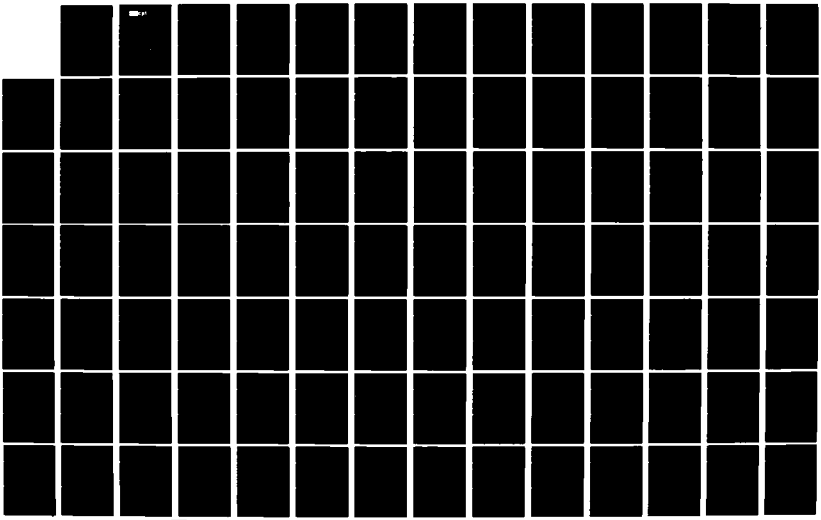
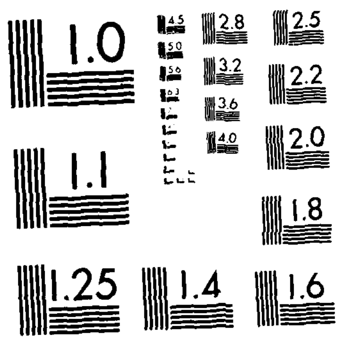
AD-A125 699 THE EVENT RELATED BRAIN POTENTIAL AS AN INDEX OF INFORMATION PROCESSING C..(U) ILLINOIS UNIV CHAMPAIGN COGNITIVE PSYCHOPHYSIOLOGY LAB E DONCHIN ET AL. NOV 82 CPL82-2 AFOSR-TR-83-0051 F49620-79-C-0233 F/G 5/10 1/2

UNCLASSIFIED NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 83-0051

COGNITIVE PSYCHOPHYSIOLOGY LABORATORY

Department of Psychology
University of Illinois
Champaign, Illinois 61820

Technical Report No. CPL82-2/
F49620-79-C-0233

November 1982

THE EVENT RELATED BRAIN POTENTIAL AS AN INDEX OF INFORMATION PROCESSING, COGNITIVE ACTIVITY, AND SKILL ACQUISITION: A PROGRAM OF BASIC RESEARCH

ANNUAL PROGRESS REPORT

Prepared for:
The Air Force Office of Scientific Research
Life Sciences Directorate

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR- 83-0051	2. GOVT ACCESSION NO. AD-A125699
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): The event-related brain potential as an index of information processing, cognitive activity, and skill acquisition: A program of basic research. ~~Annual Progress Report~~	5. TYPE OF REPORT & PERIOD COVERED: Annual Report Sept. 1, 1981-Oct. 30, 1982
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Emanuel Donchin, Christopher Wickens, and Michael G.H. Coles	8. CONTRACT OR GRANT NUMBER(s): F49620-79-C-0233
9. PERFORMING ORGANIZATION NAME AND ADDRESS: University of Illinois, Department of Psychology, Urbana, IL 61801	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102 F 2313/A4
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research (NL) Bolling AFB, DC 20332	12. REPORT DATE: November 1982
	13. NUMBER OF PAGES: 182
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Event-related brain potentials, P300, attention, workload, automation, skill, individual differences, Von Restorff effect, ocular artifact, performance, concurrent tasks, N200, reaction time, S-R compatibility, memory

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The materials assembled in this report represent work conducted with AFOSR support at the Cognitive Psychophysiology Laboratory during the reporting period. The first three items are pre-prints of papers currently in press, or under editorial review. These are followed by a set of five abstracts of papers presented at the annual meeting of the Society for Psychophysiological Research. The last three items are an abstract of a paper submitted to Human Factors, and short reports of papers presented at the Annual Conference on Manual Control, and at the Human Factors Society meeting.

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TABLE OF CONTENTS

1. Introduction

2. Cognitive Psychophysiology and Preparatory Processes - A Case Study

to appear in Kornblum, S. and Requin, J. (Eds.) Preparatory States and Processes, Erlbaum Associates, Inc., in press

3. A New Method for Off-Line Removal of Ocular Artifact

in press, Electroencephalography and Clinical Neurophysiology

4. The Performance of Concurrent Tasks: A Psychophysiological Analysis of the Reciprocity of Information Processing Resources

submitted to Science

5a. P300 and Memory: Individual Differences in the von Restorff Effect

abstract, Psychophysiology, 1982, 19(5), 558.

5b. P300 and Memory: Individual Differences in the von Restorff Effect

6. N200 Amplitude as a Function of Degree of Mismatch in a Word Categorization Paradigm

abstract, Psychophysiology, 1982, 19(5), 580.

7. P300 Latency and Reaction Time from a Visual Search Task with Varying Levels of Noise and S-R Compatibility

abstract, Psychophysiology, 1982, 19(5), 572.

8. Electrophysiology of Absolute Pitch

abstract, Psychophysiology, 1982, 19(5), 569.

9. Information Extraction and P300 Amplitude

abstract, Psychophysiology, 1982, 19(5), 573.

10. Operator Workload as a Function of the System State: An Analysis Based upon the Event-Related Brain Potential

Proceedings of the 18th Annual Conference on Manual Control, Wright-Patterson AFB, OH 1982.

11. Pseudo-Quickening: A New Display Technique for the Control of Higher Order Systems

Proceedings, 26th Annual Meeting of the Human Factors Society, October 1982.

12. An Investigation of Redundant Auditory Tracking.

Abstract of manuscript to be submitted to Human Factors.

1

Introduction

The materials assembled in this interim report represent work conducted with AFOSR support at the Cognitive Psychophysiology Laboratory (CPL) during the reporting period. The first three items are pre-prints of papers currently in press, or under editorial review. These are followed by a set of five abstracts of papers presented in October 1982 at the annual meeting of the Society for Psychophysiological Research. The last three items are an abstract of a paper submitted to Human Factors and short reports of papers presented at the Annual Conference on Manual Control, and at the Human Factors Society meeting. All but one of the abstracts report essentially completed studies which we are currently writing up for publication.

In the main, the CPL continued in this period to pursue in parallel several closely related goals. The primary mission of this research program is to develop an understanding of the Event Related Brain Potential (ERP) so that it can be used as a tool in the study of cognitive function and in the assessment of man-machine interactions. To this end, we are conducting studies that fall in three, not altogether distinct, categories, as follows:

A. The discovery of the antecedent conditions of ERP components and the elucidation of their functional significance. Much of this work focuses on the P300 component. The noteworthy findings of the current period can be briefly summarized by the following assertions:

1. We provided further evidence that the latency of the P300 is determined largely by the duration of stimulus evaluation and categorization processes, and is independent of response selection and execution processes. The study by Magliero et. al. (Item #6)

extends the results we reported earlier in McCarthy and Donchin (1981). In the earlier study we varied the discriminability of a target by introducing "noise" into the display and we varied motor demands by requiring both compatible and incompatible responses. The latency of P300 was affected solely by the introduction of noise and was unaffected by response characteristics. In the present study we replicated the previous study using different levels of noise. The data indicate that the latency of P300 is, as we would predict, a graded function of the noise level. In a related study we provided further evidence that the component in question is indeed a P300, as its amplitude varied as a function of the task-relevance of the eliciting stimulus.

2. In a study by Mane et. al. (Item #8) we show that the amplitude of P300 varies with the value of the information conveyed by the P300 regardless of the probability of the stimulus. Stimuli presented in a sequence provided varying degrees of information regarding succeeding stimuli. The greater the diagnostic value of the stimulus, the larger the P300 it elicits. These data confirm previous work from our laboratory and from other laboratories that suggest that the P300 reflects activity that is invoked in the service of updating the internal models of the environment.

3. The study by Klein et. al. (Item #7) examined the ERPs elicited during the standard odd-ball paradigm in music students that do, and do not, possess the ability to make absolute-pitch discriminations. This study was motivated by the fact that the

absolute-pitch skill is often accounted for by postulating that individuals with this skill maintain a permanent representation of reference tones in their working memory. We reasoned that if this is the case, and if P300 is invoked whenever a need to update working memory is encountered, then individuals with absolute pitch will show little, or no, P300 when performing an auditory oddball task, and a relatively normal P300 when performing a visual oddball task. The data from the first 8 subjects confirm this prediction. Data from eight additional subjects are currently being analyzed. The preliminary results of this analysis continue to confirm prediction.

4. The most direct test of our theory concerning the functional significance of the P300 was conducted in the study of the Von Restorff effect and P300 by Fabiani et. al. (Item #4a, a more extensive report of these results is presented in Item #4b which was distributed at the SPR convention). The study was designed to test the hypothesis that the larger the P300 elicited by an item the more likely is that item to be recalled in a subsequent test. We capitalized on Von Restorff's well established finding that items that differ in some respect from other items in a list will be better recalled. The main discovery of this study was that there are remarkable individual differences in the Von Restorff effect and that these are correlated with the strategy the subjects use for item storage and recall. Furthermore, the relation between P300 and recall is also related to these individual differences. Briefly, it turns out that

individuals who use rote-memory strategies show a strong Von Restorff effect, and in their case the larger the P300 the more likely the subsequent recall. Individuals who apply semantic organization procedures show almost no Von Restorff effect. For these individuals there appears to be no relationship between P300 amplitude and recall. The data provide confirmation, and force an extension, of the context-updating model of the P300.

5. Two studies in this collection focus on components other than the P300. The N200 is emerging as an interesting endogenous component that is controlled by a number of experimental manipulations that differ from those that control the P300. One of the more interesting claims made regarding the N200 is that its amplitude is proportional to the degree to which the incoming stimulus mismatches the model expected by the subject. It was difficult, however, to validate this claim as most operations that will increase the mismatch in this fashion increase the processing time associated with the stimulus. As a consequence P300 latency increases as well. As P300 often masks N200, the increase of N200 amplitude with mismatch may be due to its emergence from the P300 cover. Pritchard et. al. (Item #5) used a standard category verification paradigm to test the hypothesis. The key element in their design was the fact that when one has to decide whether an exemplar belongs to a category it takes less time to reject items that in some way match the category. For example, it take longer to decide that COPPER is not a BIRD than it does to decide that a LION is not a BIRD. Our results show that even though the latency

of P300 is shorter when the stimulus is COPPER, rather than LION, the amplitude of N200 elicited by COPPER is larger. Thus, we confirm the hypothesis the N200 amplitude is related to degree of mismatch.

In Item #1, Donchin et. al. report, in the context of a review of the nature of the Cognitive Psychophysiological paradigm, a study of the complex of ERP components that are elicited in the course of a fairly complex game-like task, in which the subject participated in a simulated automobile race. By requiring subjects to drive as fast as they can, while at the same time penalizing them for collisions we placed a high value on the processing of warning stimuli that heralded the danger of collisions and required the subject to slow down. By varying various aspects of the warning signals and their timing we were able to examine in detail the role of various "Slow Waves" in the subject's preparation for his encounter with obstacles.

B. Studies designed to determine the utility of ERP components in the analysis of the nature of man-machine interactions. The three sets of studies described in this report do not exhaust the work of the CPL in this domain. Most of our work that is designed to develop applications of ERPs in human engineering is performed within the context of contracts with AFAMRL, DARPA, SAM and similar agencies. In keeping with the basic research mission of the present contract we focus here on three studies that bear directly on the theoretical foundations of the Human Factors work.

1. The study by Wickens et. al. (Item #3) fills a serious gap in the theoretical underpinnings of the use of ERPs in the assessment of Mental Workload. We have established in our previous work that when a primary task places demands on the subject's perceptual resources the amplitude of the P300 elicited by a stimulus associated with a concurrent oddball task is diminished. Furthermore, the larger the demands of the primary task, the greater the diminution of the P300 elicited by the secondary task probes. The interpretation of these findings assumed that P300 draws on a pool of perceptual resources that are shared with the primary task. Therefore, P300 amplitude can serve as a measure of the resources remaining in the pool as a consequence of the demands of the primary task. This model implies that if we could observe P300s elicited by primary task stimuli the amplitude should increase as a function of primary task difficulty. By using a step-tracking procedure that, due to the discrete nature of the stimuli, permits the extraction of ERPs to the primary task we were able to show the predicted reciprocity of P300 amplitude. For auditory probes - P300 amplitude for secondary tasks decreases, and for primary tasks increases, with increasing demands by the primary task.

2. The study by Gill et. al. (Item 10) utilized P300, coupled with a state space analysis based upon the theory of manual control, to understand the information processing demands of manual control of higher order systems. Based upon our earlier findings that P300 amplitude was selectively sensitive to the perceptual/cognitive demands of higher order manual control, we employed a more detailed assessment of the temporal changes in this load, in an effort to choose between

perceptual and cognitive sources. Specifically, we asked whether P300 elicited by auditory probes during second order tracking was selectively attentuated as a function of the momentary perceptual state of the display, or was continuously attenuated independent of that state. Our results supported the second conclusion, thereby indicating that the processing demands of second order control were heavily central, rather than perceptual in their locus.

This investigation was part of a larger effort to develop a technique of display augmentation for second order control known as pseudo quickening. In this technique described in Item 11, information regarding the momentary state of the system, defined in terms of the variables necessary to obtain optimal control, is provided by intensification of different sides of the error cursor. The technique is compared with other display augmentation methods as well as conventional unaided displays in a transfer of training design. It is shown to provide superior training and understanding of the second order dynamics than either the other forms of augmentation or the conventional display.

3. In Mane and Wickens (Item #9) we have continued work sponsored under the earlier AFOSR contract on the use of the auditory channel as an alternate information display in manual control. In the previous investigation (Isreal's thesis), we examined the auditory display as an isolated channel to convey error information. In the present investigation we establish the value of the auditory channel as a means of redundant information display under conditions of high concurrent visual workload. While this investigation has no direct

electrophysiological component, it does bear directly on issues in dual task performance that are closely related to our P300 workload research.

C. Methodological Studies

1. The fact that movements of the eye ball generates field potentials that appear to scalp electrodes remarkably like the fields generated by brain activity has always plagued ERP research. The problem has been particularly serious for investigators of endogenous components. We have now developed as described in Gratton et. al. (Item #2) a method that allows off line subtraction of eye movement artifacts from all trials. The technique is very useful because it allows us to retain all the trials in an experiment, including those during which the subject made large eye movements. As ERP research is moving increasingly into investigations of the brain activity of subjects who perform complex tasks in interaction with elaborate displays the need to deal with eye movements has become serious indeed. The technique we described has proven most effective and is now used routinely in most of our studies.

The studies reviewed briefly above are described in some of the enclosed papers in some detail. Other studies are now being written up. Our next interim report will include the "in press" versions of these studies. Interested readers are, however, encouraged to contact us so we can provide them with the more detailed reports as they become available.

To appear in Kornblum, S. and Requin, J. (Eds.) Preparatory States and Processes, Erlbaum Associates, Inc., in press.

COGNITIVE PSYCHOPHYSIOLOGY AND PREPARATORY PROCESSES: A CASE STUDY

By

Emanuel Donchin, Michael G.H. Coles and Gabriele Gratton

Cognitive Psychophysiology Laboratory
Department of Psychology
University of Illinois
Champaign, Illinois 61820

Acknowledgements: The study described in this report was supported by contract #F49620-79-C-0233 from AFOSR, as well as by DARPA/ONR contract #N000-14-76-C-0002. Partial support was also provided by Wright Patterson AFB and Brooks AFB, through contracts #F33615-79-C-0512, and #F33615-82-C-0609, F33615-82-K-0611. The support of Al Fregly of AFOSR, Judith Daly and Jack Thorpe of DARPA and of Robert O'Donnell and John Courtright is acknowledged. The Race Car task was programmed by Ron Clapman and Dave McWilliams. Art Kramer, Narda Robinson, and Sue Kemp-Fincham made helpful comments on preliminary versions of the manuscript.

Introduction

Cognitive Psychophysiology, as its name implies, is a marriage of cognitive psychology and psychophysiology. The basic premise of this union is that the understanding of cognitive processes can be enhanced by augmenting the traditional tools of the cognitive psychologist by adding tools based on the measurement of physiological functions (Donchin, in press). The psychophysiological data are, of course, useful only to the extent that they complement and expand the view of the mind that can be developed with the use of more traditional techniques.

Psychophysiologists are psychologists who extend the range of observable aspects of behavior by developing, and using, techniques that allow the measurement of the activities of "physiological" systems. The reference is generally to the measurement of such variables as Heart Rate, the Galvanic Skin Response or the Event Related Brain Potential (ERP). When such measures are described as "physiological", and this term is used to distinguish these recordings from "behavioral" measures, one espouses a model that implies a separation between "behavior" and "physiology" that is not easily supportable. It seems better to adopt a holistic view which maintains that the organism in its entirety is involved in any act (Goldstein, 1939). Although action is ultimately manifested by specific muscular acts, a description of these acts is not an exhaustive (or even a sufficient) description of behavior. It is clearly the case that vascular, glandular, and neural activity are part and parcel of the same behavioral act. For example, when a person utters a sentence, a transcription of the sentence may for certain purposes be a sufficient record of the speech act.

But changes in cortical blood flow that accompanied the utterance may well be a necessary component of the speech act, and therefore, these blood flow changes are as much "behavior" as are the utterances. The psychophysiologist expands the study of behavior by including measures of these internal activities in the range of observation. In this fashion, it is possible to monitor behavioral subsystems whose activity cannot be observed when one restricts measurement to overt (muscular) behavior. In this way the psychophysiological enterprise, when properly deployed, enriches the study of cognition.

The range of internal, "physiological" processes which can be measured is large. In fact, any physiological function that can be measured from the human without puncturing the skin qualifies as a psychophysiological measure (Greenfield & Sternbach, 1972). However, in this chapter we shall focus on the event-related brain potential (ERP) (see Callaway, Tueting, & Koslow, 1978; Galambos & Hillyard, 1981; Donchin, in press).

The Event-Related Brain Potential

The ERP is a series of voltage oscillations that are time-locked to an event. It is derived by averaging samples (epochs) of the electroencephalogram (EEG) recorded from the human scalp with each sample having the same temporal relationship to a particular event. Note that we can look at activity preceding an event, as well as activity following an event. This is particularly important in investigations of preparatory processes. The voltage oscillations derived in this manner are regarded as manifestations of different "components". Components are defined in terms of their polarity (positive or negative voltage), latency range (temporal relationship to the event), and scalp distribution (variation in voltage with electrode location on the scalp), as well as by their relationship to experimental variables. Components can be quantified using simple magnitude measures or through the application of more advanced techniques such as Principal Component Analysis (PCA) and Vector Analysis (Gratton, Coles, & Donchin, 1981). They are labeled by a polarity descriptor (P or N for positive or negative) and a modal latency descriptor (e.g. 300, for 300 msec). Thus, the P300 is a positive ERP component with a modal latency of 300 msec. In some cases, as with Contingent Negative Variation (CNV) and Slow Wave (SW), the descriptors are omitted.

The Psychophysiological Paradigm: Assumptions and Approaches

The assumptions and the model underlying the use of ERPs in Cognitive Psychophysiology have been presented elsewhere (Donchin, 1979, 1981). In brief, we assume that the voltages we record at the scalp are the result of synchronous activation of neuronal ensembles whose geometry allows their individual fields to summate to a field whose strength can affect scalp electrodes (Galambos & Hillyard, 1981). It is convenient to parse the ERP into a set of components. The component, in our scheme of things, is characterized by a consistent response to experimental manipulations (see Donchin, Ritter & McCallum, 1978, for a discussion of components). We further assume that each component is a manifestation at the scalp of an intracranial processing entity. We are not implying that each ERP component corresponds to a specific neuroanatomical entity or that the activity manifested by the component corresponds to a distinct neural process. Rather, we assume that a consistent information processing need, characterized by its eliciting conditions, activates a collection of processes that, for perhaps entirely fortuitous reasons, have the biophysical properties that generate the scalp-recorded activity.

As a working hypothesis we postulate that ERP components are manifestations of functional processing entities that play distinct roles in the algorithmic structure of the information processing system. In other words, we believe that it is possible to describe in detail the transformations that the processing entity applies to the information stream. The goal of Cognitive Psychophysiology, within this framework, is to provide such detailed descriptions. This may be achieved by developing comprehensive descriptions of the conditions governing the elicitation and attributes of the components (the "antecedent" conditions). These descriptions can be used to support theories that attribute certain functions to the subroutine manifested by the component. In turn, the theories should lead to predictions regarding the consequences of the elicitation of the subroutines, predictions that can be tested empirically. The P300 component of the ERP has been analyzed in this manner in some detail, (see Donchin, 1981, for a discussion of the general approach, and Fabiani, Karis and Donchin, 1982, for an illustration of an empirical test of a prediction regarding an ERP's consequences).

Thus, ERPs are likely to be particularly helpful in monitoring information processing activities which lack an obvious overt manifestation. This being the case, one would have expected to see ERPs marshalled in the analysis of preparatory processes. These processes are, almost by definition, covert. One infers that an individual was prepared to perform some task when the task is performed "better", or "worse", in some circumstances than in others, and when this changed performance cannot be attributed to differences in skill, motivation or any number of other variables. We assume "preparation" to have taken place particularly when there is an antecedent event that can be presumed to have allowed the subject to launch a series of internal activities that are ultimately responsible for the change in performance quality.

While preparation may be overt, as when one places a finger on a response key or crouches at a start line, more often than not the critical "preparatory" activities are transacted in the nervous system without manifestations that can be observed by traditional means. Consider, as an example, the difference between warned and unwarned reaction time (RT). It is well established that the presentation of a warning stimulus a few hundred milliseconds before a response is required markedly shortens the RT (Woodworth, 1938). However, even though these phenomena have been well known for decades, knowledge of the processes that are active during the foreperiod, and are ultimately responsible for the shortened RT, remains limited (though see Brunia, this volume; Mountcastle, 1975; Posner, 1978; or Requin, 1978, for a variety of approaches to the problem). It is therefore reasonable to utilize the power of ERP technology in an attempt to uncover at least some preparatory processes by analyzing their ERP manifestations.

There have indeed been several attempts to address this complex of issues from the earliest days of ERP research. Some investigators focused on the differences between the ERPs elicited by stimuli associated with fast and slow responses (e.g., Donchin & Lindsley, 1965; Morrell and Morell, 1966). These studies consistently showed that fast RTs were observed for stimuli eliciting large ERPs. But these investigators did not directly observe the activities that transpired during the foreperiod and therefore their studies contributed little to our understanding of preparation. There has, however, been a direct and very promising attack on the foreperiod. Two discoveries were reported in the mid-sixties that promised to provide investigators of preparatory processes unique and powerful tools. We refer, of course, to the discovery of the Contingent Negative Variation (Walter, Cooper, Aldridge, McCallum, & Winter, 1964), and of the Readiness Potential, (Kornhuber & Deecke, 1965). These two ERP components, as will be seen below, are remarkably robust and easy to observe. In both cases the components are manifestations of activities that are patently associated with the preparation to respond, and both provide a promising window on the foreperiod. Yet, despite much research, little progress in the development of a theory of preparatory activities can be attributed to the studies of the Contingent Negative Variation (CNV) or the Readiness Potential (RP). In the following section we briefly review the current knowledge of the CNV, the RP and some related components. We shall endeavor to point to some of the difficulties in interpretation that have encumbered the CNV and the RP as useful tools in the analysis of preparation. We shall then illustrate the manner in which these, and similar, components can be used to augment the information that can be acquired about a subject's preparation to respond.

ERP Components and Preparatory Processes

Walter, Cooper, Aldridge, McCallum, and Winter (1964) first reported that when a warning stimulus signals the future occurrence of an event to which a perceptual judgment, cognitive decision, or motor response must be made, an ERP component occurs which is characterized by a slow increase in negativity (Hillyard, 1973). This "contingent negative variation" (CNV), occurring in the interstimulus interval or foreperiod, is most evident at the vertex (in the center of the scalp), although negativity is also observed at lateral electrode sites and at locations in front of, and behind, the vertex.

If the CNV and related ERP components are manifestations of preparatory processes, then we would expect that (a) they are influenced by antecedent manipulations which are linked to the level and quality of preparation, and (b) subsequent behavior which depends on preparation would be related to the components.

Before we present evidence bearing on these two points, we must consider the possibility that the CNV is, in fact, a composite of at least two ERP components. First, when the interval between warning and imperative stimuli is long (e.g., 4 sec), two distinct negative shifts are observed (e.g., Loveless & Sanford, 1974; Weerts & Lang, 1973). The first negativity reaches a maximum about 1 sec following the warning stimulus, while the second reaches a maximum at the time the imperative stimulus is presented. Loveless and Sanford (1974) refer to these two waves as the Orientation or "O" wave and the Expectancy or "E" wave. Presumably, these two components overlap when the interstimulus interval is short. McCarthy and Donchin

(1978) have shown, using Principal Components Analysis (PCA), that these two components can be identified even when the interstimulus interval is as short as 1000 msec. The two components differed in scalp distribution, and in their responsiveness to experimental manipulations (see below), making them identifiable as distinct sources of variance in the PCA.

Although both components of the CNV are associated with negativity at central electrode locations, they are distinguishable in terms of activity seen at lateral and midline sites. The early "O" wave component shows maximum negativity over frontal areas and positivity over parietal sites. On the other hand, the later "E" wave exhibits negativity at all sites (maximal at the vertex). This negativity may be lateralized depending on the anticipated response requirements following the imperative stimulus (Rohrbaugh, Syndulko, & Lindsley, 1976; McCarthy & Donchin, 1978).

Demonstration of the existence of at least two, rather than one, ERP components in the interstimulus interval led investigators to reexamine the relationship of these components to manipulations of experimental variables within the traditional CNV paradigm and to ERP components observed in other paradigms. Rohrbaugh et al. (1976) found that the distributional characteristics of the later "E" wave were precisely the same as those for the readiness potential (RP), a lateralized negative wave that precedes voluntary movements (Kornhuber & Deeke, 1965; Kutas & Donchin, 1974; Desmedt, 1977). Implication of this component in processes related to motor response preparation is also supported by the observation that it is severely attenuated when the motor response requirements inherent in the imperative stimulus are reduced (Gaillard, 1977; Loveless, 1975; Perdok & Gaillard, 1979). It is more difficult to interpret the early component.

Rohrbaugh et al. (1976) argued that its latency and distribution correspond to the ERP component elicited by unpaired stimuli requiring passive attention. They imply that this "O", or orientation wave, can be identified with processes invoked by the alerting properties of a stimulus (Loveless, 1979). This conclusion is consistent with the observation that the magnitude of the component increases with the task-relevant information provided by a warning stimulus (McCarthy & Donchin, 1978; Kok, 1978). Loveless (1979), among others, proposed that the early component is related to the "slow wave" (SW) identified by Squires, Squires, and Hillyard (1975) and Squires, Donchin, Herning, and McCarthy (1977). Others (e.g., Kok, 1978) have pointed to the similarity, in its parietal aspect, between the early component and the P300.

Thus, the early "O" wave component is often recorded in association with two other components, P300 and SW, which have been evaluated in other paradigms. The psychological processes manifested by these components have been variously identified as "surprise", "context updating", "resource allocation" (Donchin, 1981), interpretations consistent with the proposal that the components are influenced by the task relevance of the information provided by the warning stimulus.

As Donchin (1981) has suggested, questions about the significance of ERP components can only be answered by evaluating their consequences as well as their antecedents. The relationship between the CNV (late wave) and reaction time (RT) to the imperative stimulus has been the subject of much controversy (Tecce, 1972). However, recent evidence from studies where foreperiod length has been sufficient to permit a dissociation between early and late components has revealed a significant negative relationship between

RT and degree of negativity (e.g., Rohrbaugh et al., 1976). Furthermore, Perdok and Gaillard (1979) found no relationship between the magnitude of the negativity of the late component and perceptual sensitivity. Taken together, these data support the contention that the late negativity is associated with motor, rather than with sensory, preparation. The consequences of the early component have not been evaluated. However, if this component is related to the P300 or SW, then it is probable that one, or more, of its attributes will be related to aspects of the subject's performance which depend for their completion on the successful processing of information provided by the warning stimulus. In this sense, the early component (P300/SW) can also be regarded as a possible manifestation of a preparatory process, which is distinct from that manifested by the later component (traditional CNV).

It would appear, therefore, that there is an array of psychophysiological responses that can be recorded in the foreperiod of an RT experiment. Even if not every component proves "real" and only some are manifestations of interpretable processing entities, ERP recordings may contribute to the understanding of preparation. The diversity of the components is especially promising because it may allow an examination of the diversity of preparatory processes. There must be, we assume, a range of preparatory activities triggered by a warning stimulus. Having detected a stimulus and identified its significance the subject must <u>orient</u> to the source of the relevant information. This orienting may launch a process of <u>monitoring</u>, in which the environment is scanned for relevant input, which is in turn <u>processed</u>. Intermediate events that occur during the foreperiod, as well as <u>timing</u> processes, all combine to mobilize resources and to perform the necessary sensory-motor adjustments that ultimately lead to the improved performance. It would be useful indeed if different ERP components prove to be indices of these different components of preparation.

It is clearly necessary, therefore, for studies of the ERP components to be designed so that they manipulate independently likely determinants of the components and evaluate their consequences. The study we describe below seeks to do this by employing a multiple warning paradigm in which a behavioral action is variously signalled by multiple warning stimuli. The task relevance of the various warning stimuli was manipulated by changing the temporal proximity between the stimuli and anticipated behavioral action and by varying their temporal predictability. In the context of the interest of this laboratory in human engineering problems, we designed a task in which these manipulations were couched in a veridical man-machine interaction. Subjects were required to "drive" a simulated car depicted on a CRT display. The car was equipped with "radar" (warning stimuli) which detected forthcoming obstacles, giving varying degrees of advanced warning.

The Racecar Experiment

Figure 1 represents the elements of the racecar task. Six subjects performed a simulated driving task presented on a CRT display. They were required to "drive" as fast as they could but to avoid hitting obstacles which would appear from time to time. They were told that they would be warned in advance of these other obstacles but that the amount of warning would vary. Apparent speed of the car was determined by forward movement of a joystick, controlled by the right hand. When the stick was in the center position, the car moved at a minimum speed (17.5 miles per hour); in the full forward position, a maximum speed of 105 miles per hour was attained. Lateral movement of the stick controlled the lateral position of the car. At intervals which varied between 10 and 20 sec (mean = 15 sec), the digit 5, 3, or 1 (probability of each digit = .33) appeared inside the subject's car. At fixed intervals (2.25 sec) or variable intervals (1.5 - 3.0 sec) a countdown to zero was displayed in units of one. Shortly after the zero, a packet of obstacles appeared at the top of the screen. As the subject's car was "racing" towards these obstacles, collision could be avoided only by slowing down and carefully negotiating the obstacles. In each packet, there were four obstacles, two in each lane. Spacing between obstacles was such that it was always possible for the subject to avoid them.

After a practice session, subjects participated in two experimental

---

Insert Figure 1 About Here

---

sessions, one with variable intervals between countdown numbers, the other with fixed intervals. In each session, there were three runs. In each run there were 33 instances of each type of warning period (short - beginning with countdown at 1; medium - beginning at 3, and long - beginning at 5). In one session, the intervals between countdowns were fixed; in the other they were variable. To encourage fast, and <u>accurate</u> performance, subjects were paid a one dollar bonus if they traveled more than 22.5 miles and sustained less than 15 collisions in a run. Information about elapsed time, distance, and number of hits was continuously available on the screen.

The EEG was recorded from Fz, Cz and Pz referred to linked mastoids. Burden Ag-AgCl electrodes, attached with collodion, were used. In addition, Beckman biopotential electrodes were placed laterally and supra-orbitally around the right eye to record EOG. Electrode impedance did not exceed 5 Kohms/cm.

Van Gogh amplifiers (model 50000), with a time constant of 10 sec and upper half-amplitude cutoff of 35 Hz, were used to amplify EEG and EOG traces. EEG, EOG, voltage changes corresponding to vertical and lateral movements of the stick, and pulses contemporaneous with the countdown numbers and hits, were digitized at 100 Hz for off-line analysis. A PDP 11/40 computer controlled both the presentation of the task and the collection of the data.

Data Reduction. The EEG data for a 1400 msec epoch, beginning 200 msec before a digit, were averaged separately for each subject, for each countdown digit and for each series and session (fixed or variable). Because excessive EOG activity occurred following some countdown numbers particularly 0, an eye-movement correction procedure was applied to all data (Gratton, Coles, & Donchin, submitted). This procedure uses the relationship between activity in EEG and EOG traces that is not linked to known stimuli to compute regression equations describing the EOG-EEG relationship. This procedure is performed separately for blinks and saccadic eye-movements and individually for each subject, each session, and each electrode. As a result of these operations, we derived separate average ERPs for each subject, each session, and each countdown digit (i.e., 1/1, 1/0, 3/3, 3/2, 3/1, 3/0, 5/5, 5/4, 5/3, 5/2, 5/1, 5/0 - where the first element is the first digit of the countdown series and the second is the actual digit). Averages were computed with and without subtraction of a 200 msec baseline which preceded each countdown number.

For each epoch, we also recorded the vertical stick position by digitizing the voltage signal that represented that position. Movement in the vertical plane was used to evaluate the relationship between ERP components and behavioral action.

