

The Mental Prosthesis: Assessing the Speed of a P300-Based Brain–Computer Interface

Emanuel Donchin, Kevin M. Spencer, and Ranjith Wijesinghe

Abstract—We describe a study designed to assess a brain–computer interface (BCI), originally described by Farwell and Donchin [9] in 1988. The system utilizes the fact that the rare events in the oddball paradigm elicit the P300 component of the event-related potential (ERP). The BCI presents the user with a matrix of 6 by 6 cells, each containing one letter of the alphabet. The user focuses attention on the cell containing the letter to be communicated while the rows and the columns of the matrix are intensified. Each intensification is an event in the oddball sequence, the row and the column containing the attended cell are “rare” items and, therefore, only these events elicit a P300. The computer thus detects the transmitted character by determining which row and which column elicited the P300. We report an assessment, using a bootstrapping approach, which indicates that an off line version of the system can communicate at the rate of 7.8 characters a minute and achieve 80% accuracy. The system’s performance in real time was also assessed. Our data indicate that a P300-based BCI is feasible and practical. However, these conclusions are based on tests using healthy individuals.

Index Terms—ALS, brain–computer interface (BCI), locked-in syndrome.

I. INTRODUCTION

BRAIN–COMPUTER interfaces (BCI’s) that utilize the electrical activity of the brain as the carrier of the communicated signal can all be viewed as methods for providing the user with control over the variance of the EEG [1]. The methods by which such control over the variance has been achieved often focused on controlling the spectral composition of the EEG, using for the purpose some version of biofeedback [2], [3]. The other approach, illustrated in this report, controls the variance of endogenous components of event-related brain potentials (ERP’s). The ERP represents brain activity that is elicited in response to events, external or internal. The ERP is the activity that is time locked to the eliciting event. Such ERP’s consist of a sequence of components [4] some of which are exogenous, that is they are manifestations of the processing of specific external events. The exogenous components are generally obligatory responses to the presentation of physical stimuli. As long as the sensory system is intact and functional, the exogenous components are largely independent of the

role the stimuli play in the subjects’ ongoing information processing. The endogenous components, on the other hand, are manifestation of processing activities that depend on the role of the stimuli within the task the subject is performing, and on the interaction between any given event and the context in which it is presented [5]. Structuring the manner in which stimuli are presented, and controlling the interaction between external events and the subject’s task can therefore control the variance in the endogenous components.

The P300 component of the ERP is one such endogenous component [6]. It is most frequently elicited within the framework of what has come to be called the “oddball paradigm” (see [7] for a methodological review). In this paradigm [8] the subject is presented with a sequence of events that can be classified into two categories. In general, events in one of the two categories are rarely presented. Furthermore, the subject is assigned a task that cannot be performed without categorizing the events. Under these circumstances, events in the rare category elicit an ERP characterized by a P300 component; the less probable the eliciting event, the larger the P300.

Farwell and Donchin [9] described a BCI that exploited these properties of the oddball paradigm to allow a user to communicate a sequence of letters to a computer. An oddball paradigm was created by successively, and randomly, intensifying either a row or a column of a 6 by 6 matrix of characters that was displayed continually to the subject. In each “trial” the subject is “communicating” a character by focusing attention on the cell containing the character. Hence, the total sequence of events is divided into two categories. One category, which constitutes 16.7% of the intensifications (one in six), includes the cell whose content are at the focus of attention. The remaining intensifications are of rows and columns that do not include the relevant cell. If this is indeed an oddball paradigm, then the events containing the relevant cell (being the rare events in an oddball paradigm) should be the only events that elicit a P300. The communication task reduces, thus, to the detection of which row, and which column, are those eliciting a P300 on a given trial. The letter the subject is trying to communicate is at the intersection of the row and the column that elicited a P300.

Farwell and Donchin [9] demonstrated the feasibility of this version of a BCI and established, using bootstrapping techniques, that the communication speed achieved by the system allowed the subject to communicate about four characters a minute. The purpose of the study described here was to evaluate a new version of the system, ported to an Intel computer using Windows’95, and to determine if the new system increases the communication rate.

