HANDBOOK OF ELECTROENCEPHALOGRAPHY AND CLINICAL NEUROPHYSIOLOGY

EDITOR-IN-CHIEF A. REMOND

VOLUME 4

Evaluation of Bioelectrical Data from Brain, Nerve and Muscle, I

EDITOR: M. A. B. BRAZIER
Brain Research Institute, University of California Medical Center, Los Angeles, Calif. (U.S.A.

PART B

Digital Processing of Bioelectrical Phenomena

EDITOR: D. O. WALTER
Space Biology Laboratory, University of California, Los Angeles, Calif. (U.S.A.)

ELSEVIER

HANDBOOK OF ELECTROENCEPHALOGRAPHY AND CLINICAL NEUROPHYSIOLOGY

Editor-in-Chief: **Antoine Rémond**
Centre National de la Recherche Scientifique, Paris (France)

VOLUME 4

Evaluation of Bioelectrical Data from Brain, Nerve and Muscle, I

Editor: **M. A. B. Brazier**
Brain Research Institute, University of California Medical Center, Los Angeles, Calif. (U.S.A.)

PART B

Digital Processing of Bioelectrical Phenomena

Editor: **D. O. Walter**
Space Biology Laboratory, University of California, Los Angeles, Calif. (U.S.A.)

Elsevier Publishing Company – Amsterdam – The Netherlands

International Federation of Societies for EEG and Clinical Neurophysiology

HANDBOOK EDITORIAL COMMITTEE

ISBN 0-444-41035-x

Elsevier Publishing Company, Jan van Galenstraat 335, Amsterdam
Printed in The Netherlands

Sole distributor for Japan:
Igaku Shoin Ltd.
5-29-11 Hongo Bunkyo-ku
Tokyo

All other countries:
Elsevier Publishing Company
Amsterdam, The Netherlands

PART B

DIGITAL PROCESSING OF BIOELECTRICAL PHENOMENA

Editor: **D.O. Walter**
Space Biology Laboratory, University of California,
Los Angeles, Calif. (U.S.A.)

Collaborators:

Ray Cooper, *Burden Neurological Institute, Bristol (Great Britain)*

J.D. Frost, Jr., *Baylor College of Medicine, Houston, Texas (U.S.A.)*

D.O. Walter, *Space Biology Laboratory, University of California,*
Los Angeles, Calif. (U.S.A.)

Preface

This volume 4B will be of interest not only to electroencephalographers and students of evoked potentials, but also to those who wish to analyze *by computer means* such records as electromyogram, electroretinogram, electronystagmogram and perhaps others. Section I has the broadest relevance; Section III discusses checking and validation methods which should be valuable not only to EEG studies, but also to such other electrophysiological studies as become involved in elaborate methodologies and comparisons; Section II, although it deals in part with specific analysis methods, limits itself to their data-processing or validity aspects. More specific methodological counsel should be sought in other volumes.

Throughout this volume 4B the word *machine* is used as a synonym for *computer*, usually *digital computer*. A glossary of less well-known technical words and phrases is included on page 4B-56.

PART B

DIGITAL PROCESSING OF BIOELECTRICAL PHENOMENA

Editor: **D.O. Walter**
Space Biology Laboratory, University of California,
Los Angeles, Calif. (U.S.A.)

Collaborators:

Ray Cooper, *Burden Neurological Institute, Bristol (Great Britain)*

J.D. Frost, Jr., *Baylor College of Medicine, Houston, Texas (U.S.A.)*

D.O. Walter, *Space Biology Laboratory, University of California,*
Los Angeles, Calif. (U.S.A.)

Preface

This volume 4B will be of interest not only to electroencephalographers and students of evoked potentials, but also to those who wish to analyze *by computer means* such records as electromyogram, electroretinogram, electronystagmogram and perhaps others. Section I has the broadest relevance; Section III discusses checking and validation methods which should be valuable not only to EEG studies, but also to such other electrophysiological studies as become involved in elaborate methodologies and comparisons; Section II, although it deals in part with specific analysis methods, limits itself to their data-processing or validity aspects. More specific methodological counsel should be sought in other volumes.

