ERPs and Visual Signal Detection Performance: Classification Functions Based on Wavelet Decompositions

Leonard J. Trejo
University of Illinois at Urbana-Champaign
Champaign, IL
603 E. Daniel Street
Champaign, IL 61820, U.S.A.
Phone: (217) 333-8303
E-mail: trejo@uiuc.edu
WWW: http://kahuna.psych.uiuc.edu/ljt.html

Michelle M. Mullane
United States International University
San Diego, CA

Thirty-Fifth Annual Meeting of the Society for Psychophysiological Research
Oct 12-15, 1995
Toronto, Canada

Running head: ERP CLASSIFICATION

Introduction

Many important tasks require detection, recognition, and processing of visual signals. For example, air traffic controllers and radar operators must detect and respond to small computer-generated symbols and respond to subtle changes in their position or status. In tasks such as sthese, variability over time in operator arousal and attention, as well as in display clutter or signal-to-noise ratio lead to variations in overall performance.

Human perceptual and cognitive processes are indexed by electrical waves in the brain known as event-related potentials (ERP), which are time-locked to stimuli occurring in the context of a task. The amplitude, latency, and scalp distribution of ERP components often co-vary with task and subject conditions. In signal detection and classification tasks, event-related potentials (ERPs) reflect both subject- and stimulus-related sources of variance in signal detection performance (Trejo, Kramer & Arnold, 1995). Components of the ERP that covary with performance include the P1, N1, P2, N2, P300, and both positive and negative slow waves. However, component amplitude and latency measures have not provided good discrimination between high- and low performance over short periods (Trejo & Shensa, 1993).

The wavelet transform is an important new signal processing tool that is revolutionizing the analysis of signals such as the ERP and the EEG. Its most salient aspect is the replacement of frequency by scaling either in time (one-dimensional) or in space (two-dimensional). These correspond, respectively, to time-series analysis and to image processing. As a result, the bandwidth is proportional to frequency. This feature, which is common to most real broadband signals, makes wavelets particularly suitable for the analysis of transients or brief changes in the signal. This also contrasts with the Fourier transform, which has a fixed bandwidth at all frequencies. The purpose of this research was to determine the extent to which wavelet transforms of ERPs discriminate between high- and low signal detection and classification performance over periods of about 30 s.

Our approach was to classify short-term averages of ERPs using either the discrete wavelet transform or a component-oriented transform, principal components analysis (PCA). We chose PCA because it yields objective and linearly independent measures of ERP components. For both transforms we used a non-parametric bootstrapping approach to determined the classification accuracy as a function of the number of trials averaged.

Method

Subjects

The subjects were eight experienced male technicians with occupational experience in monitoring electronic displays (ages 18 to 43y). All subjects had normal or corrected vision. We trained all the subjects to stable performance levels on the task before they performed the experimental conditions .

Task

The task was presented on a radar-like display (Figure 1). Subjects pressed "T" or "NT" buttons to detect targets and nontargets, and rated their confidence in each detection response on a 3-point scale using a mouse. The task-relevant stimuli were triangles, with or without central dots, presented for 50 ms at three signal-to-noise ratios. The intertrial interval varied between 2.5 and 3.0 s. In addition to the task-relevant stimuli, an irrelevant probe (background flash) occurred in each trial, at a variable latency. However, we have not yet analyzed the ERPs elicited by this probe using wavelets. Each subject was trained to a stable level of performance on the task. Training was performed on a separate day from testing on the task.

Testing on the task was performed in two sessions spaced one week apart. In each session, about 10 blocks of 50-72 trials each were performed. The first block was a baseline run, in which ERPs were recorded but no responses were allowed. Subsequent blocks contained stimuli of varying S/N ratios. Mapping between stimuli and responses changed with each new block to enforce controlled processing.

ERP Recording

ERPs were recorded from electrodes Fz, Cz, and Pz, referred to averaged mastoids. EOGs were also recorded and artifacts were corrected off line. Visual fixation was monitored and controlled with an infrared eye tracker.

Results

Sample average ERP data for five subjects appear in Figures 2-4. Averages were computed separately for high- and low performance (median split) blocks of trials. Performance was defined as a linear composite of speed, accuracy, and confidence measures using the formula described by Trejo, Kramer, & Arnold (1995):

$$PF_1 = .33\ accuracy + .53\ confidence - .51\ reaction\ time.$$

All subjects' ERPs showed a large positivity focused parietally, which reflects P300 and slow wave. For each subject, the ERP averages of 7 to 10 consecutive artifact-free ERPs were created using running means. A corresponding series of the running mean for the PF1 measure was also computed. Both series were median-split to form low- and high-PF1 samples. We computed the discrete wavelet transform for each of the block-averaged ERPs using the Daubechies D4 wavelet (Daubechies, 1992). The high/low average DWTs for one subject appear in Figures 5-7. The ERP block averages also served as the input matrix for a covariance-based, mean-centered PCA. For each subject, six factors were retained and rotated using the varimax method. Factor scores were computed for each sample (low- and high PF1). Figure 8 shows an example of the factor loadings for one subject (S2). The best three factors were retained and combined for classifications.