Results

Our purpose in this experiment is to study preparatory processes induced by warning stimuli, when these stimuli do, and do not, provide the subject with data on the length of the interval after which they will have to take action. All the warning stimuli were presented in the form of a single digit (0-5) superimposed on the subject's car. However, the digit could be embedded in a long (5,4,3,2,1,0), medium (3,2,1,0) or short (1,0) sequence. Furthermore, the digits were sometimes separated by fixed, and other times by variable, intervals. We shall examine the degree to which the ERPs elicited by these stimuli differed and the extent to which such differences illuminate covert preparatory activities. However, we must first determine if the duration and the regularity of the warning stimulus series did indeed induce preparatory activities. To this end, we first evaluate the effects of both the length of the warning interval and the temporal predictability of warning stimuli on the subject's overt performance of the driving task. Preparation would, we assume, manifest itself in variations in subjects' driving speed depending on the different warning conditions.

Performance Data. In general, subjects followed instructions and were both quick and accurate. They consistently earned the $1 bonus by driving 22.5 miles in the 30 minute race, with less than 15 collisions (mean=7.6). Average "driving speed" during the countdown periods is shown in Figures 2 and 3. As can be seen in the Figures subjects drove at the maximum speed throughout the period preceding the first warning stimulus that began the countdown series. When the countdown digits appeared, the subjects' strategy varied with the digit and the regularity of the series. But, in all conditions, the subjects decelerated between the "1" and "0" steps of the countdown. (Recall that a "0" indicated the immediate appearance of the obstacles). The subjects tended to decelerate to the minimum speed while negotiating the obstacles and accelerated again after the last obstacle was passed.

The task was relatively easy, at least in the sense that subjects were quite able to avoid collisions. Thus, in searching for behavioral manifestations of different degrees of preparation, we chose to evaluate variations in driving speed, as indicated by joy stick movements, rather than collisions. To quantify movement, we established the point at which a criterion change in vertical stick position occurred. The criterion adopted was a movement corresponding to a decrease in speed of 10 mph in 50 msec. The latency of this change from the appearance of the penultimate warning digit ("1") was defined as the latency of movement onset. When there was no reduction in speed that satisfied the criterion prior to the expiration of the recording epoch (1200 msec after the "1"), a value of 1200 msec was assigned to the latency. Since this occurred in a substantial proportion of the trials, we sorted trials into a "fast" category including all trials for which deceleration occurred before the end of the epoch; and a "slow" category, for which deceleration occurred after the end of the epoch.

Table 1 shows the number of trials classified in each category for the "fixed" and "variable" interval conditions, over all subjects. Decelerations occurred earlier when the interval between warning digits was variable (Chi square = 116.45, $p<.001$). Table 2 shows that the number of trials classified as fast and slow varied as a function of the starting digit in the countdown series (5, 3, or 1). Subjects tended to decelerate earlier following a long countdown sequence than they did following a short sequence (Chi square = 18.57, df = 2, $p<.001$). The dependence of deceleration on the length of the countdown series was evident regardless of the regularity of the interval between digits.

These data suggest that the subjects' strategy and, by implication, the

----

Insert Tables 1 & 2 About Here

----

extent to which preparatory activity was induced by the warning stimuli, varies both with series length and with the temporal predictability of the countdown stimuli. The longer the countdown series (i.e., the greater the warning period), the earlier the deceleration. It is difficult to say whether the earlier deceleration after long countdowns was "better" or "worse" for the subject. As collisions occurred very infrequently, and as the subjects virtually always achieved their race goal, there was not enough variance in the consequences of the subject's actions to allow an evaluation of the subject's performance in terms of these actions. In subsequent studies the task will be made much more difficult. For the present discussion we can only note that there was a clear relation between the pattern of warnings and the latency of the deceleration. If we assume that deceleration was needed to avoid collisions then the earlier decelerations may have been adaptive.

The data also suggest that the less regular, and hence less predictable, the interval between countdown digits, the shorter the deceleration latency. These data can be interpreted in the following way. In the "fixed" series, the subject receives timing information from the intermediate events. The deceleration, while no doubt triggered by the "1" warning, is programmed to depend, in part, on the timing information. As this allows the subject to anticipate the appearance of the "0", deceleration is less dependent on, and occurs later than, the appearance of the penultimate warning. Not so for the "variable" series, in which timing information is not available. The deceleration must, therefore, be much more clearly a reaction to the "1" and is more likely to have a shorter latency. This analysis suggests that the subject's deceleration latency will depend on the degree to which the intermediate stimuli are used as a source of action-guiding data. In the fixed interval, attention to the intermediate stimuli leads to a useful reliance on timing cues. In the variable series these same stimuli provide relatively useless data and therefore attention to these stimuli will have a different implication for performance. We return to this point below.

Having demonstrated that our independent variables do, in fact, affect performance, and that differential preparation is induced by the different manipulations, we proceed to consider the ERP data. We first consider short term, phasic, responses to the discrete digits. We also examine long term, tonic, effects which are manifested in slow fluctuations of the "baseline" voltage levels that occur over the entire countdown interval. Second, we consider these phasic and tonic measures both in relation to our experimental manipulations and as a function of deceleration latency within each condition and series.

The ERP Data. Figures 2 and 3 show the ERP data for each countdown digit for each series, separately for the variable and fixed interval conditions. Note that there are phasic ERPs, one elicited by each digit. They vary in pattern with the digits, the sequence length, and with the regularity of the interval. That is, there are indeed scalp manifestations of intracranial activities that are affected by the same manipulations that induced preparatory processes in our subjects. To the extent that this is more than a fortuitous correlation it is possible that the processes manifested by the ERPs are identical, or related, to the preparatory activities that determined the variance in the subject's performance. The analysis of the data described below attempts to determine if this is indeed the case. Note also that the baseline voltage levels that appear in the 200 msec preceding each digit also vary. These baseline shifts are particularly evident in the longer series. To analyze the phasic ERPs, we subtracted the baseline preceding each digit from the data following each digit.

Waveforms relative to different countdown numbers, for each subject,

----

Insert Figures 2 & 3 About Here

----

condition, and electrode (9x6x2x3=324 waveforms), were subjected to a principal components analysis (PCA, Donchin & Heffley, 1978) to identify the different ERP components in the waveforms and their sensitivity to experimental manipulations. Data for the digit "0" in each series were excluded from the analysis because of large differences in baseline levels among electrodes. Two components of interest were extracted by the PCA. The first was identified as a traditional late CNV, with maximal negativity at Cz and less negativity at Fz and Pz ($F=13.92$, $df=2,10$, $P<.01$). It begins about 300 msec after the stimulus and increases gradually until the end of the epoch. Analysis of variance of the component scores revealed a significant main effect of countdown digit ($F=6.00$, $df=8,32$, $P<.01$). In each of the three countdown series, the greatest negativity was evident in response to the digit "1". This finding was supported by a significant digit x electrode interaction ($F=5.45$, $df=16,80$, $P<.01$). As can be seen in Figure 4a, the characteristic CNV distribution is evident for the countdown digit 1 but not for other digits and is largest for the series in which 1 was the first digit. Finally, the CNV tended to be larger for the variable than the fixed condition ($F=4.25$, $df=1,5$, $P<.10$).

We view this late CNV component as a manifestation of a process specifically associated with motor preparation. The component is most prominent just prior to movement (following the digit "1") and for the

----

Insert Figures 4a & 4b About Here

----

variable condition where the subject cannot rely on the intermediate countdown numbers to time the movement.

A second component that was affected by the experimental manipulations displayed a scalp distribution that made it difficult to identify it with one of the traditional ERP components. Its most noteworthy feature is a frontal negativity concurrent with a parietal positivity. This scalp distribution is characteristic of the Slow Wave (SW) described by Squires, et al. (1975). However, the component we observe is not preceded by a P300. Yet, on most other occasions in which the SW has been observed, it is preceded by a P300. Perhaps this component is equivalent to the "O" wave described by Loveless and Sanford (1974) and by Rohrbaugh, et al. (1976). However, in this paper, we label it SW. The SW reaches its maximum at about 440 msec following digit onset. Analysis of variance of component scores for this component revealed ($F=5.03$, $df=8,40$, $P<.01$) that the first digits in each countdown series were associated with a greater SW. Furthermore, the significant number x electrode interaction indicated that this component was elicited only by the first countdown digit (5/5, 3/3 and 1/1) ($F=8.95$, $df=16,80$, $P<.01$) (see Figure 4b). Thus the SW appears to be a manifestation of a process triggered by the first warning stimulus, and not by any of the intermediate events. It is noteworthy that this component's amplitude does not vary with the length of the interval indicated by the warning stimulus. There is no difference between the SW elicited by a "5", a "3", or a "1". We view this component, therefore, as a manifestation of a general orientation response. Its activation precedes the change in strategy from driving "flat out" to the initiation of preparatory processes, though these processes are not realized until deceleration occurs.

As we noted above, the two phasic components, the CNV and the SW, are superimposed on tonic shifts in baseline levels. To evaluate these tonic shifts, an analysis of variance was performed on the mean baseline activity recorded over the 200 msec period preceding each countdown number (including the 0s). The results of this analysis revealed that the polarity of the baseline varied as a function of countdown number and the countdown series ($F=3.02$, $df=8,40$, $P<.01$). For the countdown series beginning with a "5" or a "3", the baseline for all electrodes becomes positive following the first number (more so for the 5 series than for the 3 series), and remains positive (even though slowly declining) through the intermediate numbers. Then, for all series (and all electrodes), the baseline is negative prior to the "0". The distribution of baseline activity varied among electrodes for the different countdown numbers ($F=2.34$, $df=16,80$, $P<.01$). Although the positivity referred to above was equivalent at all electrode sites, the negativity preceding the 1 (in the series beginning with 5 or 3) and the 0 (in all series) was differentially distributed. It was maximal at Cz, and minimal at Fz and Pz. This distribution corresponds to the traditional late CNV.

The Relation Between Performance and ERPs. The preparatory activity manifested by these tonic shifts can best be considered in terms of two interdependent processes. Note that, as time for action approaches, the distribution of the baseline resembles that of the traditional late CNV. As above, we regard this ERP component as a manifestation of motor preparation. The earlier positive change in baseline which is particularly evident in series beginning with "5" i somewhat more difficult to interpret. In fact, there is no other ERP component that has been observed to sustain a long positive level either during a period of preparation, or as a response to a warning or an alerting stimulus. It appeared worthwhile, therefore, to determine the extent to which this, and the other components, are related to the subject's "driving". In particular, we examined the relation between the pattern of deceleration on individual trials and the preceding ERP components. We used the procedure described above to classify trials into two categories, "fast" and "slow", on the basis of deceleration that followed the digit "1". Recall that, in all trials, subjects maintained the maximal driving speed until the penultimate warning digit appeared. At this point they slowed down (decelerated). However, on some trials this was done quite soon after the appearance of the "1", on others the subject monitored for more than 1200 msec. We call the latter "slow" trials, the remainder "fast", and evaluate differences between ERPs elicited prior to "fast" and "slow" trials.

Principal Components Analysis was again applied to the phasic ERPs to identify ERP components. Of the two major components revealed in this way (CNV and SW), only the CNV was related to deceleration latency. Fast movements were associated with larger CNVs for all numbers in the preceding series, independent of series length or condition (for the speed x electrode interaction, $F(2,10)=6.85$, $P<.05$). We interpret this finding as further evidence that the late CNV is closely related to motor preparation.

For tonic shifts, we evaluated changes in baseline activity during the series beginning with both 5 and 3 as a function of deceleration to the number "1". Differences between fast and slow trials emerged well in advance of the actual movement. These differences were most evident for the frontal electrode site (see Figure 5). It is noteworthy that for the fixed condition, frontal positivity was associated with fast movements. In the

---

Insert Figure 5 About Here

---

variable condition, fast movements were preceded by greater frontal negativity. We interpret this difference between conditions in terms of the different preparatory strategies employed by the subjects as a function of temporal predictability. As we suggested above, when the stimuli occur at fixed intervals, they carry important temporal information which can be used by the subject in timing his/her deceleration. Use of this information is accompanied by tonic cortical positivity. The more attention paid to this timing information the sooner the subject decelerates after the "1". Hence, fast movements are associated with greater positivity. However, in the variable condition, reliance on the countdown numbers as purveyors of temporal information can be costly. If the subjects adjust their strategy to depend on the timing data provided by the intermediate numbers, the timing cannot but be unreliable and the decelerations will be slow.

This interpretation of the data implies that the slow baseline positivity we observe across the countdown interval is a manifestation of processing activity invested in the information arriving during the countdown interval. There is, of course, no direct support for this suggestion at this time. Naatanen (1982) has described a "processing negativity" that extends over long intervals as the subject extracts information from stimuli. But, this negativity has, thus far, been observed only over intervals shorter than one second. The reliability of the slow positivity we describe here and the reliability of its relation to performance must be examined in future experiments. However, the process by which we have examined the data and the interaction between the observation of overt performance and psychophysiological measurements illustrates the form which such future research should take.

Discussion

We assert in this chapter that psychophysiological indices enrich the analysis of Cognitive Processes. The data we present above appear to support this thesis. We challenged our subjects with a rather complex task. To maximize their gain they were forced to balance the need to drive as fast as they could with the requirement to avoid collisions. The threat of collisions demanded caution. One could not negotiate the obstacles safely without slowing down substantially as one approached them. Warning signals alerted the subject to the impending obstacles. The subjects evidently heeded the signals and prepared to face the obstacles. That they indeed prepared can be inferred from the fact that they decelerated before the appearance of the obstacles on the screen. Furthermore, the fact that the subjects managed to avoid collisions quite well can be taken as evidence that this preparation was important for the successful completion of the task.

Note, however, that if we restrict our analysis of the subjects' behavior to the measures of overt performance we have in our possession a relatively shallow source of data. When reaction time, magnitude of the deceleration, and number of collisions have been thoroughly examined, we can say little more than that subjects were, in fact, capable of using the warning stimuli to govern their performance. But as the activities triggered by the warning stimuli lack, by and large, overt "behavioral" manifestations, we do not have measures that allow an examination of the preparatory activities instigated by the radar signals. Yet, one would certainly like to know if, and how, the subjects' preparation differed when the warning was delivered well in advance of, or shortly before, the event. One would also like to know if preparation was affected by the regularity of the warning stimuli. Moreover, it is of interest to determine if subjects' adjustments are motor or perceptual, and to decide if subjects attend to the intermediate warning stimuli in the same manner as they attend to the first and the last warning. These are but a few of the questions that need to be answered in an attempt to develop a proper description of the subjects' behavior. We submit that all too often such questions are not addressed when the domain of investigation is constrained by a traditional adherence to overt performance.

We are not claiming that the incorporation of psychophysiological measures will serve as a universal panacea for problems encountered in the study of preparatory processes (or in any other area of Behavioral Science, for that matter). It does, however, seem that the data we present above illustrate the manner in which psychophysiological observations deepen and enrich the analysis of the phenomena. The analysis of the ERPs, especially when performed in terms of the separate components of the ERP, reveals a rich pattern of activities which occur in response to the different events of a trial. This pattern of activities varies with the regularity and length of the warning series and appears to be related to the subjects' overt responses.

The pattern exhibited by the data is displayed, schematically, in Figure 6. The events of the trial are indicated at the top of the figure.

Insert Figure 6 About Here

In the body of the figure we summarize the various combinations of ERPs that were observed for the different series, sorted by the speed with which the subject decelerated the car after the penultimate warning. The first warning event elicits an "orientation reaction" that appears to indicate a switch in general mode of performance that is essentially independent of the information content of the warning stimulus. It does not matter if the warning indicates an immediately impending cluster of obstacles or one that will be delayed by 10-15 seconds. It remains to be determined if the Slow Wave that manifests this switch in modes is related in some sense to the Sokolovian Orienting Reaction.

The intermediate stimuli elicit little if any phasic ERP activity of interest. None of the commonly observed endogenous ERP components appear to be associated with these stimuli. However, an examination of the shifts in the voltage levels that are observed over the interval reveal two very slow processes that are of considerable interest. In the early segment of the waiting period a slow positivity dominates. Even more interesting is the intricate relationship between this positivity and the subjects' overt performance. As we noted above, it is as if the positivity is an index of the degree to which the subject attends to the intermediate stimuli. When the warnings provide a reliable measure of elapsed time, the positivity is largest in association with early decelerations. When the warning stimuli are spaced by random intervals, the positivity is associated with slow decelerations. We believe this pattern of results is worth a much more detailed investigation.

The slow positivity gives way to a negative going process whose scalp distribution and ultimate culmination strongly suggest that it is related to the traditional CNV, and most probably to the motor aspects of the CNV. The subjects clearly begin to adjust their motor control systems for the impending manipulation of the joy stick that will bring about the deceleration. This ERP component emerges with great clarity after the penultimate warning stimuli, but its inception well before the appearance of the digit "1" indicates quite clearly that preparatory activities are established in parallel to the overt output. Even as the system is beginning to readjust its strategy and to activate the controls that allow it to negotiate the obstacles, the motor programs governing the output remain, as long as is prudent, those that yield the fastest driving. The parallelism of the human information processing system is evident from its ability both to execute ongoing programs, and to invoke and ready for execution alternate programs. This is neatly displayed in the contrast between the reckless overt driving prior to the deceleration and the range of preparatory activities hinted at by the ERP components that are mobilized sotto voce, so to speak.

Of course, the ERP data are frustratingly difficult to interpret. There is, at this time, very little data that can support an interpretation of these voltage oscillations in terms of classical neurophysiology. We do not know the intracranial origin of the potentials. We do not even know if any component that emerges from the psychophysiological analysis is a manifestation of a functionally distinct, and neuroanatomically coherent, intracranial entity. Careful analysis of intracranial potentials in humans, recorded in conjunction with clinically justified electrode implantations, reveals a fantastic complexity of patterns (Halgren, Squires, Wilson, Rohrbaugh, Babb, & Crandall, 1980; Wood, Allison, Goff, Williamson, & Spencer, 1980). Yet, in some cases, especially when psychophysiological research has mapped in some detail the parametric span of a component's relation to experimental manipulations, the intracranial results from studies of humans (Halgren, et al., 1980; Wood, et. al., 1980), the studies of evoked magnetic fields (Kaufman, unpublished) and, to some extent studies of intracranial potentials from non-human species (Deadwyler, West & Robinson, 1981) provide data that are tantalizingly consistent with theoretical speculations regarding the ERPs derived from the work in Cognitive Psychophysiology (Donchin, 1981; Fabiani, et. al., 1982). The accelerating pace with which our understanding of the P300 is progressing suggests that a similar strategy needs to be applied to other components. Much recent progress has been made in the study of negative components that occur in the interval between N100 and P300, (Hansen & Hillyard, 1980; Kutas & Hillyard, 1980; Naatanen & Michie, 1979, Pritchard, Coles & Donchin, 1982; Ritter, 1981). It seems that the components associated with preparatory activity may also prove tractable.

Cognitive Psychophysiology, as we have said, is a marriage of Cognitive Psychology and Psychophysiology. As is generally true of marriages, it will be successful only if the partners are adequately sensitive to each other's nature. Neither side ought to dominate. One need not assume that it is necessary to advocate that the methods, or measures, of either of the disciplines be displaced. It is, in fact, the great strength of the Cognitive Psychophysiological paradigm that it attempts to integrate overt and covert behavior, requiring that each set of dependent variables be used conjointly with the others. In the study we describe above, the analysis of deceleration is necessary for the proper interpretation of the ERP data. These data, in turn, open a window on preparation that promises new and interesting vistas. We hope we have tempted students of preparation to consider these vistas in the future.

References

Callaway, E., Tueting, P., & Koslow, S.H. (Eds.) Event-Related Brain Potentials In Man. New York: Academic Press, 1978.

Deadwyler, S.A., West, M.O., & Robinson, J.H. Evoked potentials from the dentate gyrus during auditory stimulus generalization in the rat. Experimental Neurology, 1981, 71, 615-624.

Desmedt, J.E. (Ed.) Attention, voluntary contraction, and event-related cerebral potentials. Progress in Clinical Neurophysiology, Vol. 1. Basel: Karger, 1977.

Donchin, E. Event-related brain potentials: A tool in the study of human information processing. In H. Begleiter (Ed.), Evoked potentials and behavior. New York: Plenum Press, 1979, Pp. 13-75.

Donchin, E. Surprise! . . . Surprise? Psychophysiology, 1981, 18, 493-513.

Donchin, E. (Ed.) Proceedings of the first annual Carmel conference on Cognitive Psychophysiology. Hillsdale, New Jersey: Lawrence Earlbaum Associates, in press.

Donchin, E., & Heffley, E. Multivariate analysis of event-related potential data: A tutorial review. In D. Otto (Ed.), Multidisciplinary Perspectives in Event-Related Brain Potential Research. EPA-600/9-77-043, Washington, D.C.: U.S. Government Printing Office, 1979, Pp. 555-572.

Donchin, E., & Lindsley, D.B. Visual evoked response correlates of perceptual masking and enhancement. Electroencephalography and Clinical Neurophysiology, 1965, 19, 325-335.

Donchin, E., Ritter, W., & McCallum, C. Cognitive psychophysiology: The endogenous components of the ERP. In E. Callaway, P. Tueting, and S.H. Koslow (Eds.), Event-related brain potentials in man. New York: Academic Press, 1978. Pp. 349-441.

Fabiani, M., Karis, D., & Donchin, E. P300 and memory: Individual differences in the von Restorff effect. Psychophysiology, 1982, 19, 558. (Abstract)

Gaillard, A.W.K. The late CNV wave: Preparation versus expectancy. Psychophysiology, 1977, 14, 563-568.

Galambos, R., & Hillyard, S.A. Electrophysiological approaches to human cognitive processing. Neurosciences Research Program Bulletin, 1981, 20, 145-265.

Goldstein, K. The Organism, A Holistic Approach to Biology Derived From Pathological Data in Man. American Book Company: New York, 1939.

Gratton, G., Coles, M.G.H., & Donchin, E. Vector analysis. Paper presented at the ERP workshop, Washington D.C., 1981.

Gratton, G., Coles, M.G.H., & Donchin, E. A new method for off-line removal of ocular artifact. Submitted for publication.

Greenfield, N.S., & Sternbach, R.A. (Eds.) Handbook of Psychophysiology. New York: Rinehart and Winston, 1972.

Halgren, E., Squires, N.K., Wilson, C.L., Rohrbaugh, J.W., Babb, T.L., & Crandall, P.H. Endogeneous potentials generated in the human hippocampal formation and amygdala by infrequent events. Science, 1980, 210, 803-805.

Hansen, J.C. & Hillyard, S.A. Endogeneous brain potentials associated with selective auditory attention. Electroencephalography and Clinical Neurophysiology, 1980, 49, 277-290.

Hillyard, S.A. The CNV and human behavior: A review. In W.C. McCallum and J.R. Knott (Eds.), Event-Related Slow Potentials of the Brain: Their Relations to Behavior. Electroencephalography and Clinical Neurophysiology, 1973, 33, 161-171.

Kok, A. The effect of warning stimulus novelty on the P300 and components of the contingent negative variation. Biological Psychology, 1978, 6, 219-233.

Kornhuber, H.H., & Deecke, L. Hirnpotential anderungen bei wilkurbewegungen und passiven bewegungen des menschen: Bereitschaftspotential und reafferente potentiale. Pflugers Archiv Fur Die Gestamte Physiologie Des Menschen und Det Tierre, 1965, 284, 1-17.

Kutas, M., & Donchin, E. Studies of squeezing: Handedness, responding hand, response force, and asymmetry of readiness potential. Science, 1974, 186, 545-548.

Kutas, M. & Hillyard, S.A. Reading senseless sentences: Brain potentials reflect semantic incongruity. Science, 1980, 207, 203-205.

Loveless, N.E. The effect of warning interval on signal detection and event-related slow potentials of the brain. Perception and Psychophysics, 1975, 6, 565-570.

Loveless, N.E. Event-related slow potentials of the brain as expressions of orienting function. In H.D. Kimmel, E.H. Van Olst, & J.F. Orlebeke (Eds.), The Orienting Reflex in Humans. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1979. Pp. 77-100.

Loveless, N.E., & Sanford, A.J. Slow potential correlates of preparatory set. Biological Psychology, 1974, 1, 303-314.

McCarthy, G., & Donchin, E. Brain potentials associated with structural and functional visual matching. Neuropsychologia, 1978, 16, 571-585.

Morrell, L. & Morrell, F. Evoked potentials and reaction time: A study of intra-individual variability. Electroencephalography and Clinical Neurophysiology, 1966, 20, 567-575.

Mountcastle, V.B. The world around us: Neural command functions for selective attention. Neurosciences Research Program Bulletin, 1975, 14, 1-47.

Naatanen, R. Processing negativity: An evoked potential reflection of selective attention. Psychological Bulletin, 1982, 92, 605-640.

Naatanen, R. & Michie, P.T. Early selective-attention effects on the evoked potential: A critical review and reinterpretation. Biological Psychology, 1979, 8, 81-136.

Perdok, J., & Gaillard, A.W.K. The terminal CNV and stimulus discriminability in motor and sensory tasks. Biological Psychology, 1979, 8, 213-223.

Posner, M. (Ed.) Chronometric Explorations Of Mind. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1978.

Pritchard, W.S., Coles, M.G.H., & Donchin, E. N200 amplitude as a function of degree of mismatch in a word categorization paradigm. Psychophysiology, 1982, 19, 580. (Abstract)

Requin, J. (Ed.) Attention and Performance VII. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1978.

Ritter, W. Cognition and ERPs: The relation of post-stimulus negative potentials and cognitive processes. In Proceedings of the EPIC VI Conference, Lake Forest, IL, 1981.

Rohrbaugh, J.W., Syndulko, K., & Lindsley, D.B. Brain components of the contingent negative variation in humans. Science, 1976, 191, 1055-1057.

Squires, N.K., Squires, K.C., & Hillyard, S.A. Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. Electroencephalography and Clinical Neurophysiology, 1975, 38, 387-401.

Squires, K.C., Donchin, E., Herning, R.I., & McCarthy, G. On the influence of task relevance and stimulus probability on event-related-potential components. Electroencephalography and Clinical Neurophysiology, 1977, 42, 1-14.

Tecce, J.J. Contingent negative variation (CNV) and psychological processes in man. Psychological Bulletin, 1972, 77, 73-108.

Walter, W.G., Cooper, R., Aldridge, V.J., McCallum, W.C., & Winter, A.L. Contingent negative variation: An electrical sign of sensorimotor association and expectancy in the human brain. Nature, 1964, 203, 380-384.

Werts, T.C., & Lang, P.J. The effects of eye fixation and stimulus and response location on the contingent negative variation. Biological Psychology, 1973, 1, 1-19.

Wood, C.C., Allison, T., Goff, W.R., Williamson, P., & Spencer, D.C. On the neural origin of P300 in man. Progress in Brain Research, 1980, 54, 51-56.

Woodworth, R.S. Experimental Psychology. New York: Holt, 1938.

TABLE 1

Analysis of Stick Movement

Number of "fast" and "slow" trials as

a function of condition

	Latency of Movement		
Condition	"Fast" < 1200 msec	"Slow" > 1200 msec	Total
Fixed	659	1040	1699
Variable	974	739	1713
	1633	1779	3412

NOTE: Differences between conditions in total numbers of trials are the result of artifacts or missing data due to equipment failure.

TABLE 2

Analysis of Stick Movement

Number of "fast" and "slow" trials as

a function of series length

Latency of Movement

Initial Series Number	"Fast" < 1200 msec	"Slow" > 1200 msec	
1	520	654	1174
3	526	604	1130
5	587	521	1108
	1633	1779	3412

NOTE: Differences among series in total numbers of trials are the result of artifacts or missing data due to equipment failure.

Figure Legends

Figure 1: Schematic representation of the task.

Figure 2: ERPs and stick movements (associated with deceleration) for each countdown digit, for the variable condition. For each series, ERPs are shown with the baseline intact.

Figure 3: ERPs and stick movements (associated with deceleration) for each countdown digit, for the fixed condition. For each series, ERPs are shown with the baseline intact.

Figure 4: Component scores for CNV (a) and SW (b) factors as a function of electrode, countdown digit and series.

Figure 5: Change in baseline at Fz preceding countdown digits as a function of deceleration latency following the digit "1". Data for fixed and variable conditions are shown separately.

Figure 6: Summary Figure: Idealized phasic and tonic ERP data for the different conditions of the Racecar experiment.

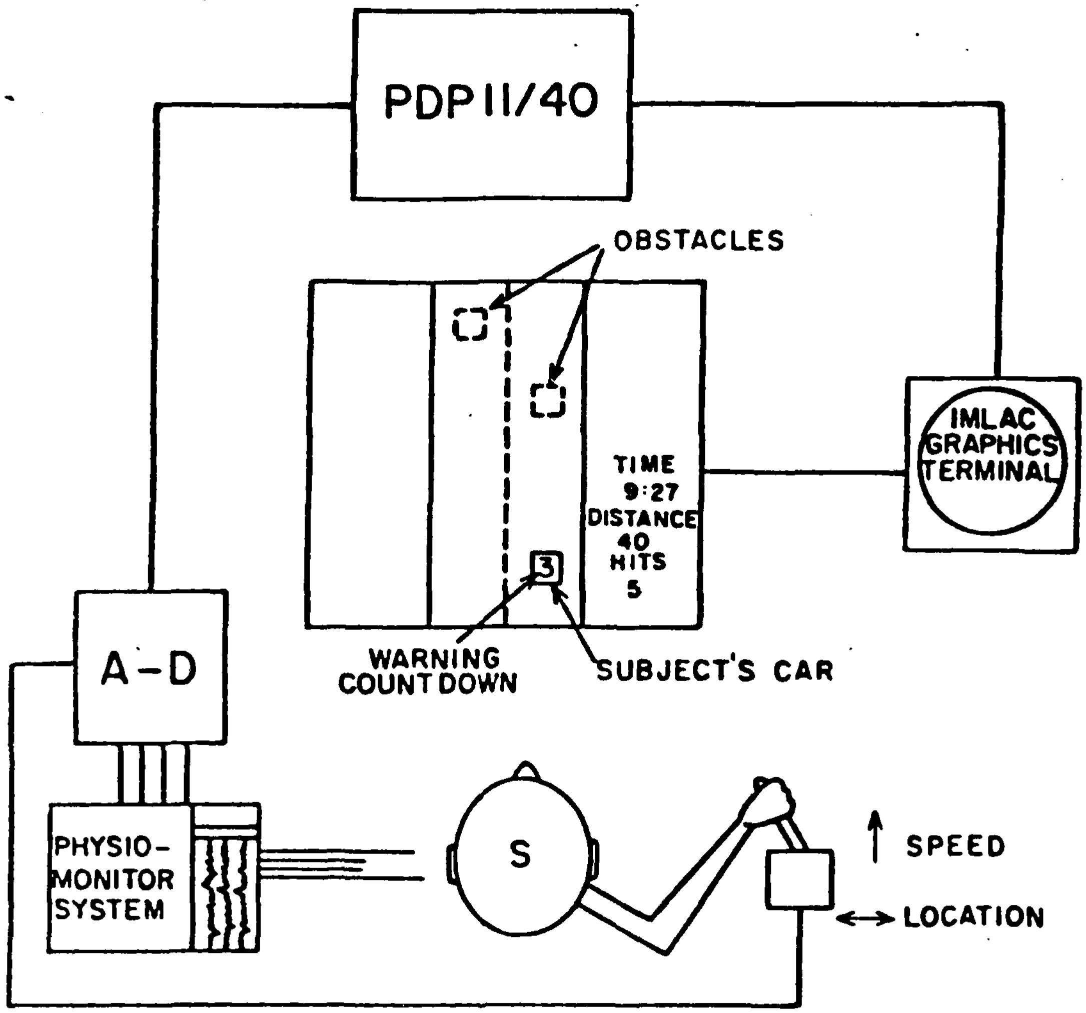
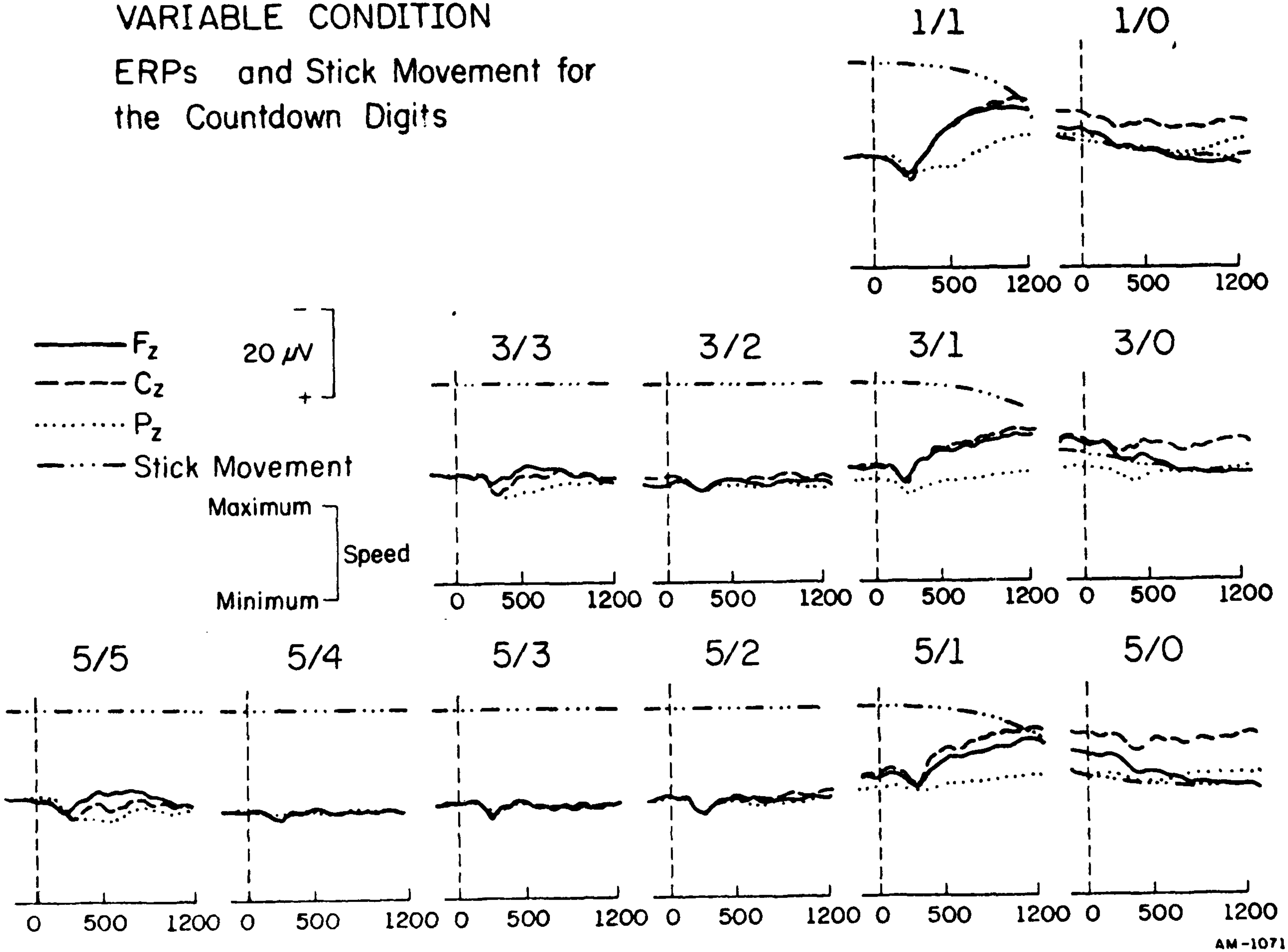

