Statistical Practice

A Graphical Technique for Displaying Correlation Matrices

mes A. KOZIOL, Joel E. ALEXANDER, Lance O. BAUER, Samuel KUPERMAN, Sandra MORZORATI,
an J. O'CONNOR, John ROHRBAUGH, Bernice PORJESZ, Henri BEGLEITER, and John POLICH

recent multilaboratory study to investigate phenotypic d genotypic markers for alcoholism is described. A eliminary investigation of the reliability of electroen- phalographic data amassed at six laboratories was under- ken; data from each laboratory could be usefully summa- ed with sample correlation matrices. A graphical method r depicting these correlation matrices is presented. The rntz–Perlman procedure for assessing the equality of rrelation matrices can immediately be incorporated into is graphical technique. The experimental data from the eliminary investigation are illustrated with the graphical ethod, which readily conveys large amounts of correla- nal information and reveals meaningful patterns in the ctroencephalographic responses.

EY WORDS: Electroencephalography (EEG); Event- ated brain potential (ERP); Fisher z transformation; rntz–Perlman procedure; P300.

1. INTRODUCTION

As part of a large-scale, multilaboratory program de-

signed to investigate phenotypic and genotypic markers for alcoholism, six identical neuroelectrophysiologic laborato- ries have been established at various locations in the United States (CA, CT, IA, IN, MO, NY). The laboratories were designed to acquire 19 channels of electroencephalography (EEG) data (Scharbrough et al. 1990) in subjects under a variety of event-related brain potential (ERP) paradigms. Each laboratory was constructed with identical hardware, computer systems, and software so that EEG and ERP data from large numbers of subjects could be obtained in a stan- dardized and replicable fashion.

A preliminary study was undertaken to assess the reliabil- ity of collecting ERP data from these laboratories (Alexan- der et al. 1994). A total of 90 young adult males from the six sites ($N = 15$/laboratory) were recruited; subjects had no personal or familial psychiatric or neurologic disorders, and were neither alcohol nor drug users. EEG activity was recorded at 19 electrode sites located over all major cor- tical areas, with additional electrodes placed near the eyes to assess electroocular activity. Each subject was presented with two different auditory stimuli binaurally to elicit the P300 ERP component by using a typical procedure. The overall average waveforms computed across the $N = 90$ subjects for each electrode site are illustrated in Figure 1. The large, positive-going deflection occurring at approxi- mately 300 ms (and labeled at the Pz electrode site) is the P300 from the "target" stimulus to which the subject re- sponded (and did not react to the "standard" stimulus so that no P300 was produced).