A classification method similar to that described by Kramer, Trejo, & Humphrey (in press) was applied to the DWT and PCA samples. Block averages of *n*=1, 2, 4, ..., 64 blocks were created by drawing 1000 random sub-samples of size *n* from each of the samples. Each block average was classified by its multivariate Euclidean distance from the mean vector for the high or low PF1 samples (Figures 9 & 10).

A Kruskal-Wallis rank sum test was used to test the difference in the number of blocks averaged required to correctly classify a block average with 70% or 90% accuracy. For both accuracy levels, the DWT required significantly fewer blocks than the PCA (70%: $\chi^2(1)=8.98$, $p=.003$; 90%: $\chi^2(1)=9.78$, $p=.002$, Figure 11).

Conclusions

As compared to PCA scores, the DWT transform of short-term ERP averages provided for better classification of performance states in a signal detection and classification task. In eight of eight subjects, the DWT classification functions exceeded the PCA functions at all averaging levels. At the two levels tested, 70% and 90%, the DWT required fewer trials to correctly classify ERP averages than the PCA.

The average number of blocks required to correctly classify 70% of the averages was 1.6 for the DWT. This corresponds to a time on task of 48 seconds. Thus a DWT-based algorithm may provide for on-line ERP-based assessment of human performance. Such measurements could provide for analysis of dynamic changes in task performance in experimental or applied settings. For example, adaptive interfaces for complex systems could use such data to trigger decision aids or other levels of automation.

Future research will seek to identify the DWT coefficients that provide the highest sensitivity to performance differences. By examining the inverse DWT transforms of these coefficients, new insights about the physiological correlates of performance variability may be gained.

References

Daubechies, I. (1992). Ten Lectures on Wavelets. Philadelphia: Society for Industrial and Applied Mathematics.

Kramer, A. F., Trejo, L. J., & Humphrey, D. (in press). Psychophysiological measures of human information processing activities: Potential applications to automated systems. In R. Parasuraman & J. Mouloua (Eds.), Automation and Human Performance: Theory and Applications. Hillsdale: Lawrence Erlbaum.

Trejo, L. J., Kramer, A. F., & Arnold, J. A. (1995). Event-related potentials as indices of display-monitoring performance. Biological Psychology *40,* 33-71.

Trejo, L. J., & Shensa, M. J. (1993). Linear and neural network models for predicting human signal detection performance from event-related potentials: A comparison of the wavelet transform with other feature extraction methods. Proceedings of the Fifth Workshop on Neural Networks: Academic/Industrial/NASA/Defense, SPIE Volume 2204 (pp. 153-161). San Diego: Society for Computer Simulation.