Manuscript received January 18, 2000; revised February 18, 2000 and March 1, 2000. This work was supported by a grant from the Campus Research Initiatives Program at the University of Illinois at Urbana-Champaign (UIUC) and by a STTR Grant 1 R41 MH56319-01 to Biologic Systems, Inc., in collaboration with UIUC.

The authors are with the Department of Psychology and Beckman Institute, University of Illinois at Urbana-Champaign, IL 61820 USA.

Publisher Item Identifier S 1063-6528(00)04116-1.

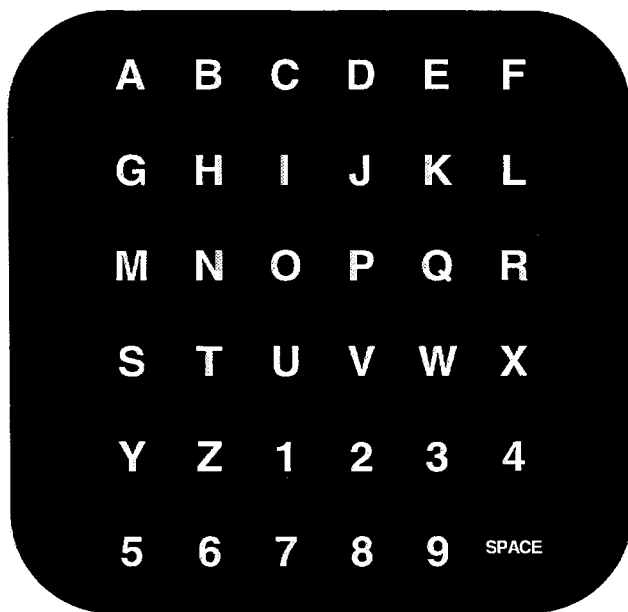


Fig. 1. The stimulus matrix monitored by the subject. Every 125 ms one of the rows, or one of the columns of the matrix was intensified.

II. METHODS

A. Subjects

Ten able-bodied (six female) and four disabled subjects (wheelchair-bound; three with complete paraplegia, one incomplete paraplegia; two female) from the university community participated in the experiment.

B. Stimuli and Procedure

The subjects viewed a display of the matrix exhibited in Fig. 1. The characters were presented as white characters on a black background, using a moderate and easily visible intensity. The matrix used in the present study differed in a number of minor details from the display used by Farwell and Donchin. Intensifying, in a random sequence, each of the 6 rows and 6 columns of the matrix produced an oddball sequence. Each intensification lasted 100 ms, with an SOA of 125 ms. The interval between trials (6 row and 6 column intensifications) was 1500 ms.

The subjects sat 50 cm from the display and were instructed to observe the display and to count the number of times the row, or the column, containing the designated target letter “P” was intensified. Thus, $\text{Prob}(\text{Target}) = 2/12$ or 0.167. The rows and the columns were intensified in a random sequence in such a manner that all 6 rows and 6 columns were intensified before any was repeated. A “trial” in the study is thus defined as the intensification of all 12 elements of the matrix. The total duration of such a trial was 1500 ms. The specific implementation of the BCI, using two separate computers, one for data acquisition the other for controlling the display, forced an interval of 1000 ms between trials. In a BCI implemented as a special purpose device the intertrial interval could be made arbitrarily short, except for the time required shifting gaze between characters. Each subject performed blocks of 15 trials each.

1) *Data Acquisition and Processing:* The EEG was recorded from tin electrodes in an electrode cap (Electro-Cap International) at the Fz, Cz, Pz, O1, O2, and right mastoid sites, referenced to the left mastoid. The data were referenced off-line to averaged mastoids. The EEG was amplified using Biologic amplifiers (0.01–100 Hz passband) and digitized at the rate of 200 Hz. Vertical and horizontal EOG artifacts were removed from the EEG by an eye-movement correction method [10].

The P300 detection method followed the procedure developed by Farwell and Donchin. Single-trial EEG epochs were derived in association with each intensification, beginning 300 ms prior to the intensification and lasting for 1100 ms. Thus, each trial yielded 12 such epochs, each associated with a specific row and a specific column. The method assumes that the epochs associated with the relevant column and the relevant row will contain a detectable P300, while the other epochs will not. The data submitted to the detection algorithms were obtained by averaging together each combination of row and column single-trial epochs. Thus, there were 6 rows by 6 columns = 36 epochs for each trial.

