

Computer-managed inferential statistical analysis of EEG data

F. J. BREMNER, M. YOST, and R. MCKENZIE
Trinity University, San Antonio, Texas 78212

Much of the neuroelectrical/psychophysiologic research of the past has involved the use of "group-type" statistics. Often, this has resulted in large variances that make it difficult, if not impossible, to find significant differences. Using the INNOVATE technique of single-subject design, we are able to consistently demonstrate significant differences in EEG spectra related to experimental and clinical effects previously masked.

Analyzing EEG data has always been difficult. Prior to the advent of the computer, EEG analysis was done by visual inspection. In the early 1960s, EEG data analysis changed rapidly as fast digital computers became available (Adey, Dunlop, & Hendrix, 1960). In the 1960s, the analog EEG signal was converted by hand to a numerical series (A/D) acceptable to digital computers (Bremner, 1964). This hand digitizing was very laborious. Fortunately, electronic A/D conversion was soon available (Elazar & Adey, 1967).

Once the EEG was A/D converted, various descriptive statistics were applied (Adey & Dunlop, 1960; Bremner, 1964). It is now standard practice to use the fast Fourier transformer (FFT) to describe EEG data (Benignus & Mueller, in press). The FFT computes the squared voltage (i.e., power) for each frequency in Hertz as requested (Walter, 1963). The result of applying the FFT is a voltage by frequency plot of the EEG (spectrogram) that is sampled. In practice, the A/D conversion and FFT of necessity are computer-resident tasks. However, the result of this level of analysis is still equivalent to a descriptive statistic. In order to make comparisons between psychological tasks, an appropriate inferential statistic is needed. This paper presents an argument for the applications of standard computer-based multivariate statistics to the EEG data of a single subject. The rationale of applying multivariate statistics to single-subject designs has been presented by Kelly, McNeil, and Newman (1973), Kratochwill (1978), and Shine (in press-a, in press-b). Therefore, this paper will concentrate on the special case of applying computer-based multivariate statistics to the EEG data of a single subject.

Our data are gathered from individual subjects, and the subject is considered to be the individual unit of statistical analysis. Once data are gathered from a single subject in a single session, they are submitted to the FFT.

This research report was supported by USAFOSR Contracts F 49620-79-C-0139 and F 33617-78-D-0629-0026. The authors would like to express their appreciation to Gary Macomber and Robert Woods for their assistance in gathering and analyzing the data used in this paper.

The spectra resulting from the FFT are further computer analyzed by applying the multivariate statistics available in the *Statistical Package for the Social Sciences* (SPSS) computer programs (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975). We have referred to this technique as INNOVATE (individual nominal variate analysis technique). To our knowledge, this is the first time this form of statistical analysis has been applied to neurological/psychophysiologic data.

METHOD

Subjects

The subjects in this study were four male college students who were paid for their participation. Each subject had a right and left occipital scalp electrode attached.

Apparatus

Standard electrophysiological amplifiers (Data Corporation 2124) were used to amplify the EEG signal before it was sent to an analog tape deck. The EEG data were sent from the tape deck to an A/D converter controlled by a Texas Instruments (TI) 960A process-controlled computer. The A/D-converted EEG signal was then transmitted by hard wire from the TI 960A computer to an IBM 3031 computer. The IBM 3031 performed the FFT on the A/D-converted EEG data. The IBM 3031 was managed by a VM operating system. Under this operating system, the FFT program, the multivariate statistics program, and the SPSS program are disk resident. The multivariate and SPSS programs are controlled from a terminal in the laboratory or our offices.

Procedures

The subjects were trained on two tasks: meditation and biofeedback. The description of these tasks is detailed elsewhere (Bremner, Benignus, & Moritz, 1972; Hyman, 1978). Twenty seconds of EEG data were collected on each subject for each task, using an A/D sampling rate of 256 counts/sec. We maintain all of our data for the entire time interval in which they were gathered and do not merge across time intervals. That is, if we collect 20 1-sec pieces of information, we maintain them as separate seconds of information and do not merge them in any way across the 20 sec. Following application of the FFT, the data are indexed and filed. An index is attached to each second's data. Typically, this index contains five parts: subject number, session number, experimental condition, side, and lead location. Once the data have all been indexed, they are filed in three locations. First, we make a historical IBM card file. Second, we make a file in our conversational monitoring system (CMS)

machine within the IBM 3031. The CMS machine is accessed through any one of the terminals in the laboratory or in our offices. The powerful CMS commands permit us to easily transform, merge, correct, delete, or manipulate the data. The third file is an SPSS file. The SPSS program is one of a number of commercially available computer programs that can be used to analyze EEG data. The SPSS file is completely labeled to insure that all output is easily read and interpreted. SPSS permits a great deal of statistical power in analyzing data. For example, if a particular file contains data from 4 to 14 Hz and we are interested in analyzing four 2-Hz windows between 4 and 11 Hz, we can use compute statements and develop subfiles that enable us to perform this analysis. Also, if data are missing, the program permits us to either omit those data points or substitute estimated values.