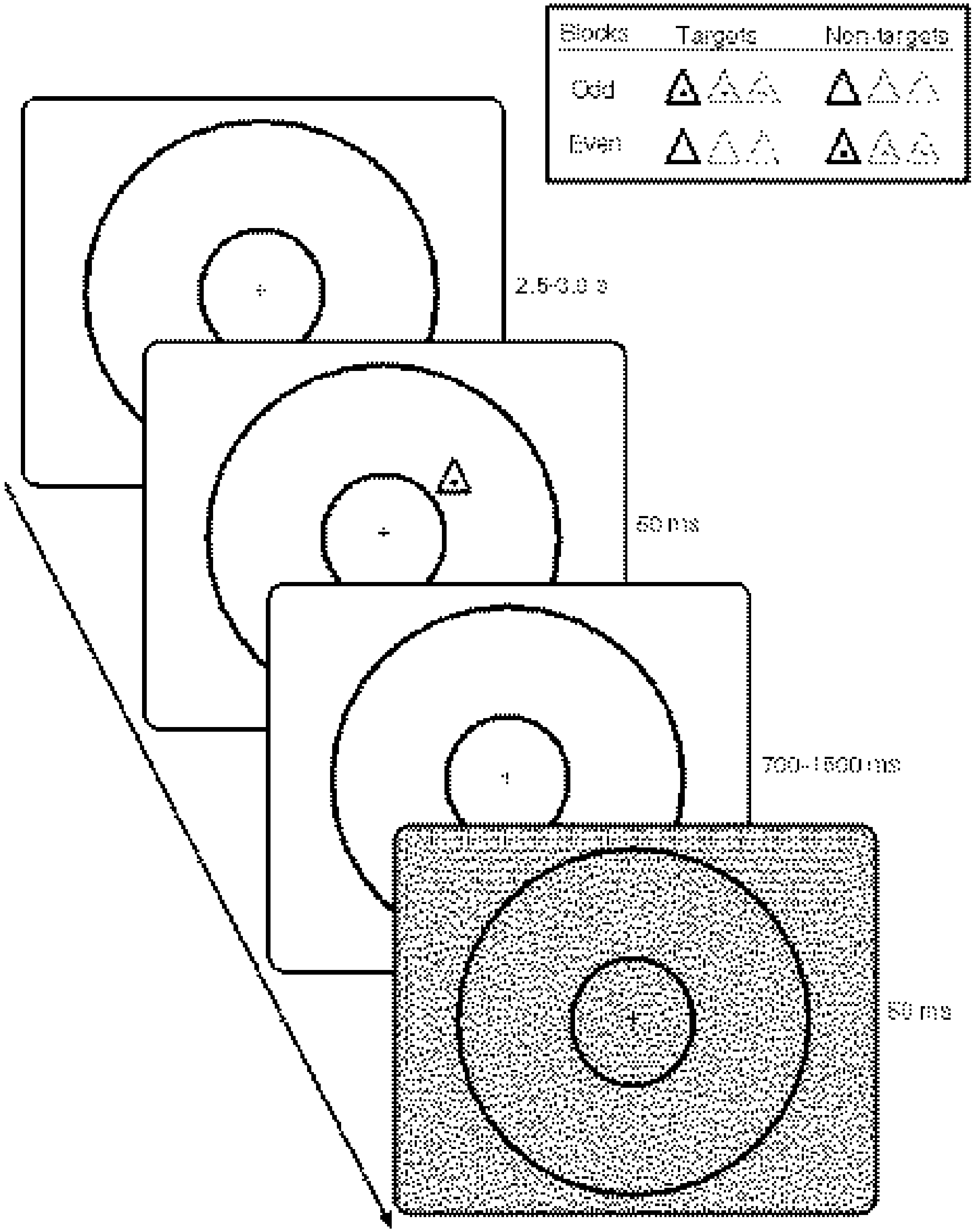

Figure 1. Signal detection task. Targets were triangles with (odd blocks) or without a dot (even blocks). Symbols appeared at one of eight locations around the central ring (2° from fixation). Contrast varied over three levels from trial to trial. Subjects pressed “T” or “NT” buttons to classify then signalled their confidence on a three-point scale. Inter-trial intervals ranged from 2.5-3.0 s. The task-relevant stimulus appeared for 50 ms. ERPs were recorded from -200 to 1500 ms relative to this stimulus. An irrelevant probe stimulus (50-ms background flash) occurred at a random time during the response period.

Figure 2. Average ERPs recorded at electrode Fz referred to average mastoids. The high-PF1 average is the mean of the trial blocks in which PF1 exceeded the median value of PF1 for a given subject. The low-PF1 average is the mean of the trial blocks in which PF1 was below the median.

Figure 3. Average ERPs recorded at electrode Cz referred to average mastoids. Details are as in Figure 2.

Figure 4. Average ERPs recorded at electrode Pz referred to average mastoids. Details are as in Figure 2.

