A comparison of single-subject and traditional statistical designs in steady-state evoked potential research

DOUGLAS R. EDDY

United States Air Force School of Aerospace Medicine, Brooks Air Force Base, San Antonio, Texas

and

SAMUEL L. MOISE

Nova Technology, Inc., Los Gatos, California

An experiment was conducted to assess the relationship between level of auditory noise and a measure derived from the steady-state evoked potential (SSEP) called relative transmission time (RTT). The visual stimulus consisted of a 150-fL light that was modulated at 45, 48, or 51 Hz to a depth of 30%. The auditory stimulus was bandpass-limited white noise at 65, 75, or 85 dB. The outcomes of single-subject and group statistical analyses were compared. Neither statistic indicated a significant effect for cross-modal white noise on the SSEP. Whereas the more traditional group approach indicated only large between-subject variability, the single-subject approach clearly indicated deviant data of one of the subjects.

Traditionally, researchers using physiological dependent measures have been plagued by two major problems: (1) Dependent variables often exhibit very large standard deviations, especially between subjects. (2) Because of the complexity of the measures, the research designs suffer from small subject sizes. The purpose of this article is to demonstrate a single-subject analysis technique on steady-state evoked potential (SSEP) data. In addition, the outcome will be compared with more traditional statistical analysis.

Visual SSEPs are elicited by presenting a continuously flickering light stimulus to the observer and recording the voltage changes in the region of the visual cortex. Regan (1972) provided an excellent review of the clinical and research uses of the SSEP, as well as the theory of its generation and relationship to other evoked potential measures. In general, when flickering unpatterned stimuli between 5 and 60 Hz are presented to a person, three ranges of enhanced EEG amplitudes are found around 10, 18, and 50 Hz. The 50-Hz range is of special interest, since this frequency range is usually above the fusion point of flicker. Therefore, using high-frequency stimulation provides the possibility of measuring SSEPs with minimal intrusion on a primary task.

One of the two primary measures of the SSEP is phase lag; amplitude is the other. Phase lag is the angular differ-

ence between sine-wave components of the stimulus input and the EEG at each frequency. If more than two stimulating frequencies are used, the relative transmission time (RTT) of the signal can be computed from the phase lag. The RTT is computed by dividing the slope of the phase-lag/frequency function by 360° . The validity of this measure is based on the assumption that the lag time is constant for all frequencies in the given range.

SSEPs have been shown to be sensitive to different mental workload manipulations (Wilson, 1980). However, simple stimulus conditions such as light color and brightness and other modes of stimulation such as auditory noise may also affect SSEPs. An experiment was conducted to assess the relationship between auditory noise level and RTT of the SSEP.

METHOD

Subjects

Six males ranging in age from 27 to 44 years served as subjects. All subjects underwent a photically driven EEG examination that was interpreted by a qualified neurologist. Some subjects suffered from mild astigmatism and nearsightedness, but all were able to easily identify the small symbol used as an eye fixation point during data collection. The voluntary informed consent of the subjects used in this research was obtained in accordance with AFR 169-3.

Apparatus

A schematic diagram showing the functional connec-

This study was supported in part by AFOSR Contract F49620-83-C-0102 to Samuel L. Moise. D.R.E. was supported by a research associateship from the National Research Council. Requests for reprints should be sent to Douglas R. Eddy, USAFSAM/VN, Brooks AFB, San Antonio, TX 78235-5000.

Figure 1. A schematic diagram of the subject and the apparatus.

tions and relationship of the subject to the apparatus is shown in Figure 1.

Stimuli. The steady-state visual stimulus was produced by a pair of flickering fluorescent light tubes mounted horizontally. The center-to-center distance between the tubes was 2.75 in., and the distance from the center of either tube to the surface containing a fixation point (a "plus" character) was .625 in. The cool white 8-W tubes were manufactured by Sylvania (Model F8T5/CW). A vertically mounted 2 × 2 ft gray board covered the tubes except for a 5-in.-square opening at the center through which the subject viewed the lights. The gray board was 2.625 in. from the surface containing the fixation point, and the subject's eye was 24.625 in. from the board.

The fixation point (0.125 in. square) was centered within the 5-in.-square opening and between the tubes. The fixation point subtended a visual angle of 0.27° at 27.25 in. from the subject's eye. The 5 × 5 in. opening was covered by a NDO.30 neutral-density Kodak filter. The border around the filter further restricted the opening to 4.5 × 4.5 in., which subtended 10.4° of visual angle at 24.625 in. from the subject's eye.