Each of the three files is developed at different points in the data gathering sequence. The historical IBM card file and the initial CMS file are developed immediately after each experimental session. After the last data are gathered and entered into the two initial files, they are reviewed in the CMS file. Once they are "cleaned up" in the CMS, the SPSS file is built. This sequence of file development insures that all data are reviewed as they are gathered and that the data that are analyzed have been "cleaned up."

RESULTS

A tutorial example of the concepts described above in a single-subject statistical design follows. We recently completed this single-subject experiment four times. Each of the four subjects was run under two different treatment conditions on each of 3 days. Two EEG leads were used on each side of the head. For purposes of example, analyzed data are described for only one EEG lead.

Figure 1 represents a block design for the study. Twenty 1-sec samples were taken for each of the two treatment conditions, on each of 3 days, and the indexing system described above was employed with the data. Data for 4 through 14 Hz were gathered. The hypothesis tested involved the EEG theta band, and we analyzed four 1-Hz bands (4, 5, 6, and 7 Hz).

The individual subject was considered as the unit of analysis. Initially, we set out to determine whether there were significant differences for a subject between days within a condition (intrasubject reliability). A discriminant analysis was used to compare Day 1 with Day 2 and Day 2 with Day 3 within each condition. Using the SPSS program, each analysis was computed using a series of "SELECT IF" and "COMPUTE" statements. Using these two types of statements with the indexing system permitted us to analyze the data for each subject for each condition for each day without having to physically handle the data. The results of these analyses indicated that three of the four subjects produced consistent responses within the four 1-Hz windows across days. The subject who did not respond consistently was omitted from further analysis. The rationale for omitting this subject was that we did not feel it appropriate to include an individual whose reaction under a

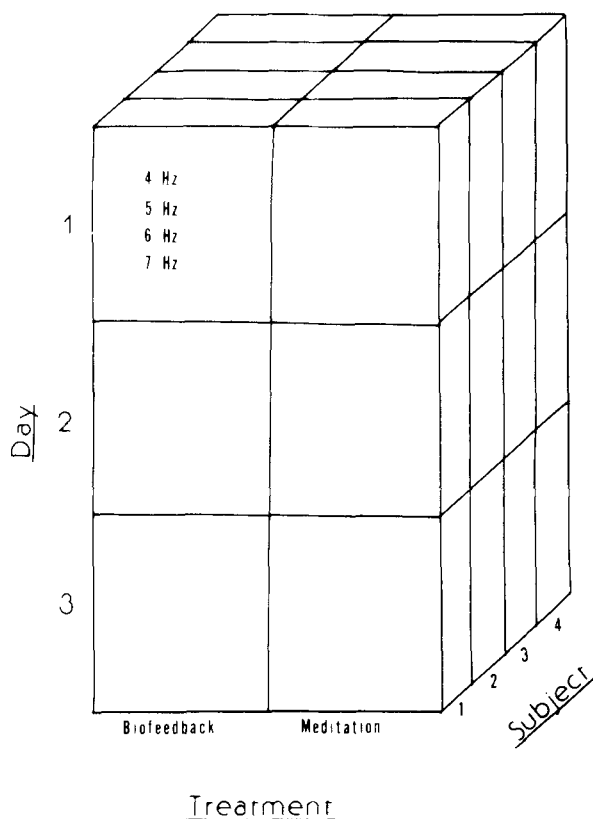


Figure 1. Experimental design.

treatment condition may have been idiosyncratic to a day. That is, this subject did not exhibit intrasubject reliability.

The next level of the analysis involved the use of a multivariate analysis of variance (Hughes, LaRue, & Yost, 1969), with a single subject as the unit of analysis. The efficacy of this approach has been presented by Shine (in press-a, in press-b). The multivariate statistical test employed initially tested the interaction of condition and days within each subject. In this test, the difference in EEG spectra between conditions was consistent across days for each of the three subjects. The main effect for differences in EEG spectra between treatment conditions was also tested. For each of the three subjects, the difference between conditions across days was significant. We interpreted this as indicating that each of the subjects was responding differently within the theta EEG band under the two different conditions.