The primary analyses were conducted on the P300 ERP component, which is most prominent over the center of the scalp (Fz, Cz, Pz electrodes). This cognitive brain po- tential is believed to reflect the neural activity associated with attentional and memory mechanisms (Donchin, Karis, Bashore, Coles, and Gratton 1986; Pritchard 1981) and has proven quite useful in the assessment of neurologic and psychiatric disorders (e.g., Pfefferbaum, Ford, White, and Roth 1989; Polich 1991; Polich, Ladish, and Bloom 1990; Polich, Pollock, and Bloom 1994; Pritchard 1986). The ERP data employed here were limited to consideration of P300 waveform morphology from each electrode site for each subject. It is important to determine whether ERP record- ings in subjects drawn from the same population (normal young adult males) under identical experimental conditions produce similar component measures at the various lab- oratories, and if so, to characterize these measures. Al- though substantial interindividual variability in the ampli- tudes of P300 waveforms is to be expected, it was hypoth-
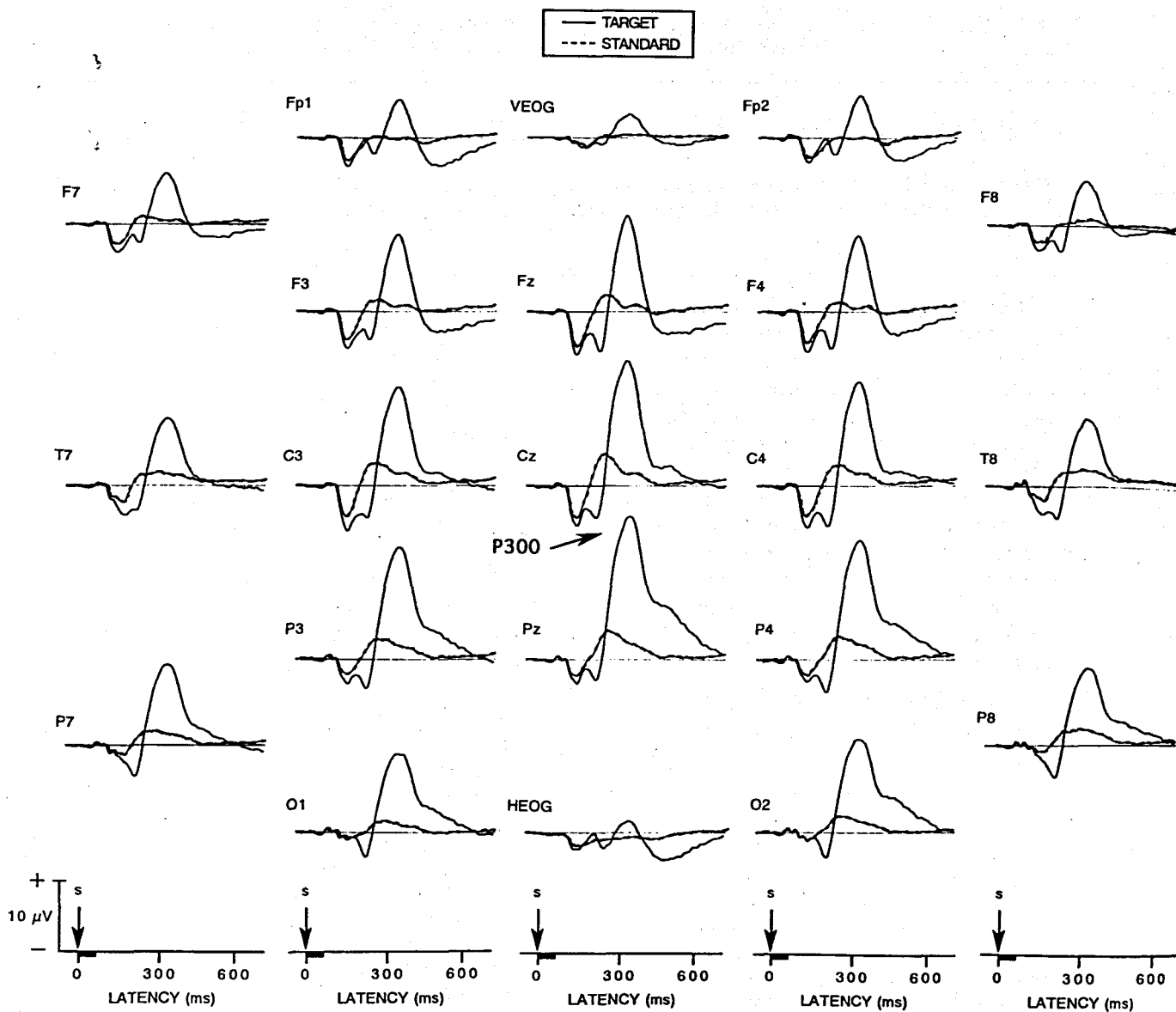
Figure 1. *Grand Averaged Event-Related Brain Potentials Taken Over N = 90 Normal Subjects from 19 Electrode Locations Over the Scalp. The amplitude and time base are illustrated at the bottom for the ERP data from each electrode site for the target and standard auditory stimuli. Subjects were required to respond only to the target stimuli, and these produced the P300 component, as labeled at the Pz recording site. The amplitude (strength or height) of each P300 component from each subject was obtained and used to construct the correlation matrices obtained from each of the six laboratory locations. See text for additional details.*

esized that correlations in waveform morphology (which are scale-invariant) should be relatively homogeneous. To examine this hypothesis, data from each laboratory were summarized in a 19 × 19 correlation matrix that reflected the strength of intersubject associations at the 19 electrode sites. Assessment of whether the correlation matrices derived at each laboratory are equal can be accomplished using a procedure proposed by Larntz and Perlman (1988); if homogeneity is not rejected, data can be pooled across the six laboratory sites to characterize comprehensively an overall pattern of waveform morphology. In the next section the Larntz–Perlman (LP) procedure is reviewed, and a graphical procedure for displaying correlation matrices is described. It turns out that the particular structure of the LP test allows for its graphical depiction by this procedure

also. In the concluding section the experimental data are reexamined using the methods of Section 2.

2. METHODS

2.1 The Larntz–Perlman Procedure

Let $\mathbf{R}^{(1)}, \ldots, \mathbf{R}^{(k)}$ denote the $p \times p$ sample correlation matrices from $k \geq 2$ independent p-variate normal distributions with underlying population correlation matrices $\mathbf{P}^{(1)}, \ldots, \mathbf{P}^{(k)}$, respectively. It is of interest to test the null hypothesis

$$H_0 : \mathbf{P}^{(1)} = \cdots = \mathbf{P}^{(k)} = \mathbf{P} \qquad (2.1)$$

where $\mathbf{P} = (\rho_{ij})$ is an unspecified nonsingular correlation matrix, against the general alternative of nonequality.