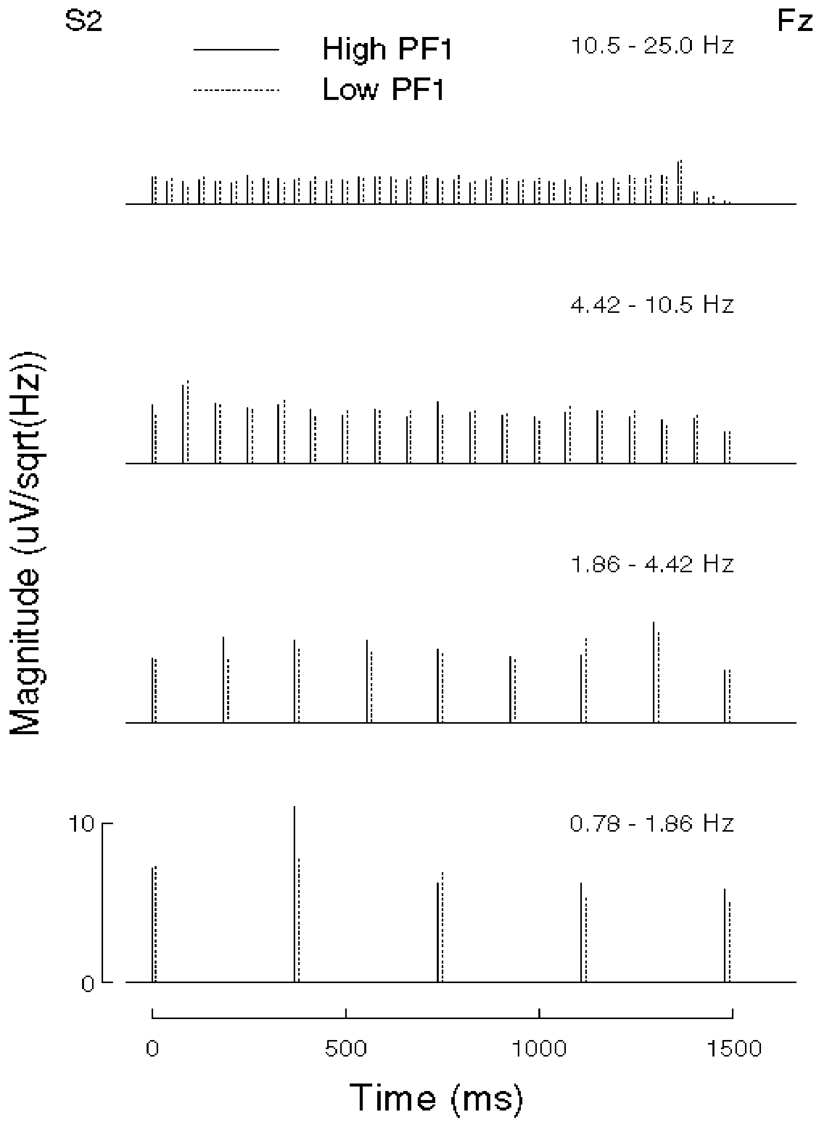

Figure 5. Average power of the discrete wavelet transforms (DWT) of the10-trial block average ERPs at Fz for the low- and high-PF1 samples for subject 2. At each scale the DWT coefficients represent the average DWT power for the high (solid) and low (dashed) samples at the indicated post-stimulus time. Each scale decomposes the ERP into power in a specific frequency band as indicated. Temporal resolution increases with the scale frequency, whereas frequency resolution decreases with the scale frequency.

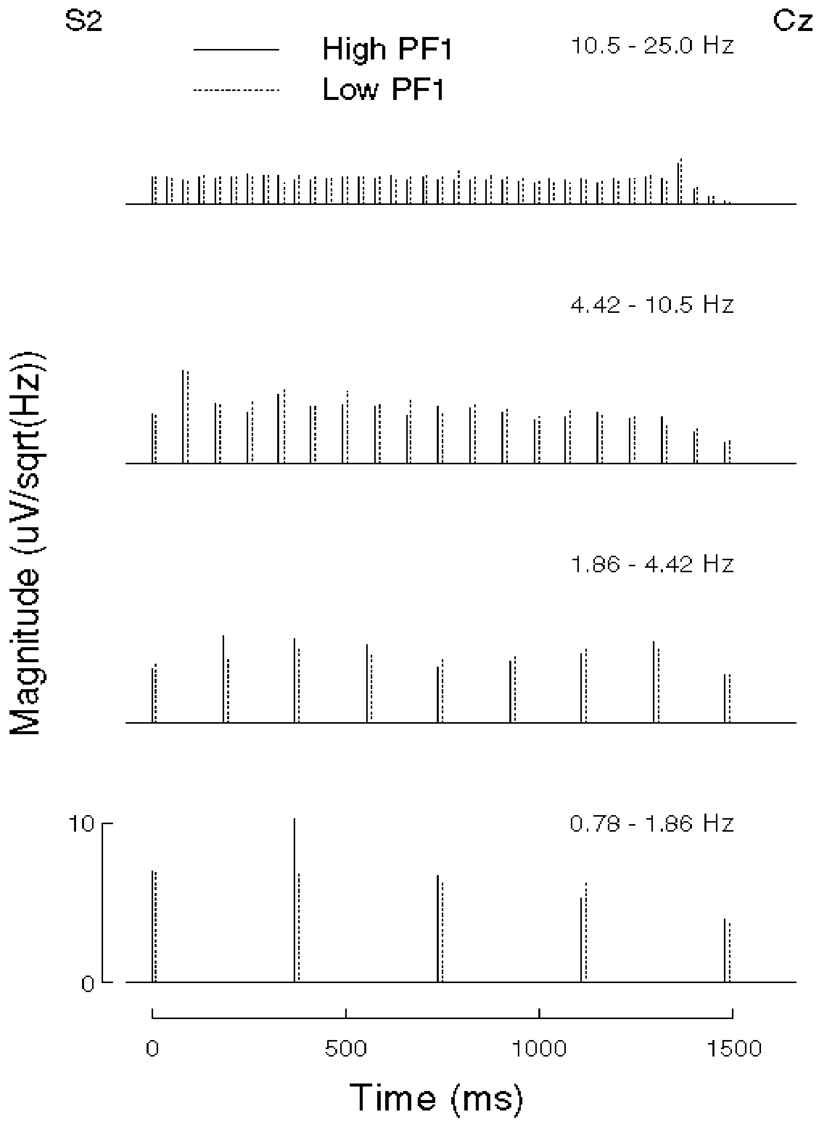

Figure 6. Average power of the discrete wavelet transforms (DWT) of the10-trial block average ERPs at Cz for the low- and high-PF1 samples for subject 2. Details are as in Figure 5.