PDP11/40
OBSTACLES
IMLAC
GRAPHICS
TERMINAL
TIME
9:27
DISTANCE
40
HITS
5
3
WARNING
COUNT DOWN
SUBJECT'S CAR
A-D
PHYSIO-
MONITOR
SYSTEM
S
SPEED
LOCATION

Figure 1

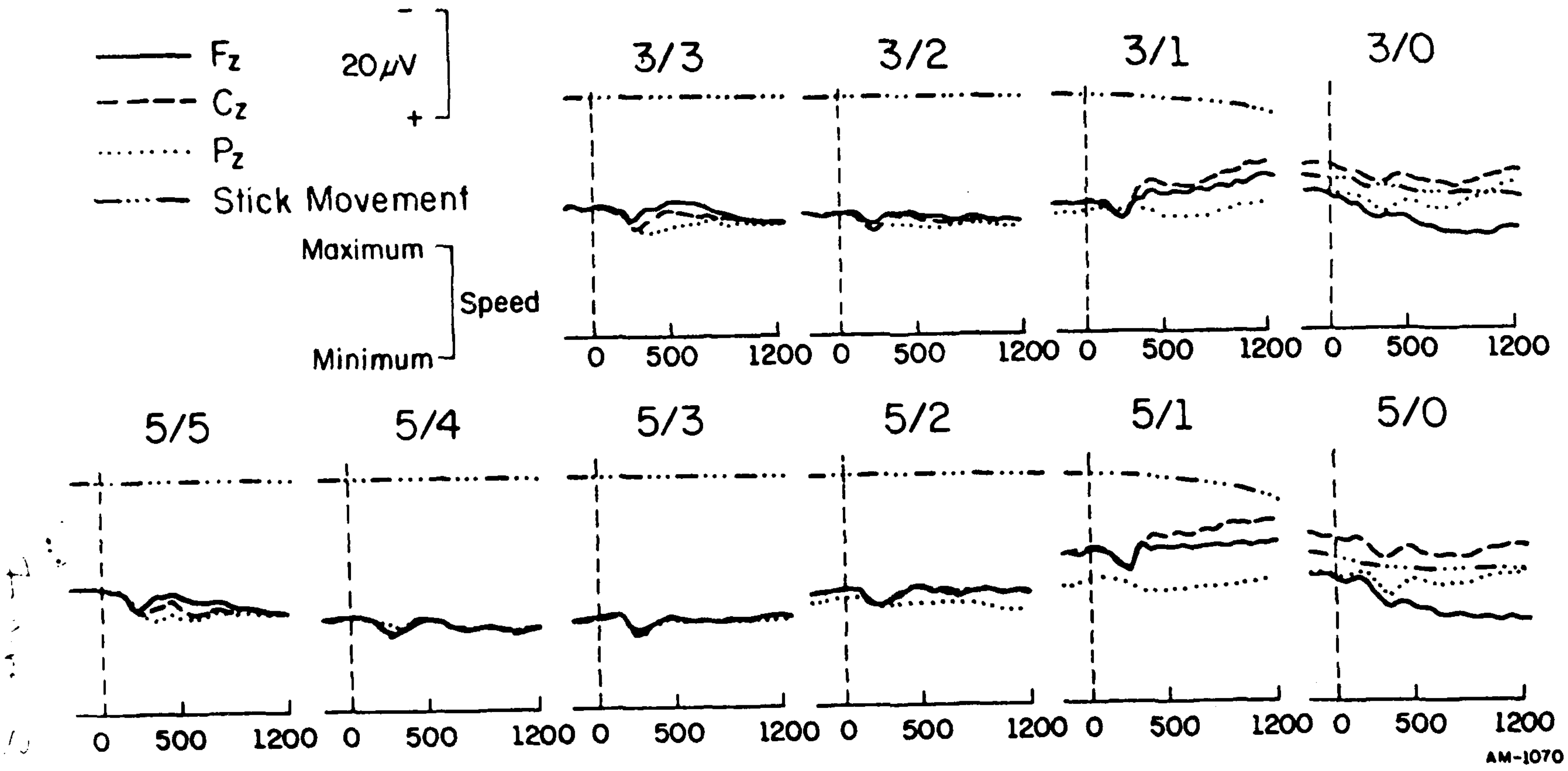
VARIABLE CONDITION

ERPs and Stick Movement for the Countdown Digits



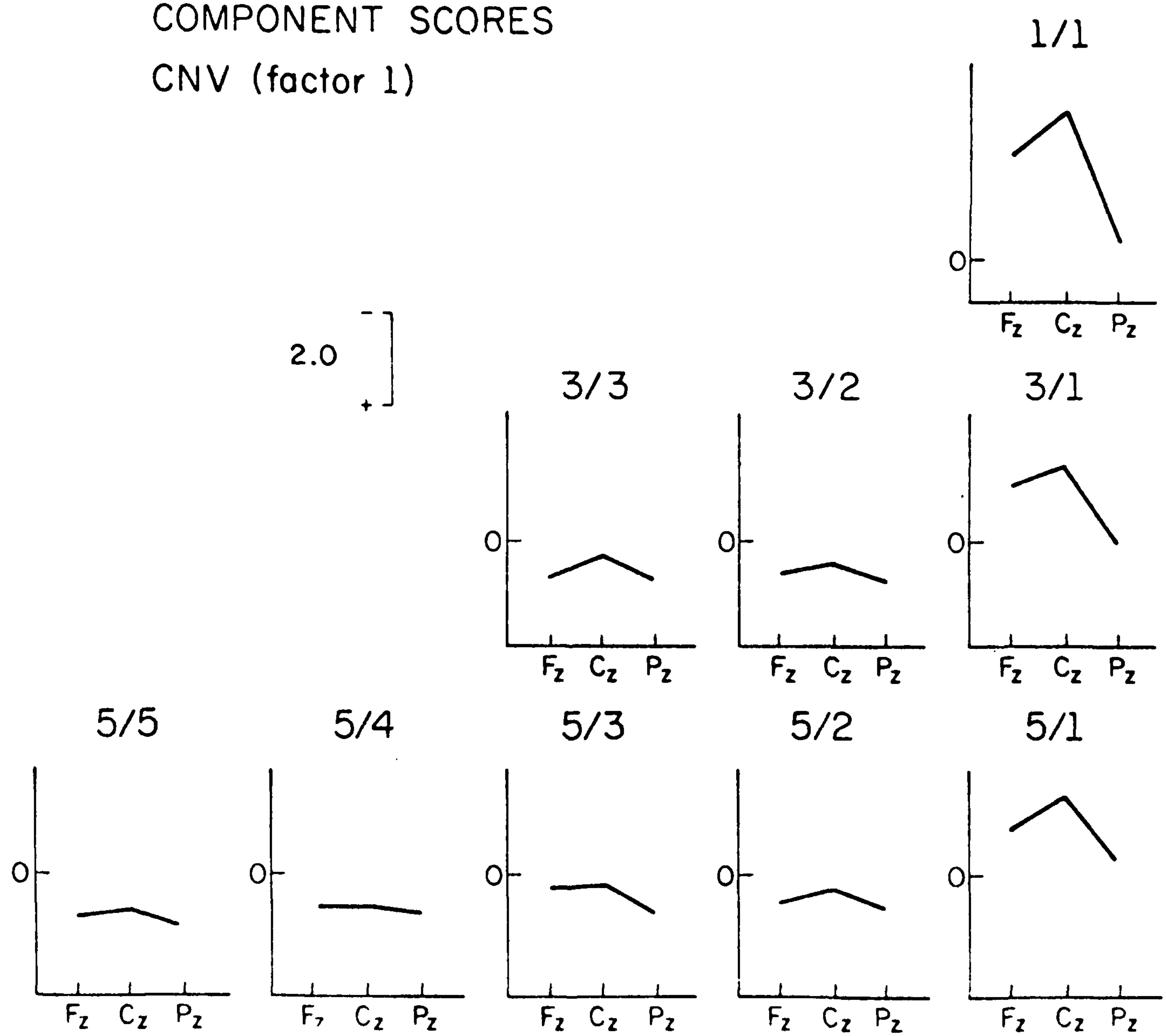
FIXED CONDITION

ERPs and Stick Movement for the Countdown Digits



COMPONENT SCORES

CNV (factor 1)

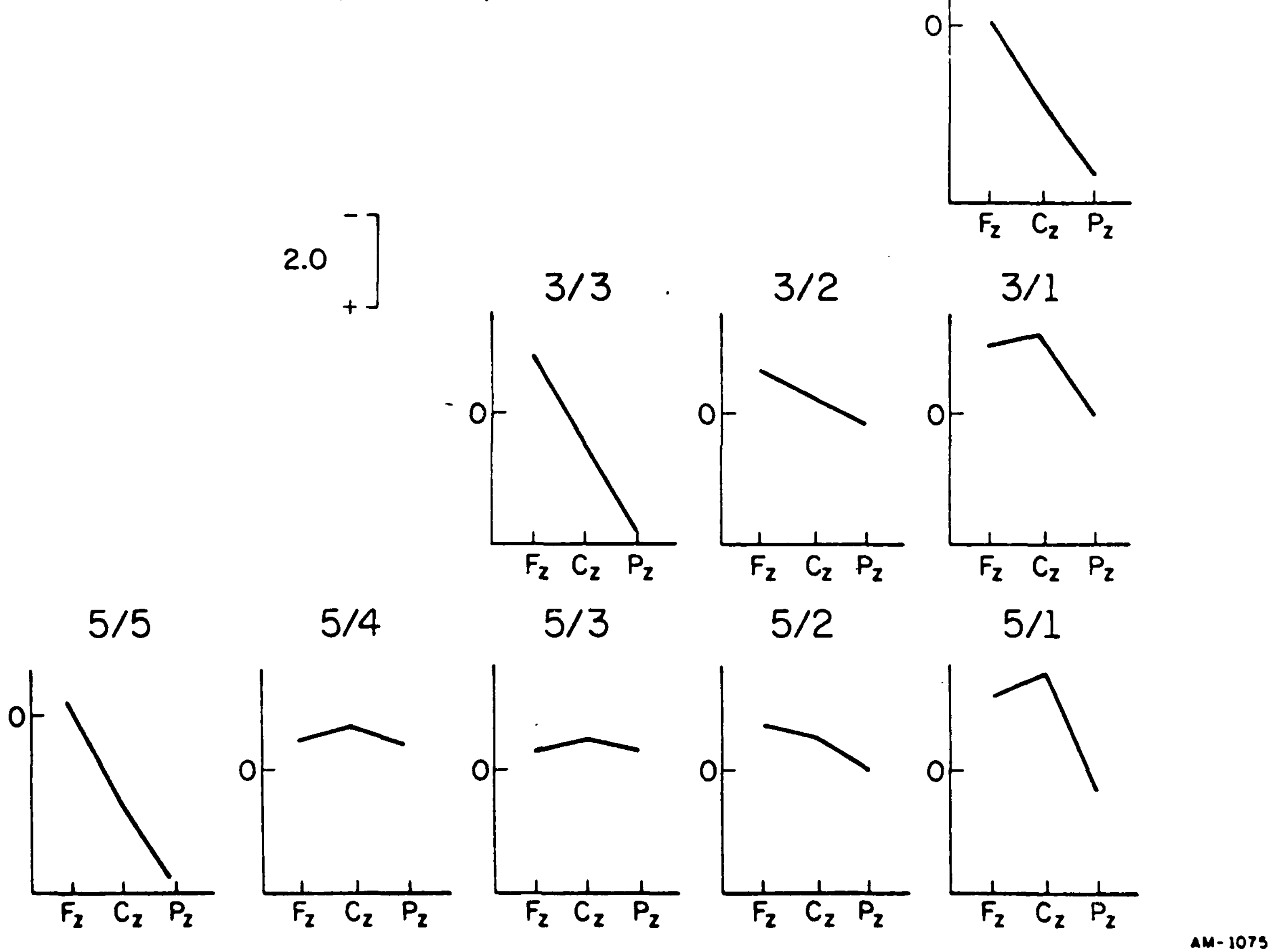


AM-1

Figure

COMPONENT SCORES

Slow Wave (factor 2)



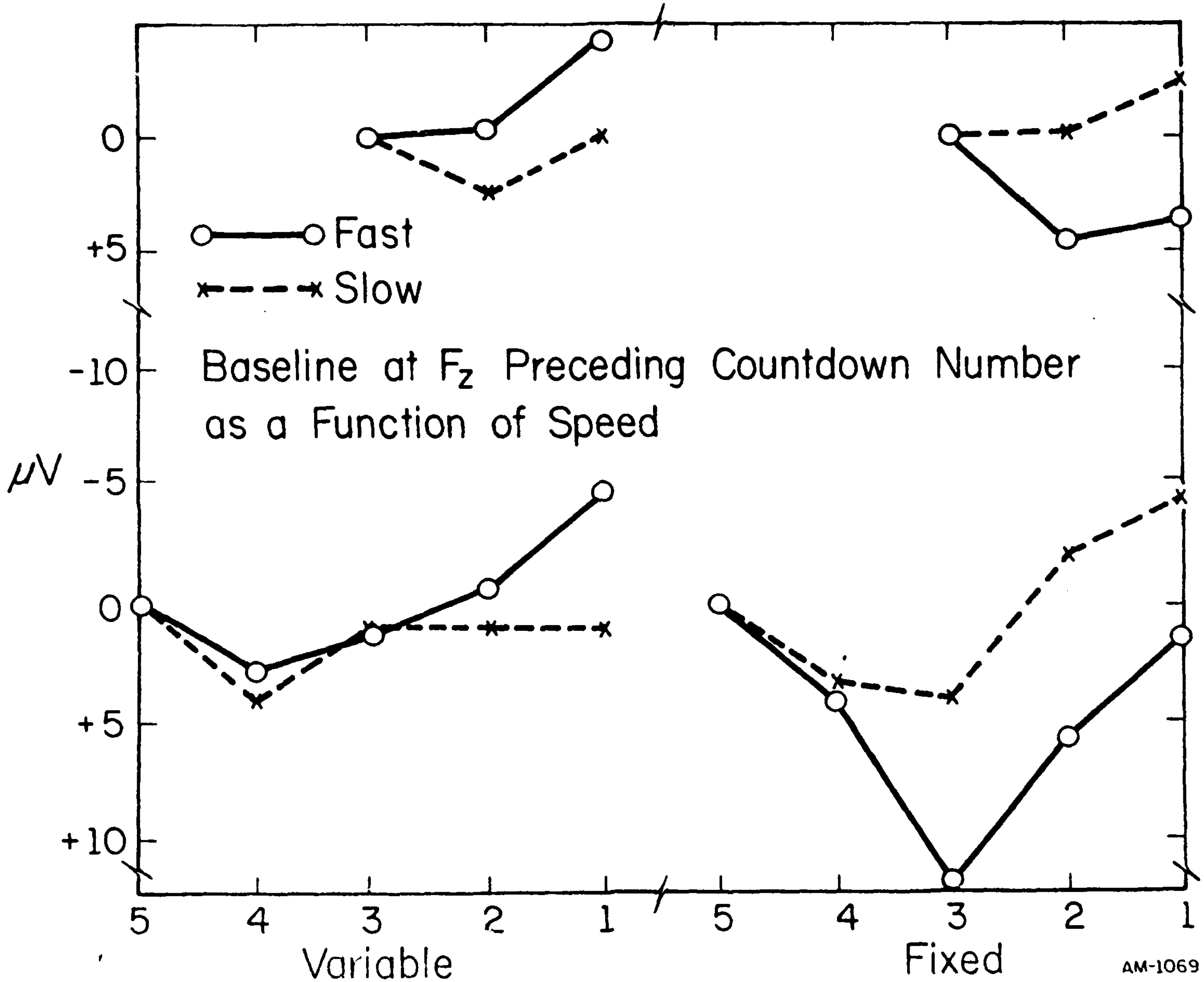

Fast
Slow
Baseline at F_z Preceding Countdown Number
as a Function of Speed
μV
0
+5
-10
-5
0
+5
+10
5
4
3
2
1
5
4
3
2
1
Variable
Fixed
AM-1069

Figure 5

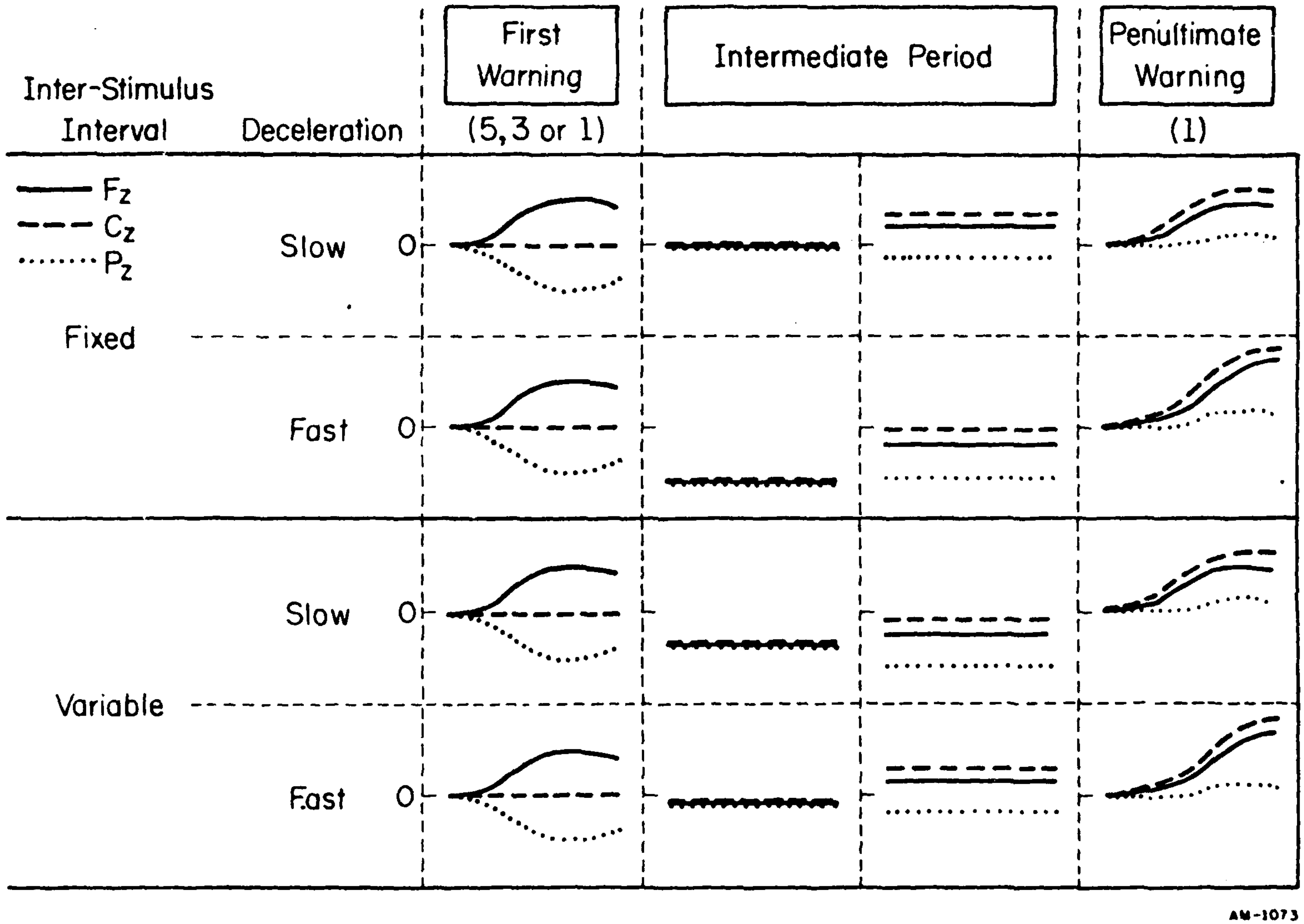

First Warning (5,3 or 1)
Intermediate Period
Penultimate Warning (1)
Inter-Stimulus Interval
Deceleration
Fz
Cz
Pz
Fixed
Variable
Slow
Fast
Slow
Fast
0
0
0
0
AM-1073

Figure 6

Electroencephalography and Clinical Neurophysiology, in press

MA2367A020782 (Revised)

A NEW METHOD FOR OFF-LINE REMOVAL OF OCULAR ARTIFACT[1]

GABRIELE GRATTON, MICHAEL G.H. COLES, AND EMANUEL DONCHIN
COGNITIVE PSYCHOPHYSIOLOGY LABORATORY
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF ILLINOIS
CHAMPAIGN, ILLINOIS 61820

A New Method for Off-line Removal of Ocular Artifact

Gabriele Gratton, Michael G. H. Coles, and Emanuel Donchin

Ocular potentials are clearly a nuisance in investigations of Event-Related Brain Potentials (ERPs). Eye-movements, or blinks, that occur in synchrony with the events eliciting the ERP introduce unwelcome artifacts into the data. This is especially the case because the frequency spectra of endogenous ERP components (such as the CNV and P300) are similar to the spectra of the oculographic potentials (e.g., Hillyard & Galambos, 1970; Girton & Kamiya, 1973). Treating oculographic artifacts has therefore become a required procedure in ERP research (Donchin, Callaway, Cooper, Desmedt, Goff, Hillyard, & Sutton, 1977).

Movements of the eye and of the eye-lid generate changes in electrical fields that interact with fields generated by intracranial generators and thus affect the electrical activity at the scalp. Two different mechanisms have been proposed to account for these ocular potentials. First, rotation of the eye ball modifies scalp recorded electrical activity because movement of the dipole between cornea and retina results in electrical field changes which are propagated through the skull (Mowrer, Ruch, & Miller, 1935; Overton & Shagass, 1969). Second, the upper and lower eye-lids act as "sliding electrodes" as they move across the eye-ball creating an electrical field which can also be propagated through the skull (Barry & Jones, 1965).

The procedure commonly used to eliminate ocular artifacts requires the investigator to detect and discard data recorded on trials in which excessive eye-movements, or blinks, occur. To this end, ocular activity is monitored using the electro-oculogram (EOG). All trials are rejected for which some attribute of the EOG trace, (such as amplitude, variance, area, or slope) exceeds a criterion value.

Unfortunately, this simple procedure is somewhat problematic. First, eye-movements and blinks may have important effects on those perceptual and cognitive processes that are under investigation (Carpenter, 1948; Drew, 1950). Thus, rejection of trials associated with ocular activity may lead to the use of unrepresentative samples of trials in the computation of the average ERP. Second, certain classes of subjects cannot, or will not, cooperate and avoid ocular movements during the trials. In these cases (e.g., children and psychiatric or neurological patients) it is difficult, if not impossible, to obtain a sufficient sample of trials free of ocular movement. Third, some tasks of interest to ERP researchers involve complex tracking or scanning of the visual field. In these tasks, ocular activity is inherent and therefore the proportion of "clean" trials is likely to be small indeed. Further, these "clean" trials are, by definition, unrepresentative.

The trial rejection procedure is often accompanied by instructions to the subjects to "keep their eyes still" during the experiment, or to confine blinks and movements to designated times when experimental stimuli are not being presented. These instructions effectively assign a "secondary task" to the subject, requiring the division of resources between the experimental task and self-monitoring of ocular activity. Note that, if there are individual differences in ocular activity, the secondary task will vary in difficulty from subject to subject.

Alternative procedures seek to correct the ERP data for the effect of ocular activity. These procedures assume that the scalp potential is given by the linear summation of brain and ocular potentials. Therefore, if we subtract the ocular potentials from the EEG record, we should obtain a "clean" brain potential. According to the simplest version of this procedure, the EOG is subtracted from each trial or from the average (either on-line or off-line) to remove that portion of the raw EEG that can be attributed to ocular activity. As the EOG signal is presumed to propagate by volume conduction across the skull, it is reasonable to expect that the contribution of EOG varies across the scalp. Thus, subtraction of the _same_ EOG signal at each electrode site is likely to yield erroneous estimates of the residual ERP. A more reasonable approach would attempt to estimate a propagation factor for each electrode site and scale the EOG signal by that factor prior to subtraction.

There have been several attempts to estimate these propagation factors. Hillyard and Galambos (1970) and Girton and Kamiya (1973) estimated the propagation factor by recording the EOG elicited when subjects moved their eyes by known amounts during a pre-experimental calibration session. The effects of these movements at different electrode locations were used to estimate specific correction factors, and regression techniques (Hillyard & Galambos, 1970) or analog procedures (Girton & Kamiya, 1973) were used to derive corrected values of EEG. This technique, then, depends for its validity both on the linearity assumptions and on the validity of the estimation procedure for the propagation factor.

The assumption of linearity is supported by data obtained by Hillyard and Galambos (1970) and by Overton and Shagass (1969). Slight deviation from linearity is found for vertical movements, due probably to differential involvements of the eye-lids in upward versus downward movements (Overton

& Shagass, 1969). But these deviations are not sufficient to violate the assumption of linearity.

However, the assumptions underlying the estimation of the correction factor may be untenable. Thus, it was assumed that the fields generated by voluntary movements made during a pre-experimental calibration session are identical to the fields generated by task contingent eye-movements. There is evidence that voluntary and involuntary eye-movements are different (Records, 1979). These differences may lead to differences in the nature of the ocular artifact. It would seem prudent to estimate the propagation factor under the <u>same</u> conditions that obtain when it is applied. A second assumption is that fields generated by blinks and eye-movements are the same. Overton and Shagass (1969) have demonstrated convincingly that this is not the case. While both blinks and eye-movements affect the ocular dipole, there is considerable eye-lid activity during blinks which creates characteristic fields. Thus, a correction factor for blinks should be estimated separately from a correction factor due to eye-movements. The procedure recently proposed by Verleger, Gasser, and Mocks (1982) involves the computation of propagation factors from a subset of the data which needs correction. In this sense, it avoids the first problem mentioned above. However, the procedure fails to differentiate between blinks and saccades, and, thus may yield invalid correction factors.

We describe here an eye-movement correction procedure (EMCP) for computing the correction factors that avoids both problems. First, estimates of correction factors are derived from EOG and EEG records obtained during the experiment rather than during a pre-experimental calibration period. Second, correction factors are computed separately for blinks and eye-movements. A third and novel feature of the current procedure is that the degree to which signals at the EOG electrodes are propagated to any given EEG electrode pair is estimated after removal of event-related EOG and EEG activity from the data. If there are event-related effects in both EEG and EOG records the propagation of EOG signal to EEG electrode would be overestimated. Our estimates of correction factors are computed on data at each timepoint on every trial after event-related activity has been subtracted.

After a description of the algorithm we present several tests that confirm the validity and reliability of this off-line procedure for removing ocular artifacts from the EEG.

Procedure

We begin by describing the computational procedures that are used to derive the correction factors. A formal derivation is presented in Appendix

---

Insert Figure 1 About Here

A. Figure 1 gives a graphic representation of these procedures. Assume, for the sake of this description, that a subject is run in a study with P conditions and that N trials, and thus N records, are available for each condition. The record for each trial consists of data, obtained at K electrode pairs, one of which abuts the eye-ball yielding an EOG trace for each of the trials.[2]

(a) Raw Averaging. For each of the P conditions all the N trials are averaged, for each of the K electrode pairs, separately. These ensemble averages estimate the stimulus related variation for the EEG and the EOG channels.

(b) Subtraction of the Raw Average From the Single Trial Data. The raw averages are subtracted from the corresponding single trial records. That is, the average is subtracted from each of the N trials used in its own computation. After subtraction of a "baseline" computed as the mean of all points in an epoch, we have a set of waveforms that are considered an estimate of the activity at an electrode site, on each trial, that is not event-related.

(c) Correction Factor. Step (b) yields estimates of the electrical activity on each trial at the EOG and the EEG electrodes which is not stimulus related. The EEG data are then considered as a dependent variable in a regression computation in which the EOG data serve as the independent variable. The regression equations computed in this fashion are used to derive the correction factors.

(d) Separation of Blinks and Eye-Movements. These correction factors are computed separately for blinks and eye-movement data. Blinks are detected first by locating all timepoints in the EOG record where the rate of change in electrical activity exceeds a criterion. These points are used to obtain a correction factor as described above. The raw (pre-subtraction) EEG data are then "corrected" by subtracting from the EEG, the EOG value scaled by the correction factor. A new correction factor for eye-movements is then computed using the data remaining after the blinks have been removed. Note that although the correction factors are computed using only activity that is not event-related, it is applied to all activity in the epoch.

(e) Averaging. After subtracting the scaled EOG from all single trials the data can be reaveraged as in (a) yielding the corrected ERPs.

Tests of the Procedure

Three series of tests has been conducted to evaluate the validity and the reliability of EMCP.

Series 1

The first series of tests involved data obtained from a study on the relationship between P300, Slow Wave and CNV (Donchin, Coles & Gratton, in press). Subjects were required to perform a simulated driving task in which they controlled a "car" with a joystick as it appeared to move along a road depicted in a CRT display. From time to time, numbers appeared within the subject's car which warned the subject of the upcoming appearance of obstacles which had to be avoided. EEG measures, obtained from Fz, Cz, and Pz, were used to derive ERP responses to the warning numbers. EOG measures[3] were also obtained from the same epoch (1400 msec). For each of the subjects, data obtained on about 100 trials were used. The nature of the task and the display we used led to a sufficient number of trials being associated with eye-movements, providing a reasonable data base for these validity tests. Two types of tests are presented: (A) Deviation from "true" ERP, and (B) Reduction in variance.

For the purpose of these tests we note the following terminology:

a. True ERP. For a given subject the average of the records obtained for trials during which EOG activity did not exceed a strict variance criterion.

b. Dependent and Independent Samples. Two samples of trials were used to provide the ERPs to be compared with the "true" ERP.

(i) A sample consisting of the trials excluded from the sample used to compute the "true" ERP. Thus, this sample contained all trials for which EOG activity exceeded our strict criterion. Because the sample contains a different set of trials than that used to compute the "true" ERP, we use the term "independent" sample.

(ii) A sample consisting of all the trials. Because this sample included those trials used to compute the "true" ERP, we refer to it as the "dependent" sample.

c. Raw ERP. For a given subject, the raw ERP is defined as the average of all trials in either the independent or dependent samples. The records are not corrected and are used in averaging regardless of the occurrence of EOG artifacts.

d. Corrected ERP. For a given subject, the corrected ERP is defined as the average of all trials in either the independent or dependent samples after the trials have been corrected by EMCP.

e. Randomly Corrected ERP. This ERP is computed in the same manner as the corrected ERP except that the correction factor is obtained by substituting a random number between -1 and +1 (from a rectangular distribution) for the correlation coefficient used in the correction procedure. In this way, ten "random regression coefficients" were generated by multiplying these randomly selected values by the ratio of the standard deviations of EEG and EOG. This random-correction procedure was used to estimate a sampling distribution of deviations within which we could evaluate deviations yielded by the EMCP.

A. Deviation from "True" ERP. The first tests of validity of EMCP involve comparisons of corrected ERPs with an estimate of the "true" ERP. While we had no direct measure of the "true" ERP, we assumed that the ERP obtained from trials selected by the traditional rejection procedure can serve as a reasonable substitute. This appears plausible because this type of ERP is accepted as a "clean" ERP by most investigators. If we can show that the ERP corrected by EMCP is more similar to this "true" ERP than uncorrected or randomly corrected ERPs, then we can consider our procedure to be valid.

The degree to which an ERP corresponds to (or differs from) the "true" ERP is expressed by means of a deviation index, computed as follows:

$$\text{Deviation} = \sqrt{\frac{\sum_{t=1}^{n} (X_t - ERP'_t)^2}{n}}$$

where

Xt = value of the evaluated ERP at time t

ERP't = value of "true" ERP at point t

n = number of points in an epoch

In the first test of this series, the deviation index from the true ERP was computed for the corrected ERP, the raw ERP, and each of ten randomly corrected ERPs for each of five subjects, for each electrode.

Figures 2 and 3 show the values of these deviation indices. Figure 2 shows results for the independent trial sample; Figure 3 gives data for the dependent trial sample. The deviation for the raw ERP is marked by an asterisk, the deviation for the corrected ERP with an arrow and that for the ten randomly corrected ERPs with bars. As a combined measure of total deviation across the three electrodes, we computed the square root of the mean of the squared individual deviations. Values for total deviation are also shown in Figures 2 and 3.

It can be seen that deviations for randomly corrected ERPs are

----

Insert Figures 2 & 3 About Here

----

generally larger than those for ERPs obtained using the derived correction factor. For both independent and dependent trial samples, the deviation for the corrected ERP is always smaller than the deviation for the raw ERP. In a few cases randomly corrected ERPs yield smaller deviations, but these can be explained as follows. First, for subject 4, the EOG trace associated with the "true" ERP showed a small but consistent EOG response. Thus, the estimate of the "true" ERP for this subject may not be a good estimate. Secondly, subject 2 showed little evidence of eye-movements. Third, subject 1 had very few trials for which EOG variance was less than the strict criterion. Thus, the estimate of the "true" ERP for this subject may have been contaminated with noise.