Suppose each $\mathbf{R}^{(m)} = (r_{ij}^{(m)})$ is based on a random sample of size $n_m, 1 \leq m \leq k$. For fixed i and j let

$$z_{ij}^{(m)} = \frac{1}{2} \ln \left[\frac{1 + r_{ij}^{(m)}}{1 - r_{ij}^{(m)}} \right] \qquad (2.2)$$

which is the Fisher z transformation of the sample correlation coefficient $r_{ij}^{(m)}$, and let

$$\mu_{ij}^{(m)} = \frac{1}{2} \ln \left[\frac{1 + \rho_{ij}^{(m)}}{1 - \rho_{ij}^{(m)}} \right] . \qquad (2.3)$$

As $\min\{n_1, \ldots, n_k\} \to \infty$ the random vector $\mathbf{Z}_{ij} = (z_{ij}^{(1)}, \ldots, z_{ij}^{(k)})$ has an asymptotic k-variate normal distribution with mean vector $\boldsymbol{\mu}_{ij} = (\mu_{ij}^{(1)}, \ldots, \mu_{ij}^{(k)})$ and covariance matrix $\Delta = \mathrm{diag}[(n_1 - 3)^{-1}, \ldots, (n_k - 3)^{-1}]$.

univariate hypotheses

$$H_{ij} : \mu_{ij}^{(1)} = \cdots = \mu_{ij}^{(k)}, \qquad 1 \leq i < j \leq p; \qquad (2.4)$$

note that H_0 in (2.1) is the intersection of the $p(p-1)/2$ distinct H_{ij}. The hypothesis H_{ij} can be tested with the statistic

$$S_{ij} = \sum_{m=1}^{k} (n_m - 3)[z_{ij}^{(m)}]^2 - \frac{\left[\sum_{m=1}^{k} (n_m - 3) z_{ij}^{(m)} \right]^2}{\sum_{m=1}^{k} (n_m - 3)} ; \qquad (2.5)$$

under H_{ij}, S_{ij} has an asymptotic χ_{k-1}^2 distribution, and suitably large values of S_{ij} would lead to its rejection (Rao 1973). If $n_1 = \cdots = n_k = n$, then (2.5) simplifies to

$$S_{ij} = (n - 3) \sum_{m=1}^{k} (z_{ij}^{(m)} - \bar{z}_{ij})^2 \qquad (2.6)$$

where

$$\bar{z}_{ij} = k^{-1} \sum_{m=1}^{k} z_{ij}^{(m)}. \qquad (2.7)$$

Larntz and Perlman proposed the statistic

$$T = \max_{1 \leq i < j \leq p} S_{ij} \qquad (2.8)$$

for testing H_0: H_0 would be rejected at level α if $T > \chi_{k-1,\varepsilon(\alpha)}^2$, where $\chi_{k-1,\varepsilon}^2$ is the upper ε percentage point of the χ_{k-1}^2 distribution and $\varepsilon(\alpha) = (1 - \alpha)^{2/p(p-1)}$. From an inequality of Khatri (1967), this procedure yields a possibly conservative level α test of H_0. Larntz and Perlman (1988) suggest that the nominal significance levels of tests based on T should be reliable for small sample sizes and for singular sample correlation matrices, and that the power properties of tests based on T should be quite satisfactory relative to other tests of the same hypothesis.

2 The Graphical Procedure

A $p \times p$ correlation matrix $\mathbf{R} = (r_{ij})$ can be viewed in p graphs, where the ith graph displays the ith row of $\mathbf{R}$, namely, the $r_{ij}, 1 \leq j \leq p$. Clearly, there is much flexibility in the choice of graphical display for the rows. The most effective graphical displays are those that bring out the essence of the underlying data in some transparent manner. Because the data in the present study are derived from a clinical experiment with underlying spatial characteristics, the proposed graphical representation here is chosen to reflect this application. Hence for the present study the individual graphs are drawn so as to represent the spatial positions of the individual electrodes on the skull rather than, say, in a conventional rectangular grid. Furthermore, variation in colors is adopted so as to indicate the magnitudes of the correlation coefficients. This graphical technique is immediately applicable also to display the components $S_{ij}, 1 \leq j \leq p$, of the LP statistic from (2.5) or (2.6) because the S_{ij} can be displayed in the form of a $p \times p$ matrix. In this regard note that the S_{ij} can be considered (proportional to) variance estimates of the correlation coefficients in the z scale; moreover, the LP statistic (2.8) may be found from examination of these graphs. (The row-wise correlations might also be depicted in the monotonically transformed z-scale. However, most clinical investigators seem more accustomed to correlation coefficients, and willingly accept the notion that the S_{ij} are appropriate variance estimates, or the $\sqrt{S_{ij}}$, standard errors.) Additional details concerning the graphical procedure and its utility are described next in the context of the experimental data.