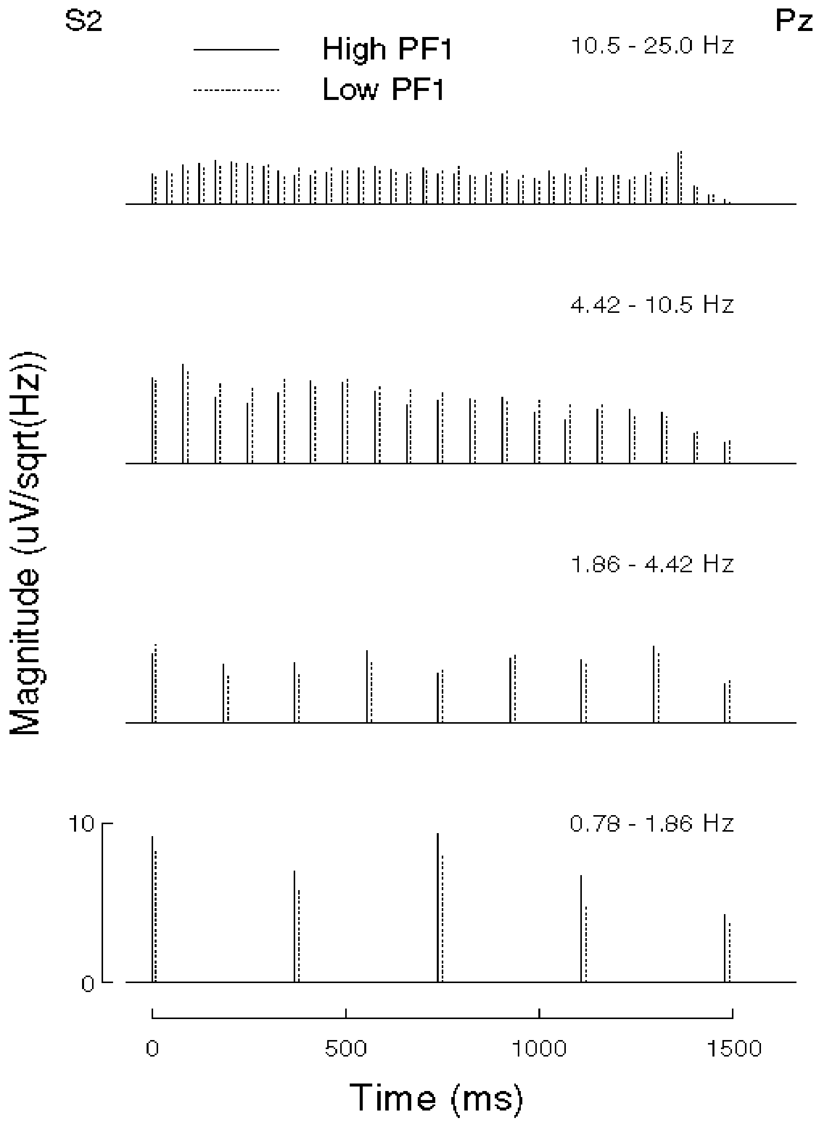

Figure 7 Average power of the discrete wavelet transforms (DWT) of the10-trial block average ERPs at Pz for the low- and high-PF1 samples for subject 2. Details are as in Figure 5