The standardized and nonstandardized discriminant function coefficients indicate that the differences in EEG spectra between treatment conditions for each of the three subjects are not always attributable to differences in the same 1-Hz bands. That is, the differences between the treatment conditions is consistent across subjects but not consistent in terms of the 1-Hz bands

on which the differences are dependent. From an electrophysiological point of view, these results are interpretable, indicating that there are both task and individual subject differences in the reactions of individuals under these treatment conditions. Had we used a more traditional statistical approach, we would have sought to show that there were significant differences between conditions when all three (or four) subjects were combined.

For tutorial purposes, we performed the analysis using all three subjects. The results indicate that there are no significant differences in EEG spectra between treatment conditions when the data are merged across subjects. The differences that were found between conditions within the individual subjects were "washed out" in the merging of subjects.

In much of the research conducted in recent years, we have used "group-type" statistics. The experimental designs were used to investigate differences between the EEG spectra of subjects under different treatment conditions. The differences in EEG spectra, although large, were always accompanied by large standard deviations. In many instances, the standard deviations were considerably larger than the means. These large standard deviations made it difficult to find significant differences in EEG spectra, even though the differences in the means were large. Using a single-subject design, we obtain consistency in EEG spectra of a subject under a particular treatment condition across days and significant differences in EEG spectra between treatments. Using these designs and statistical approaches, the results are consistent with the predictions of neuropsychological theory. Also, we now feel that some of the multivariate statistical applications with groups of subjects that we used in the past may have masked significant effects that were present in the data.

DISCUSSION

Since the statistics applied to EEG data have often not been powerful, and in many cases only descriptive, brain researchers have not been rigorous in controlling situational variables (Gevins, 1981). Now that computer-based mathematical and statistical rigor are readily available (SPSS, SAS, etc.), researchers must be more diligent than ever before in controlling independent and potentially confounding variables. Brain researchers now have a measuring stick as reliable as the data collection technique.

The techniques for increasing data reliability are to design homogeneous baseline tasks and to use repeated-measures designs. With single-subject experiments, this means testing the same subject on multiple occasions on a task (e.g., meditation) that produces homogeneous EEGs. Once the homogeneous baseline has been established, an experimental task or treatment can be added and repeated on multiple occasions. This methodological approach works well not only in the laboratory, but also in the clinic.

REFERENCES

- ADEY, W. R., & DUNLOP, C. W. The action of certain cyclohexamines on hippocampal system during approach performance in the cat. *Journal of Pharmacology and Experimental Therapeutics*, 1960, **130**, 418-426.
- ADEY, W. R., DUNLOP, C. W., & HENDRIX, C. E. Hippocampal slow waves. *A.M.A. Archives of Neurology*, 1960, **3**, 74-90.
- BENIGNUS, V., & MUELLER, K. *Computer systems requirements for information flow in EEG*. Chapel Hill, N.C.: Environmental Protection Agency, in press.
- BREMNER, F. J., BENIGNUS, V. H., & MORITZ, F. EEG correlates of attention in humans. *Neuropsychologia*, 1972, **10**, 307-312.
- ELAZAR, Z., & ADEY, W. R. Spectral analysis of slow frequency components in the electrical activity of the hippocampus during of attention in humans. *Neuropsychologia*, 1972, **10**, 307-312.
- ELAZAR, Z., & ADEY, W. R. Spectral analysis of slow frequency components in the electrical activity of the hippocampus during learning. *Electroencephalography and Clinical Neurophysiology*, 1967, **23**, 225-240.
- GEVINS, A. S. The use of brain electrical potentials (BEP) to study localization of human brain function. *International Journal of Neuroscience*, 1981, **13**, 27-41.
- HUGHES, E. F., LARUE, R., & YOST, M. *Multivariate analysis of variance on small computers*. Miami: Clyde Computing Service, 1969.
- HYMAN, P. S. *Hemispheric specialization and ideomotor activity*. Unpublished master's thesis, Trinity University, 1978.
- KELLY, F. J., MCNEIL, K., & NEWMAN, I. Suggested inferential statistical models for research in behavior modification. *Journal of Experimental Education*, 1973, **41**, 4, 54-63.
- KRATOCHWILL, T. R. *Single subject research*. New York: Academic Press, 1978.
- NIE, N. H., HULL, C. H., JENKINS, J. G., STEINBRENNER, K., & BENT, D. H. *Statistical package for the social sciences*. New York: McGraw-Hill, 1975.
- SHINE, L. C. Effects of serial dependencies in single-subject data. *Educational and Psychological Measurement*, in press. (a)
- SHINE, L. C. Multivariate single-subject ANOVA. *Educational and Psychological Measurement*, in press. (b)
- WALTER, D. O. Spectral analysis for electroencephalograms: Mathematical determination of neurophysiological relationships from records of limited duration. *Experimental Neurology*, 1963, **8**, 155-181.