The fluorescent tubes were driven by a mixing circuit that received a sine-wave signal from a Hewlett-Packard

low-frequency oscillator (Model 202C), operating at 45, 48, or 51 Hz. A luminance of 150 fL was measured for 1° of visual angle by a Minolta luminescence meter focused on the fixation point from the subject's eye position. The modulation depth was 30% for each frequency. All data were recorded with the subject and the visual display inside an Industrial Acoustics Corporation (IAC) sound- and electricity-shielded chamber that measured 6 ft 11 in. × 7 ft 4 in. × 6 ft 6 in. high. Light level within the chamber was less than 1 fL, except for the stimulus light.

The auditory noise was presented over a standard Air Force Astrocom headset. The noise generator was an HP HOI-3722A set for Gaussian 0-15 kHz noise. The RMS amplitude of the noise was selected to be 65, 75, or 85 dB using a rotary selector. Each rotary setting was calibrated with an octave band noise analyzer and a high-precision microphone. The microphone was located in the same position as the subject's tympanic membrane relative to the headset.

Recording system. The EEG was recorded bipolarly from the midline occipital (Oz) and the midline parietal (Pz) sites of the 10/20 Electrode System using Beckman silver/silver-chloride electrodes. An electrode over the

right mastoid served as the patient ground. The Oz-Pz bipolar lead directly connected to a Beckman dynograph recorder (Model R711). The patient box was bypassed when it was discovered that it permitted some interaction of channels and provided some filtering of the signal. Amplification of the EEG was 20,000 to 1, with a 100-Hz low-pass filter and a 0.16-Hz high-pass filter. No 60-Hz notch filter was used.

The light energy was monitored by a photodetecting circuit built from two Radio Shack photodetectors. The detector circuit was mounted 1.5 in. below the lower tube. The output of this circuit was used to calibrate the depth of modulation using a Tektronix oscilloscope (Model TM515) at the beginning of each session. The output of the photodetector circuit was also directly connected to the Beckman amplifier, with all filter settings identical to those for the EEG. Amplification was 5 to 1.

The Beckman output was sent to a Hewlett-Packard Fourier Analyzer (Model 5441B), where the signals were digitized and analyzed for their frequency content. The signals were visually monitored on the Tektronix oscilloscope. No amplification of the signal was performed by the Fourier analyzer. A program inside the Fourier analyzer computed the average power spectra for each signal and the average cross-power from 10 1-sec epochs. The sample rate was 256 per second. The phase lag, the transfer function, and the coherence were computed and then forwarded to a DEC PDP-11/34 computer for storage in disk files. Further analysis of the data was performed with a set of programs that culled the phase-lag and coherence information from the data files on the PDP-11/34. The data were then moved to a PDP-11/70, where they were processed with the BMDP statistical package.

Experimental Procedure and Design

Subjects arrived at the laboratory rested and ready for a 1.5- to 2-hr session. A consent form was read and signed, electrodes were attached, and the subjects were seated before the stimulus display and fit with the ear-phone headset. Electrodes were removed and reattached if impedances were greater than 5 k Ω . Subjects were directed to be attentive to instructions given by the experimenter prior to each data-collection period. Each period lasted approximately 15 sec, and the data were sampled during the last 10 sec. In the first and last periods of the session, data were collected with the subjects' eyes closed, but with the subjects facing in the direction of the display lights. This procedure provided the control data for possible transients in the EEG record. During all other data-collection periods, the subjects were asked to fixate the plus between the two fluorescent tubes while remaining relaxed. An auditory noise level and a light modulation frequency were pseudorandomly selected for each of 27 periods, yielding 3 periods for each treatment condition. Subjects were instructed to look away from the lights or to close their eyes for the 30 sec between periods. Electrode impedance was verified at the end of the session.

Some periods were rerun when the subject reported difficulty maintaining the eye fixation throughout the 10 sec of data collection.

Data Manipulation and Statistical Analysis

Using BMDP, a two-step process was used to produce three RTT values for each decibel level (65, 75, 85 dB) for each subject. Initially, a slope from the phase lags at each of the three frequencies for a particular decibel level was computed. For example, to compute the first 65-dB slope, the phase lags at the first 45-, 48-, and 51-Hz periods were used. The second 65-dB slope was computed from the second 45-, 48-, and 51-Hz periods. The third 65-dB slope was computed from the last 45-, 48-, and 51-Hz periods recorded at 65 dB. Finally, the RTT values were produced by dividing the slopes by 360°. Each RTT value under a decibel level is considered a trial. The RTT values were analyzed with a typical ANOVA with all variables within subjects, and an ANOVA was computed for each of the subjects with trials as the random variable.