Throughout this volume 4B the word *machine* is used as a synonym for *computer*, usually *digital computer*. A glossary of less well-known technical words and phrases is included on page 4B-56.

Section I. General Design of Data Processing for a Proposed Study

A. OVERALL GOALS OF THE STUDY

Probably no technical innovation in history has been so widely discussed and yet so poorly understood as has the modern computer. While most medical investigators undoubtedly have a greater insight into the nature of computers than does the general public, some of the mystery and aura surrounding them finds its way into nearly everyone's thoughts. When a problem is encountered, the computer is often visualized as an answer instead of a tool; it is spoken of as if it possessed insight and inherent problem-solving abilities. "Let's feed the data into the computer." "What does the computer say?" This vision is of course rapidly shattered when programming and debugging begin (see page 4B-34), as all who have made the attempt agree; yet, we seem to forget.

This mental bias, while seemingly harmless, does have an adverse effect in many instances: the investigator may overlook some of the most basic rules of research. The most outstanding omission may be the most obvious one: definition of the problem at hand. In any undertaking, including neurophysiology, one must first set his goal—what he hopes to extract from the given or selected quantity of data. This initial step determines the future course of action, and the proper procedure is entirely dependent upon it.

There are many goals achievable by computer methodology, but within the framework of neurophysiology and particularly electroencephalography, most problems seem to fall into two major categories: automation and research analysis. These categories may not always be mutually exclusive, but if the investigator tries to fit his problem within one or the other, at least his objectives will become clearly defined.

In *automation*, the computer performs a precisely defined task which can be and usually has been done by other, probably less efficient, means. An example would be the automatic scoring of sleep records. As performed by EEGers, relatively precise criteria exist for judging each epoch in an all-night recording of EEG, EOG (electrooculogram) and EMG (electromyogram) (Rechtschaffen and Kales 1968). However, the job is very time-consuming, relatively boring, and not economical in terms of man hours. Currently being approached by a number of data-processing techniques (*e.g.*, Frost 1970) the *goal* is to achieve a final output which is the *same* as that of expert (human) evaluation, at a saving in terms of human time and effort. Because EEGers disagree to some extent in spite of standardization, a machine which agrees as well with most of them as they agree among themselves is a more realistic goal.

The investigator seeking computer assistance in *analysis* of research results, on the other hand, hopes to extract new and useful information from the data—information not necessarily known to be present, but which might be. The limitations of human senses are recognized, and the computer's manipulative and mathematical abilities are substituted in the attempt to correlate data with other factors. An example would be the determination of coherence, or degree of relationship, between several data channels within selected frequency bands (see Volume 5A). Such information is not obtainable by visual inspection of the record (*i.e.*, by human analysis), and the computations and measurements involved in its determination are far too lengthy to be accomplished without the aid of a computer.

It is most important that the investigator himself make the distinction between automation and analysis for each case, and then set his goals within these limits. The computer scientist, mathematician or programmer, to whom the investigator may turn for help, must be made aware of these goals as well. If not, the approach chosen for a particular problem may turn out to be more costly and prolonged than either party imagines at the start. As between automation and research analysis, automation has thus far been much less exploited. This should not discourage its development, for it holds great potential because many more clinical than research EEGs are recorded and reported.

B. INITIAL APPROACH

Once the investigator has established his general goals and indicated the information he hopes to gain from computer use, he must study the various alternative approaches.

The most basic consideration is, simply, does the problem really require a computer? This is primarily a matter of economics, as it is taken for granted that the problem is worthy of solution; but is a *computer-derived* answer really worth the investment in time and money? The answers depend upon several factors, each to be determined for the specific case: (1) the generality of the problem, (2) duration of need and (3) anticipated difficulty in establishing the data-processing procedures.