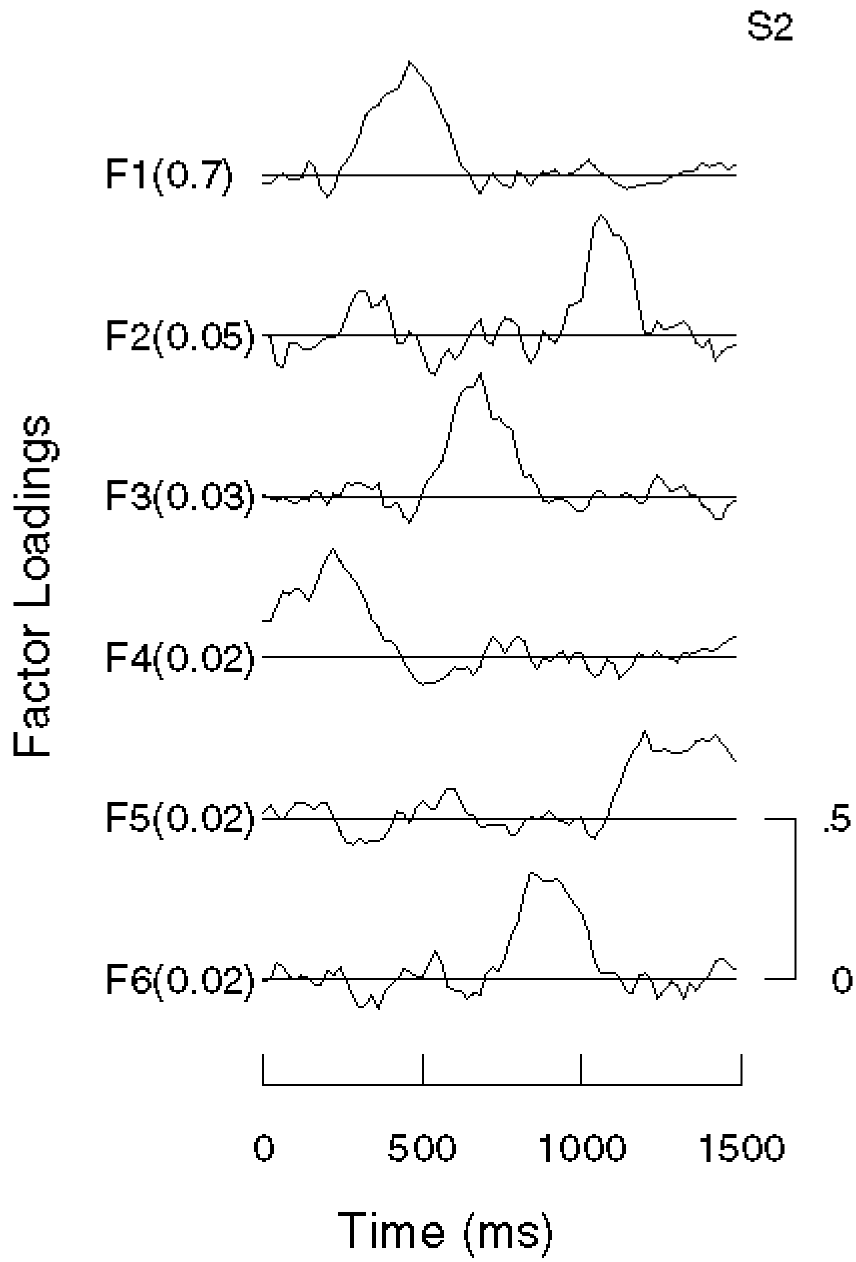

Figure 8 Principal components (n=6) extracted from the entire sample of block-average ERPs for subject 2 using a covariance-based, mean-centered method and varimax rotation. Proportions of variance accounted for by each component are indicated. Scores computed for low- and high-PF1 block averages based on these factors were used to perform the PCA-based classification. Each subject's PCA was performed separately.