As is generally the case for ERP components, it is virtually impossible to visualize, or to detect numerically, the presence of an ERP in the epoch following a single event. The ERP is substantially smaller than the ongoing EEG activity; hence detecting ERP’s requires a method that extracts the ERP signal from the EEG “noise.” While averaging over all the trials obtained in a given study provides a very clear picture of the pattern of the ERP’s, it cannot be relied on for the purposes of communication at a relatively acceptable speed. Hence, while we will examine the grand-average ERP’s to obtain a clear view of the pattern of the signal we seek, much of the effort of developing the BCI consists of determining the smallest number of trials that must be averaged to insure reliable detection.

III. RESULTS

1) *Grand Average ERP’s:* The epochs associated with the target and nontarget stimuli were averaged over all trials used with each of the subjects, for each of the electrode sites used in the study. These data, averaged over the subjects, are presented in Fig. 2. It is quite evident that the rare “targets” elicit a large P300 whose scalp distribution is that used to define the P300, with the largest amplitude elicited at centro-parietal electrode sites. Thus, ERP’s for “Target Letter” were associated with the cell at the intersection of the correct row and correct column. The ERP’s for “Target Row/Column” were associated with the cells at the intersection of the correct row or column and an incorrect column or row, respectively. The ERP’s for “Standards” were associated with the cells at the intersection of incorrect rows and columns.

It is evident that a communication system relying on an average of 40 trials can achieve perfect accuracy. That is, target rows and columns definitely elicit a large P300 that, given 40 trials, can be easily detected. As such an average requires a total of 60 s to communicate each character (1.5 s per trial for 40 trials = 60 s) the system, while perfectly reliable, is unacceptably slow. There is, of course, a direct relationship between