A problem which appears quite difficult to solve by manual methods may still be unworthy of the effort required for computer use. For example, sorting the reports of 1000 previously recorded EEGs by age and presence of various clinical abnormalities would appear an impossible task to the lone electroencephalographer. Yet if this is the extent of the project and no further or continuing effort utilizing the same format is anticipated, he is best advised to hire extra, temporary help for a few weeks, rather than invest part of his own time and probably more total expense in a computer program. If, on the other hand, reports of only 100 records per week need to be classified but the need is a continuing one, the investment in a data-processing scheme would be profitable in the long run.

However, if the project is not one of processing *reports* on EEGs, but of analyzing the EEGs themselves, the project is one for which the basic methodology is unknown. Only for the restricted area of stages of sleep in normal subjects do we know how to automate the basic recognition of the phenomena sought. Inappropriate formalization or quantification of more complex or variable phenomena can lead to results more misleading than would appropriate non-formalization or non-quantification.

In some cases, the question is not so much whether a computer is necessary as whether the proposed analysis is likely to yield sufficiently valuable results. In these instances the decision to go ahead rests with the investigator and his insight (and his budget), but in all cases the point must be honestly explored.

If the decision is to proceed with computation, another primary consideration is choosing the type of computer to be used, *i.e.*, special purpose, general purpose, analog, digital, etc. Although the specific advantages and disadvantages of the various possibilities are discussed in detail below (Section II, B and C), a tentative decision will guide the overall study. Even though the specifics of programming or other detail development may be delayed, the investigator should outline the hardware as a preliminary step. This permits an accurate economic estimate and also provides the basis for specific evaluation of the alternative processing schemes or programs.

If the anticipated data processing is an example of automation, the researcher must scrutinize his current procedures and methodology. Can he explain it to himself in a precise manner? If not, as is often the case, the road ahead is likely to be a long one. He must expect a lengthy period of trial and error, when the results of preliminary programming attempts are compared to his present method.

In a case of automation, the result sought is inflexible. But when a new type of analysis or a new application of previously developed analytic techniques is the goal, the preliminary results are likely to be more interesting since the precise result is an unknown quantity. However, extreme care must be taken with controls and the like, simply because the result *is* unknown; proper validation and checking procedures must be instituted at each level of complexity (see Section III).

The time factors involved in computer usage must also be recognized in the planning stage, including that for programming and debugging (see Section II, B). If a large, central computer facility is involved, the investigator must learn the administrative and technical delays inherent in the system. What input rate can he reasonably hope to achieve? What are the anticipated delays in receiving the output? Will these delays and limitations influence the manner of experimental data acquisition?

Finally, consideration must be given to procedures for checking the validity of the approach and for evaluating the progress being made. This should be discussed in the overall preliminary planning, since in many cases it will require preparation of special data or simulated data and, especially, an investment of time (see Section III).

C. SPECIFIC PROBLEM FORMULATION

1. *Computer Input*

While specific details of the nature, form and value of the input to a data processor are discussed in other sections, a number of basic considerations common to all systems should be dealt with under the heading of initial formulation.

Probably the most important factor in terms of potential trouble later is the problem of artifactual signals mixed with the true data. These signals can, of course, be introduced at any stage of the data-acquisition and storage manipulations, from the electrodes to the tape recorder. Indeed, though they are not usually classified with artifacts, computer errors, plotter errors, or printer errors can equally be regarded as artifacts of these later stages. The point of this discussion is not how to *deal* with early artifacts specifically, but rather to emphasize how much simpler the data-processing scheme may be if the presence of such extraneous signals is *prevented* before reaching the computer. The presence of obvious artifacts is usually known to the investigator but too often their importance is underestimated. Man is extremely efficient as a "pattern recognizer", and thus he quickly learns a number of criteria, many relatively subtle, which permit him to disregard automatically most artifacts as long as they do not obscure "too much" of the real data. Pattern recognition, however, is extremely difficult for the computer, and the task of programming is made tremendously more complex.