Fisher (1950) described the procedure for testing whether the aggregate of a set of independent tests of significance should be regarded as significant. The test is based on the fact that the sum of a number of chi-square (χ^2) values is χ^2 -distributed with the appropriate number of degrees of freedom. A probability can be transformed into its χ^2 equivalent by taking its natural logarithm, changing its sign, and doubling it. This χ^2 value has two degrees of freedom. In the present case, each subject's χ^2 computed from the probability of each subject's ANOVA can be summed to determine the significance of the aggregate.

RESULTS AND DISCUSSION

A traditional completely factorial ANOVA was run on the RTT values with fixed effects of trials and decibels, with subjects as the random variable. Neither the trials variable nor the trials $\times$ decibels interaction was significantly different from chance [$F(2,10) = 0.93$, $p = .43$, and $F(4,20) = .27$, $p = .895$, respectively]. The decibel variable was nearly significant [$F(2,10) = 3.83$, $p = .058$]. A single-subject ANOVA was computed using trials as the random variable. Table 1 lists the F values computed for each subject, along with the probabilities,

Table 1
Single-Subject Analyses of Slope Data:
Effect of Decibels

Subject	F	p	$-\log_e p$	df
1	23.3	.006	5.116	2
2	1.2	.387	.949	2
3	.2	.864	.146	2
4	.9	.468	.759	2
5	3.9	.116	2.154	2
6	1.4	.339	1.082	2
			10.206	12

Note— $\chi^2(12) = 20.412$, $p = .059$.

χ^2 transforms, degrees of freedom, and the probability of the aggregate's being different from chance. The F values in Table 1 had two and four degrees of freedom for each subject. In order to use the trials variable as a random variable in the ANOVA, it was assumed that there was no correlation between trials. Typically, this assumption is tested by correlating the data at different lags. In the present case, with the smaller number of trials, this technique was not attempted. However, the lack of a trials effect in the grouped ANOVA was weak support for the independence of trials. Furthermore, G. F. Wilson (personal communication, December 18, 1984) found that the phase lag of the SSEP is stable over weeks, months, and even 2 years when the stimulus conditions and the subject's state were held constant.

Note that both tests of significance produced similar probabilities for the subjects as a group. However, the within-subject analysis did not indicate the cause of near significance. From the single-subject analysis of Table 1, it was evident that the data of Subject 1 were the major contribution to the decible effect. Because it is a necessary requirement that RTT be computed from data that approximate a straight line, each subject's phase-lag data were plotted. Figure 2 shows that the data for Subjects 1 and 3 clearly violate the requirement of linearity. Each point represents an average of nine values, three periods under three decibel conditions. Whereas regression around three points is tenuous at best, the proportion of the variance accounted for by a straight line for Subjects 1 and 3 was .8 and .06. For the other four subjects, values of .99, .98, .93, and .95 were computed for a linear fit. The observation that two subjects did not meet the assumptions of calculating RTT could also have been discovered by carefully graphing the data for each subject.

From an analysis of the SSEP data of the current study, several conclusions can be drawn. First, cross-modal auditory noise level has little or no effect on RTT values. Second, not all subjects provide phase-lag data that meet the requirements for computing RTT. Third, single-subject analyses can be used to give a researcher greater insight into their data. Fourth, Fisher's test of significance for independent tests yields results compatible with traditional statistical tests. With regard to the last point, researchers must guard against the following point: It may

Figure 2. Mean phase lags for each subject at each frequency collapsed across both periods and decibels.

be that one subject changes in one direction significantly, whereas another subject changes in the opposite direction, also significantly. Combining probabilities may indicate a significant difference, whereas the grouped ANOVA would not because the changes would cancel. In this case, the researchers may be examining either random fluctuations in one population or two separate populations. If the separate-populations explanation is supported with additional research, then the researchers may be able to show the effects of their independent variables on each of the subpopulations in future research.

REFERENCES

- FISHER, R. A. (1950). *Statistical methods for research workers* (11th ed.). New York: Hafner.
- REGAN, D. (1972). *Evoked potentials in psychology, sensory physiology and clinical medicine*. London: Chapman and Hall.
- WILSON, G. F. (1980). Steady-state average evoked potentials as a measure of tracking difficulty. In G. E. Corrick, E. C. Haseltine, & R. T. Durst (Eds.), *Proceedings of the Human Factors Society—24th Annual Meeting* (pp. 678-680). Santa Monica, CA: Human Factors Society.