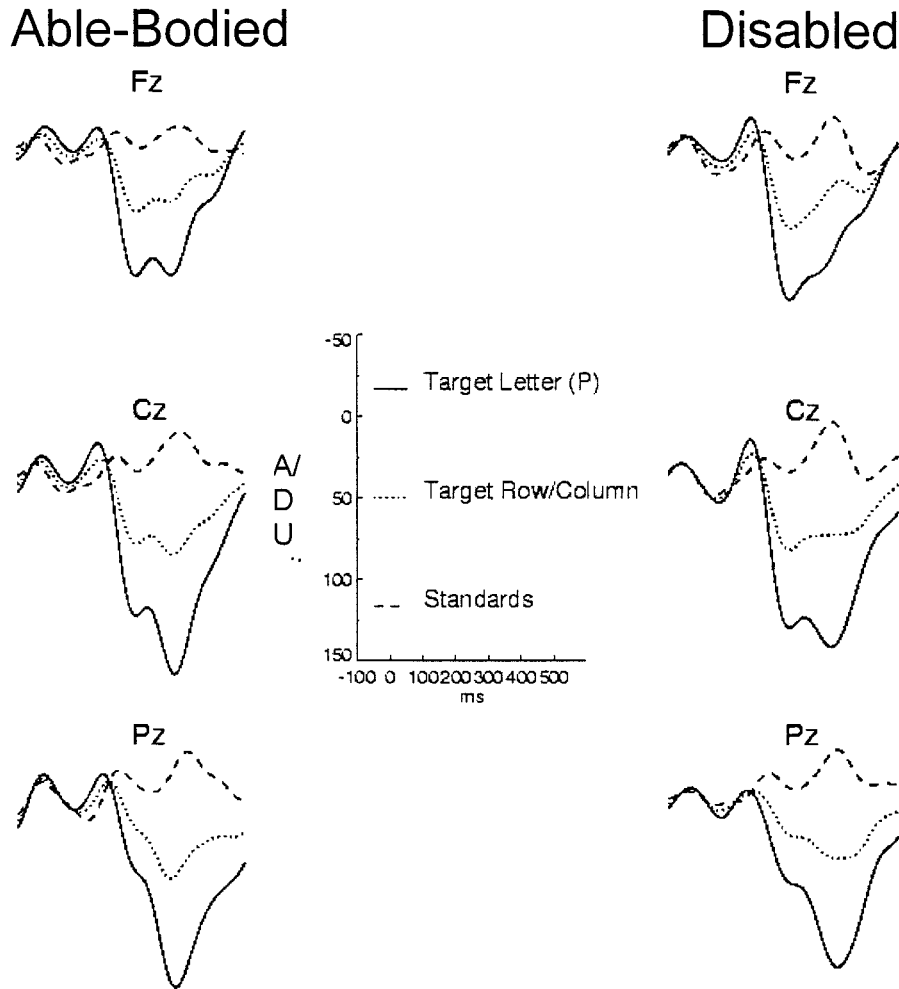


Fig. 2. ERP's at the midline electrode sites recorded from the wheel-chaired subject and the able-bodied subjects. The data associated with the three types of trials are superimposed. Positivity is indicated by a downward deflection.

the number of trials required for reliable transmission and the speed of communication. If detection could be achieved using just one trial, the system would allow communication at the rate of 40 characters/min. Assuming, of course, that no time is lost on gaze switching and letter selection. If reliable transmission requires four trials the communication rate drops to ten characters/min and so on. To assess the limit on the communication speed achieved with the data recorded in this experiment, we resorted to bootstrap analysis as used previously by Farwell and Donchin, to assess the effectiveness of the BCI at different levels of signal averaging.

2) *Bootstrap Analyzes*: To conduct this assessment we followed the standard bootstrapping method [11] of obtaining random samples with replacement from the sample of data on hand. The data set used for the bootstrapping consisted of 75 trials, for each of which there were 12 events, 6 for the rows and 6 for the columns. We assessed the reliability of the communication, that is the percent of correct detections of the communicated character, for values of N between 2 and 40. The following procedure was executed 1000 times for each value of N assessed:

- 1) obtain a random sample of N trials for each cell by sampling w/replacement from the set of 75 trials;

- 2) compute the average of N trials for each cell;
- 3) apply stepwise discriminant analysis (SWDA) to the set of cell averages;
- 4) compute the discriminant score for each cell;
- 5) select the cell with the maximum discriminant score;
- 6) if the selected cell is the defined target cell count a hit, otherwise count a miss.

When done, record the percentage of hits among the 1000 samplings. This is the percent accuracy at the communication speed determined by the N trials.

SWDA was applied to a data set constructed by bootstrapping to assess the accuracy with which the target cell was detected as a function of the number of trials used for averaging. This procedure was applied with two preprocessing methods.

- 1) SWDA: Single-trial cell epochs were filtered at 0–8 Hz and resampled at 50 Hz, yielding 30 time points for the 0–600 ms period of each epoch.
- 2) SWDA/DWT: Single-trial cell epochs were filtered at 0–50 Hz and resampled at 50 Hz, yielding 32 time points for the 0–640 ms period. These time points were converted to wavelet coefficients with the discrete wavelet transform (DWT) using a Daubechies wavelet with four coefficients.

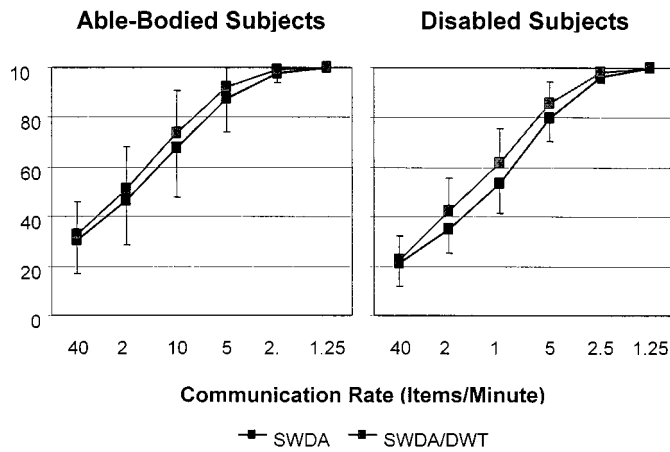


Fig. 3. The percent of correct bootstrapped trials obtained at each level of communication speed. Note that the speed of communication is inversely related to the number of trials averaged.