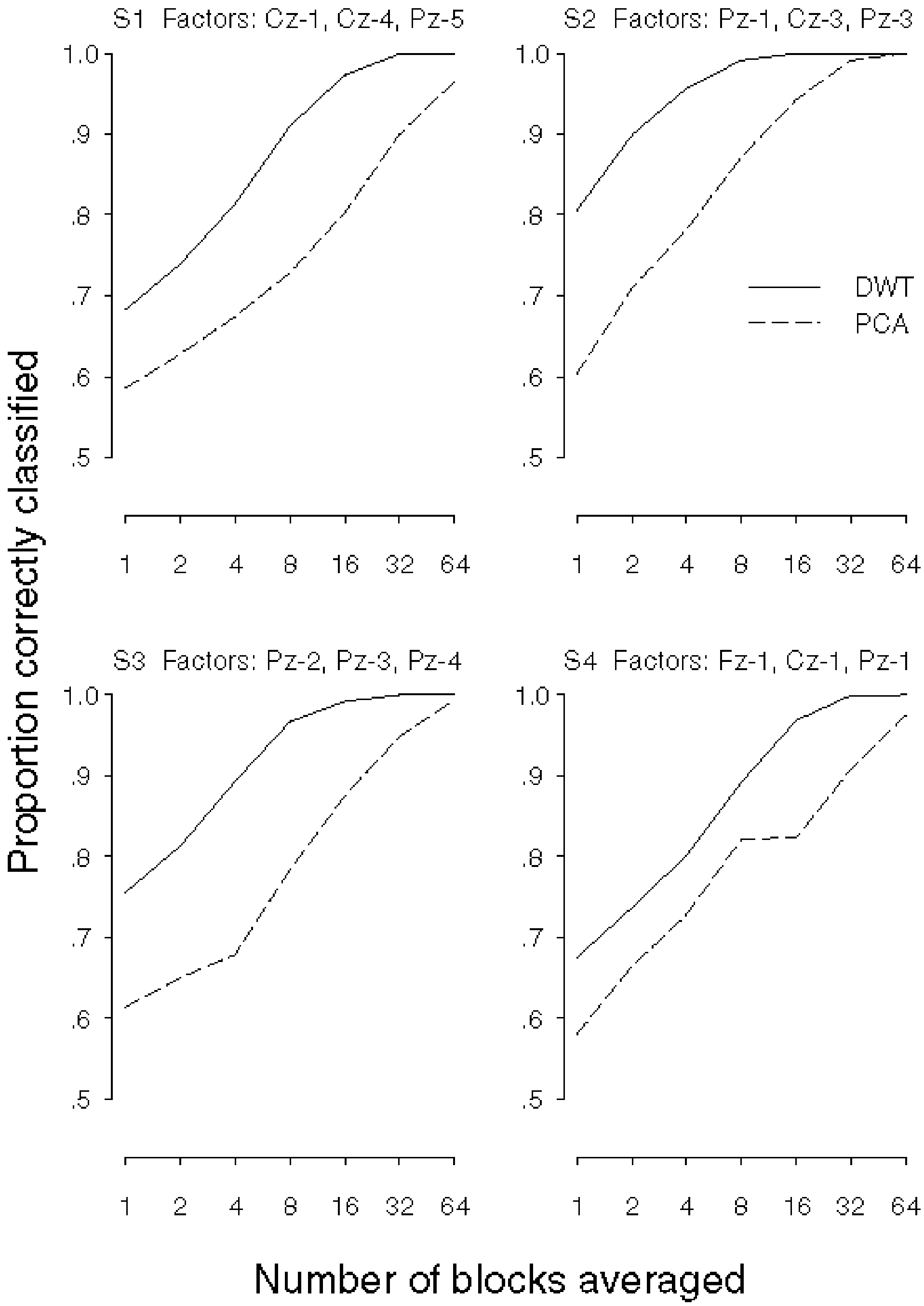

Figure 9 Classification functions for subjects 1-4 based on either the DWT (solid lines) or PCA (dashed lines) transforms. For each level (2, ..., 64), 1000 averages of randomly drawing sets of blocks were classified, except for level 1, where each block was classified. Averages from the low- and high-PF1 samples were compared to the means for those samples and classified using a Euclidean distance measure (square root of sum of squared differences).

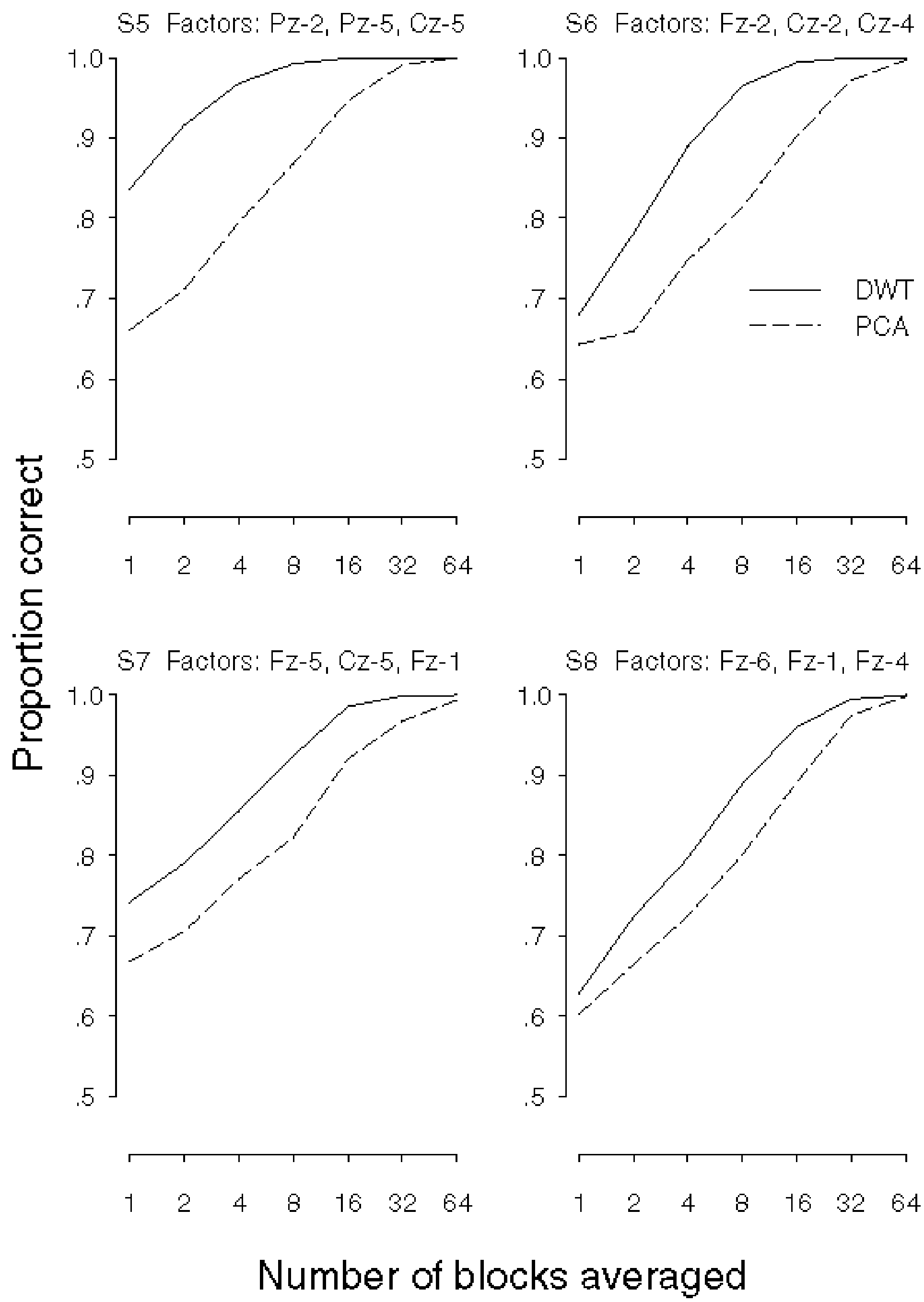

Figure 10 Classification functions for subjects 5-8 based on either the DWT (solid lines) or PCA (dashed lines) transforms. Details are as in Figure 10.