A simple illustration arises in the processing of data from extracellular micro-electrodes. In most instances, the investigator is only interested in the *times* of occurrence of the single-unit discharges and desires such things as interspike-interval measurements and distributions, poststimulus-latency measurements, etc. These tasks may be done quite economically by utilizing small, special-purpose laboratory computers, with the spikes being detected by simple amplitude-discrimination procedures, providing that it can safely be *assumed* that all spikes entering the computer are real data. If care is not taken, spurious spikes may appear in the recorded tracings (*e.g.*, from dirty tapes or dirty recorder heads). When the unprocessed data are examined visually (oscilloscope or photograph), the artifactual signals are usually easily distinguished from true neuronal spikes by their non-uniformity, different durations, and unusual configurations. The computer, however, looks only at amplitude excursions and treats the artifacts as neuronal spikes. Pattern recognition is not possible on the small, special-purpose devices and would even tax the capabilities of a much larger general-purpose device, in addition to greatly increasing the programming time and skill involved. Thus, the only alternative is usually simply to do the experiment over, taking more precautions to guard against the presence or recording of unwanted signals.

The point is that the problems associated with artifacts should ideally be recognized before the data gathering begins, but if they are not, they must be honestly faced and dealt with accordingly, as soon as possible in the stage of planning the data processing.

"Hidden" artifacts are much more difficult to handle, and often cannot be planned against but must constantly be sought at all stages. An example is described in Section III, A, and illustrates the requirement for test runs on controlled or synthetic data to detect such possibilities.

The *form* of the data input to the computer plays a role in planning the specific programming approach, and this should also be considered early. Direct input to the computer is not without its problems (see page 4B-42). Problems are also associated with all forms of data storage: analog tapes, digital tapes, reproduced paper traces and preprocessed data (see Section II, C for further comments on each of these); however, it is generally true that analog tapes offer the best solution for neurophysiological problems, and for the similar problems of other clinical recordings requiring off-line analysis. When used properly, the data are preserved in essentially their original form, thus permitting playback in real time and editing by conventional "human" means (*i.e.*, EEG recorder, oscilloscope) and, for the computer, playback at high speed during the analog–to–digital conversion process.

2. *Computer Output*

Another important point to be considered initially is the best form and amount of the anticipated output. This aspect requires careful evaluation, since the *output alone* conveys the information content or result of the *entire data-processing procedure*—hence, often of the entire experiment. No specific advice can be given unless the particular problem is considered, but in general it seems that most mistakes made result in too much complexity. The computer user, in an attempt to get all the "information", tends to overburden himself with computer-generated material, which may as a result greatly exceed in volume the original, raw data.

The processor's capabilities should be used to their utmost to condense and rearrange results into a form easily comprehended by the human senses and intellect. For example, a compact, single-page, three-dimensional, contour map can convey the information contained in several hundred conventional, two-dimensional graphs or histograms (in which the individual plots are sequentially related to one another), and the end result is not only smaller but more readily understood. Prints-outs of useless (to the human) information should be avoided, except in the validation stages when this may assist the testing procedure. Output should contain only the answer, stated in simple terms, especially when automation is the goal.

Section II. Specific Means of Data Processing

A. ON-LINE DATA PROCESSING: SOME SPECIFIC METHODS AND SOME SPECIAL MACHINES

The methods of analysis of the EEG and EMG can be separated into two kinds, the first being the processing of time-locked signals, especially averaging evoked potentials, and the second the analysis of ongoing activity. There is some overlap between the two groups, because the so-called ongoing activity is sometimes related to a particular event in time, such as the increase of alpha rhythm on eye closure. For this section, processing will be considered in the first category if the effect of the stimulus persists only a short time after the occurrence, *e.g.*, measurement of evoked potentials, wherein the effect on the EEG usually persists at most for a few seconds after the stimulus. Analysis of ongoing activity is considered in paragraphs 3–6 of this Section (pp. 4B-20 et seq.) and in Volume 5A.