TABLE I
WORDING NEEDED

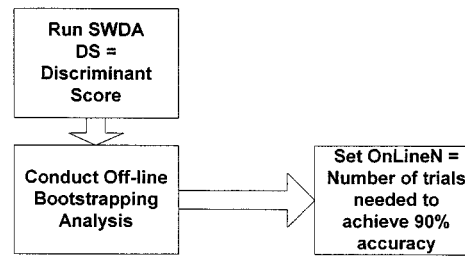
Subject Group	Pre-processing	Accuracy Level	
		80%	95%
Able-Bodied	SWDA	6.3 items/min	3.4 items/min
	SWDA/DWT	7.8 items/min	4.3 items/min
Disabled	SWDA	4.8 items/min	2.8 items/min
	SWDA/DWT	5.9 items/min	3.2 items/min

The results of these analyzes are exhibited in Fig. 3 in which are plotted the percent of the 1000 trials at a given communication speed in which the SWDA yielded the maximal score for the cell containing the letter “P.” The communication speed is a function of the number of trials used for the average computed on each of the 1000 bootstrap samples. Note that these values assume that the BCI can proceed with no delay between trials. In the current implementation of the BCI, technical considerations dictated a 1000-ms pause between trials.

In Table I it can be seen that, as expected, the quality of the detection increases as the number of trials in each average increases. However, the results indicate that under the conditions of the present study, the communication speed at which the system allows communication at the 80% level of accuracy is 7.8 characters/min. The speed is decreased to 4.8 characters/min at the 90% accuracy level. This represents a substantial improvement relative to the rates reported by Farwell and Donchin [9]. As virtually all elements of the system were modified for the current study, using higher quality displays, better digitization hardware and software, better common mode rejection in the amplifiers and, it would seem, improved algorithms in the packaged SWDA procedures, it is difficult to identify the precise reason for the improvements. However, the most likely cause for the improved communication rate in this implementation of the BCI is that the data submitted for SWDA analysis were the ERP averages for each combination of row and column rather than the individual row and column averages. An analysis of variance of the bootstrap communication rates with the design subject group (Able-bodied vs. Disabled) \times method (SWDA vs. SWDA/DWT) \times accuracy (80% vs. 95%) revealed that application of SWDA to the wavelet transform of the data resulted in a

On-Line, Real Time, Implementation of BCI

1. Off-line Preparations



2. On line events per test

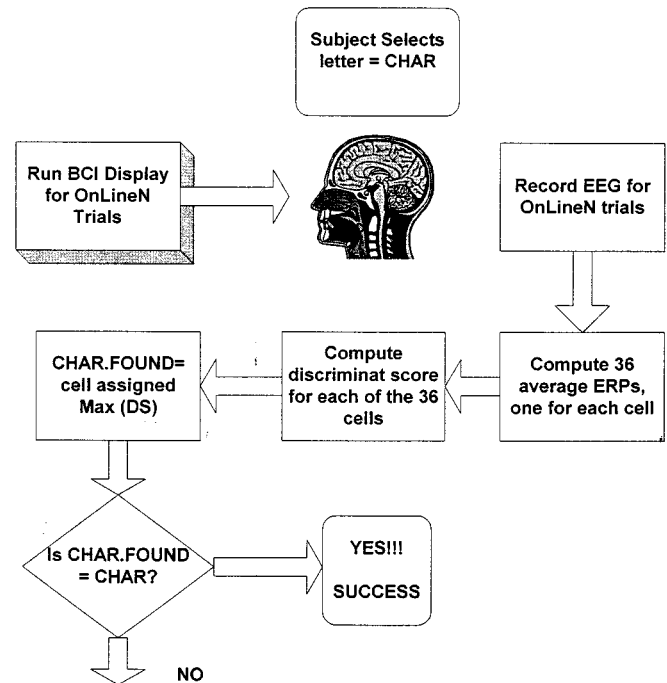


Fig. 4. The data flow, and logical structure, of the on-line, real time, application of the BCI.

small but statistically reliable improvement of ~ 1.4 items/min (main effect of method, $F[1,12] = 7.51, p < 0.05$).

