the burst and the tone was fixed, the relative phases of the two differed at each frequency. At each frequency, the levels were varied in 5 dB steps over a 25 dB range encompassing the anticipated threshold. The number of repetitions was varied depending on the response magnitude obtained from each animal. In general, 200 to 400 repetitions for each level at each of four frequencies were run in each session. Each measurement session lasted approximately 45 minutes.

The experiment was designed so that the same data could be used to evaluate both of the statistical threshold determination methods under study. Data were collected beginning 3 ms after the onset of each tone burst for a period of 25.6 ms and for a like period of time beginning 51.2 ms after cessation of stimulation. This yielded background activity interleaved with responses to the tone bursts and with the tone bursts plus the long-duration tones. Each of the 25.6 ms data intervals contained 128 points. Data for each of the intervals were stored separately for off-line analysis.

The data for each stimulus condition were averaged later in blocks of 200 to 4000 repetitions. Each data epoch consisted of 128 points; however, adjacent points are not independent. The partial correlation between points has the effect of inflating the degrees of freedom used in the statistical tests. Since the lack of independence between points is a function of the frequency bandwidth used in the recording, autocorrelations were calculated for random noise samples generated with the same sampling rate and bandwidth used in data collection. The correlations approached zero when every fourth point was used in the calculations. Therefore, every fourth point was used in the computations of the correlations and variance ratios during data analysis.

Pairs of poststimulus waveforms, the results of averaging odd and even numbered responses to the stimulus presentations, were correlated (Pearson’s r) with one another for each stimulus condition. The variance ratio for the entire set of data used in the correlation computations was calculated as \( F = \frac{\text{VAR}(S)}{\text{VAR}(N)} \). \( \text{VAR}(S) \) is the variance of the averaged postsignal data taken during the temporal interval presumably containing the response. \( \text{VAR}(N) \) is the background noise variance calculated for a single point across all responses. This method of estimating the noise
variance has been shown to be an accurate measure of the variance of the true background noise (Elberling and Don, 1984). A response was deemed to be present if the 0.05 level of statistical significance was reached in 4000 or fewer trials. Threshold was defined as the mean of the lowest level at which a response was present and the level 5 dB below it at which no response was present.

Results and discussion

Tone-burst condition

With the exception of one frequency in one animal, there was perfect agreement between the thresholds obtained by the correlation and variance ratio methods for the traditional tone burst data. Since the two thresholds in that one case differed by only 5 dB, they were averaged and the audiograms produced by the two methods were considered the same. The tone burst thresholds are compared with a database of 118 behavioral thresholds obtained from the chinchilla by Patterson, et al. (1991) in Figure 1. The greatest differences occur at 0.5 and 8 kHz, where the separation between the two is only 3.8 dB.

The durations of the stimuli used in obtaining the two sets of thresholds were not the same, however. The evoked-potential data were obtained using stimuli with an effective duration of 4 ms (Dallos and Olsen, 1964), while the behavioral data are based on stimuli with durations exceeding the upper limit for temporal integration in the chinchilla. Henderson (1969), as well as others (Wall and Ferraro, 1981; Davis and Ferraro, 1983) have shown the chinchilla integrates the energy of an auditory stimulus in the same manner as human subjects, at least for durations up to 100 ms. This would place the 4 ms threshold stimuli used in the present study some 14 dB below those used for obtaining the behavioral thresholds. Thus, the sensitivity of either of the objective statistical methods, when applied to conventional evoked potential data, is greater considerably than that of behavioral methods.

A study by Arnold (1985), who used supra-threshold click-evoked data obtained from human subjects, also compared the correlation and variance ratio methods with a subjective, visual method for determining the presence of a response. At 10 dB above behavioral threshold, the lowest level used in that study, the correlation method was slightly more sensitive than the visual method and the variance ratio was the least sensitive. Her method for estimating the variance of the background activity was different from that used in the present study, however. In addition, two replications of 2000 averaged repetitions of stimulus-evoked activity interleaved with two runs of 2000
Figure 1. Behavioral thresholds compared with conventional tone-burst evoked thresholds of present study. Tone-burst evoked thresholds represent combined data from correlation and variance ratio methods.

averaged repetitions of background activity were used in those calculations, so that changes in the signal-to-noise ratio between samples could not be ruled out. Finally, no frequency-threshold values were obtained. Therefore, the results of that study cannot be compared directly with those of the present study.