The need for processing of evoked potentials arises because the amplitude of the response is usually similar to or less than the amplitude of the background activity, thus making it impossible to separate the two in the primary trace. Most methods of extraction of the evoked potential from this ongoing activity depend upon the repeated presentation of the stimulus and the summation or visual correlation of the time-locked activity.

1. *Superimposition*

Dawson (1951) was one of the first to use the radar technique of superimposition to study evoked potentials. Superimposition can be done photographically using a cathode ray oscilloscope (Fig. 1), or using a Y–T plotter (Fig. 2). Few plotters have the necessary frequency response for the direct writing of the evoked potentials, so that some form of storage and slow playback (by magnetic tape, for example) is usually necessary.

When only a few trials (less than 20) are superimposed, the evoked potentials are emphasized more by the line–to–line correlation (agreement of successive curves in general shape) than the photographic integration (increased photographic density due to exact superimposition of values). This integrative property can be used to detect small responses in wideband noise, but this method is rarely used in evoked potential studies. The line–to–line correlation is most effective when amplifiers with short time-constants are used, because then the starting points are all on the same baseline. If time-constants greater than 1 sec (or DC coupling) are used, the baseline of the trace can vary considerably, and the visual effect of the superimposition is much reduced.

The big advantages of superimposition, apart from its simplicity and cheapness, are: (a) that estimates of the variability can be made and (b) that the occasional far-different trace (due, perhaps, to movement artifact) is easily ignored. This is particularly important when the response to the stimulus is a blocking of the ongoing activity (lower EEG trace, Fig. 1). As will be seen later, this effect is not easily detected using simple averaging techniques.