In Figure 4 we display ERP waveforms for two subjects based on independent trial samples. Of particular interest are those for the "true" ERP, and those for corrected and uncorrected ERPs based on trials with EOG variance greater than the criterion. Note that, for both subjects, the

----

Insert Figures 4 & 5 About Here

----

similarity between corrected and "true" ERPs is greater than that for raw and "true" ERPs. Figure 5 shows similar waveforms for Subject 1 for the total trial sample. Note that, again, the effect of correction is to make the corrected ERP more similar to the "true" ERP.

B. Reduction in Variance. We assumed above that the true ERP may be estimated by computing an average ERP based on trials selected for low EOG variance. This assumption, although made by almost every ERP researcher, may not be valid. Either because of residual consistent EOG activity, or because of inadequate sample size, the estimate of the true ERP may be inaccurate. For this reason, we conducted two additional validity tests that were based on the prediction that our correction procedure should reduce the variance between ERPs, or among trials at each timepoint. This is because some of the variance between trials is due to eye-movements. The first of the additional tests examines the prediction that if EMCP increases the similarity between the corrected and the "true" ERP, then corrected ERPs, derived from the same subject, condition and electrode, should be more alike than corresponding uncorrected ERPs.

The samples used in this test were the same as the independent samples used above. That is, for each subject, two sets of trials were derived, those for which EOG variance was greater, or less, than the strict criterion. Comparison between average ERPs based on these two samples, both before and after EMCP, was accomplished using the deviation index. Table 1 gives the relevant deviation indices. Note that, in every case, corrected ERPs are more similar to each other than are the raw ERPs. This is also

---

Insert Table 1 About Here

---

illustrated in Figure 4.

The second additional test examines the prediction that, if the values at each timepoint in the trial correspond more to those associated with the "true" ERP as a result of correction, then the variance across trials for each timepoint should be reduced by correction. To test this prediction we computed, for each subject, electrode and timepoint, the variance across trials before and after correction. Then, differences in variance between uncorrected and corrected single trials were derived for each timepoint.[4] The frequency distributions of these differences are shown in Figure 6.

Insert Figure 6 About Here

For two subjects (out of five) all the timepoints for all the electrodes show a reduction in the variance between trials following the application of EMCP. For subjects 4 and 5, about 1% of timepoints are associated with an increase in variance following correction. For subject 1 there is an increase following correction on 13% of timepoints. As a whole, about 99% of the timepoints show decrease in the variance between trials following correction.

Series 2

In the experiment which provided the data for the tests in Series 1, the EOG electrodes were placed above and to the side of the eye. This oblique derivation is affected by horizontal, as well as by vertical eye movements. However, we proposed (Footnote 3) that because eye-movements in this task were almost exclusively in the vertical direction, the EOG would reflect only vertical movements. Nevertheless, because of the possible contamination of the EOG record by horizontal movements in the first series of tests, we conducted a second series using data from an experiment in which a vertical placement was used for EOG. These tests involved data from a simple visual oddball task (Fabiani, Karis, and Donchin, 1982). The word "Count" was presented 100 times to each of four undergraduate female subjects, in each of two sessions, one to two weeks apart. On twenty occasions the word was larger than it was on the other 80 occasions. The subjects were instructed to count the number of "rare" words (those with larger size). EEG was recorded at Fz, Cz, and Pz, referred to linked mastoids, and vertical EOG was recorded from electrodes placed above and below the right eye.

Data for each session were analyzed independently and averages were computed for "rare" and "frequent" stimuli and for each electrode separately. Three kinds of averages were computed:

A. True ERP: Average of the records for trials during which EOG activity did not exceed a strict variance criterion.

B. Raw ERP: Average of the records for all the trials, regardless of the occurrence of ocular artifacts.

C. Corrected ERP: Average of the records for all the trials, regardless of the occurrence of ocular artifacts, but corrected according to EMCP.

The correction factors yielded by EMCP are presented in Table 2. Separate correction factors are presented for each subject, session, scalp electrode and for blinks and saccades. Means and standard deviations for each scalp electrode and for blinks and saccades are also shown.

As can be seen in Table 2, the correction factors for saccades are consistently larger than those for blinks, in accordance with the propagation factors reported in the literature (Corby & Kopell, 1972). The correction factors both for blinks and saccades appear to be stable, although small differences between subjects can be observed. The correlations between the correction factors for the first and second sessions were .97 for blinks and .91 for saccades ($n = 12$, $p < .01$). Note that the correction factor may change between experimental sessions if electrodes are not placed in exactly the same position because propagation characteristics will change with electrode position. This is particularly true if there are variations in the placement of the EOG electrodes. Furthermore, variation in impedances betweeen sessions might also lead to instability in the estimate of correction factors. We also would not necessarily expect to find identical values for different subjects because of morphological differences.

Insert Table 2 About Here

Waveforms for the first subject are presented in Figure 7. In this figure the averages from each session both for rare and frequent stimuli are presented. "True", "Raw", and "Corrected" ERPs for Fz, Cz, and Pz are

----

Insert Figure 7 About Here

----

presented, together with the vertical EOG trace. This figure shows that a large ocular artifact is present at a latency of about 800 msec for the rare stimuli. It is particularly evident at the frontal electrode. Note that the artifact has been entirely corrected by EMCP.

To quantify the "goodness" of the correction, deviation indices from the "true ERPs" (defined as above) were computed for both the "raw" and the "corrected" ERPs, for each subject, session and electrode. The deviation indices are presented in Table 3.

Insert Table 3 About Here

As can be seen in this Table, EMCP generally reduces the deviations from the "true" ERPs. In a few cases the difference between the deviations for the "corrected" and "raw" ERP is slightly negative. In this case the subjects made very few eye-movements, and almost no trials were actually rejected in the computation of the "true" ERP. To illustrate this case, waveforms from the first session of the fourth subject are presented in Figure 8.

Insert Figure 8 About Here

Series 3

In the third series of tests, we assessed the constancy of the correction factor over different blocks of trials obtained from the same subject in the same experimental session as well as between sessions. The data used for this test series came from another experiment (Coles, Bashore, Gratton, & Eriksen, in preparation). Four subjects performed a total of 2160 trials of a choice reaction time task in blocks of 108 trials over three or four experimental sessions. Each trial was initiated by depression of a foot pedal by the subject. One sec later a visual display containing target letters S or H appeared on a projection tachistoscope and the subject made a left or right-hand squeeze response. EOG and EEG, recorded from Fz, Cz, and Pz, were obtained as in the previously described experiment. The epoch for analysis began 1280 msec before the foot press and extended through 1280 msec after the stimulus, making a total duration of 3560 msec.

For the purposes of analysis, each block of 108 trials was treated

----

Insert Figures 9 & 10 About Here

----

separately and individual correction factors for each electrode were computed. Figure 9 shows the mean and inter-quartile range for correction factors for each subject and electrode separately over all trial blocks. Note that there is no overlap between electrodes, with correction factors for Fz > Cz > Pz. Note also that there is stability in correction factors between subjects within electrodes. The within-subject variability seen in Figure 9 is mostly due to between-session variability. This is illustrated in Figure 10 where the data for Subject 3 are shown for each session (day) separately. Note that there is some variability across days, but that the variability within a session is smaller than that between sessions. As noted above, the variability across days may be due to slight variations in electrode placement and/or impedance between sessions. This variability points to the importance of computing separate correction factors at least for each experimental session.

Discussion

The results of the first and second series of tests reveal that trials corrected by EMCP yield an ERP which corresponds more closely to the "true" ERP than ERPs derived from "randomly" corrected or uncorrected trials. This is true for trials selected on the basis of large eye-movement or blink contamination and for unselected trials. For both trial samples, EMCP reduces deviation from the "true" ERP, but some deviation remains. This residual deviation is not surprising. As we have noted, our estimate of the "true" ERP may not correspond to the "true" ERP. First, this estimate may be computed on too small a number of trials to eliminate random noise from the average. Secondly, any traditional rejection criterion, such as that used here, can permit trials associated with small, and perhaps consistent, eye-movement artifact to be included in the average. A rejection criterion of zero variance in the EOG channel is clearly unrealistic. Thus, some contamination may be present even in an average based on trials selected according to a strict EOG variance criterion.

A second reason for the persistence of some deviation from the "true" ERP even after correction is, of course, that the correction factors are themselves estimates, although based, in our procedure, on a large sample. Noise in the EOG channel could significantly reduce the magnitude of the correction factor.

Other tests of validity also support our claim that EMCP is a valid procedure. The difference between ERPs derived from trials with different degrees of artifact decreases following correction. Furthermore, variance between individual timepoints across trials decreases following application of EMCP. In both cases, the procedure reduces, but does not eliminate, variance between ERPs or among single trials. Again, this is not surprising. Eye-movement or eye-blink artifacts are not the only source of variance contributing to differences between ERPs or among single trials.

The validity of EMCP is also supported by the data obtained from the reliability tests (second and third series). Because of the propagation characteristics of the head, we would expect that ERPs obtained from frontal electrode sites would be more contaminated than those obtained from more posterior deviations (cf. Overton & Shagass, 1969). Figures 9 and 10 and Table 2 indicate clearly that the magnitude of the correction factor decreases from frontal to parietal electrodes. The second series of tests also show that the correction factors are consistently larger for saccades than for blinks (Corby & Kopell, 1972).

The data from the third series reveal that values of the correction factor are stable within subjects and sessions, although there is some variability between sessions.

Thus, we believe that EMCP is both a valid and reliable procedure for dealing with the problem of eye-movement and eye-blink artifact. Furthermore, the procedure has clear advantages. First, it distinguishes between blink and eye-movement artifact. Second, by deleting stimulus-related EOG and EEG activity before computing the correction factors it provides corrections that are insensitive to stimulus-locked activity at the EOG electrodes. Third, all trials are retained for use in subsequent analyses. Fourth, it is a general procedure which does not require special data to be collected. In this respect, it also reduces the problem of a possible difference between voluntary and involuntary movements and blinks. Fifth, the subjects need not be instructed to "control" or minimize their eye-movements. As we have noted, such instructions may impose an unwanted secondary task. Sixth, the estimate of artifact is based on a large sample (trials x timepoints) rather than a few points obtained from a few prescribed eye-movements (cf., Hillyard & Galambos, 1970; Girton & Kamiya, 1973).

The EMCP has been implemented in a Fortran program on the Harris/7 computer at the Cognitive Psychophysiology Laboratory. For the data used in the validity tests described above (Donchin, Coles & Gratton, in press; Fabiani, Karis, & Donchin, 1982), computation of correction factors for each subject and subsequent correction takes approximately five minutes of CPU time. These data consisted of 100 trials, 140 points/trial, for three EEG channels, 1 EOG channel, and 1 event channel. For comparison, we note that simple averaging of the same data set takes 1.5 minutes of CPU time, while latency adjustment using the Woody technique (3 iterations) and step-wise discriminant analysis take 4 minutes and 2.5 minutes, respectively. Thus, the EMCP technique does consume considerable computer resources. However, the gains in precision and data preservation, in our opinion, more than justify the costs.

The procedure could be made both more simple or more complicated. If blinks are not a particular problem, then trials on which blinks occur could be eliminated and EMCP performed only to correct for movements. This would result in a 50%-75% reduction in computer time. On the other hand, the procedure could be more complex if up and down movements were considered individually. This would also require monopolar recordings of vertical EOG.

The advantages of the procedure listed above suggest that EMCP could be fruitfully used in any situation in which ERPs are recorded. However, there are certain situations where it seems particularly appropriate. These situations include those where eye-movements are part of the task or are characteristic of the population being studied. When the procedure is used, our tests suggest that individual correction factors should be computed for each subject, electrode, and experimental session. Where ERPs from midline sites are of interest, only vertical artifacts need to be corrected. However, if lateral derivations are used, then both vertical and horizontal EOG should be recorded and used to compute separate correction factors for blinks and vertical and horizontal eye-movements.

Finally, we should note that the procedure has been used by us in the analysis of the data from Donchin, Coles and Gratton (in press, see above). The benefits gained were very apparent. First, all trials could be used in analysis - even though, for one condition, only 5% of trials were free of eye-movements and blinks. Second, condition differences which were evident, but not reliable, in analyses of data selected by traditional rejection procedures, became reliable when the analyses were based on corrected trials.

Summary

A new offline procedure for dealing with ocular artifacts in ERP recording is described. The procedure (EMCP) uses EOG and EEG records for individual trials in an experimental session to estimate a propagation factor which describes the relationship between the EOG and EEG traces. The propagation factor is computed after stimulus-linked variability in both traces has been removed. Different propagation factors are computed for blinks and eye-movements.

Tests are presented which demonstrate the validity and reliability of the procedure. ERPs derived from trials corrected by EMCP are more similar to a "true" ERP than are ERPs derived from either uncorrected or randomly corrected trials. The procedure also reduces the difference between ERPs which are based on trials with different degrees of EOG variance. Furthermore, variability at each time-point, across trials, is reduced following correction.

The propagation factor decreases from frontal to parietal electrodes, and is larger for saccades than blinks. It is more consistent within experimental sessions than between sessions.

The major advantage of the procedure is that it permits retention of all trials in an ERP experiment, irrespective of ocular artifact. Thus, studies of populations characterized by a high degree of artifact, and those requiring eye-movements as part of the experimental task, are made possible. Furthermore, there is no need to require subjects to restrict eye-movement activity. In comparison to procedures suggested by others, EMCP also has the advantage that separate correction factors are computed for blinks and movements and that these factors are based on data from the experimental session itself rather than from a separate calibration session.

ACKNOWLEDGEMENTS

Earle Heffley, Amir Mane, and Noel Marshall made helpful comments on earlier versions of the manuscript. In developing the procedure for distinguishing between eye-movements and blinks, we relied on unpublished work by Ron Clapman in our laboratory.

REFERENCES

Barry, W., and Jones, G.M. Influence of eyelid movement upon electro-oculographic recording of vertical eye movements. Aerospace Med., 1965, 36:855-858.

Carpenter, A. The rate of blinking during prolonged visual search. J. Exp. Psychol., 1968, 38:587-591.

Corby, J.C., and Kopell, B.S. Differential contributions of blinks and vertical eye movements as artifacts in EEG recording. Psychophysiology, 1972, 9:640-644.

Donchin, E., Callaway, E., Cooper, R., Desmedt, J. E., Goff, W. R., Hillyard, S. A., and Sutton, S. Publication criteria for studies of evoked potentials (EP) in man. Report of the Methodology Committee. In: J.E. Desmedt (Ed.), Attention, voluntary contraction and event-related cerebral potentials. Prog. clin. Neurophysiology, Karger, Basel, 1977:1-11.

Donchin, E., Coles, M.G.H., and Gratton, G. Cognitive psychophysiology and preparatory processes: A case study. In S. Kornblum and J. Requin (Eds.), Preparatory states and processes. Erlbaum, N.J., in press.

Drew, G.C. Variations in reflex blink-rate during visual motor tasks. Q. J. Exp. Psychol., 1950, 4:78-88.

Fabiani, M., Karis, D., & Donchin, E. P300 and memory: Individual differences in the von Restorff effect. Psychophysiology, 1982, 19:558. (Abstract)

Girton, D.G. and Kamiya, J. Simple on-line technique for removing eye movement artifacts from the EEG. Electroenceph. clin. Neurophysiol., 1973, 34:212-216.

Hillyard, S.A. and Galambos, R. Eye movement artifact in the CNV. Electroenceph. clin. Neurophysiol., 1970, 28:173-182.

Mowrer, O.H., Ruch, T.C. and Miller, N.E. The corneo-retinal potential differences as the basis of the galvanic method of recording eye movements. Amer. J. Physiol., 1935, 114:423.

Overton, D.A. and Shagass, C. Distribution of eye movement and eyeblink potentials over the scalp. Electroenceph. clin. Neurophysiol., 1969, 27:544-549.

Records, R.E. Eyebrows and eyelids. In: R.E. Records (Ed.), Physiology of the Human Eye and Visual System. Harper and Row, Hagerstown, 1979:1-24.

Verleger, R., Gasser, T., and Mocks, J. Correction of EOG artifacts in ERP of the EEG: Aspects of reliability and validity. Psychophysiology, 1982, 19:472-480.

FOOTNOTES

1. (title page) This research reported here has been supported under the Office of Naval Research (Contract #N00014-76-C-0002) with funds provided by the Defense Advanced Research Projects Agency; AFOSR under Contract Number F49620-79C-0233; Wright Patterson AFB under Contract Number F33615-79C-0512; and the Environmental Protection Agency under Contract Number R805628010.

2. (pg. 5) If primary interest is in ERPs from mid-line derivations only vertical EOG need be obtained. If interest is also in ERPs from lateral electrode sites, then both horizontal and vertical EOG should be obtained and the correction procedure applied separately for each EOG channel. Picton (personal communication) indicated that spuriously high correction factors, for either horizontal or vertical movements, could be obtained if there are consistent oblique movements. A solution to this problem is to detect timepoints in which such movements occur using vertical and horizontal EOG, and then either compute a separate correction factor for these movements, or discard these timepoints from the data bases used to compute vertical or horizontal correction factors. It should be noted that correction for blinks is accomplished using vertical EOG only. EEG and EOG should be recorded using the same filtering parameters.

3. (pg. 7) An oblique electrode derivation was used to derive the EOG data in this experiment. Thus, variations in the EOG trace could be due to either vertical or horizontal movements (or both). In fact, task requirements were such that vertical (and not horizontal) movements

occurred. Therefore, the EOG trace reflected predominantly vertical eye-movement activity.

4. (pg. 12) Note that, for each trial, the baseline, defined as the mean value of the timepoints across the epoch, is subtracted from each timepoint before the variance across trials is computed.

TABLE 1

Deviations Between ERPs Based On Large And Small*

EOG Variance Trials Before and After Correction

Subject #	Deviation Before Correction				% of Trials With Small EOG Variance
	Fz	Cz	Pz	Tot	
1	81.91	48.02	37.81	59.01	25
2	51.75	41.26	28.79	41.67	77
3	20.00	14.24	12.53	15.91	52
4	43.31	33.69	24.00	34.58	56
5	37.07	25.13	18.69	28.02	50

Subject #	Deviation After Correction			
	Fz	Cz	Pz	Tot
1	28.84	19.89	16.84	22.44
2	41.23	36.56	25.95	35.17
3	14.21	11.86	10.68	12.34
4	31.62	29.15	20.55	27.52
5	21.24	23.49	16.79	20.70

*Note: ERPs based on small EOG variance trials before correction correspond to the "true" ERPs.

TABLE 2

Correction Factors from

the Visual Oddball

Experiment

Subject #	Session #	Blinks Fz	Blinks Cz	Blinks Pz	Saccades Fz	Saccades Cz	Saccades Pz
1	1	.20	.07	.04	.29	.11	.07
	2	.19	.07	.04	.23	.10	.05
2	1	.13	.08	.05	.18	.10	.06
	2	.11	.06	.07	.10	.05	.03
3	1	.21	.09	.04	.23	.09	.03
	2	.19	.07	.02	.23	.10	.04
4	1	.16	.05	.03	.20	.07	.03
	2	.18	.04	.00	.25	.07	.02
Mean		.17	.07	.04	.21	.09	.04
Sd		.03	.01	.02	.05	.02	.02

TABLE 3

Deviations from the "true"

ERP before and after correction

Subject #	Session #	Deviation before Correction Fz	Cz	Pz	Total
1	1	69.29	38.02	33.48	49.56
	2	61.94	38.98	27.23	45.08
2	1	28.74	20.44	17.67	22.78
	2	6.04	5.61	4.51	5.43
3	1	30.19	23.42	15.12	23.73
	2	59.87	60.52	39.86	54.27
4	1	7.81	8.20	8.78	8.28
	2	24.66	19.90	17.44	20.88

Subject #	Session #	Deviation after Correction Fz	Cz	Pz	Total
1	1	31.72	33.75	24.93	30.37
	2	31.20	34.84	27.76	31.40
2	1	19.82	17.91	15.47	17.83
	2	5.63	5.20	4.38	5.10
3	1	21.16	24.50	12.85	20.11
	2	35.38	28.21	27.38	30.53
4	1	7.07	9.04	9.04	8.43
	2	15.27	19.35	17.27	17.38

FIGURE LEGENDS

Figure 1. Schematic representation of the correction procedure.

Figure 2. Deviations (in log arbitrary units) from "true" ERP of ERPs derived from trials with EOG variance larger than a criterion (independent sample). Deviations relative to waveforms corrected by EMCP (Arrows), uncorrected (Stars), and corrected according to a random procedure (Bars) are shown separately for each subject and electrode. The total deviation computed over all electrodes is also shown.

Figure 3. Deviations (in log arbitrary units) from "true" ERP of ERPs derived from all trials (dependent sample). Deviations relative to waveforms corrected by EMCP (Arrows), uncorrected (Stars), and corrected according to a random procedure (Bars) are shown separately for each subject and electrode. The total deviation computed over all electrodes is also shown.

Figure 4. ERPs and associated EOG for Subjects 1 and 2 for trials with EOG variance larger (a) and smaller (b) than the criterion. Note that (a) corresponds to the "independent" sample and (b) to the "true" ERP sample. The Figure shows the waveforms obtained from trials before correction (left) and after correction by EMCP (right). Note the ERP derived from uncorrected trials with EOG variance smaller than the criterion corresponds to the "true" ERP.

Figure 5. "True" ERP (a), and ERP derived from all the trials before (b) and after (c) correction by EMCP for Subject 1. Associated EOG is also shown.

Figure 6. Frequency distributions of the differences between the variance of corrected and uncorrected trials for each timepoint. Separate histograms are shown for each subject and electrode.

Figure 7. "True" ERPs, ERPs based on all trials, and corrected ERPs, for rare and frequent stimuli in the visual oddball experiment. Data for Subject 1, Session 1.

Figure 8. "True" ERPs, ERPs based on all trials, and corrected ERPs, for rare and frequent stimuli in the visual oddball experiment. Data for Subject 4, Session 1.

Figure 9. Mean and interquartile range of the correction factors for four subjects for each electrode.

Figure 10. Subject 3: Mean and standard deviation of the correction factors for each electrode as a function of session (day).

APPENDIX

Derivation of the EMCP Procedure

Between any given electrode pair, on any trial, we record for each of the time-points a voltage V_t, which can be expressed as

$$V_t = ERP_t + EMA_t + N_t \text{ -- } (1)$$

where ERP_t = the "true" value of the ERP at time-point t'.

EMA_t = the contribution to V_t of signals generated at the eye ball.

N_t = is the "noise" which includes all electrical activity not time locked to the eliciting event and not associated with oculographic signals. It is assumed that the expected value of N_t is 0. Note that $\overline{V}_t = E(V_t) = ERP_t + \overline{EMA}_t$ where EMA_t is the ensemble average of the individual EMAs. As there is no reason to believe that $E(EMA_t) = 0$, the EMA_t is an artifact that must be removed from V_t or $E(V_t)$.

We assume that EMA_t is a scaled value of the signal recorded at the eye ball by the oculographic electrode pair.

Thus

$$EMA_t = K(EOG_t) \text{ -- } (2)$$

Where EOG_t is the value recorded at the ocular electrodes at time point t.

K is the scaling constant that determines the contribution of the EOG signal at V_t.

It is the goal of our procedure to estimate K. This value can then be used to "correct" V_t to remove $\overline{EMA}_t$ and to allow proper interpretation of the data. The estimation problem is of course complicated by the fact that EMA_t is unknown. We solve this problem in the following way.

We assume that the EOG is composed of two components as follows

$$EOG_t = EMR_t + EMN_t \text{ -- } (3)$$

Where

EMR_t = is the voltage generated by an eye-movement triggered, and time-locked, to the eliciting event.

EMN_t = is voltage at time-point t that is not time locked to the eliciting event. It is assumed that $E(EMN_t) = 0$.

We can now express V_t as follows:

$$V_t = ERP_t + K\,(EMR_t + EMN_t) + N_t \quad (4)$$

Therefore

$$E(V_t) = ERP_t + K(EMR_t) \quad (5)$$

Of course, $\overline{V}_t$, the ensemble average of the V_t over the n trials, is an estimate of $E(V_t)$.

We can now obtain the following relations.

From (5) and (4), we get

$$V_t - \overline{V}_t = K(EMN_t) + N_t \quad (6)$$

from (6) and (3) we get

$$V_t - \overline{V}_t = K(EOG_t - EMR_t) + N_t \quad (7)$$

As EMR_t can be estimated from the ensemble average of the EOG_t, by the same reasoning that estimates the ERP_t by $\overline{V}_t$, we get

$$V_t - \overline{V}_t = K(EOG_t - \overline{EOG}_t) + N_t \quad (8)$$

Note that all elements in equation (8), with the exception of K and N_t, are given by the data. However, as $E(N_t) = 0$, we can estimate K using standard least-squares techniques. The correction factor K is therefore obtained by

(1) Computing $\overline{V}_t$ and $\overline{EOG}_t$, for all p time points, over the trials.

(2) Computing, for each trial, $(V_t - \overline{V}_t)$ and $(EOG - \overline{EOG}_t)$.

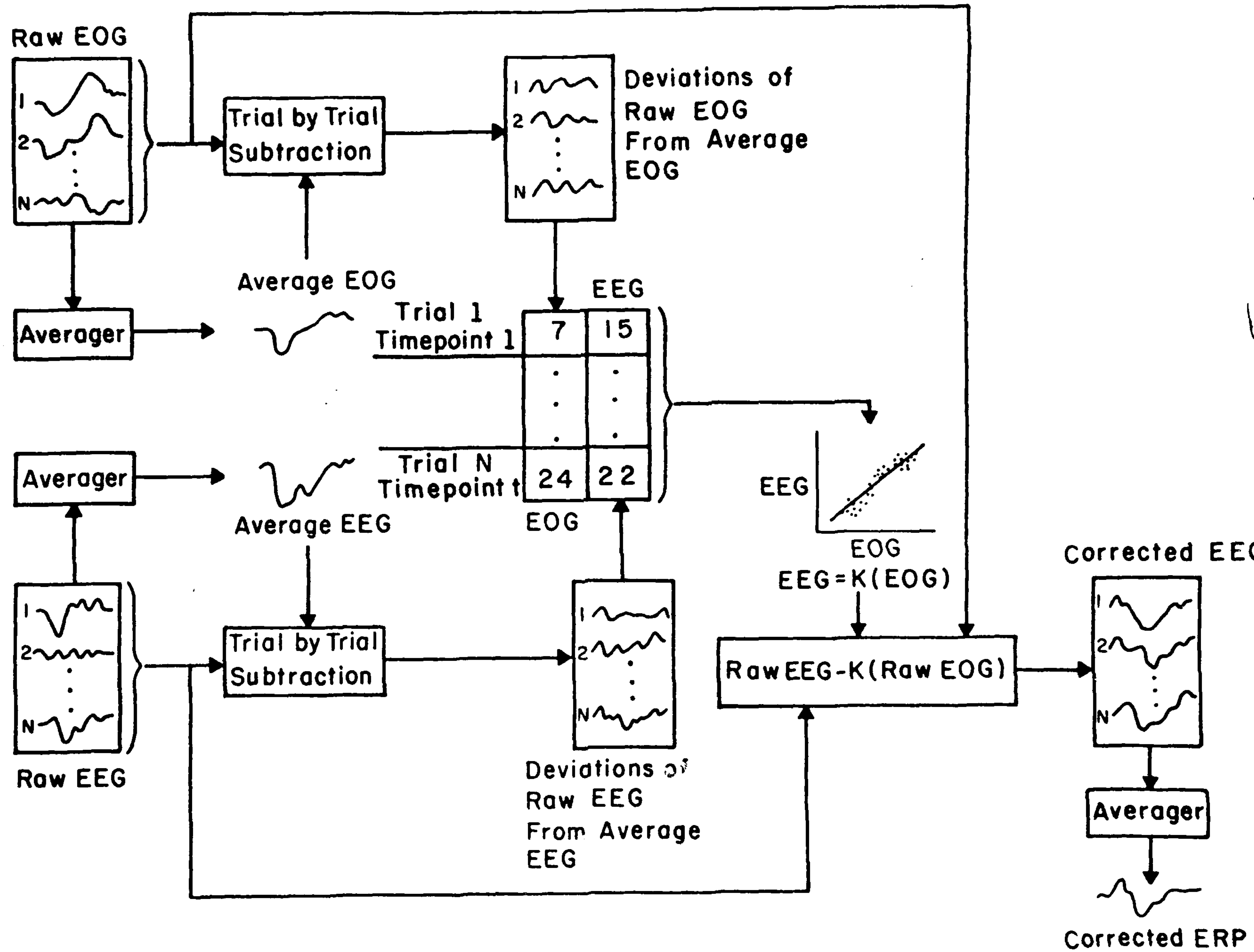
(3) Solving, using least-squares regression, K in the equation,

$$(V - \overline{V}_t) = K\,(EOG - \overline{EOG}_t),$$

using all the pairs $((V_t - \overline{V}_t), (EOG_t - \overline{EOG}_t))$ generated by all trials and all time points.

The equation is solved separately for time points affected by blinks and those affected by eye movements. A blink is assumed to have occurred at time point t, whenever the EOG signals exceed a preset criterion within a 20 msec interval bracketing time point t. Thus, if

$(EOG_t - EOG_{t-10}) + (EOG_t - EOG_{t+10}) \geq$ criterion, a blink is assumed to have occurred.