Derived condition

The thresholds obtained from derived data using the two statistical methods did not show the high degree of correspondence found for the conventional tone-burst data. Differences of 5 dB between the correlation and variance ratio methods for at least one frequency in each of four animals were found. Because of these discrepancies, the thresholds for the two methods were treated separately. The two sets of thresholds are compared with the one obtained from the conventional tone burst method in Figure 2. Although others (Hood, et al., 1991) have reported lower thresholds for the chinchilla at the higher frequencies using the derived evoked-potential method, no such advantage was found in the present study. The thresholds based on the derived waveforms are parallel to and slightly higher than those obtained with the conventional method.
Figure 2. Conventional tone-burst evoked thresholds compared with derived thresholds. Tone-burst evoked thresholds represent combined data from correlation and variance ratio methods. Correlation and variance-ratio derived thresholds are plotted separately.

Several procedural differences between the two studies may explain the differences in the results. First, the thresholds reported by Hood, et al. (1991) were based on subjective, visual inspection of the data. Second, their evoked potentials were recorded from surface electrodes and may represent activity from different brain regions than those of the present study. Third, their continuous tones and tone bursts were not phase-locked as were those in the present study.

Conclusions

The objectively determined audiograms based on evoked-potential methods are parallel to those based on behavioral methods.

When the difference in stimulus duration between the evoked-potential and behavioral methods is taken into account, the evoked-potential method proves to be the more sensitive.

The derived-potential method, as implemented in the present study, is less sensitive than the conventional evoked-potential method.
References


Initial distribution

Commander, U.S. Army Natick Research, Development and Engineering Center
ATTN: SATNC-MIL (Documents Librarian)
Natick, MA 01760-5040

Library
Naval Submarine Medical Research Lab
Box 900, Naval Sub Base
Groton, CT 06349-5900

Commander
10th Medical Laboratory
ATTN: Audiologist
APO New York 09180

Executive Director, U.S. Army Human Research and Engineering Directorate
ATTN: Technical Library
Aberdeen Proving Ground, MD 21005

Chairman
National Transportation Safety Board
800 Independence Avenue, S.W.
Washington, DC 20594

Commander
Man-Machine Integration System
Code 602
Naval Air Development Center
Warminster, PA 18974

Commander
Naval Air Development Center
Technical Information Division
Technical Support Detachment
Warminster, PA 18974

Commander
Technical Support Detachment
Warminster, PA 18974

Commanding Officer, Naval Medical Research and Development Command
National Naval Medical Center
Bethesda, MD 20814-5044

Commanding Officer
Armstrong Laboratory
Wright-Patterson
Air Force Base, OH 45433-6573

Deputy Director, Defense Research and Engineering
ATTN: Military Assistant for Medical and Life Sciences
Washington, DC 20301-3080

Director
Army Audiology and Speech Center
Walter Reed Army Medical Center
Washington, DC 20307-5001

Commander, U.S. Army Research Institute of Environmental Medicine
Natick, MA 01760

Commander/Director
U.S. Army Combat Surveillance and Target Acquisition Lab
ATTN: SFAE-IEW-JS
Fort Monmouth, NJ 07703-5305
Commander
USAMRDALC
ATTN: SGRD-UMZ
Fort Detrick, Frederick, MD 21702-5009

Italian Army Liaison Office
Building 602
Fort Rucker, AL 36362

Directorate of Training Development
Building 502
Fort Rucker, AL 36362

Chief
USAHEL/USAAVNC Field Office
P. O. Box 716
Fort Rucker, AL 36362-5349

Commander, U.S. Army Aviation Center
and Fort Rucker
ATTN: ATZQ-CG
Fort Rucker, AL 36362

Chief
Test & Evaluation Coordinating Board
Cairns Army Air Field
Fort Rucker, AL 36362

Canadian Army Liaison Office
Building 602
Fort Rucker, AL 36362

German Army Liaison Office
Building 602
Fort Rucker, AL 36362

French Army Liaison Office
USAAVNC (Building 602)
Fort Rucker, AL 36362-5021

Australian Army Liaison Office
Building 602
Fort Rucker, AL 36362

Dr. Garrison Rapmund
6 Burning Tree Court
Bethesda, MD 20817