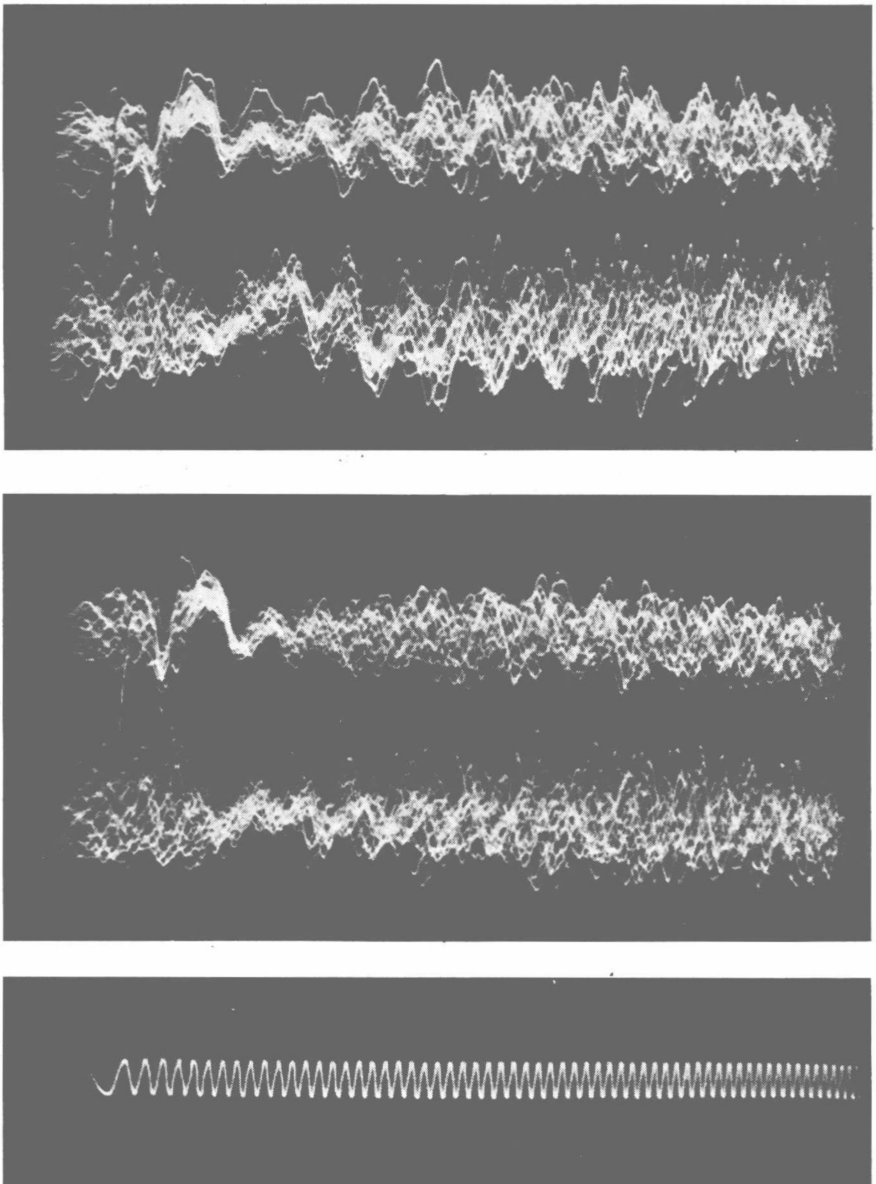

Fig. 1. Superimposition photograph of 20 traces showing evoked potentials to a flash of light at intracerebral occipital electrodes. Calibration 100 μV at 50 Hz.

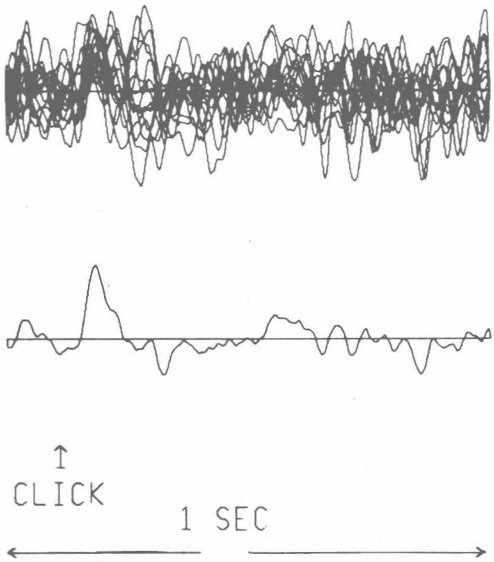

Fig. 2. Superimposition of 16 traces on Y–T plotter; average shown below. This and all subsequent figures (except Fig. 10) were produced by the LINC 8 computer at the Burden Neurological Institute, Bristol (Great Britain).

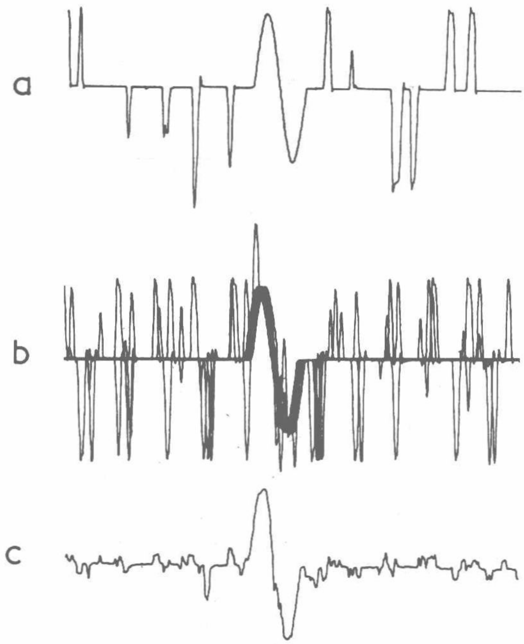

Fig. 3. A large signal-to-noise improvement is obtained when the noise is composed of brief transients. *a*: Isolated sine wave and noise. *b*: Superimposition of 9 traces. *c*: Average of 9 traces (from Cooper *et al.* 1969).