3. APPLICATION TO THE EXPERIMENT

As outlined above, six different laboratories recorded EEG activity at 19 electrode sites from a total of $N = 90$ normal individuals using the same methods and a typical ERP paradigm. The data were then analyzed such that P300 waveforms yielded amplitude (i.e., component size) measures for each electrode site and subject. A matrix of pairwise correlations of the P300 amplitudes (i.e., the amount of strength of the electrophysiological response) at the 19 sites across the subjects was then calculated for each laboratory. The hypothesis of interest is whether the underlying population correlation matrices for each laboratory are identical. The overall Larntz–Perlman test statistic T [see (2.8)] is found to be 15.53, $p = .47$, for assessing this hypothesis. Thus there appear to be no systematic differences in P300 variation over the scalp at the six different laboratory sites.

Because the null hypothesis of equality of the underlying correlation matrices was not rejected, it is reasonable to take $\bar{\mathbf{R}}$, the average of the six sample correlation matrices, as an estimate of the common $\mathbf{P}$. In Figure 2 are displayed three rows of $\bar{\mathbf{R}}$: namely, the average correlations of the P300 amplitudes in relation to the three midline electrode sites, the Fz (frontal), Cz (central), and Pz (posterior) where P300 is largest, with the amplitudes at the other electrode sites. The color figures were constructed by averaging across an arbitrary rectangle using linear interpolation to estimate the interelectrode correlation values. This figure suggests that P300 amplitude: 1) is highly consistent across the laboratory recording sites at the point of maximum amplitude,

CORRELATIONS   **DISPERSIONS**

Fz

Cz

Pz

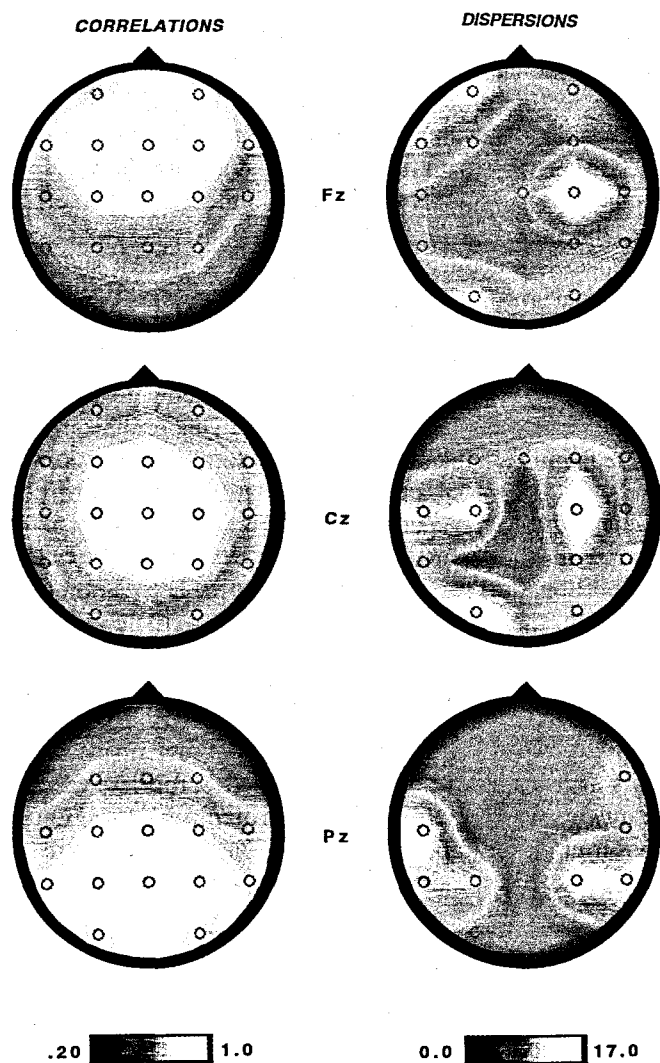.20 ▮▮▮▮ 1.0   0.0 ▮▮▮▮ 17.0

Figure 2. Graphical Representations of Three Rows of the Average Correlation Matrix (Left) and Corresponding Test Statistics (Right) of the Sample P300 Component Data. Each circle depicts a correlation coefficient (left) or measure of dispersion, Equation (2.6) (right), over the 19 different recording locations spaced equally across the scalp. The Fz, Cz, and Pz designations refer to the major midline electrodes over the frontal, central, and parietal portions, respectively, of the scalp. Note that the correlational strength moves from the frontal to the parietal sites just as P300 amplitude also varies. See text for further details.