A. On Line Assessment Phase

In the tests reported above, all subjects focused on the same character, chosen by the experimenter. Furthermore, the data reported above assess the quality of the communication retrospectively, analyzing off line the database acquired during the sessions. To obtain a preliminary assessment of the effectiveness of the P300-based BCI when employed on line, and when the detection is made in real time, we conducted an on-line test in which five of the ten able-bodied subjects participated. The process is summarized in Fig. 4. The results from the bootstrap analysis described above yielded for each subject the discriminant function required to achieve performance at the 90% accuracy level, and the specific number of trials that were necessary for this subject to achieve this level of accuracy. In the online

test, each subject selected successively five individual characters, chosen by himself, and focused attention on these characters for the selected number of trials for that subject. The discriminant function was applied to the averages and the resulting discriminant scores were used by the BCI to identify the character selected by the subject.

Using this procedure the BCI identified correctly the cell selected by the subject on 56% of the trials. On 36% of the cells, the BCI chose either the correct row or the correct column but erred on the other element. Thus the BCI was incorrect with respect to both row and column on only 8% of its choices.

IV. DISCUSSION

The data reported above confirm the report by Farwell and Donchin [9] that it is possible to construct a Brain Computer Interface that, using the P300, allows an individual to operate a virtual keyboard without using or requiring any activation of skeletal muscles. As noted above the BCI described here performed at a faster communication rate than that described by Farwell and Donchin. One noteworthy difference introduced in the current version of the BCI is that the discriminant analysis was applied to the 36 individual cells rather than to the rows and columns. This approach seems to execute the detection with higher sensitivity. However, the factors underlying this improvement in speed need further investigation.

In general, the P300 based BCI has the very interesting property that there does not seem to be any need to train the subjects to generate a P300 in response to the rare events in the oddball sequence. The data acquired in literally hundreds of studies of the P300 since Sutton's original report [6] has established that virtually all subjects when challenged with the oddball paradigm will generate a P300. The reliability of the phenomenon has been examined in some detail by Fabiani *et al.* [7] who were able to establish the considerable inter-subject and intra-subject reliability of the P300.

The learning-free nature of the procedure is purchased at the cost of its relying on a structured environment in which the subject must, when using the system, interact with a closely controlled physical environment that is used to generate the events in the oddball sequence. Our approach differs from that adopted by those developing BCI's that are based on the modulation of the spectral composition of the EEG (see for example [2]). These biofeedback dependent devices use a "control" metaphor to describe the nature of the BCI they develop. Their goal is to provide the subject with a control signal that can be deployed in a voluntary manner, triggered only by the subject's intentions and control goals. The metaphor underlying the BCI described in this report is the metaphor of the keyboard. Our intent is to provide the subjects with a substitute for the keyboard. The structural nature, and confining nature of keyboards does not render them useless. It merely limits their utility to a particular domain. Their ubiquity suggests that the domain of application of keyboards is usefully rich.

It worth noting that our results tend to underestimate the potential speed of the P300 based BCI. Our tests assume that each decision made by the system is independent of all previous choices. We are also assuming that correct communication

requires perfect spelling. Both assumptions are incorrect. It is well known that there are substantial sequential dependencies in English. It is our intent to incorporate information about the sequential structure of the language in the next phase of the development of the BCI. Similarly, it is possible to incorporate spelling correction software so that spelling mistakes can be managed even as increases in the operating speed may be associated with an increased number of errors. Finally, we note that it is possible to gain a considerable increase in the system's speed by replacing the individual characters in the 36 cells by words, so that rather than spelling out words the user will be selecting items from a menu. While this increase in speed is purchased at the cost of a limit on the flexibility of communication, such limits may be highly acceptable.