Raw EOG
1
2
N
Trial by Trial Subtraction
Deviations of Raw EOG From Average EOG
Averager
Average EOG
Trial 1 Timepoint 1
EEG
7
15
24
22
Trial N Timepoint t
EOG
Averager
Average EEG
Raw EEG
Trial by Trial Subtraction
Deviations of Raw EEG From Average EEG
EEG
EOG
EEG=K(EOG)
Raw EEG-K(Raw EOG)
Corrected EEG
Averager
Corrected ERP

Figure 1

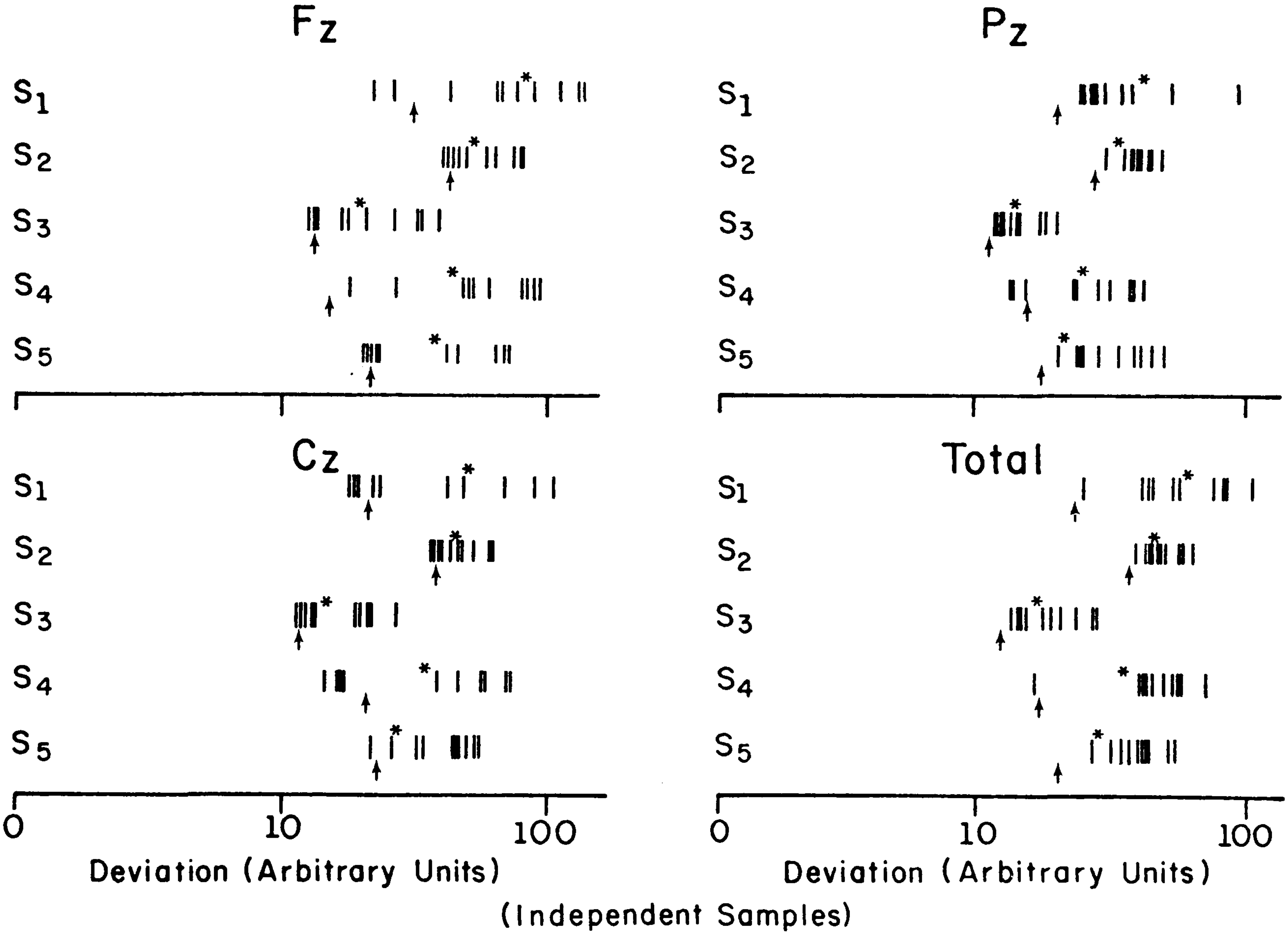
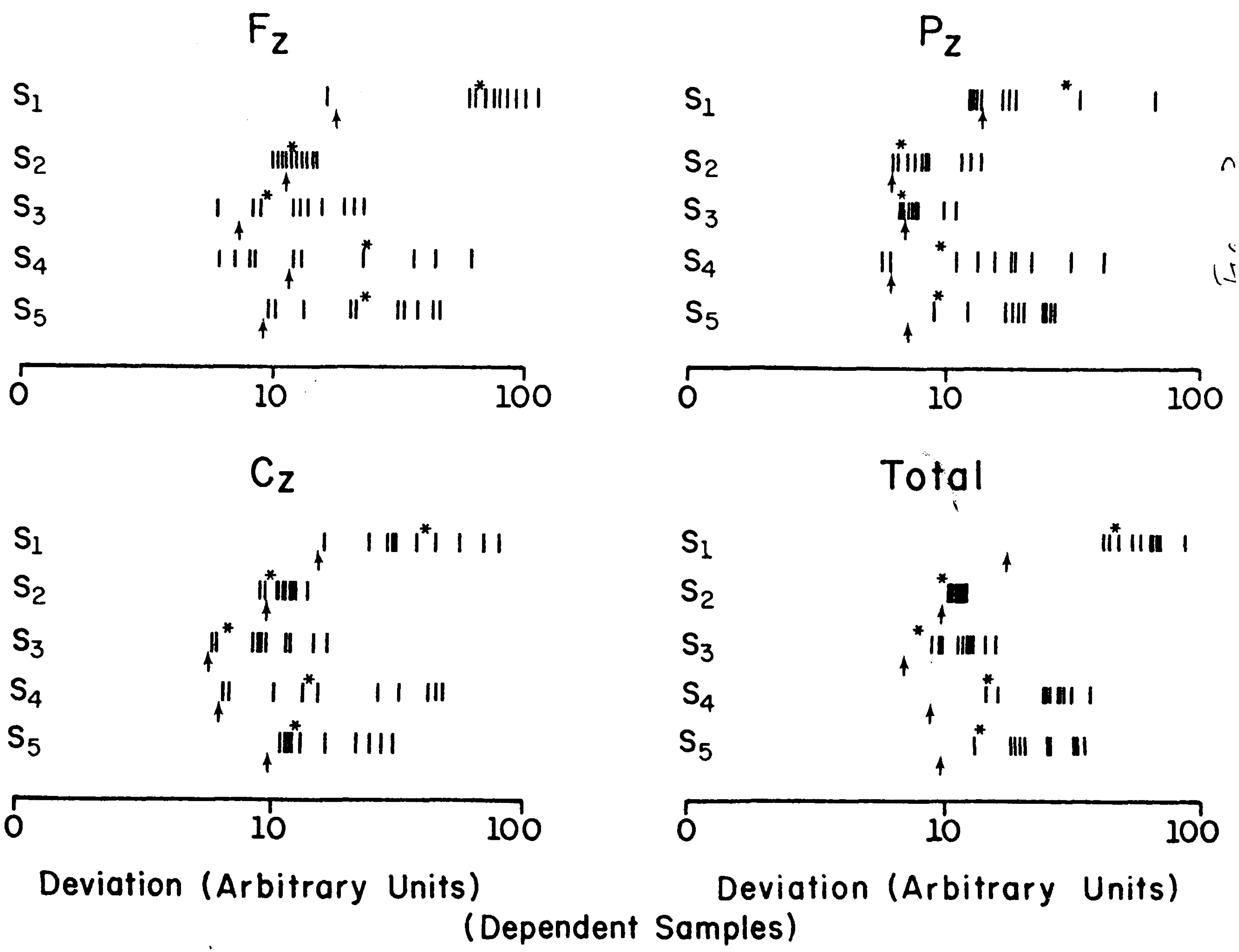

Fz
Pz
Cz
Total
S1
S2
S3
S4
S5
0
10
100
Deviation (Arbitrary Units)
(Independent Samples)

Figure 2



F_z
P_z
C_z
Total
S_1
S_2
S_3
S_4
S_5
0
10
100
Deviation (Arbitrary Units)
Deviation (Arbitrary Units)
(Dependent Samples)

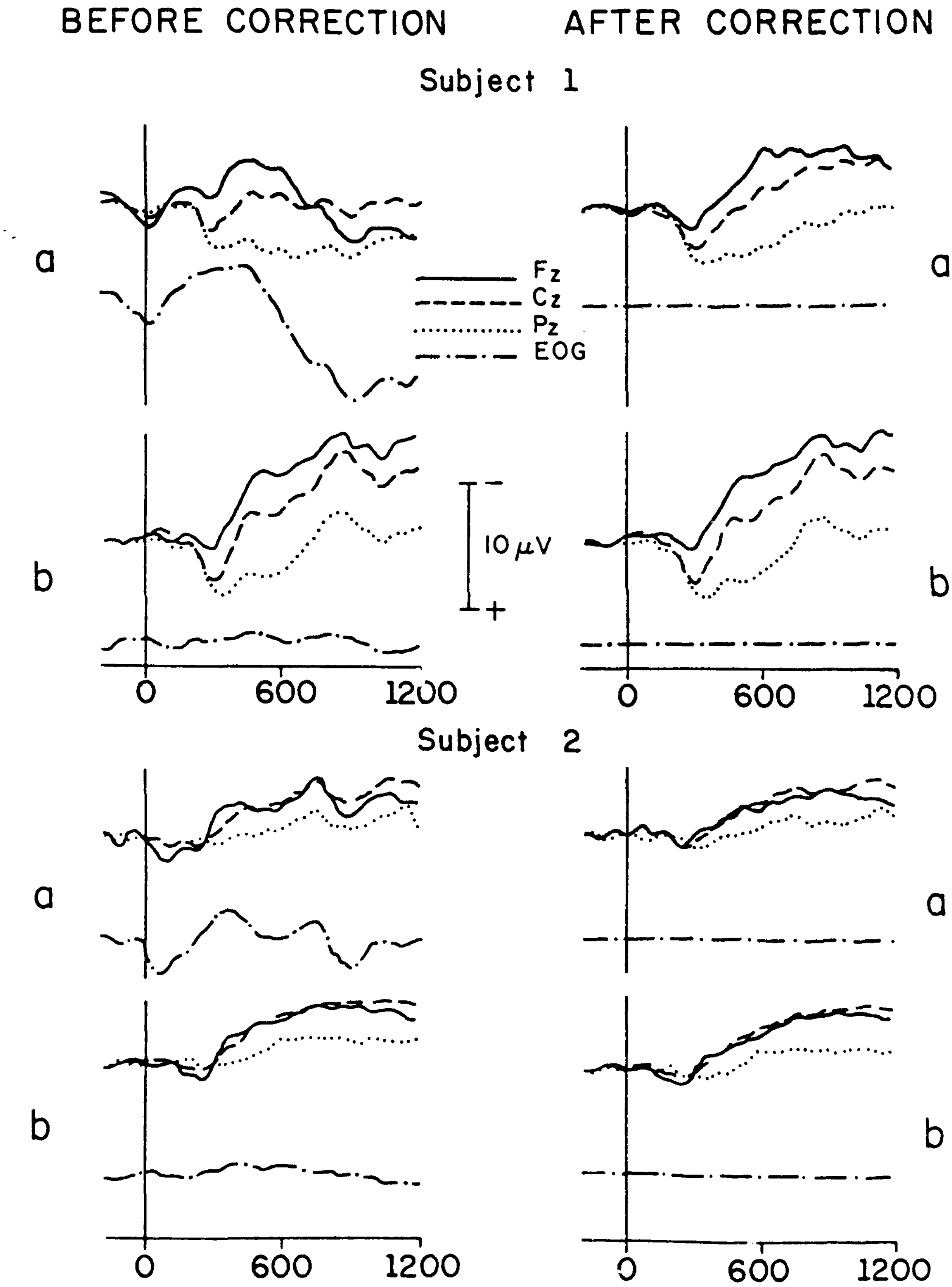
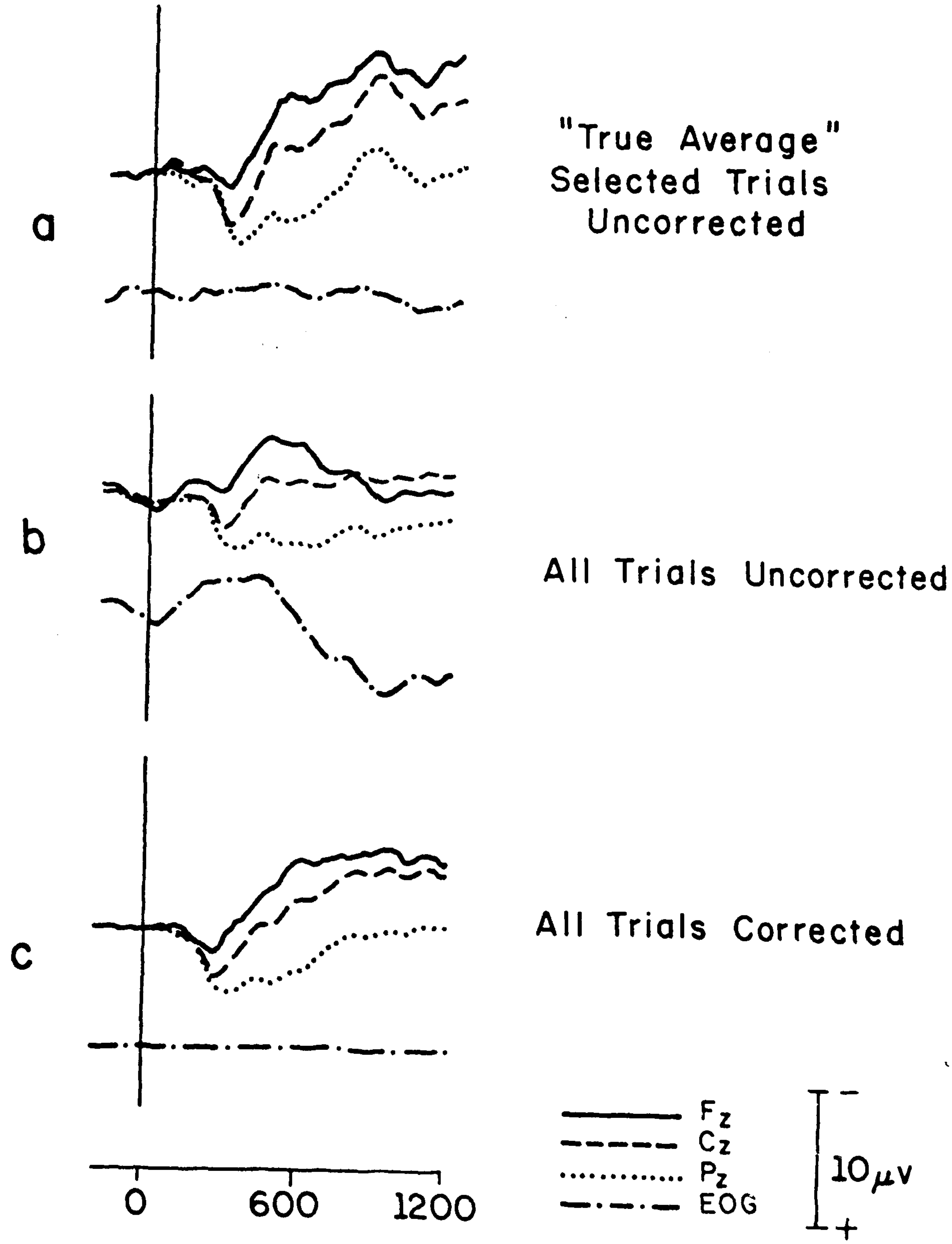

BEFORE CORRECTION
AFTER CORRECTION
Subject 1
a
a
Fz
Cz
Pz
EOG
b
b
10 μV
−
+
0
600
1200
0
600
1200
Subject 2
a
a
b
b
0
600
1200
0
600
1200

Figure 4



a
"True Average"
Selected Trials
Uncorrected
b
All Trials Uncorrected
c
All Trials Corrected
0
600
1200
F_z
C_z
P_z
EOG
–
$10 \mu v$
+

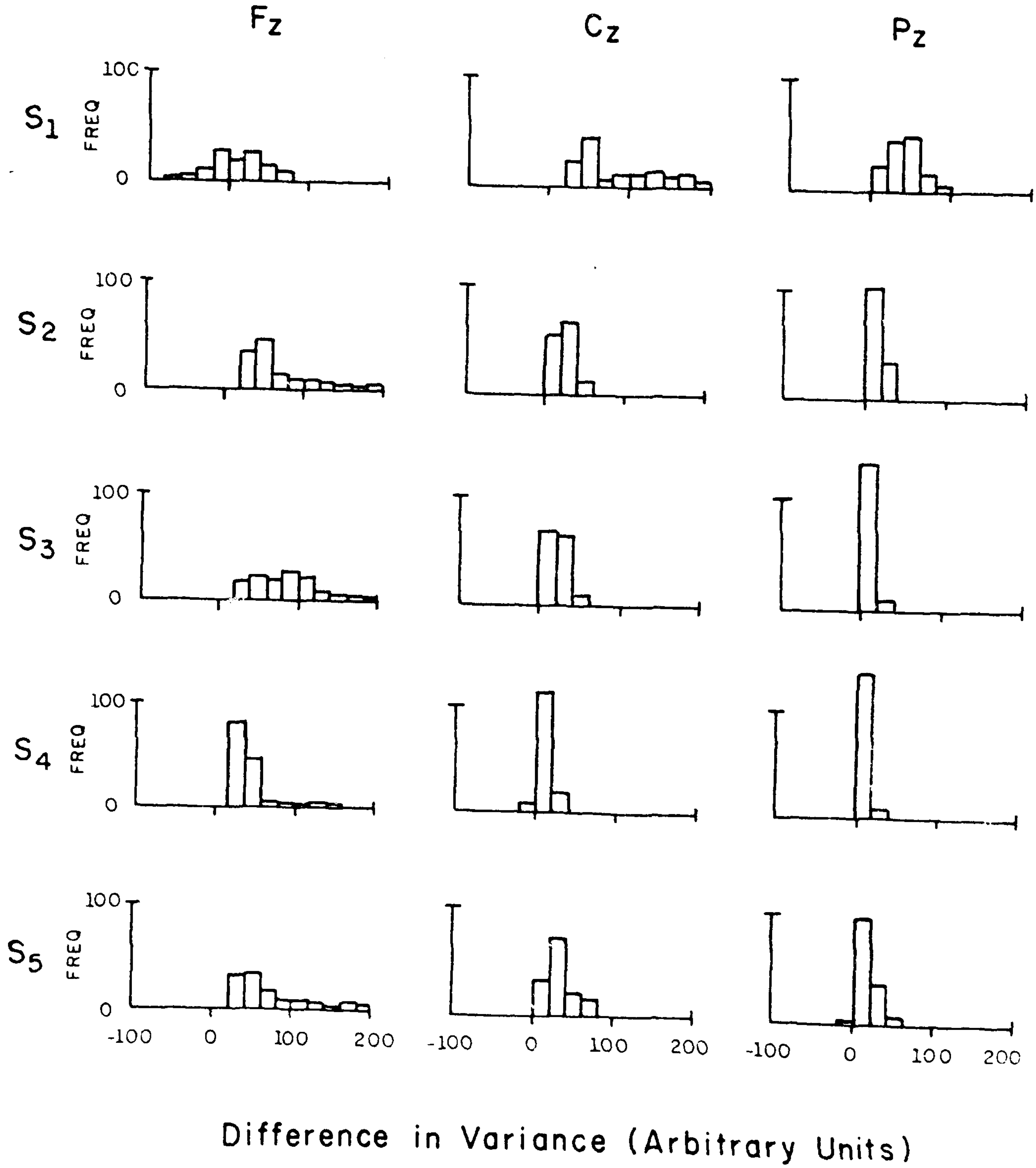

Fz
Cz
Pz
S1
S2
S3
S4
S5
FREQ
100
0
-100
0
100
200
Difference in Variance (Arbitrary Units)

Figure 6

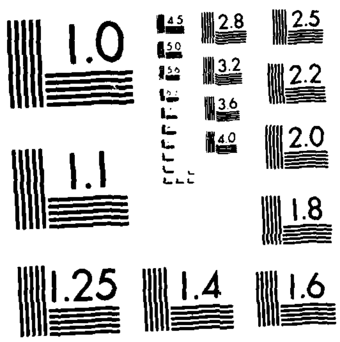
AD-A125 699 THE EVENT RELATED BRAIN POTENTIAL AS AN INDEX OF INFORMATION PROCESSING C..(U) ILLINOIS UNIV CHAMPAIGN COGNITIVE PSYCHOPHYSIOLOGY LAB E DONCHIN ET AL. NOV 82 2/2

UNCLASSIFIED CPL82-2 AFOSR-TR-83-0051 F49620-79-C-0233 F/G 5/10 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

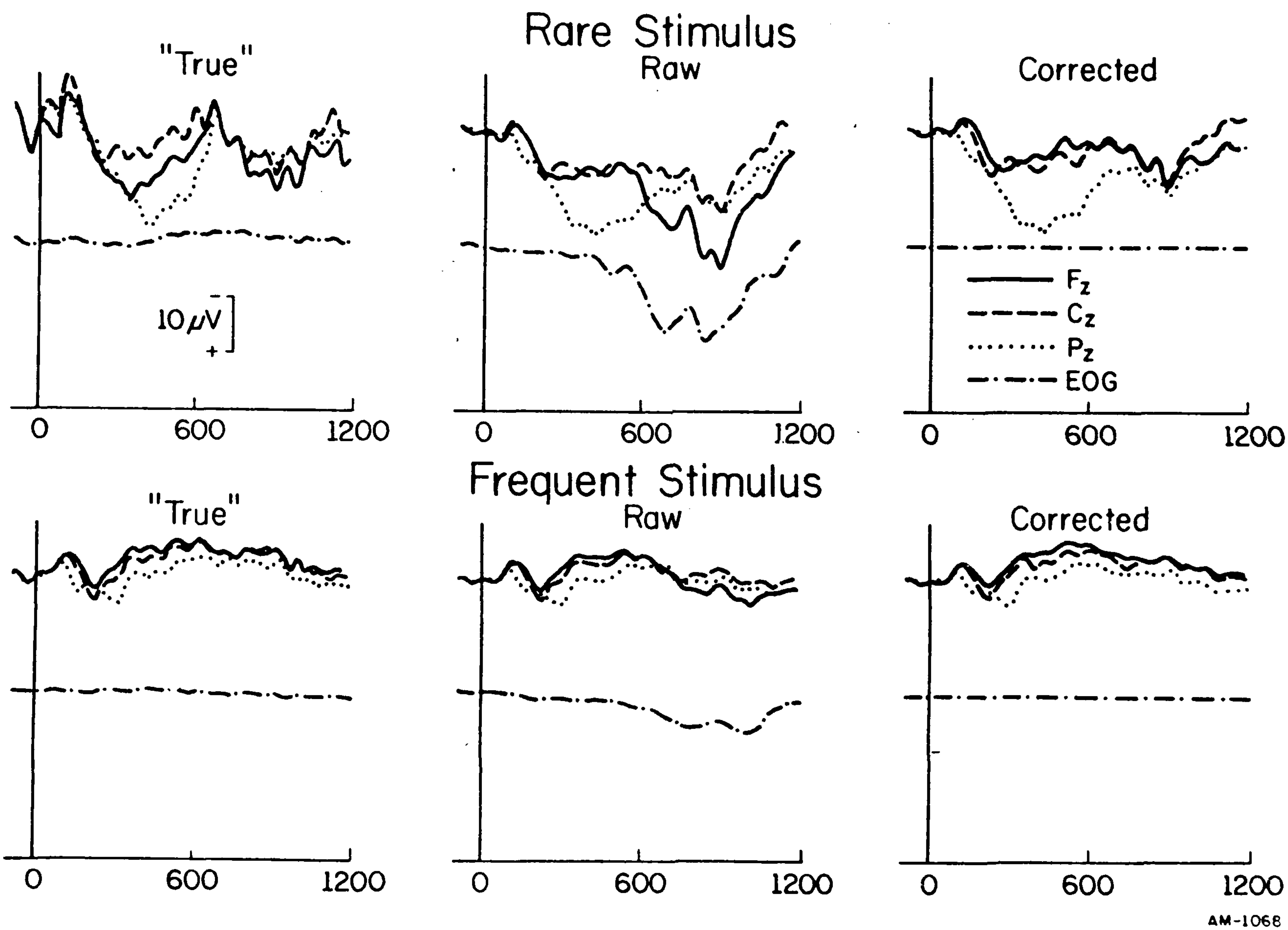

Rare Stimulus
"True"
Raw
Corrected
10 µV
F_z
C_z
P_z
EOG
0
600
1200
Frequent Stimulus
"True"
Raw
Corrected
AM-1068

Figure 7

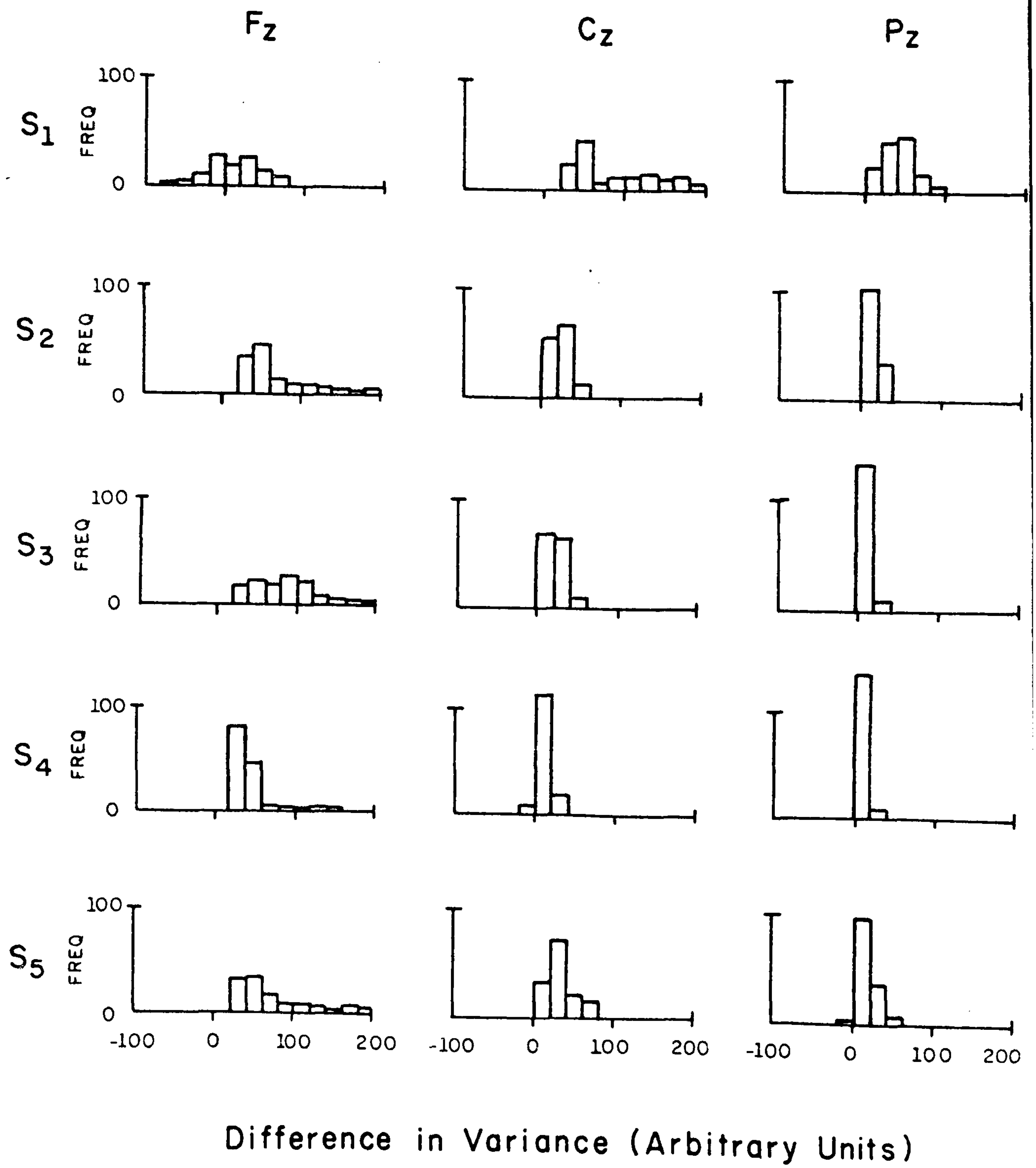
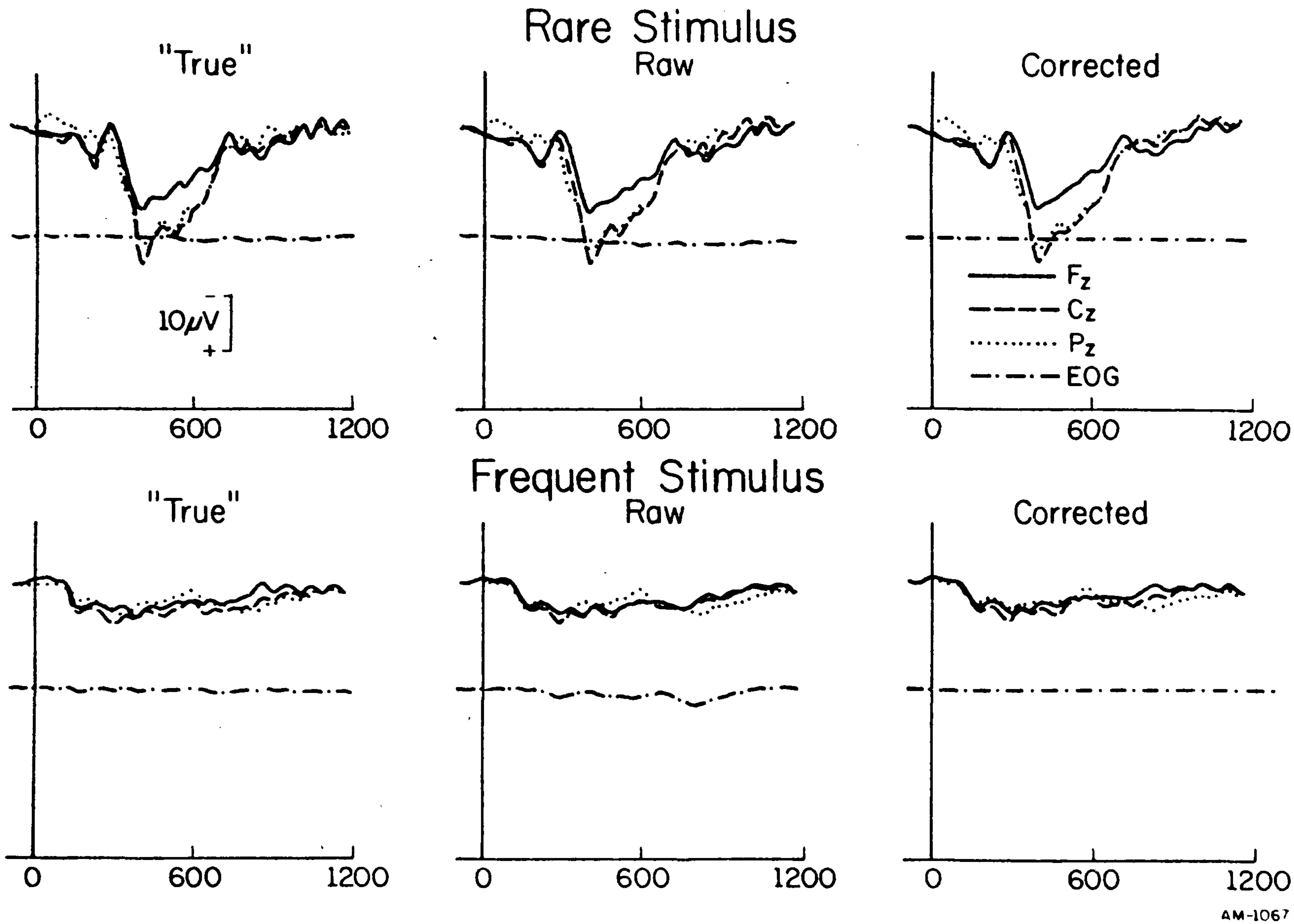

Fz
Cz
Pz
S1
S2
S3
S4
S5
FREQ
100
0
-100
0
100
200
Difference in Variance (Arbitrary Units)

Figure 6



Rare Stimulus
"True"
Raw
Corrected
10μV
−
+
Fz
Cz
Pz
EOG
0
600
1200
Frequent Stimulus
"True"
Raw
Corrected
0
600
1200
AM-1067

Figure 8

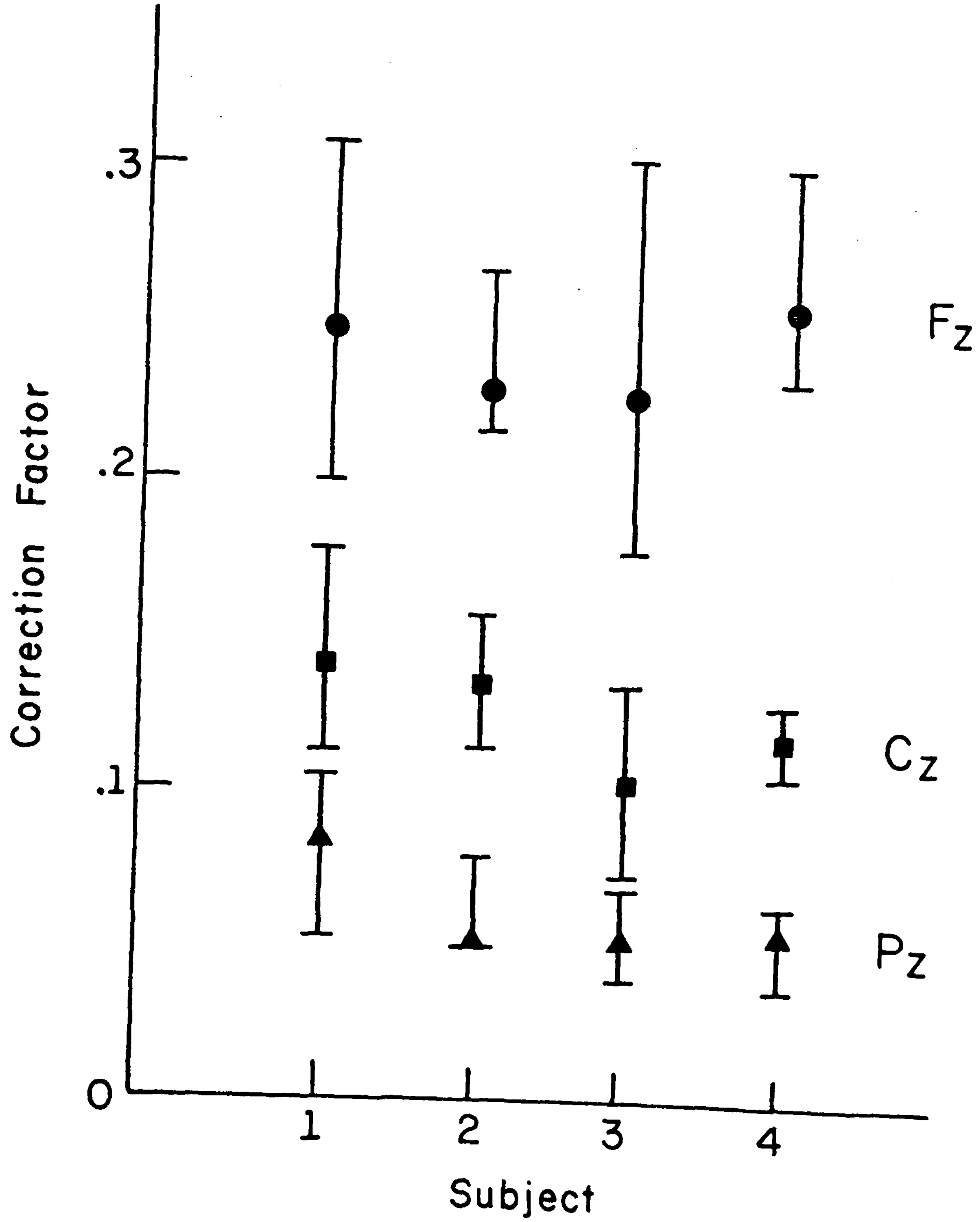


Figure 9

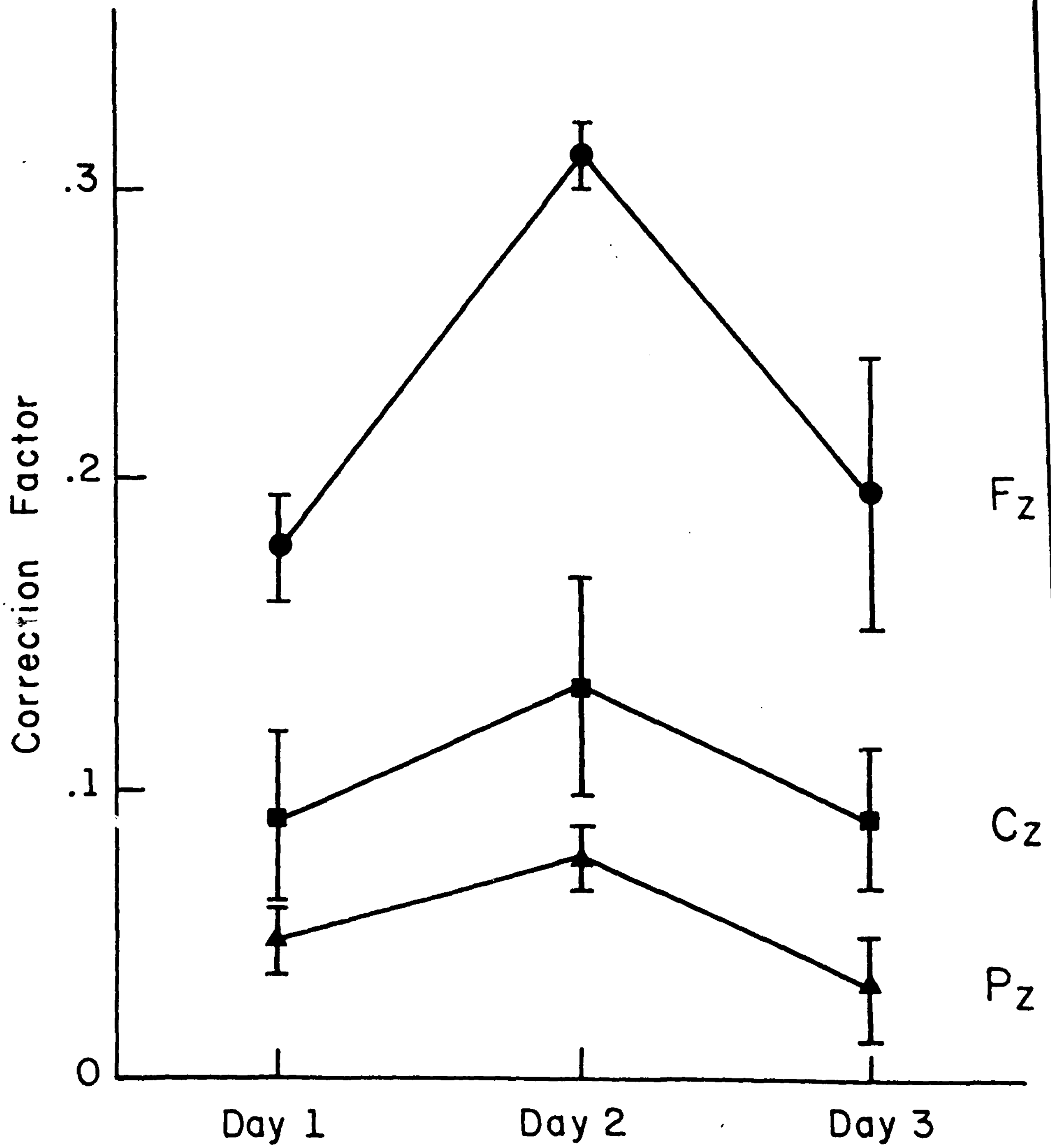


Figure 10

THE PERFORMANCE OF CONCURRENT TASKS: A PSYCHOPHYSIOLOGICAL ANALYSIS OF THE RECIPROCITY OF INFORMATION PROCESSING RESOURCES

C. Wickens, A. Kramer, L. Vanasse, E. Donchin

Submitted to *Science*

Abstract

This experiment demonstrates that the resources allocated to a primary and secondary task are reciprocal. Subjects perform a tracking task in which the discrete displacements of the tracking cursor may be used to elicit Event-Related Brain Potentials. As the resource demands of the tracking task are increased, ERPs elicited by the task-defined events increase in amplitude, while those elicited by secondary task auditory stimuli decrease.