ent electrode sites on the skull. Depicting the spatial relationships of the electrode sites in this manner transparently conveys to the investigators the spatial information inherent in the standard EEG recording paradigm. Of course, it is to be expected that different graphical displays would be adopted in other contexts. The selected technique of color variation was adopted so as to capitalize on one's general visual capability of easily apprehending general color patterns, and hence patterns of large differences. (Scales of gray tones might also be used, although in our experience colors more effectively convey both small as well as large differences.) Broad correlational patterns readily emerge from even cursory inspection of the 19 individual graphical displays, of which the three shown in Figure 2 are representative. The ability of such displays to readily convey large amounts of correlational information suggests that these methods hold considerable promise for the rapid evaluation of correlation matrices.

REFERENCES

Alexander, J. E., Polich, J., Bloom, F. E., Bauer, L., Kuperman, S., Rohrbaugh, J., Morzorati, S., O'Connor, S. J., Porjesz, B., and Begleiter, H. (1994), "P300 from an Auditory Oddball Task: Interlaboratory Consistency," *International Journal of Psychophysiology*, 17, 35–46.

Donchin, E., Karis, D., Bashore, T. R., Coles, M. G. H., and Gratton, G. (1986), "Cognitive Psychophysiology and Human Information Processing," *Psychophysiology: Systems, Processes, and Applications*, eds. M. G. H. Coles, E. Donchin, and S. W. Porges, New York: Guilford Press, pp. 244–267.

Khatri, C. G. (1967), "On Certain Inequalities for Normal Distributions and Their Applications to Simultaneous Confidence Bounds," *Annals of Mathematical Statistics*, 38, 1853–1867.

Larntz, K., and Perlman, M. D. (1988), "A Simple Test for the Equality of Correlation Matrices," in *Statistical Decision Theory and Related Topics IV, Vol.2*, eds. S. S. Gupta and J. O. Berger, New York: Springer-Verlag, pp. 289–298.

Pfefferbaum, A., Ford, J. M., White, P., and Roth, W. T. (1989), "P3 in Schizophrenia is Affected by Stimulus Modality, Response Requirements, Medication Status, and Negative Symptoms," *Electroencephalography and Clinical Neurophysiology*, 46, 1035–1044.

Polich, J. (1991), "P300 in the Evaluation of Aging and Dementia," in *Event-Related Brain Research* (EEG Suppl. 42), eds. C. H. M. Brunia, G. Mulder, and M. N. Verbaten, Amsterdam: Elsevier Science, pp. 304–323.

Polich, J. Ladish, C., and Bloom, F. E. (1990), "P300 Assessment of Early Alzheimer's Disease," *Electroencephalography and Clinical Neurophysiology*, 77, 179–189.

Polich, J., Pollock, V. E., and Bloom, F. E. (1994), "Meta-Analysis of P300 Amplitude from Males at Risk for Alcoholism," *Psychology Bulletin*, 115, 55–73.

Pritchard, W. S. (1981), "Psychophysiology of the P300," *Psychology Bulletin*, 89, 506–540.

———(1986), "Cognitive Event-Related Potential Correlates of Schizophrenia," *Psychology Bulletin*, 100, 43–66.

Rao, C. R. (1973), *Linear Statistical Inference and its Applications* (2nd ed.), New York: Wiley.

Scharbrough, F., Ghatrian, G.-E., Lesser, R. P., Luders, H., Nuwer, M., and Picton, T. W. (1990), "Guidelines for Standard Electrode Position Nomenclature," American Electroencephalography Society, Bloomfield, CT.

2) varies systematically with midline scalp location, and 3) is correlated most strongly with values obtained from the more adjacent electrodes. The present display is consistent with more traditional statistical analyses (Alexander et al. 1994).

Also depicted in Figure 2 are the corresponding measures of dispersion (2.6) for electrode sites Fz, Cz, and Pz. The S_{ij} are minimal for adjacent electrode sites, and tend to increase (but not in a monotonic fashion) as physical distances between the electrode sites increase.

The particular graphical displays illustrated here are not conventional rectangular grids, but were specifically chosen to delineate the clinical context of EEG recording at differ-