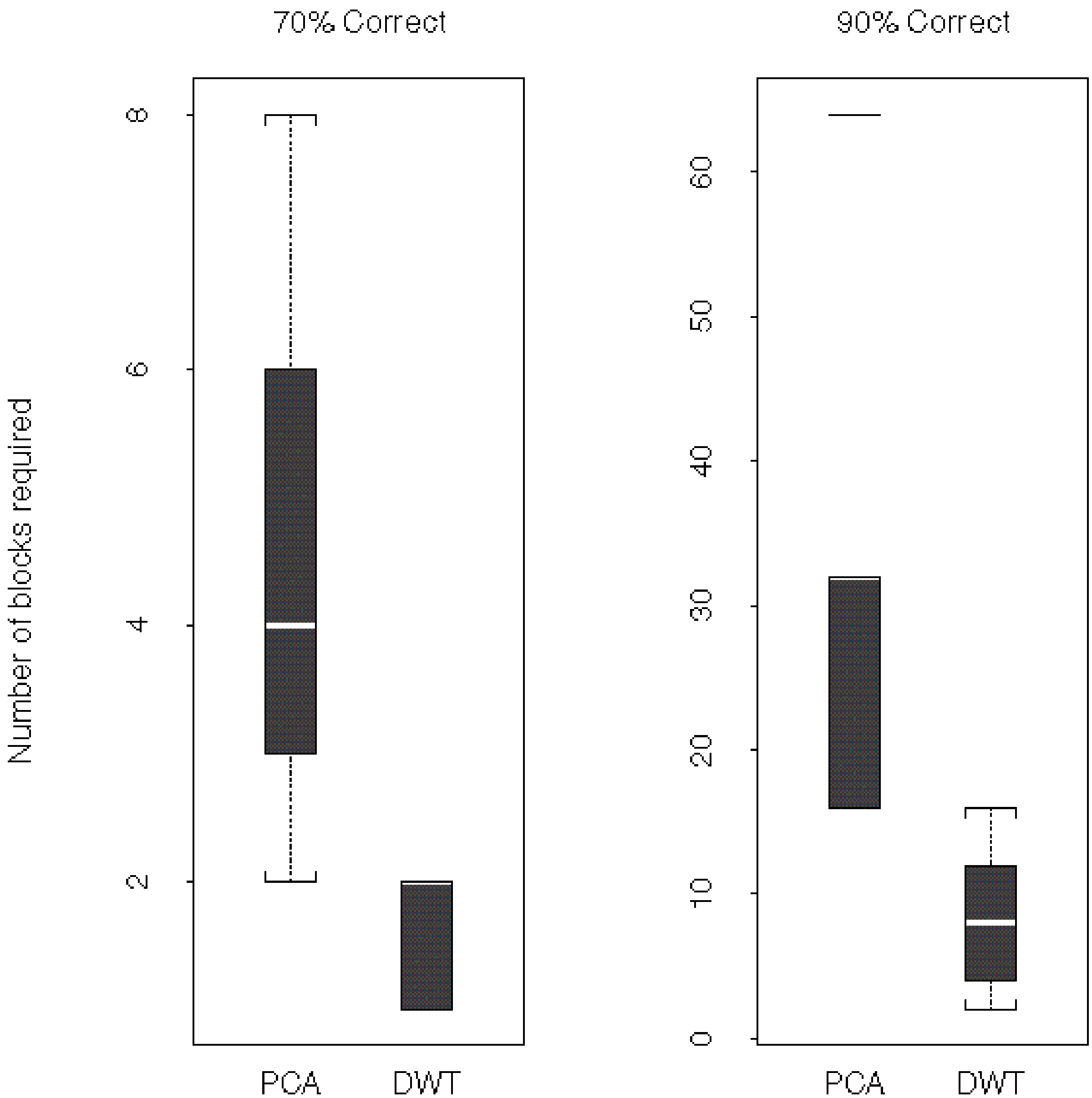

Figure 11. Boxplots of the distributions of classification accuracy requirements for the eight subjects. Each bar shows the mediam (white line), inter-quartile range (black bar) and range (whiskers) of the number of blocks required to classify ERP averages for either the PCA or DWT measures with accuracies of 70% (left) and 90% right.