The Performance of Concurrent Tasks: A Psychophysiological Analysis of the Reciprocity of Information Processing Resources

The limitations on the capacity of the human information processing system have long been recognized. Dante, for example, begins the fourth Canto of Purgatorio as follows:

> When we receive a pleasure or a pain
> Through any one of all our faculties,
> The soul is concentrated on that sense
> And seems to heed no other faculty;
> And this belies the theory which maintains,
> That there are several souls in each of us.
> And therefore when a thing is heard or seen
> That keeps the soul intently turned to it,
> The time goes by and we perceive it not:
> The sense of time is as it were suppressed,
> The other sense possesses all the soul. (Mackenzie, 1)

These lines capture concisely and with surprising precision a central, practical and theoretical, problem of human information processing. When the soul, or the mind, as is the modern preference, is occupied in performing one task it often lacks capacity for performing other tasks (2). This, of course, is not always the case. Under some circumstances, many tasks can be performed concurrently. Much ingenuity has been devoted to defining the limits of processing capacity, and to elucidating the manner with which capacity varies under different conditions to influence the success of dual task performance. In a conceptualization that owes much to Kahneman (3), and has been developed by others (4), performance of a task is assumed to consume resources that are in limited supply. The degree to which two tasks interfere with each other depends on the extent to which they compete for the common supply of resources. Thus, when the mind is "intent" on some tasks resources may not be available for other tasks. In such cases the subject's primary task does, as Dante put it, "possess all the soul".

Any attempt to study the limitation on the human's capacity for time-sharing must contend with severe measurement problems. It is not possible to assess capacity directly. Furthermore, capacity cannot be inferred from the quality of task performance: performance <u>can</u> remain constant in quality despite changes in the demands on the operator. For example, operators may change their strategy and approach to a task to compensate for limits on their capacity. Such difficulties have lead investigators who are interested in the practical study of mental workload, and theorists interested in the attentional system, to devise techniques for indirectly assessing the resources allocated to tasks. A frequently used technique for measuring the mental workload associated with a specific task is to assign to the operator yet another, "secondary", task that must be executed concurrently with the "primary" task. The operator is instructed to perform the primary task in the best possible manner and to perform the secondary task as well as possible under the circumstances. The level of performance of the <u>secondary</u> task is used to assess the workload imposed by the primary task, on the assumption that secondary task performance reflects the amount of resources not consumed by the primary task (5).

This rather widely used technique involves the implicit assumption that the performance of the secondary and primary tasks do not interact. However, as the assessment of performance requires observable, overt responses, it is not surprising that this assumption is not entirely tenable (6). The need to press buttons, manipulate levers or otherwise exercise the skeletal musculature or the speech system in service of the secondary task, of necessity interferes with the same actions made on behalf of the primary task. In partial solution of this problem it has been proposed (7) that the amplitude and latency of the P300 component of the human event related potential (ERP) can serve as a dependent variable in studies of workload. Donchin, Wickens and their coworkers presented subjects who were performing a variety of complex tasks, with a concurrent Bernoulli sequence of stimuli with unequal probabilities. It is well established that if such a sequence is attended, then the rare stimulus elicits a P300 whose magnitude is inversely related to the probability of the eliciting event (8). We were able to show that when this "oddball" task is performed concurrently with a "primary" task, the amplitude of the P300 elicited by the rare stimuli depends on the difficulty of the primary task, provided that difficulty is manipulated in the perceptual-cognitive domain. For example, Wickens, Kramer, and Donchin (9) required subjects to count tones while they were performing a complex multidimensional target acquisition task. The tones elicited a P300 component whose amplitude varied inversely with the perceptual difficulty of the target acquisition task.

These data suggested that the amplitude of the P300 can serve as an index of the resources available to the operator from a primary task. This hypothesis implies that when the primary and secondary tasks tap common resources there is a reciprocity in the availability of these common resources to the primary and the secondary task (10). As a consequence, there will be a reciprocity in the amplitude of the P300 associated with the two tasks as primary task difficulty is varied. In other words, while P300 associated with the secondary task decreases in amplitude with increasing difficulty of the primary task, it should be possible to observe corresponding increase in the amplitude of the P300 elicited by stimuli associated with the primary task. This hypothesis has not been tested before and the study reported here was designed to provide such a test.

Twelve subjects (11), performed a pursuit step-tracking task. In this task a visually displayed target executes a series of discrete horizontal displacements. The magnitude of each displacement is determined by a random process. The displacements occurred at 3 sec intervals. The subjects manipulated a control stick and attempted to superimpose a cursor on the target (12). Changes in the difficulty of the task were accomplished by manipulating two variables: (a) The directional regularity of the sequence of the step changes had two levels. Under the "high-predictability" condition the step displacements alternated in a regular left-right sequence. Only the magnitude of the step change remained unpredictable. In the "low-predictability" condition, both magnitude and direction were uncertain so that two successive steps in the same direction were possible. (b) The relationship between the movement of the joy stick and the movement of the cursor were varied. In the "first-order control" conditions constant displacement of the control stick caused the cursor to move at a constant velocity in the direction of the movement. In the second-order control conditions, constant movement of the stick accelerated the cursor's movement (13). These two manipulations of difficulty were combined to create 3 conditions of increasing difficulty: First-order control with predictable input (1P), first order control of unpredictable input (1U), and second order control of unpredictable input (2U).

Concurrently with the tracking tasks the subjects were assigned one of three "oddball" tasks designed to elicit the ERPs (14). Two of these were clearly secondary while stimuli for the third were embedded in the primary task itself. In the auditory probe condition subjects heard a Bernoulli series of tones of high and low pitch. The two tones occurred with equal probability (15). Subjects were instructed to count the number of occurrences of the low pitch tone. In the visual probe flash condition a horizontal bar was imposed along the course traversed by the target. The bar was flashed for 100 ms. On a randomly selected half of the trials the flashes were brighter than on the other trials; the subject was instructed to count the dimmer flashes. These two secondary tasks are thus, equivalent to the tasks employed in our previous work. The step-tracking task which is characterized by discrete events, afforded an opportunity to create a third type of secondary task using events associated with the primary task. Therefore, in the visual probe step condition subjects counted all steps in which the target moved in a given direction. In an additional control condition subjects tracked the step changes in the absence of a secondary task. In the control condition, and in the visual probe step condition, the averaging computer was triggered by the step changes so that ERPs elicited by the change in the target's direction were recorded. Finally, three additional conditions were included in which each of the three "secondary" counting tasks (auditory probe, visual probe flash, visual probe step) was performed without the concurrent tracking task, though the moving target remained on the screen (16). After performing each condition, subjects were asked to rate, on a 7-point scale, the subjective difficulty of the task just performed.

The data indicate that the primary tasks did, in fact, present the subject with an increasingly greater challenge. As Figure 1 shows, the tracking error (17) was lowest in the 1P condition, and greatest in the 2U condition: There was an interaction between the primary and secondary tasks in that the RMS error was larger when the counting task was performed. Figure 1 also presents the average difficulty rating assigned by the subjects to the different tasks. Clearly, the subject's experience was consistent with the order of difficulty indicated by the tracking errors. Thus, the subjects not only performed more poorly on the tasks we assumed to be more difficult, they also related these tasks as more difficult (18).

The ERP waveforms, recorded at the parietal electrodes are shown in Figure 2a. Figure 2b depicts the base-to-peak measure of P300 amplitude for each of these conditions. Statistical analysis of the data was accomplished by submitting the digitized waveforms to four principal component analyses (19): One for the auditory condition, one for the visual flash and one for each of the visual step conditions. Separate analyses were necessary due to variability in the latency of the P300 component across experimental conditions. In the upper left panel of Figure 2a are the ERPs elicited by the secondary task, auditory probes. These data replicate the previously reported results in similar studies. A large P300 is elicited by the rare event when counting is the subject's sole assignment. The introduction of concurrent tracking leads to a substantial reduction in P300 amplitude, and there is a further reduction in P300 amplitude with increasing difficulty of the primary task (20). The reduction in P300 amplitude cannot be attributed to the subjects' failure to count probes during the more difficult dual-task conditions since the counting accuracy remained uniformly high across the experimental conditions and difficulty manipulations.

As in the auditory probe condition, the amplitude of the P300 in the two visual probe conditions is also significantly attenuated by the introduction of the tracking task. This decrement or "cost of concurrence" (21) results from reallocation of a certain portion of resources from the count task to the tracking task, independent of the probes used in the former or the difficulty of the latter. In the visual flash condition, increases in primary task difficulty failed to produce any further attenuation (22). In both of the step probe and control conditions, increasing the resources demanded by the primary task resulted in a concomitant enhancement in the amplitude of the P300 elicited by primary task stimuli (23). All of the difficulty levels were significantly different from each other in both of the step probe conditions.

These results provide additional support for the hypothesis that the reduction in the amplitude of the P300 elicited by secondary-task stimuli results from a depletion of the resources deployed in the service of the secondary task by competition for these resources from the primary task (24). As predicted, when the secondary task utilizes stimuli embedded within the primary tracking task, these stimuli elicit P300s with increasing magnitude as tracking difficulty increases. Of course, the functional role of these depleted resources remains obscure. We assume that the P300 manifests the activation of some information processing activity that is invoked by the appearance of task-relevant events, its amplitude inversely related to its expectancy. It has been suggested that this "subroutine" is involved in updating, or revising, the model of the environment maintained in working memory (25). The resources upon which this updating activity depends would appear to be limited in their availability and, when deployed in the service of one task, their availability to be of service to other tasks is reduced.

Whether this particular interpretation of the P300 is valid or not, the data we present provide additional support for the assertion that the P300, and perhaps other endogenous components of the ERP, may be of use in the assessment of workload. While the P300 is not the ever-elusive universal measure of workload, it does provide a reliable index of the demands, in the perceptual-cognitive domain, that tasks impose upon a subject. An important implication of the data we report here is that it is possible, in the assessment of workload, to utilize stimuli that are integral to the primary task. In selecting psychophysiological workload measures, investigators can hereafter choose between two modes. They can rely on purely mental, relatively standard, irrelevant probes -- gaining standardization and ease of implementation at the cost of adding a completely extraneous secondary task. Alternately, when it is possible to interface the primary task to the averaging computer, a measure of workload can be obtained without extraneous tasks. No doubt, circumstances will arise that would make one or the other the method of choice.

Christopher Wickens

Arthur Kramer

Linda Vanasse

Emanuel Donchin

Cognitive Psychophysiology Laboratory

University of Illinois,

Champaign-Urbana

Footnotes

(1) Dante, Alighiere. The Divine Comedy. Purgatorio, Canto IV., translated by McKenzie, Folio Society Edition, London, 1980.

(2) Knowles, W.B. Operator loading tasks. Human Factors, 1963, 5, 155-161. Moray, N. Where is capacity limited? A survey and a model. In A.F. Sanders (Ed.), Attention and Performance I. Amsterdam: North Holland Publishing Company, 1967. Sanders, A.F. Some remarks on mental load. In N. Moray (Ed.), Mental Workload: It's Theory and Measurement. New York: Plenum Press, 1979.

(3) Kahneman, D. Attention and Effort. Englewood Cliffs, NJ: Prentice Hall, 1973.

(4) Navon, D. and Gopher, D. On the economy of the human processing system. Psychological Review, 1979, 86, 214-255. Norman, D.A. and Bobrow, D.G. On data-limited and resource-limited processes. Cognitive Psychology, 1975, 7, 44-64. Schneider, W. and Shiffrin, R.M. Controlled and automatic information processing: Detection, search, and attention. Psychological Review, 1977, 84, 1-66. Wickens, C.D. The structure of attentional resources. In R. Nickerson and R. Pew (Eds.), Attention and Performance, VIII, Englewood Cliffs, NJ: Erlbaum, 1980.

(5) Moray, N. _Mental workload: Its Theory and Measurement_. New York: Plenum Press, 1979. Ogden, G.D., Levine, J.W., and Eisner, E.J. Measurement of workload by secondary tasks. _Human Factors_, 1979, _21_, 529-548. Rolfe, J.M. The secondary task as a measure of mental load. In W.T. Singleton, J.G. Fox and D. Witfield (Eds.), _Measurement of Man at Work_. London: Taylor and Francis, 1971.

(6) Ogden et al. Ibid. Sheridan, T. and Stassen, H. Definitions, models and measures of human workload. In N. Moray (Ed.), _Mental Workload: Its Theory and Measurement_. New York: Plenum Press, 1979.

(7) Isreal, J.B., Wickens, C.D., and Donchin, E. The dynamics of P300 during dual-task performance. In H.H. Kornhuber & L. Deecke (Eds.), _Motivation, Motor and Sensory Processes of the Brain Electrical Potentials, Behavioral and Clinical Use_. Netherlands: North Holland Biomedical Press, 1980. Isreal, J.B., Wickens, C.D., Chesney, G.L., and Donchin, E. The event-related brain potential as an index of display monitoring workload. _Human Factors_, 1980, _22_, 211-224. Wickens, C.D., Gill, R., Kramer, A., Ross, W., and Donchin, E. The cognitive demands of second order manual control: Applications of the event-related brain potential. _Proceedings of the 17th Annual NASA Conference on Manual Control_, NASA TM, 1981.

(8) Donchin, E., Ritter, W., and McCallum, W.C. Cognitive psychology: The endogenous components of the ERP. In E. Callaway, P. Tueting, &

S.H. Koslow (Eds.), Event-Related Brain Potentials in Man. New York: Academic Press, 1978. Duncan-Johnson, C.C. and Donchin, E. On quantifying surprise: The variation in event-related potentials with subjective probability. Psychophysiology, 1977, 14, 456-467. Tueting, P. Event-related potentials, cognitive events and information processing: A summary of issues and discussion. In D. Otto (Ed.), Multidisciplinary Perspectives in Event Related Brain Potential Research. EPA 60019-77-043, Washington, D.C. U.S. Government Printing Office, 1978.

(9) Wickens, C.D., Kramer, A., and Donchin, E. The event-related potential as an index of the processing demands of a complex target acquisition task. In the Proceedings of the 6th International Conference on Event-Related Slow Potentials of the Brain. (EPIC VI). R. Karrer (Ed.), in press.

(10) Such a reciprocity will not hold if functionally different processing resources are employed for the primary and secondary task. (Navon and Gopher, 1979; Wickens, 1980; Isreal, Chesney, Wickens, and Donchin, 1980).

(11) Twelve (4 females) right-handed students were paid for their participation. None had previous experience with the tracking task. The subjects had either normal or corrected-to-normal vision.

(12) The tracking elements (2 cm x 2 cm) were presented on a Hewlett Packard CRT which was positioned approximately 70 cm from the subjects. The position of the target along the horizontal axis was varied once every 3 sec. Subjects controlled the position at the cursor with a single axis joystick.

(13) Second order control has been validated to interfere more with concurrent tasks, generate greater ratings of subjective difficulty and produce poorer performance (Wickens, Gill, Kramer, Ross and Donchin, Op. Cit.; Gopher, D. and Navon, D. Acta Psychologica, 1980, 46, 161-180). The source of increased demands is attributable to the perceptual demands of prediction and processing higher derivatives of the error signal, the increased central processing demands of maintaining a more complex internal model, and increased demands associated with executing more complex responses (Wickens, Gill, Kramer, Ross, and Donchin, Op. Cit.).

(14) EEG was recorded from three midline sites (Fz, Cz and Pz according to the 10-20 system; Jasper, 1958) and referred to linked mastoids. Two ground electrodes were positioned on the left side of the forehead. Burden Ag-AgCl electrodes affixed with collodion were used for scalp and mastoid recording. Beckman Bipotential electrodes, affixed with adhesive collars, were placed laterally and sopra-orbitally to the right eye to record electro-ocologram (EOG) and this type of electrode was also used for ground recording. Electrode impedances did not exceed 5 Kohms/cm.

The EEG and EOG were amplified with Van Gogh model 50000 amplifiers (time constant 10 sec and upper half amplitude of 35Hz, 3dB octave roll-off). Both EEG and EOG were sampled for 1280 msec, beginning 100 msec prior to stimulus onset. In different experimental blocks the stimuli were defined as two tones differing in frequency, two intensifications of a horizontal bar varying in brightness or changes in the spatial position of the cursor towards the right or left. The data were digitized every 10 msec. ERPs were filtered off-line (-3dB at 6.29Hz, 0dB at 14.29 Hz) prior to statistical analysis.

(15) Auditory stimuli were presented in a Bernoulli series of low pitched (1200 Hz) and high pitched (1400 Hz) sinosodial tone bursts (60 dB SPL, re 20 un/m) which were delivered binaurally through TDH-39 headphones. Tones were 60 msec in duration (including 10 msec rise and fall) and were presented every 2.8 sec. The probability of a high tone occurring on any trial was .50.

(16) A repeated measures, four-way factorial design, was employed. The factors were primary task difficulty (count only, 1P, 1U and 2U), secondary task (counted tones, flashes or steps), stimuli (counted or uncounted) and scalp recording site (Fz, Cz and Pz). The subjects also performed three single task tracking conditions.

Subjects participated in two experimental sessions. The first session consisted of 30 blocks of practice trials, ten with first order control dynamics and twenty with second order control dynamics. RMS error, count accuracy and subjective difficulty ratings were collected during the practice session. In the second session subjects performed two practice blocks (first and second order) prior to participation in the experimental conditions. The 30 experimental blocks were composed of two replications of the fifteen conditions (four difficulty manipulations x three secondary tasks and three single task tracking blocks). The order of presentation of the conditions was counterbalanced across subjects.

(17) The RMS error means for the three difficulty conditons (1P, 1U, 2U) were 137, 208, and 249 [$F(2,22) = 289.7$, $p < .0001$]. Paired contrasts between adjacent levels also revealed these to be reliable ($p < .01$ in both cases). The interaction between experimental condition and primary task difficulty was also statistically significant [$F(6,66) = 9.5$, $p < .001$]

(18) The means for the subjective difficulty ratings in the count only, 1P, 1U and 2U conditions were 1.46, 3.07, 3.54, and 4.71 [$F(3,33) = 99.2$, $p < .0001$].

(19) Donchin, E. A multivariate approach to the analysis of average evoked potentials. IEEE Transactions on Bio-Medical Engineering, 1966, BME-13, 131-139.

Donchin, E. and Heffley, E. Multivariate analysis of event-related potential data: A tutorial review. In D. Otto (Ed.), Multidisciplinary Perspectives in Event-Related Brain Potential Research. EPA 600/9-77-043, Washington, D.C.: United States Government Printing Office, 1979.

The data bases submitted to the PCA's were composed of 288 average waveforms (12 subjects x 4 primary task difficulty manipulations x 2 stimuli x 3 electrodes) containing 128 (1.28 sec) points each. Five components were derived from each of the PCA's. The component scores obtained from the PCA's were analyzed in repeated measures ANOVAs.

(20) ERP components are customarily defined in terms of their latency relative to a stimulus or response, electrode distribution and sensitivity to experimental manipulations. Based on these criteria, a component was identified in each of the PCA's which would qualify as the P300. Other components are beyond the scope of this paper. In the auditory condition, component 2 loadings were maximal in the temporal range associated with the P300 (300 to 500 msec), the amplitude of the component was largest at the parietal electrode [$F(2,22) = 37.8$, $p < .0001$] and the component score was larger for the counted than the uncounted stimulus [$F(1,11) = 21.4$, $p < .001$].

The main effect for the difficulty factor in the auditory condition was statistically significant [$F(3,33) = 4.28$, $p < .01$]; progressively smaller amplitude P300s were elicited as the difficulty of the primary task increased. P300 amplitude was also reliably different ($p < .05$) between each adjacent difficulty level.

(21) Navon and Gopher, Op. Cit.

(22) The main effect for the difficulty factor in the visual flash condition was statistically significant [$F(3,33) = 3.99$, $p < .02$]. However, this effect is due to a decrement in the amplitude of the P300 from the single task count only condition to the dual task conditions. No further attenuation of P300 amplitude was found in the dual task blocks.

(23) The main effect for the difficulty factor in both visual probe step conditions was statistically significant [Count: $F(2,22) = 6.59$, $p < .005$, Control: $F(2,22 = 9.56$, $p < .001)$]; progressively larger amplitude P300s were elicited as the difficulty of the primary task decreased.

(24) The reason why P300 did not decline with tracking difficulty in the visual flash condition cannot be stated with certainty, although lack of a difficulty effect on secondary task visual P300s is consistent with the results of another unpublished study in our laboratory. Two hypotheses may be tentatively offered. (a) Tracking error is higher in the flash than in the auditory probe condition. Thus subjects may have biased their allocation of resources toward the probes to a greater extent when the probes were visual than when they were

auditory. (b) The visual flash condition may be placed in the middle of an ordered continuum defining the degree of separation between primary and secondary task stimuli. Like the step probes, the visual flash probes share the common visual modality; but like the auditory probes they are independent events from stimuli in the tracking task. Since the step probes produce increasing amplitude with primary difficulty, and the auditory probes a decreasing amplitude, we could anticipate that the visual flash probes, in the middle of this continuum, would reflect the compromise of these two trends, i.e., no effect.

(25) Donchin, E. Surprise! . . . Surprise! Psychophysiology, 1981, 18, 493-513. Donchin, E., Kramer, A., and Wickens, C. Probing the cognitive infrastructure with event-related brain potentials. Proceedings of the AIAA Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics. California, 1982.

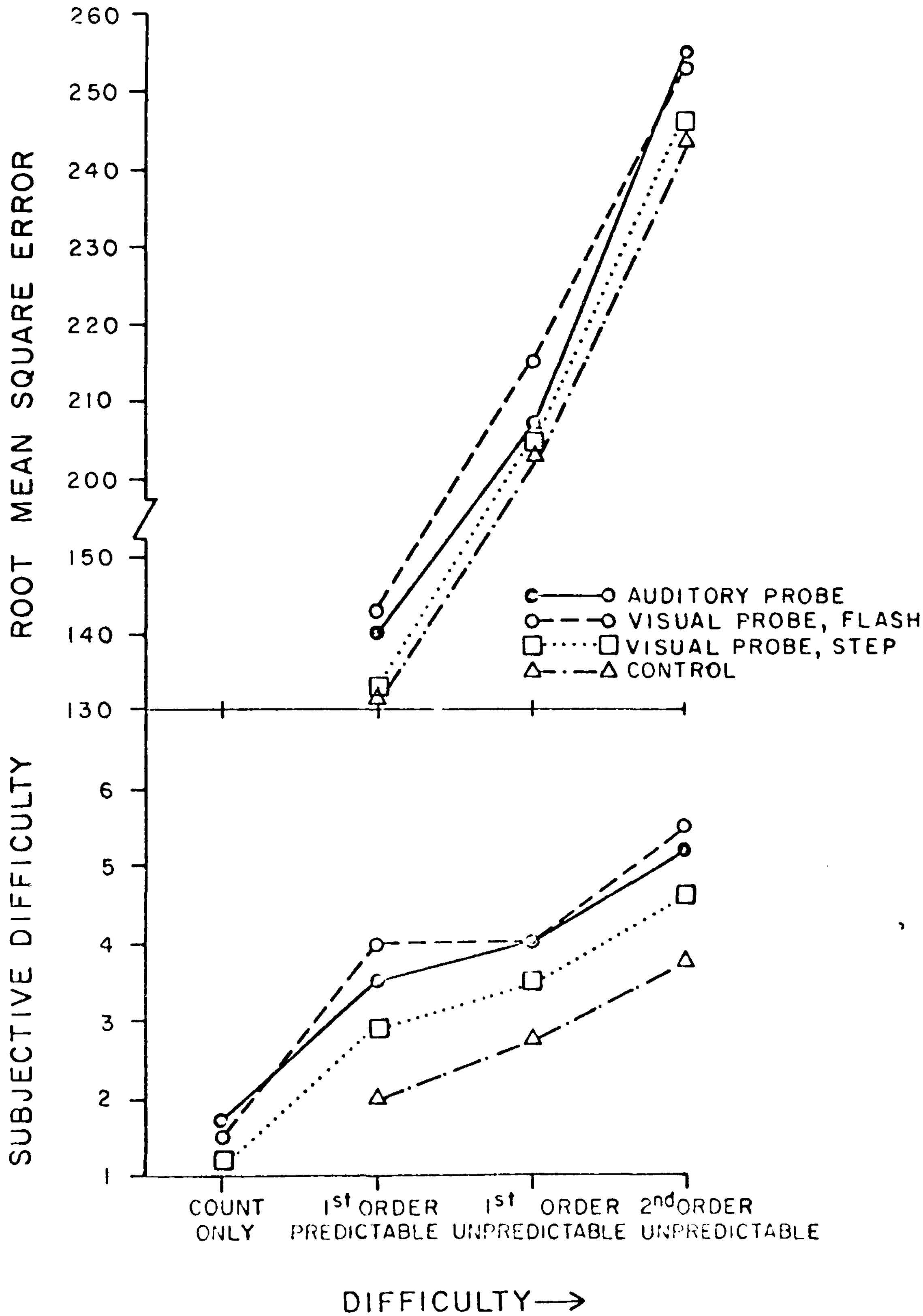
(26) This research was supported by the Air Force Office of Scientific Research under contract F49620-79-C-0233 with Dr. Alfred Fregly as Technical Monitor.

Figure Captions

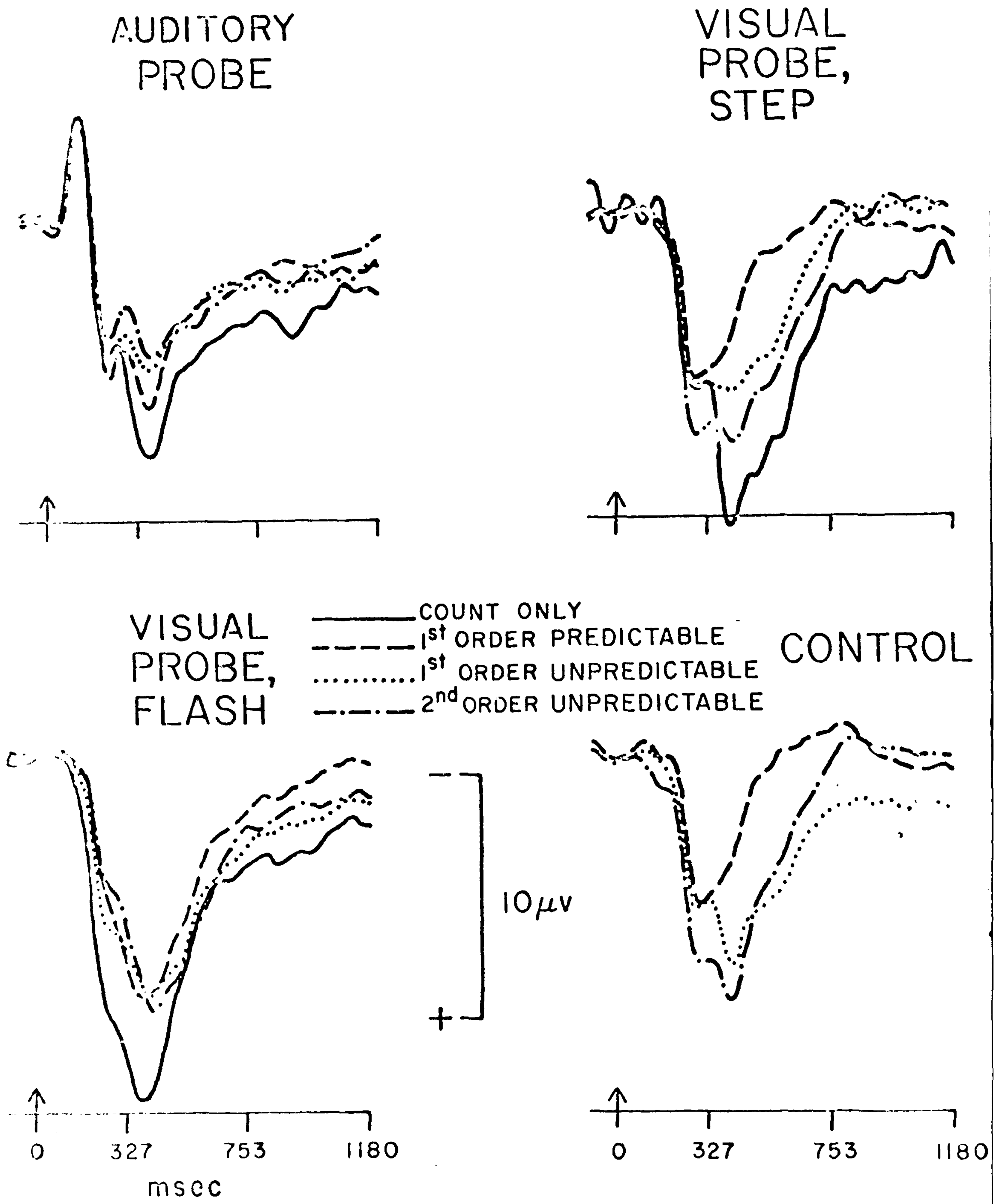
Figure 1. Average root mean square error and subjective difficulty ratings recorded for each of the experimental conditions.

Figure 2. Average parietal ERPs elicited by visual, auditory and spatial probes presented concurrently with the pursuit step tracking task at each level of difficulty.

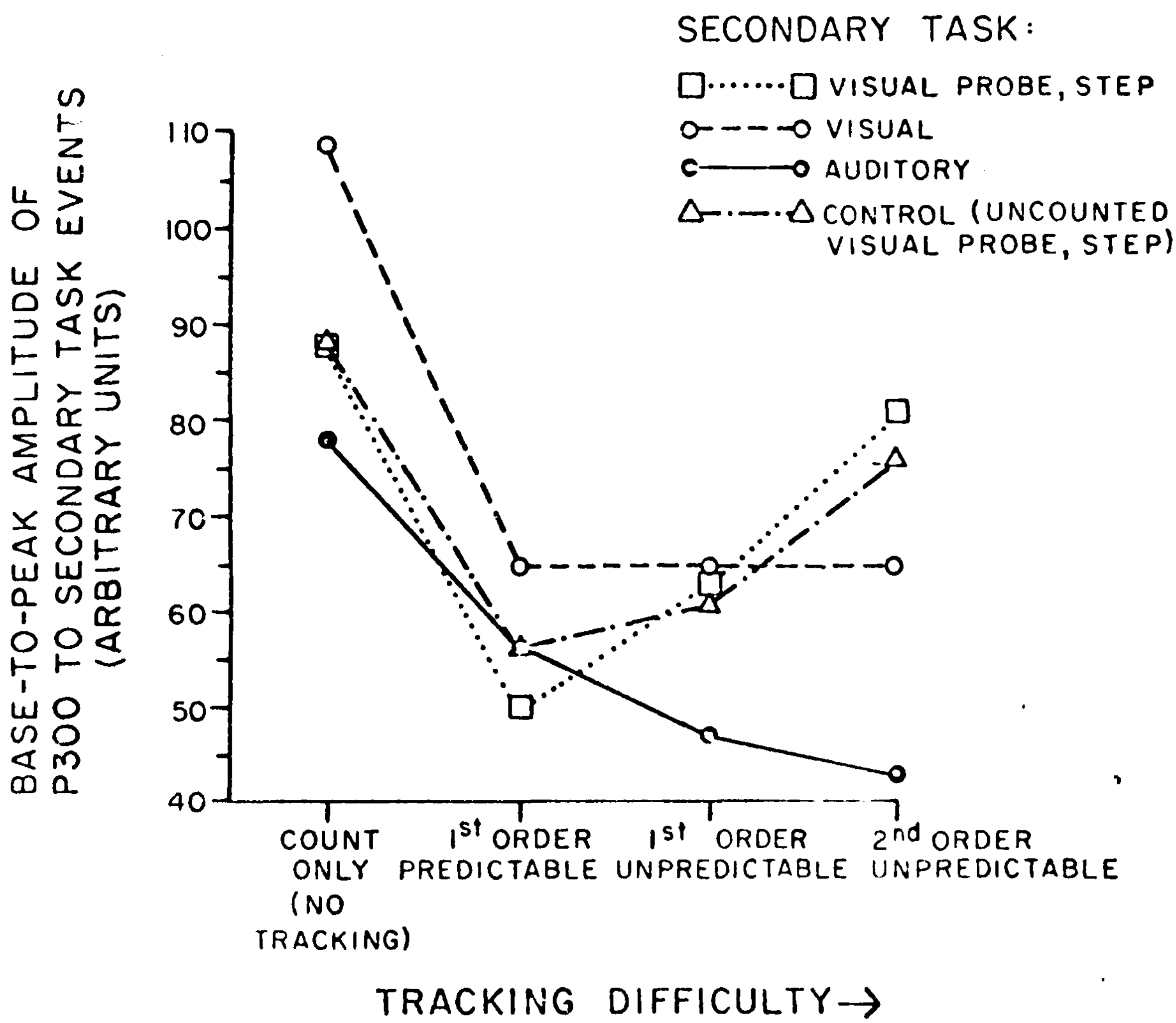
Figure 3. Normalized base to peak measure of P300 amplitude for each of the experimental conditions.