We are aware of the fact that at this time the system we described has not been used with the truly disabled, locked-in patients, or patients suffering from ALS. Until such tests are conducted what we report is merely a feasibility study, testing the validity of the concept and an assessment of its functionality under rather optimal conditions. We do not know at this time if the locked-in patients will be capable of maintaining their attention on one cell long enough to communicate their choice given the nature of the display. Studies in the realistic conditions of the patients' bedside are now in preparation. We nevertheless believe that such a test of the concept and the development of a procedure for assessing the validity of the concept is important at this time. We are encouraged by the success of Birbaumer and his colleagues [3] to use the subject's slow waves for another keyboard emulator. They were able to demonstrate that ALS patients can use a rather specialized procedure for conducting a binary choice through the alphabet. In order to achieve their goal their patients had to fixate and monitor a screen over fairly complex and slow moving conditions. In evaluating the speed of the P300-based BCI it is important to recall that the device is intended for use by individuals who are completely disabled. As a base of comparison one needs to use the communication method used by Bauby [12], a "locked-in" patient, to write his book, "The Diving-Bell and the Butterfly."

ACKNOWLEDGMENT

The authors would like to thank O. Karni who helped in running the subjects. The BCI system was programmed by B. Foote, and M. Anderson was invaluable in the design of the hardware. They would also like to acknowledge the strong support that was provided by our colleagues in Biologic Systems, Inc., and in particular, by G. Raviv and I. Pal.

REFERENCES

- [1] T. M. Vaughan, J. R. Wolpaw, and E. Donchin, "EEG-based communication: Prospects and problems," *IEEE Trans. Rehab. Eng.*, vol. 4, pp. 425-430, 1996.
- [2] J. R. Wolpaw, D. Flotzinger, G. Pfurtscheller, and D. J. MacFarland, "Timing of EEG-based cursor control," *J. Clin. Neurophysiol.*, vol. 14, pp. 529-538, 1997.
- [3] N. Birbaumer, N. Ghanayim, I. Iversen, B. Kotchoubey, A. Kobler, J. Perelmouter, E. Taub, and H. Flor, "A brain-controlled spelling device for the completely paralyzed," *Nature*, vol. 398, pp. 297-298, 1999.

- [4] E. Donchin, W. Ritter, and C. McCallum, "Cognitive psychophysiology: The endogenous components of the ERP," in *Brain Event-Related Potentials in Man*, E. Callaway, P. Tueting, and S. H. Koslow, Eds. New York: Academic, 1978.
 - [5] M. G. H. Coles and M. D. Rugg, "Event-related brain potentials: An introduction," in *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition*, M. D. Rugg and M. G. H. Coles, Eds. New York: Oxford University Press, 1995.
 - [6] S. Sutton, M. Braren, J. Zubin, and E. R. John, "Evoked-potential correlates of stimulus uncertainty," *Science*, vol. 150, pp. 1187–1188, 1965.
 - [7] M. Fabiani, G. Gratton, D. Karis, and E. Donchin, "Definition, identification, and reliability of the P300 component of the event-related brain potential," in *Advances in Psychophysiology*, P. K. Ackles, J. R. Jennings, and M. G. H. Coles, Eds. New York: JAI Press, 1987, vol. 2.
 - [8] E. Donchin and M. G. H. Coles, "Is the P300 component a manifestation of context updating?," *Behav. Brain Sci.*, vol. 11, pp. 355–372, 1988.
 - [9] L. A. Farwell and E. Donchin, "Talking off the top of your head: A mental prosthesis utilizing event-related brain potentials," *Electroencephalogr. Clin. Neurophysiol.*, vol. 70, pp. 510–523, 1988.
 - [10] G. Gratton, M. G. H. Coles, and E. Donchin, "A new method for off-line removal of ocular artifact," *Electroencephalogr. Clin. Neurophysiol.*, vol. 75, pp. 468–484, 1983.
 - [11] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. New York: Chapman and Hall, 1993, p. 436.
 - [12] J.-D. Bauby, *The Diving Bell and the Butterfly*. Fourth Estate, London, U.K., 1997, p. 132.
- Emanuel Donchin**, photograph and biography not available at the time of publication.
- Kevin M. Spencer**, photograph and biography not available at the time of publication.
- Ranjith Wijesinghe**, photograph and biography not available at the time of publication.