ROOT MEAN SQUARE ERROR
260
250
240
230
220
210
200
150
140
130
AUDITORY PROBE
VISUAL PROBE, FLASH
VISUAL PROBE, STEP
CONTROL
SUBJECTIVE DIFFICULTY
6
5
4
3
2
1
COUNT ONLY
1st ORDER PREDICTABLE
1st ORDER UNPREDICTABLE
2nd ORDER UNPREDICTABLE
DIFFICULTY⟶


AUDITORY PROBE
VISUAL PROBE, STEP
VISUAL PROBE, FLASH
CONTROL
COUNT ONLY
1st ORDER PREDICTABLE
1st ORDER UNPREDICTABLE
2nd ORDER UNPREDICTABLE
10 µv
+
0
327
753
1180
msec

STEP TRACKING EXPERIMENTS



P300 and Memory: Individual Differences in the von Restorff Effect

Monica Fabiani, Demetrios Karis,
and Emanuel Donchin
University of Illinois at Urbana-Champaign

We tested the hypothesis that the amplitude of the P300 component elicited by a stimulus will be positively correlated with the probability that the stimulus will be recalled in a subsequent test. This hypothesis was derived from an interpretation of the P300 as a manifestation of processes invoked when schemas (in working memory) require revision by the occurrence of unexpected events. We used the well-known von Restorff paradigm in which a series of items that must be memorized includes an item which is distinctly different from the rest. These "isolated" items are generally better recalled. Such isolates should elicit a P300 component. We predicted that the larger the P300 elicited by an isolated item the more likely it is to be recalled.

Twelve females were presented with 40 lists of 15 words in each of two sessions. A word was presented every 2 seconds. Seven seconds after the last word in a list was presented the subject wrote down as many words as she could remember. In 30 of the lists one word (in positions 6 thru 10) was displayed in a larger, or smaller, type. The ERPs elicited by each of the words was recorded from Fz,Cz and Pz.

Subjects differed in the extent to which they displayed the von Restorff effect. Those who recalled the isolates better than the corresponding control words also showed, among the isolated words, a strong relationship between the amplitude of P300 and recall (i.e., the recalled isolates elicited a larger P300 on their initial presentation). Other subjects recalled the isolates as well as they recalled other words. In these cases an equally large P300 was elicited by both recalled and unrecalled isolates. Debriefings revealed that subjects differed in their mnemonic strategies. "Good" von Restorff subjects (greater recall of isolates) employed rote rehearsal strategies, while "poor" von Restorff subjects tended to use complex, elaborative strategies.

The data suggest that the process manifested by P300 is indeed activated whenever a novel attribute requires registration in working memory, and this activation facilitates recall when complex strategies (which involve extensive post-stimulus processing) are not used.

Presented at annual meeting of Society for Psychophysiological Research, October, 1982.

"P300" and Memory: Individual Differences in the von Restorff Effect

Monica Fabiani, Demetrios Karis, & Emanuel Donchin
Cognitive Psychophysiology Laboratory
University of Illinois, Champaign, Illinois

The Society for Psychophysiological Research
Minneapolis, October 21-24, 1982

INTRODUCTION

We have argued previously that the amplitude of the P300 component of the human event-related brain potential (ERP) is a manifestation of processes that are involved when memory is updated (Donchin, 1981; Karis, Bashore, Fabiani, & Donchin, 1982). If the amplitude of P300 is proportional to this updating process, then changes in P300 amplitude should be related to encoding and storage processes which affect recall. We tested this hypothesis in a von Restorff paradigm by recording ERPs elicited when words were presented during a study period. We then compared ERPs elicited by words which were later recalled with ERPs elicited by words which were not recalled. We predicted that the larger the P300 elicited by a word the greater the likelihood of subsequent recall, and recognition.

Von Restorff has shown that "isolated" items are better recalled (von Restorff, 1933). A word is "isolated" when it is distinctly different from the other items in a list (e.g., because of color, size, meaning, or class). As the "isolated" words are both novel and task relevant they should elicit a large P300 (Donchin, Ritter, and McCallum, 1978). Furthermore, we predict that the variance in the amplitude of the P300 elicited by the words will be related to the probability the words will be recalled in a subsequent test.

SUBJECTS: Twelve right handed female undergraduates were run in two sessions about a week apart. They were paid $3.00 per hour, with a $5.00 bonus when they completed the second session.

WORD LISTS: To form the word lists a computer program randomly picked words with 3 to 6 letters from a master list derived from Toglia and Battig (1978). Each word was presented no more than once to each subject.

ERP RECORDING: EEG was recorded from Fz,Cz, and Pz (referred to linked mastoids) using a 10 second time constant and an upper half amplitude cutoff of 35 Hz. EOG was recorded using a 1 second time constant and an upper half-amplitude cutoff of 15 Hz. EEG and EOG were digitized at 100 samples/sec for 128 points beginning 100 msec prior to the presentation of a word. Eye movement artifacts were corrected off-line using a procedure described in Gratton, Coles, and Donchin (submitted for publication).

PROCEDURE

The experimental sequence in depicted in Figure 1. The major task involved free recall (Figure 1A). Forty lists of 15 words were presented to the subject during each session. Thirty out of the 40 lists contained one "isolated" word that appeared at random in position 6 through 10. In these experimental lists isolation was produced by displaying the word using either larger, or smaller, characters than were used for the other words. The other 10 control lists, interspersed randomly with the experimental lists, did not include an isolated word. That is, all words in the control list were displayed with the same sized characters. The subject's task was to recall as many words from a list as she could, writing them down 7 seconds after the list ended.

After the 40 lists there was an "oddball" task, followed by grand recall and a recognition task (see Figure 1 B, C, & D). At the end of each session the subjects were asked about the strategies they used to remember the words.

FREE RECALL: We computed two indices to summarize the subjects' performance on the free recall task: A measure of the von Restorff effect (von Restorff index, or VRI) and an index of overall performance (performance index, PI).

VRI = Percentage of isolated words recalled (position 6-10) minus
the percentage of non-isolated words recalled (position 6-10)

PI = Overall percentage of words recalled from
all positions (isolates and non-isolates)

There were no systematic differences between the control and experimental lists in the recall of non-isolated words, and both classes of lists were used to compute the VRI.

The von Restorff index is plotted against performance (PI) in Figure 2. There is a high negative correlation ($r = -.837$, $p < .05$) between overall recall performance and the von Restorff index, and subjects cluster into three groups: low general recall (PI) and high von Restorff index (VRI) (Group 1, N = 3), high PI and low VRI (Group 3, N = 3), and those intermediate on both indices (Group 2, N = 6). Thus, it appears that in subjects who recall more words in general, the effect of the size of the font on recall is relatively small. Subjects who are relatively poor in recall, however, tend to benefit more from the isolation effect.

Figure 3 shows serial position curves for the three clusters, or groups, defined in Figure 1. The magnitude of the von Restorff effect is represented by the elevation in recall of the isolated items (the red triangles) relative to the rest of the curve. Clearly, the von Restorff effect is largest in Group 1, moderate in Group 2, and absent in Group 3.

In Figure 4 we have superimposed the serial position curves for the three groups, with the isolated items (which were presented in position 6 through 10) presented separately on the bottom. The performance difference among groups is clearly visible (top figure), while there is no significant difference in the recall of isolates (bottom figure).

STRATEGIES: The strategies reported by subjects after each session were judged blindly by nine undergraduates. They ranked the strategies from the most simple (rote memorization) to the most complex (elaborative strategies). Interjudge reliability (Cronbach's alpha) was .98, and the correlation between VRI and mean rank was -.57 ($p < .05$). Subjects who used simple strategies tended to be in Group 1, while subjects who used complex strategies tended to be in Group 3.

Examples

Rote Strategy, Group 1:	" ...I repeated the words in a row. I also tried to repeat each word three times."
Complex Strategy, Group 3:	" ...I tried to connect words into a story or a picture. I tried to make the story or picture ridiculous."

ERP RESULTS: ERPs elicited by words in the free recall were sorted by word type (isolates, non-isolates in the experimental lists, and control words), by position (position 6-10, other positions), and by subsequent recall (recalled, not recalled). Our main interest is in examining differences between ERPs elicited by words subsequently recalled vs. those not recalled, especially within the class of isolates.

The twelve individual averages for isolates later recalled vs. not recalled (at Pz) are presented in Figure 5. In Figure 6 ERPs elicited by the isolates are presented again, but are averaged by group, and waveforms for each electrode site are presented. In Figure 7 group averages for all three word types are presented: isolates, non-isolates from the experimental lists, and control words. All were presented in position 6 to 10. In all three figures each average is further divided into words subsequently recalled and those not recalled.

From the waveforms it is clear that the largest effect of recall is for isolated words in Group 1, although there is a small effect of recall in all groups for the non-isolated and control words.

ERP ANALYSIS: A Principal Component Analysis (PCA) was performed on all the EEG records of words in position 6 through 10. Four components (explaining 92% of the variance) were rotated using a varimax rotation procedure. From latency and scalp distribution we labeled component 1 "P300". Component 2 is more positive frontally than centrally and

In Figure 4 we have superimposed the serial position curves for the three groups, with the isolated items (which were presented in position 6 through 10) presented separately on the bottom. The performance difference among groups is clearly visible (top figure), while there is no significant difference in the recall of isolates (bottom figure).

STRATEGIES: The strategies reported by subjects after each session were judged blindly by nine undergraduates. They ranked the strategies from the most simple (rote memorization) to the most complex (elaborative strategies). Interjudge reliability (Cronbach's alpha) was .98, and the correlation between VRI and mean rank was -.57 ($p < .05$). Subjects who used simple strategies tended to be in Group 1, while subjects who used complex strategies tended to be in Group 3.

Examples

Rote Strategy, Group 1:	" ...I repeated the words in a row. I also tried to repeat each word three times."
Complex Strategy, Group 3:	" ...I tried to connect words into a story or a picture. I tried to make the story or picture ridiculous."

ERP RESULTS: ERPs elicited by words in the free recall were sorted by word type (isolates, non-isolates in the experimental lists, and control words), by position (position 6-10, other positions), and by subsequent recall (recalled, not recalled). Our main interest is in examining differences between ERPs elicited by words subsequently recalled vs. those not recalled, especially within the class of isolates.

The twelve individual averages for isolates later recalled vs. not recalled (at Pz) are presented in Figure 5. In Figure 6 ERPs elicited by the isolates are presented again, but are averaged by group, and waveforms for each electrode site are presented. In Figure 7 group averages for all three word types are presented: isolates, non-isolates from the experimental lists, and control words. All were presented in position 6 to 10. In all three figures each average is further divided into words subsequently recalled and those not recalled.

From the waveforms it is clear that the largest effect of recall is for isolated words in Group 1, although there is a small effect of recall in all groups for the non-isolated and control words.

ERP ANALYSIS: A Principal Component Analysis (PCA) was performed on all the EEG records of words in position 6 through 10. Four components (explaining 92% of the variance) were rotated using a varimax rotation procedure. From latency and scalp distribution we labeled component 1 "P300". Component 2 is more positive frontally than centrally and parietally, appears late (540 msec) and slowly increases until the end of the epoch. We labeled it "frontal-positive slow wave", to distinguish it from the frontal negative slow wave usually found.

An analysis of variance (ANOVA) with repeated measures and unequal ns was performed on the PCA component scores. Of greatest interest were the following significant differences ($p < .05$):

For Component 1 (P300):

1. Isolated words show a larger P300 than control and experimental words.
2. Words that were recalled (regardless of type) elicited larger P300s than words not recalled.
3. Group 1 (poor memorizers with a high von Restorff index) shows a larger P300 for words recalled than not recalled. The amplitude difference is smaller for Group 2 and virtually absent for Group 3.
4. Group 1 shows, at Pz, a larger P300 for the isolated recalled than not recalled than the other two groups.

For Component 2 (frontal-positive slow wave):

1. Isolated words show a larger frontal-positive slow wave than control and experimental words.
2. Words that were recalled (regardless of type) elicited a larger frontal-positive slow wave than words not recalled.
3. Group 3 (good memorizers with a low von Restorff index) show more evidence of this component than the other two groups.
4. Group 3 shows, at Fz, a larger frontal-positive slow wave for the isolates recalled than not recalled than the other two groups.

SUMMARY OF ANOVA: P300 is sensitive to both isolation and probability of recall. The effect of memory (P300 to words recalled larger than to words not recalled) is greater in Group 1 than in the other two groups. It is interesting to note that the frontal-positive slow wave component is sensitive to isolation and memory as well, but that the larger amplitude of this component for isolates recalled vs. not recalled is more evident in subjects of Group 3.

GRAND RECALL: ERPs recorded to words presented during free recall were reaveraged on the basis of grand recall performance into three groups (for each word type, and for position 6-10): 1. words never recalled; 2. words recalled during the first free recall but not in the grand recall; and 3. words recalled in both the first free recall and the grand recall. Another PCA was performed and an ANOVA revealed a significant memory effect for component 1 (P300): the largest amplitude P300 belongs to words recalled in both the free recall and the grand recall, while the smallest belongs to words never recalled (see Figure 8).

When ERPs to isolates were sorted, taking recognition performance into account (Figure 9), the graded effect of memory remained (although no statistical analysis was performed due to the small number of trials).

The P300 to the isolates, in order of increasing amplitude, was as follows:

1. Neither recognized nor recalled
2. Recognized but not recalled
3. Recognized, recalled in the first free recall, but not in the grand recall
4. Recognized and recalled in both free recalls

DISCUSSION

A coherent picture emerges from the recall data, the ERPs, and the strategies subjects report. Subjects who use simple rote strategies do not recall many words, but isolating a word improves the likelihood of its recall. For these subjects P300 predicts recall (particularly for isolates), because recall depends primarily on the quality and nature of encoding processes performed on the initial attributes of the words. Subjects who use complex strategies recall many more words, and isolation does not aid recall. For these subjects P300 amplitude is not a good predictor of recall, because retrieval depends on very effective organizational processes that proceed well beyond the coding of surface attributes of the word. However, the frontal-positive slow wave does predict recall for these subjects, particularly for the isolates. This component starts very late and continues long after P300.

Note that isolated items elicited large P300s in all subjects. There is variance in the process manifested by P300, and this variance is related to recall. However, this relationship can be overshadowed, as in our Group 3, by other memory processes occurring after P300. It is possible that the frontal-positive slow wave may represent the initiation of these other processes.

CONCLUSION

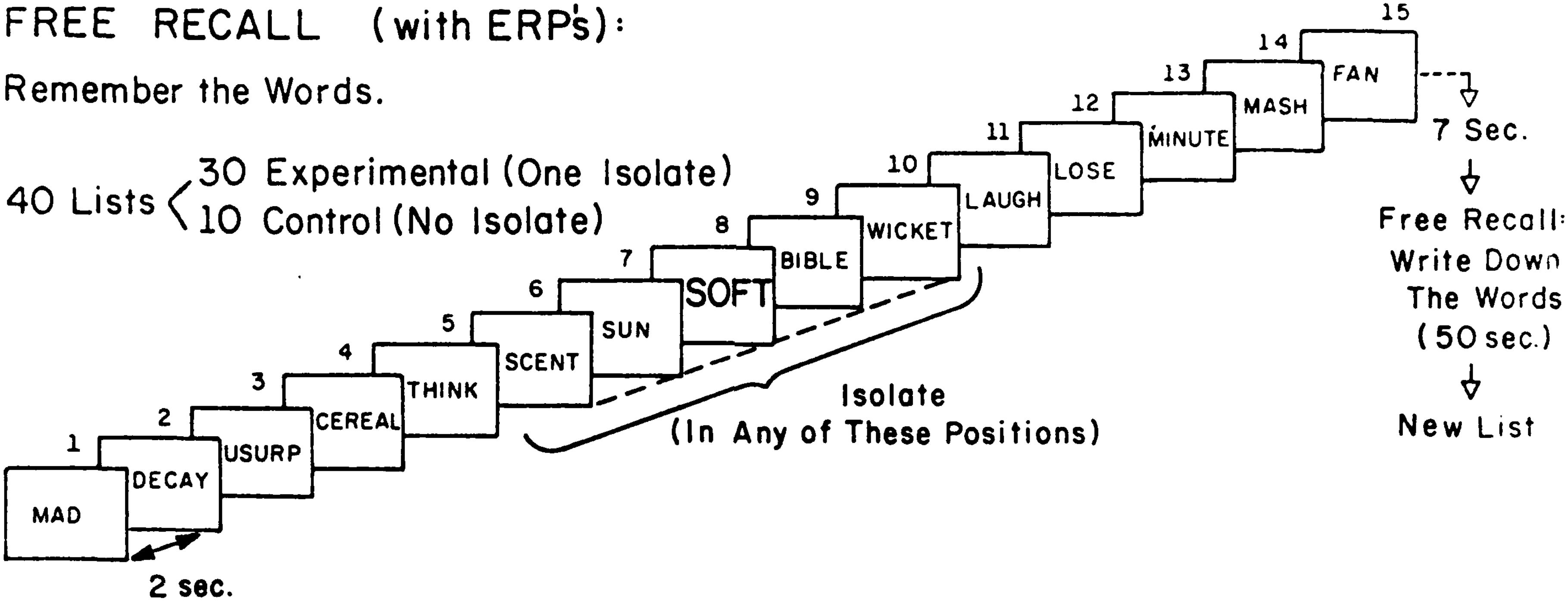
P300 is a manifestation of processing involved when we update, or modify, schemata related to environmental events. Updating involves the "tuning" of a schemata, and this will result in a more accurate representation of the environment, and improve memory. If P300 is proportional to this updating process, then the P300 elicited by an event should be related to the future recall or recognition of that event. We have demonstrated this relationship in a von Restorff experiment and, by examining individual differences, have clarified the circumstances under which it will emerge.

Figure 1

Ⓐ FREE RECALL (with ERP's):

Remember the Words.

40 Lists: 30 Experimental (One Isolate); 10 Control (No Isolate)



Ⓑ "ODDBALL" (with ERP's)

Count the Large (or Small) Words. Counted Words Are Rare, p = .20.

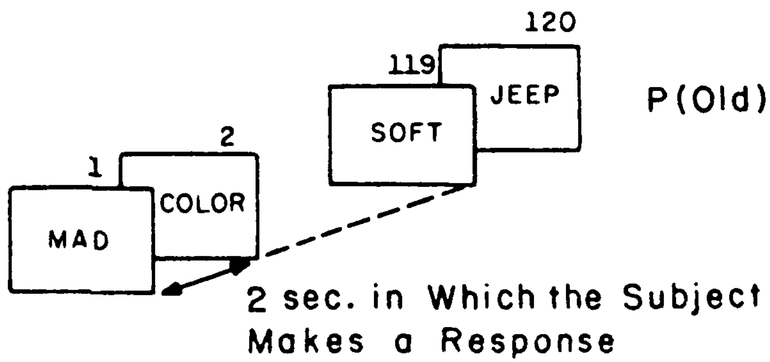
Ⓒ GRAND RECALL (No ERP's)

Write the Words You Remember From all 40 Word Lists, (10 min.)

Ⓓ RECOGNITION (with ERP's)

Respond: Right - Old Words
Left - New Words

120 Words: 30 Isolates, 30 Non-Isolates (Old Words); 60 New Words



P(Old) = P(New) = .50

Figure 2

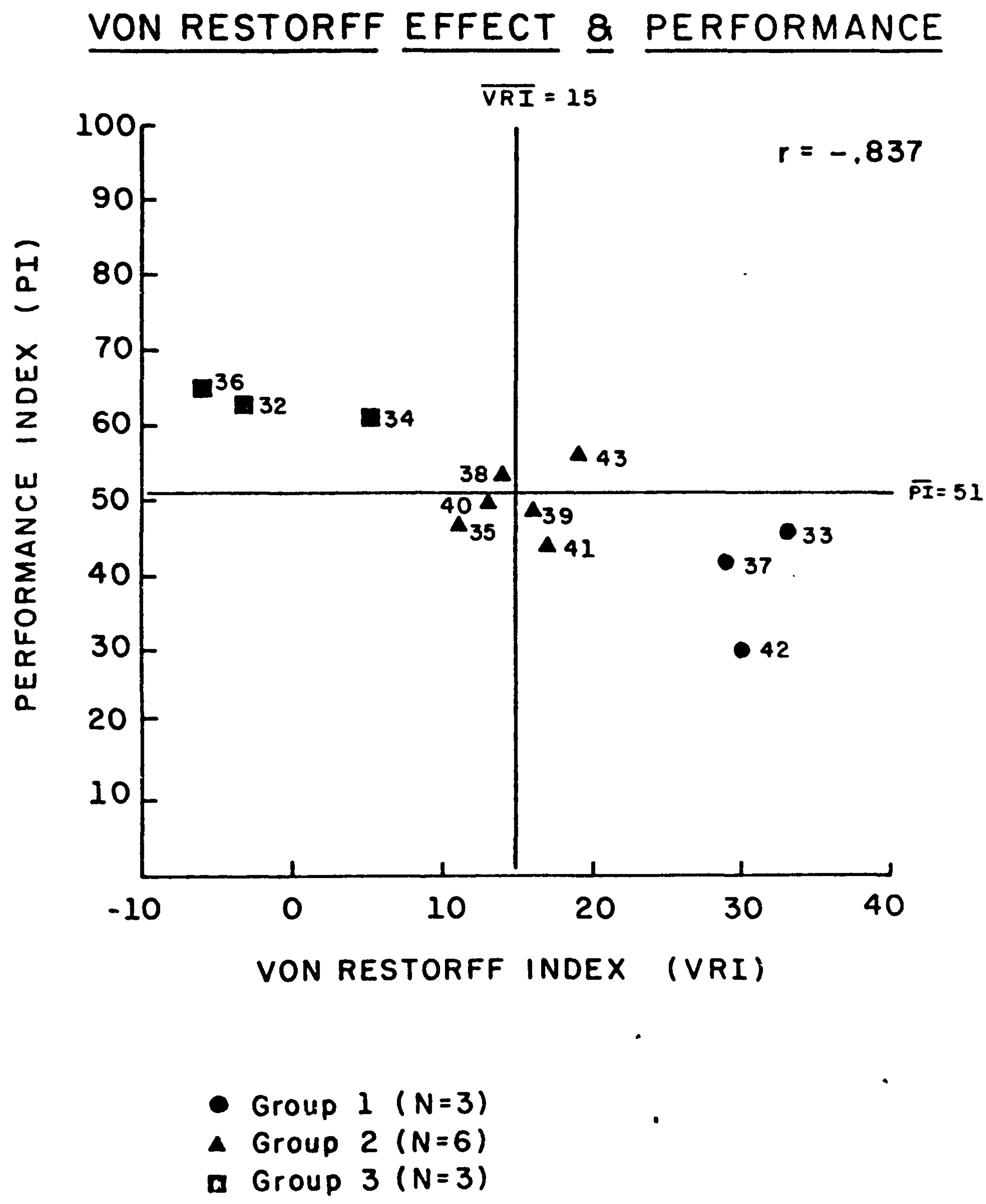

VON RESTORFF EFFECT & PERFORMANCE
$\overline{VRI} = 15$
r = −.837
$\overline{PI} = 51$
100
90
80
70
60
50
40
30
20
10
PERFORMANCE INDEX (PI)
-10
0
10
20
30
40
VON RESTORFF INDEX (VRI)
36
32
34
38
43
40
39
35
41
33
37
42
Group 1 (N=3)
Group 2 (N=6)
Group 3 (N=3)

Figure 3

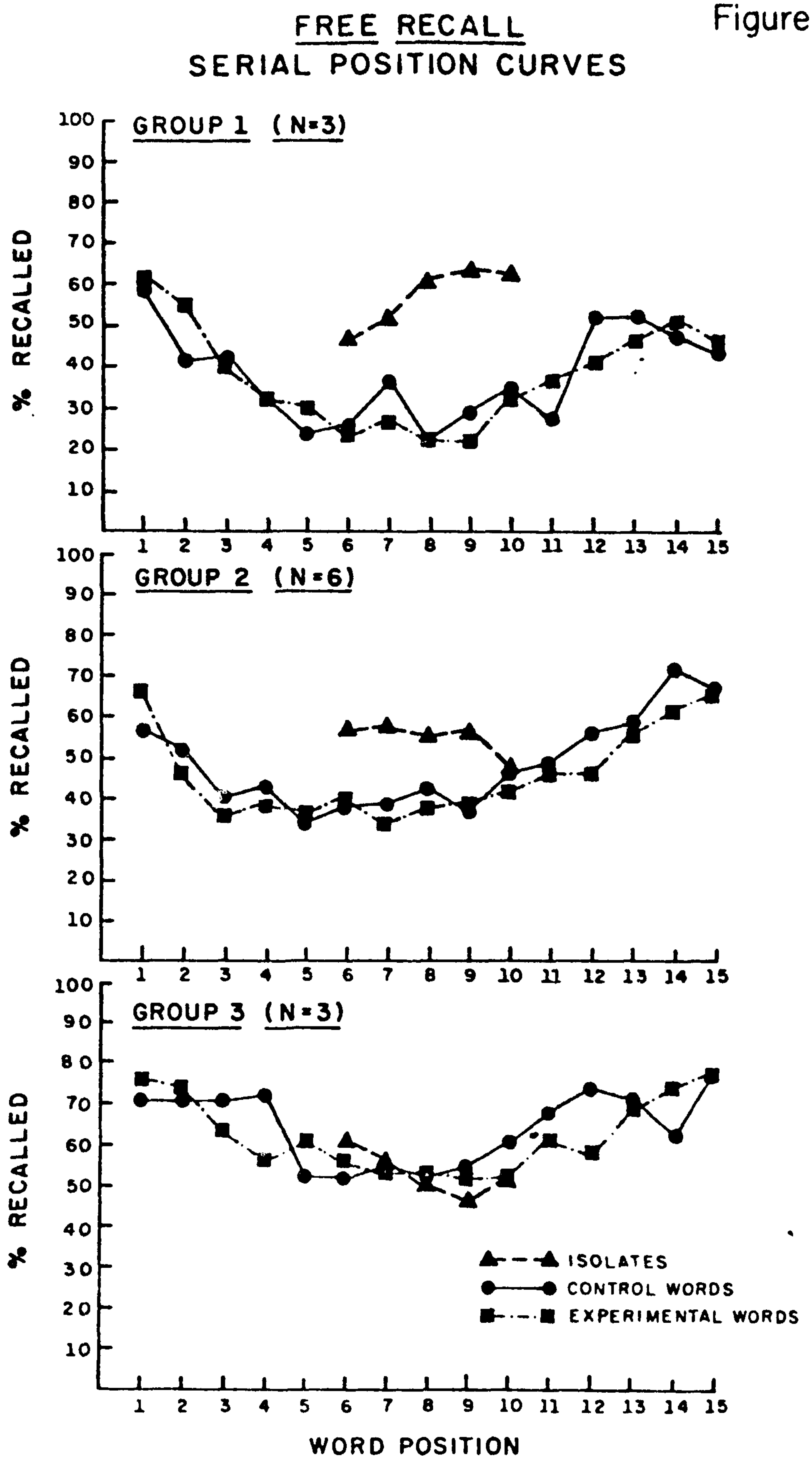
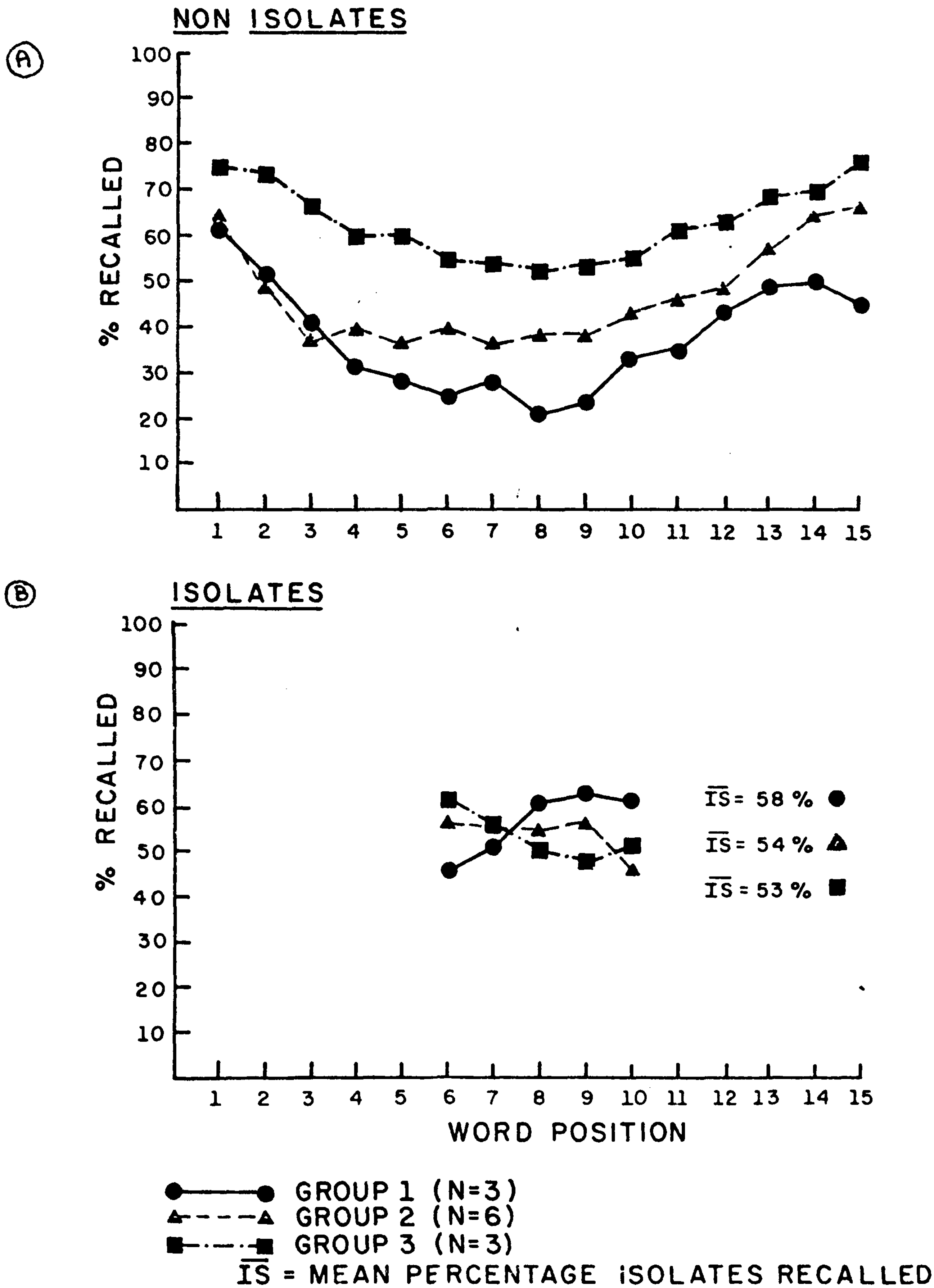

FREE RECALL
SERIAL POSITION CURVES
GROUP 1 (N=3)
GROUP 2 (N=6)
GROUP 3 (N=3)
% RECALLED
WORD POSITION
ISOLATES
CONTROL WORDS
EXPERIMENTAL WORDS

Figure 4

FREE RECALL
SERIAL POSITION CURVES



FREE RECALL

Figure 5

ISOLATES

(INDIVIDUAL SUBJECTS, P_Z)

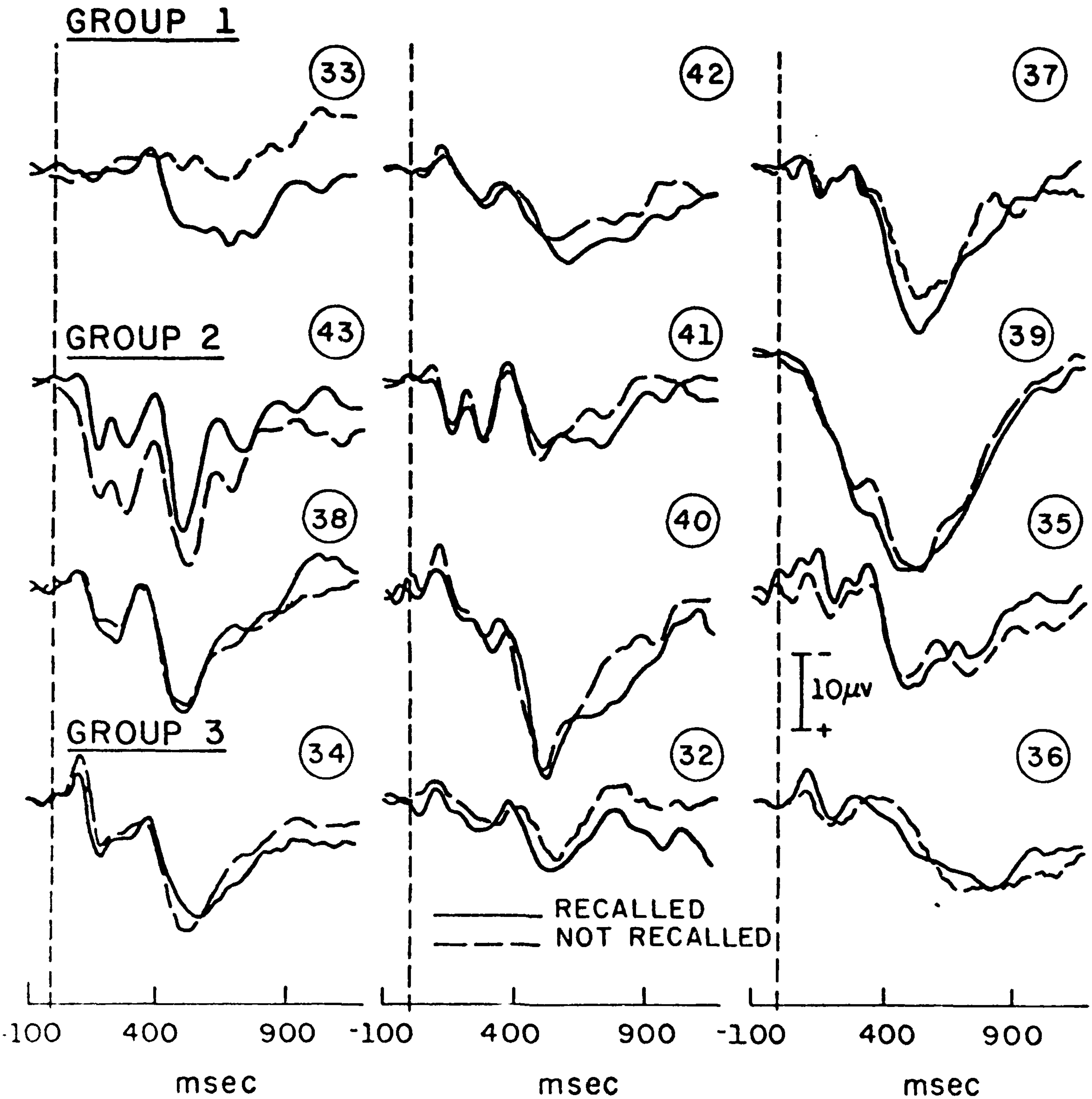


Figure 6

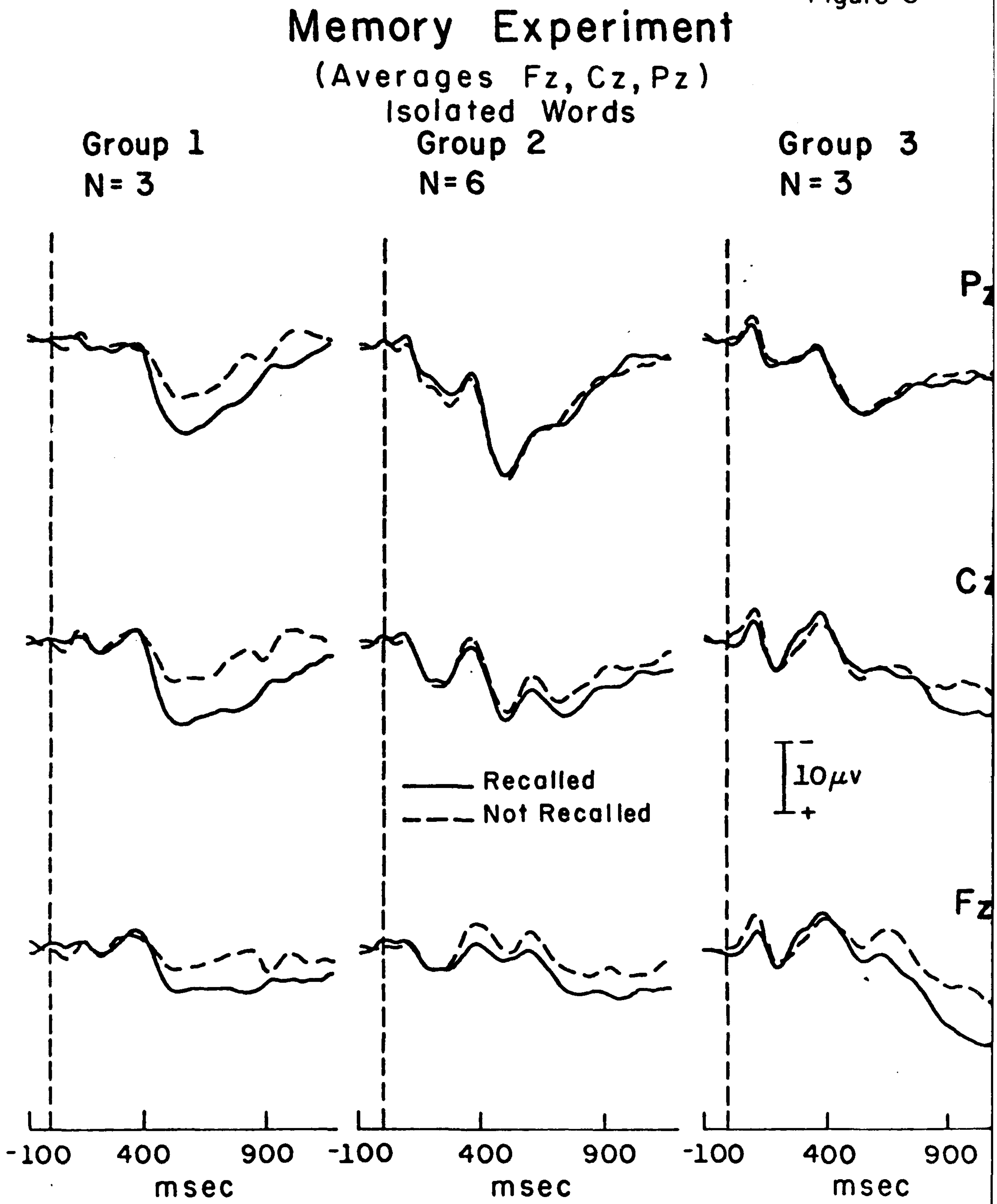
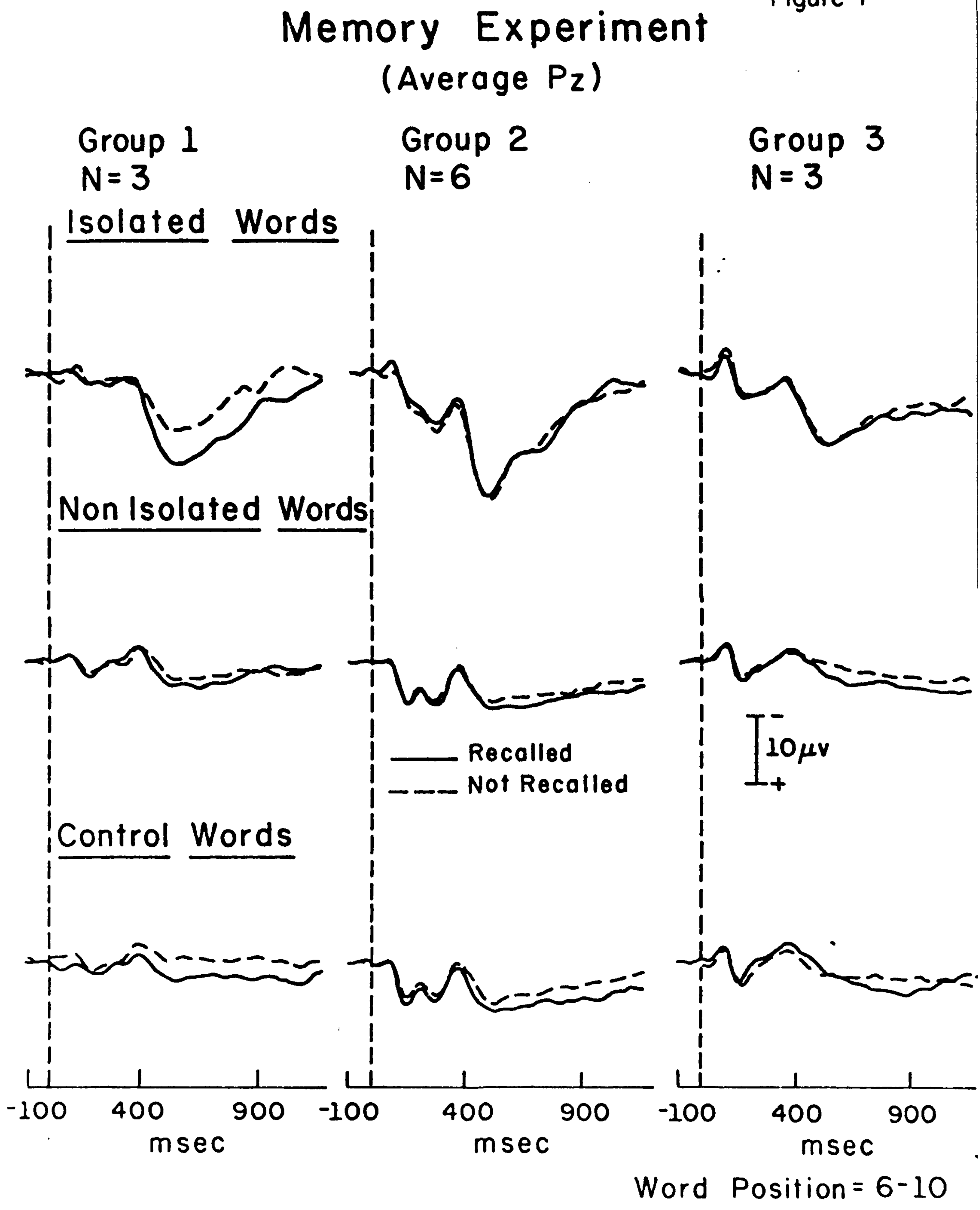
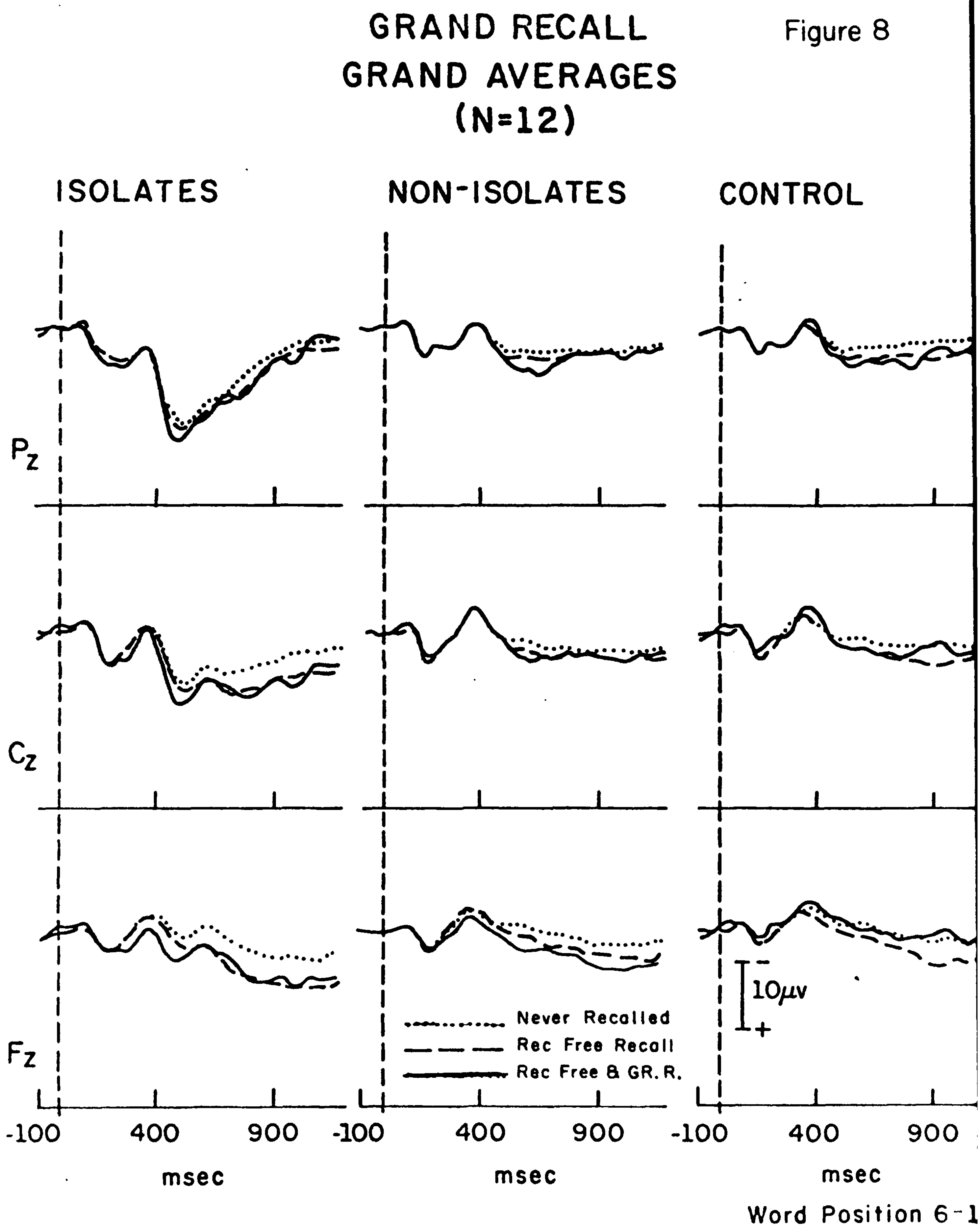

Memory Experiment
(Averages Fz, Cz, Pz)
Isolated Words
Group 1
N = 3
Group 2
N = 6
Group 3
N = 3
Pz
Cz
Fz
Recalled
Not Recalled
10μv
-
+
-100
400
900
msec

Figure 7



Memory Experiment
(Average Pz)
Group 1
N=3
Group 2
N=6
Group 3
N=3
Isolated Words
Non Isolated Words
Control Words
Recalled
Not Recalled
10μv
-100
400
900
msec
Word Position = 6-10

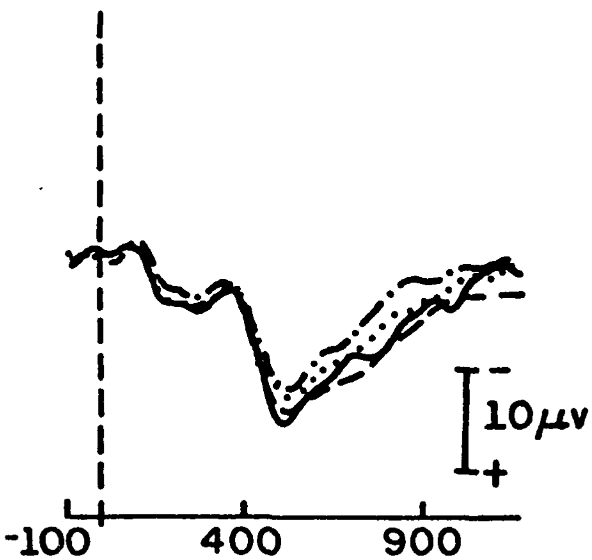
Figure 8



GRAND RECALL
GRAND AVERAGES
(N=12)
ISOLATES
NON-ISOLATES
CONTROL
P_z
C_z
F_z
10µv
Never Recalled
Rec Free Recall
Rec Free & GR. R.
-100
400
900
msec
Word Position 6-1

Figure 9

RECOGNITION, FREE & GRAND RECALL GRAND AVERAGES-ISOLATES (N=12)



———— Recognized, and Recalled in Both Free and Grand Recall
– – – Recognized and Recalled in Free Recall Only
.... Recognized Only
—.—.— Neither Recognized Nor Recalled

N200 Amplitude as a Function of Degree of Mismatch in a Word Categorization Paradigm

Walter S. Pritchard, Michael G.H. Coles, and Emanuel Donchin

University of Illinois at Urbana-Champaign

Research indicates that the amplitude of the N200 component of the ERP appears to be a function of the degree to which a stimulus mismatches a subject's ongoing information-processing set. However, experimental manipulations that increase stimulus mismatch usually also increase the latency of the subsequent P300 component. As a result of these latency shifts, N200 may become more visible and its amplitude may appear to increase. The present study employed a word categorization paradigm which allowed the degree of stimulus mismatch to be increased without a corresponding increase in P300 latency. Thus, in this paradigm, increases in the amplitude of N200 cannot be attributed to the fact that possible obscurring effects of P200 or P300 are being attenuated.

In two experiments, N200 amplitude on later trial blocks was found to increase as a function of degree of semantic mismatch. This result held true regardless of whether degree of mismatch was varied within each block (Experiment 1) or between counterbalanced blocks (Experiment 2). Further, this increase in N200 amplitude was not the result of increased P300 latencies. For negative exemplars in both experiments, an inverse relationship existed between P300 latency and degree of mismatch.

The relation between N200 amplitude and degree of mismatch was not observed in the early trial blocks of both experiments. A model attributing this change across blocks to different subject categorization strategies will be discussed. In general, the data of the present experiments are consistent with the view that N200 is a consequence of the detection of a mismatch between an event and a subject's information-processing set.

Presented at annual meeting of Society for Psychophysiological Research, October 1982.

P300 Latency and Reaction Time From a Visual Search Task with Varying Levels of Noise and S-R Compatibility

Anthony Magliero, Theodore Bashore, Michael G.H. Coles and Emanuel Donchin

Cognitive Psychophysiology Laboratory
University of Illinois at Urbana-Champaign

Previous studies (e.g. McCarthy and Donchin, 1981) have concluded that P300 is related to stimulus evaluation, but not to response selection. The present experiment tested this hypothesis by manipulating discriminability and stimulus-response compatibility.

Target words, "RIGHT" or "LEFT", were embedded in a matrix of noise characters in a visual display. Discriminability was manipulated by varying the noise characters. On "no noise" trials the background was filled with the # symbol. On "noise" trials, the background characters were all As, or characters from the letter sets A to D, A to G, or A to Z.

Response compatibility was varied by presenting a cue word ("same" or "opposite") before target presentation. The cue indicated whether the hand specified by the target word, or the opposite hand, was to be used.

Ten subjects were run for four days, with a different noise set used on each day. On each session, each trial could be, with equal probability, a "noise" or a "no noise" trial. The compatible and incompatible responses were also called for with equal probability.

RT to the noise trials increased with increasing noise set size. Incompatible responses were given with a longer RT than were compatible responses.

P300 latencies were derived from each subject's average ERP. P300 latency increased with increasing size of the noise set, but was not affected by the response requirement. The data are consistent with the report by McCarthy and Donchin (1981) who argued that P300 latency is affected by the amount of perceptual noise, but is independent of response selection and execution.

Presented at annual meeting of Society for Psychophysiological Research, October 1982.

Electrophysiology of Absolute Pitch

Mark Klein, Michael G.H. Coles,
and Emanuel Donchin

Cognitive Psychophysiology Laboratory
University of Illinois at Urbana-Champaign

Individuals who have absolute pitch (AP) can identify by name the pitch and the octave of a tone with near perfect accuracy. They appear to do so without requiring a named external reference pitch. Apparently, individuals with AP maintain a permanently stored comparison standard for tones and, therefore, they do not have to update their internal representations of the tonal inputs. The interpretation of the P300 component as a manifestation of a context updating process suggests that AP subjects will show a small P300 in response to rare tones embedded in a Bernoulli series of tones. Subjects who do not have AP will, of course, display a large P300 in response to rare stimuli. We further predict that rare visual stimuli will elicit a P300 in subjects with AP. The present study then compared the response of subjects with and without AP when challenged with auditory and visual Oddball tasks.

Four AP and four control subjects were required to count the number of times the rarer ($p=.20$) of two events occurred. The ERPs were recorded from Fz, Cz and Pz referred to linked-mastoids. Each subject was presented with 40 auditory and 40 visual series of 100 stimuli each. The auditory stimuli were 60 msec tones (10 msec rise/fall time) with frequencies 1000 and 1100 Hz and the visual stimuli were the letters H and S (visual angle 1.3 degrees) presented on a screen for 60 msec. Order of presentation of the modalities was counterbalanced. Subjects without AP displayed the usual pattern of ERPs with large P300 components elicited by the rare events regardless of modality. There was no appreciable difference between the P300 elicited by visual and by auditory Oddballs. The AP subjects showed a marked difference in the ERPs elicited by visual and auditory rare stimuli. The auditory rares elicited, as predicted, a markedly smaller P300 than the visual rares. In general, AP subjects showed a smaller P300 than the control subjects.

Presented at annual meeting of Society for Psychophysiological Research, October 1982.

Information Extraction and P300 Amplitude

A.M. Mane, C.D. Wickens, L. Vanasse
and E. Donchin

Cognitive Psychophysiology Laboratory
University of Illinois at Urbana-Champaign

Previous research on event related potentials has implicated event probability and task relevance as determinants of P300 amplitude. One interpretation of these findings views P300 as reflecting the information extracted from a stimulus. We report here an experiment in which information delivered by a stimulus is manipulated independently of stimulus probability. The effects of this manipulation on P300 amplitude and reaction time (as a manipulation check) were evaluated.

Ten subjects were presented with a sequence of three stimuli: informative, warning, or imperative. The informative stimulus was either a square or a circle in which was displayed a digit (either 0, 4 or 8). The probability of any one of the six shape by digit combinations was equal. The informative stimulus predicted the shape of a future imperative stimulus with a reliability of 50%, 70%, or 90%, depending on the digit. After a time interval, which varied between 1.5 and 2.1 secs, subject was presented with a warning signal (the letter 'X') which was followed 400 msec later by the imperative stimulus. Subjects were required to respond as quickly as possible to the imperative stimulus by pressing one of two buttons depending on shape.

Reaction time measures revealed that five of the subjects extracted the information provided by the informative stimulus. These subjects' reaction time was shorter following the more reliable informative cues. For these subjects, P300 amplitude following the informative stimuli was directly related to the predictive value of the stimuli.

Presented at annual meeting of Society for Psychophysiological Research, October 1982.

Proceedings 18th Annual Conference on Manual Control. Wright Patterson AFB 1982

OPERATOR WORKLOAD AS A FUNCTION OF THE SYSTEM STATE: AN ANALYSIS BASED UPON THE EVENT-RELATED BRAIN POTENTIAL

by
Richard T. Gill
Department of Engineering
Wright State University
Dayton, Ohio 45432

Christopher Wickens
Department of Psychology
University of Illinois
Champaign, Illinois 61801

ABSTRACT

While the processing demands of second order manual control are known to be greater than those of 1st or 0 order, the precise nature or locus of these increased demands is not well established. The purpose of this research is to determine if the demands are perceptual, related to the perception of higher derivatives of the error signal or characteristics of the system state, and thereby fluctuating with changes in these variables, or central. In the latter case, we assume the demand to be constant over time, a consequence of the increased demands of activating a more complex internal model. Event-related brain potentials -- more specifically, P300 amplitude -- were employed to assess operator workload while controlling a second order system. The ERP waveforms were categorized according to the system state at the time of the eliciting probes. Statistical analyses revealed no differences in P300 amplitude among the categories. Thus, it was concluded that the increased level of operator workload remained constant rather than fluctuating with changes in the system state. These results identify central processing rather than perception as the locus of higher order load.

INTRODUCTION

The difficulty in manually controlling higher order systems has long been recognized (1, 2, 3). The basis for this inherent difficulty may be attributed to several different sources. For example, the operator of a higher order system must act as a differentiator, generating lead in response to the perceived error signal; also there exists greater uncertainty in control because the required control input is not a one-to-one mapping of system error. Thirdly, double impulse control of second order systems requires a greater demand for accurate measurement of timing. The potential source of increased demand relevant to this research relates to the fact that the controller of higher order systems must perceive a greater number of system states and then combine them via a more complex internal model (2, 3, 4). For example, in order to optimally control a second order system, the operator must estimate both position and velocity, and then combine them via the internal model to determine the appropriate control input.

Furthermore, this increased difficulty in controlling higher order systems has also been shown, as would be expected, to result in greater operator workload (5, 6). More specifically, these investigations have suggested that the locus of this increased workload has been shown to reside in the perceptual and central processing stages of the operator and not in the response stage. The fundamental issue addressed in the present research concerns the temporal nature of this increased workload and the precise localization between perception and central processing. We hypothesize that if the source of higher workload relates to problems encountered in perceiving higher derivatives of the error signal, or particular combinations of state variables from the display, then workload should vary from moment to moment as a function of those variables. If, on the other hand, the source is attributed to the requirement to activate and maintain a more complex internal model in working memory, we anticipate that these increased demands will be more stable and constant for the duration of the tracking trial.

In this research, event related brain potentials (ERPs) were used to assess changes in operator workload. The ERP is a transient series of voltage fluctuations in the brain in response to some discrete stimulus or cognitive event. ERPs are recorded by the use of skin electrodes which are located at specific sites on the scalp. By time locking the EEG recording to a discrete eliciting stimulus event and then ensemble averaging, a very consistent pattern of positive and negative peaks or components will be observed (7). Extensive research has shown the amplitude of the P300 component to be inversely proportional to the operator's perceptual and central processing workload (5, 7, 8). P300 amplitude was used as a dependent measure of operator workload in the research described below.

EXPERIMENT

Five subjects were recruited from the university community and paid for their services. Their primary task was to minimize the RMS tracking error of a pure second order system in the presence of random noise. System error was displayed on a CRT (HP model 3010) and control inputs were made with the right hand via a spring loaded joystick. All subjects were requested to use a bang-bang or double-impulse control strategy (3) as such a strategy was essential for later analyses.

The experiment was conducted in two phases: Phase 1 - Training, and Phase 2 - Data Collection. The primary purpose of Phase 1 was to train subjects until they reached asymptotic performance. It consisted of 50 two-minute tracking trials, 25 trials per day for two consecutive days. Phase 2 was conducted on the third day during which subjects participated in 20 two-minute tracking trials. In addition to performing the primary tracking task, subjects were also required to perform a secondary task of counting discrete auditory tones or probes. The purpose of these probes was to provide the discrete event for eliciting the ERPs. They were presented, via headphones, in a Bernoulli series of low ($P = .33$) and high ($P = .67$) intensity. The subject's task was to maintain a covert mental count of the number of low intensity probes presented and report this count at the end of each block.

ANALYSIS AND RESULTS

Each ERP waveform was selectively tagged according to the momentary state (error, error velocity, control position, control velocity) at the time of stimulus presentation. This enabled us to create selective averages of ERPs (with their associated P300 component amplitudes) according to any particular characteristics of system state. Two separate analyses were performed: one based solely upon system error and error velocity, each considered independently, and the other based upon a state-space representation.

Analysis 1:

Two dichotomies were employed to categorize ERPs, one based upon system error, and one upon system velocity. All waveforms were ordered in terms of the absolute value of system error at the time of the eliciting stimulus. Then waveforms above and below the median of this ordering were separately pooled and contrasted. A similar procedure was followed with regard to system error velocity.

A separate ANOVA was performed on the component loading scores, from the PCA, for each of the dichotomies. That is, P300 amplitude (as measured by its component loading score) elicited from probes when the absolute value of system error was above its median value was compared with the P300 amplitude when system error was below its median value. Similarly, the same comparison was made for the case when absolute system velocity was above and below the median value.

For both comparisons the results were the same, with no significant difference between high and low values. That is to say operator workload (as measured by P300 amplitude) did not vary in a systematic manner with the error variables. These results are consistent with our previous findings (5).

Analysis 2:

The rationale for the second analysis is based upon the subjects' employment of a bang-bang control strategy. Such a strategy requires the operator to maintain the maximum positive control input when attempting to eliminate a given error and then at the proper instant (as specified by an analytically determined optimal switching line) reverse the input to its maximum negative value, a relatively natural strategy often employed in 2nd order control (2).

If operator workload were to fluctuate during such a task, it is clear that it would be the greatest when the system state was in the "vicinity" of the operator's empirical switching line when a decision to control is impending. That is, workload would be higher when the operator was required to: (1) perceive the magnitude of the system error; (2) perceive the magnitude of the system velocity; (3) combine them via his internal model; and (4) decide if a control reversal was required. Alternatively, the workload would be lower when the system state was such that the control input was obvious.

For example, when system error and velocity are both of the same sign, a control reversal is obviously required since error is diverging. In short, the objective of the second set of analyses was to compare operator workload when the system state was in the vicinity of the operator's empirical switching line to that when the required control input was obvious.

An algorithm was developed that analyzed the subject's continuous control input and identified the system state at the time each control reversal was initiated. The locus of these control reversals was defined to be the "vicinity" of the subject's empirical switching line. Figure 1 is a graphical representation of the system state space on which these control reversals are plotted for a typical subject, along with the theoretically optimal switching line. After reviewing such plots for each subject, it was concluded that the general locus of control reversals for each subject was sufficiently similar that the same ERP categorization algorithm could be used for all subjects.

This algorithm is illustrated graphically in Figure 2. The rationale for each area is discussed below.

Area 1. This represents the region in the vicinity of the subject's empirical switching line. The boundaries were chosen conservatively to select probes that occurred when the system state was such that the operator would be required to sense both position and velocity, and then combine them (via his internal model) to determine the proper control input. This was assumed to be the area defining a state of greatest cognitive demands.

Area 2. This region represents a conservative selection of probes when the state was such that the subject's control input was obvious; that is, when the cursor was moving away from the target line. The response in this region should be a relatively automatic reversal.

Area 3. This represents a system state similar to Area 1, but in which the velocity, relative to the position, is too low to warrant a control reversal. This is, the cursor is moving towards the target line, but not fast enough to warrant a control reversal to prevent overshoot.

Area 4. This region is comprised of the remainder of the system state space which fails to meet the requirements of any of the preceding areas.

Areas 1 and 2 were defined in a conservative manner in order to maximize the potential differences between the two, thereby increasing the sensitivity of the subsequent analyses. Area 3 was selected to provide a condition in which workload, if workload does indeed fluctuate, would lie between that of Areas 1 and 2.

This sorting algorithm was applied to the single trial ERP data of each subject. Initially, the parameters were chosen very liberally, such that Area 4 was non-existent. The resultant ERPs from the three areas were then averaged across subjects. As before, an ANOVA was performed on the component loading scores for P300; no significant differences were observed. In order to increase the sensitivity of the test, subsequent analyses were done in which the boundaries of Area 4 were systematically expanded by decreasing the parameter A_3 (see figure 2). That is, Areas 1 and 2 were defined in an increasingly conservative manner. The results remained unchanged, however, as no differences emerged with this manipulation.

CONCLUSIONS

The results of these two separate analyses were consistent. No matter whether ERP data were categorized via the magnitude of each state variable, relative to the mean, or whether the categorization was based on the location of the system state relative to the empirical switching line, differences in P300 amplitude between categorical levels failed to be observed. In fact, there were not even nonsignificant trends in this direction. These results provide strong support for the hypothesis that the increased resource demands of controlling second order systems remain relatively constant.

They are, of course, consistent with the hypothesis that it is the continuous demands of activating the complex internal model, rather than the momentary demands associated with changing perceptual variables that influence workload. There is, however, a second possibility that cannot be discounted altogether. It is possible that resource demand of control did, in fact, wax and wane, but that the "cost" associated with switching the allocation of resources to the auditory task was greater than the benefits derived. That is, it was more efficient to allocate a given amount of resources continuously to the tracking task, even though they would not all be used 100% of the time, than it was to vary the allocation of resources based on the instantaneous task demands. (It should be noted in this regard that, while P300 was consistantly attenuated, subjects did not ignore the tone counting. Their counts remained accurate.) It is impossible to determine from the current data which of the two hypotheses is correct. To do so would require running an additional condition in which greater priorities on the auditory task are stressed. When payoffs or costs thereby impell the operator to devote greater resources to the ERP task, we would now expect differential allocation (and therefore P300 amplitude) to the extent that the second hypothesis was in effect. If the first hypothesis were true, we would predict an overall depression in tracking performance and an increase in P300 amplitude, as the internal model was maintained with less fidelity. However, P300 would still fail to discriminate between system states.

ACKNOWLEDGEMENTS

This research was supported by a grant from the Air Force Office of Scientific Research Life Sciences Program F49620-79-C-0233. Dr. Alfred Fregly was the Technical Monitor.

REFERENCES

1. Birmingham, H.P., and Taylor, F.V.A. A human engineering approach to the design of man-operated continuous control systems. U.S. Naval Research Laboratory Report 4333. Washington, D.C., 1954.
2. Pew, R.W. Human perceptual-motor performance. In B. Kantowitz (Ed.), Human Information Processing: Tutorials in Performance and Cognition, Wiley and Sons, 1974.
3. Sheridan, T.B., and Ferrell, W.R. Man-Machine Systems: Information, Control, and Decision Models of Human Performance. Cambridge, MA, 1974.
4. Kelley, C. Manual and Automatic Control. NY: Wiley, 1965.
5. Wickens, C.D., Gill, R., Kramer, A., Ross, W., and Donchin, E. The cognitive demands of second order manual control: Applications of the event-related brain potential. Proceedings of the 17th Annual NASA Conference on Manual Control, NASA TM, 1981.
6. Wickens, C.D. and Derrick, W. The processing demands of higher order manual control: Application of additive factors methodology. University of Illinois, Technical Report EPL-81-1/ONR-81-1, 1981.
7. Donchin, E. Event-related brain potentials: A tool in the study of human information processing. In H. Begleiter (Ed.), Evoked Potentials and Behavior. New York: Plenum Press, 1979, pp. 13-75.
8. Donchin, E., Kramer, A., and Wickens, C. Probing the cognitive infrastructure with event-related brain potentials. Technical Report CPL82-2/AFOSR82-1, 1982.
9. Donchin, E. and Heffley, E. Multivariate analysis of event-related potential data: A tutorial review. In D. Otto (Ed.), Multidisciplinary Perspectives in Event-Related Brain Potential Research. Washington, D.C. U.S. Government Printing Office, EPA-600/9-77-043, 1979, pp. 555-572.

FIGURE TITLES

Figure 1: System state space -- typical plot of control reversals and theoretical switching line

Figure 2: Algorithm for categorizing ERP probes based on location relative to empirical switching line.

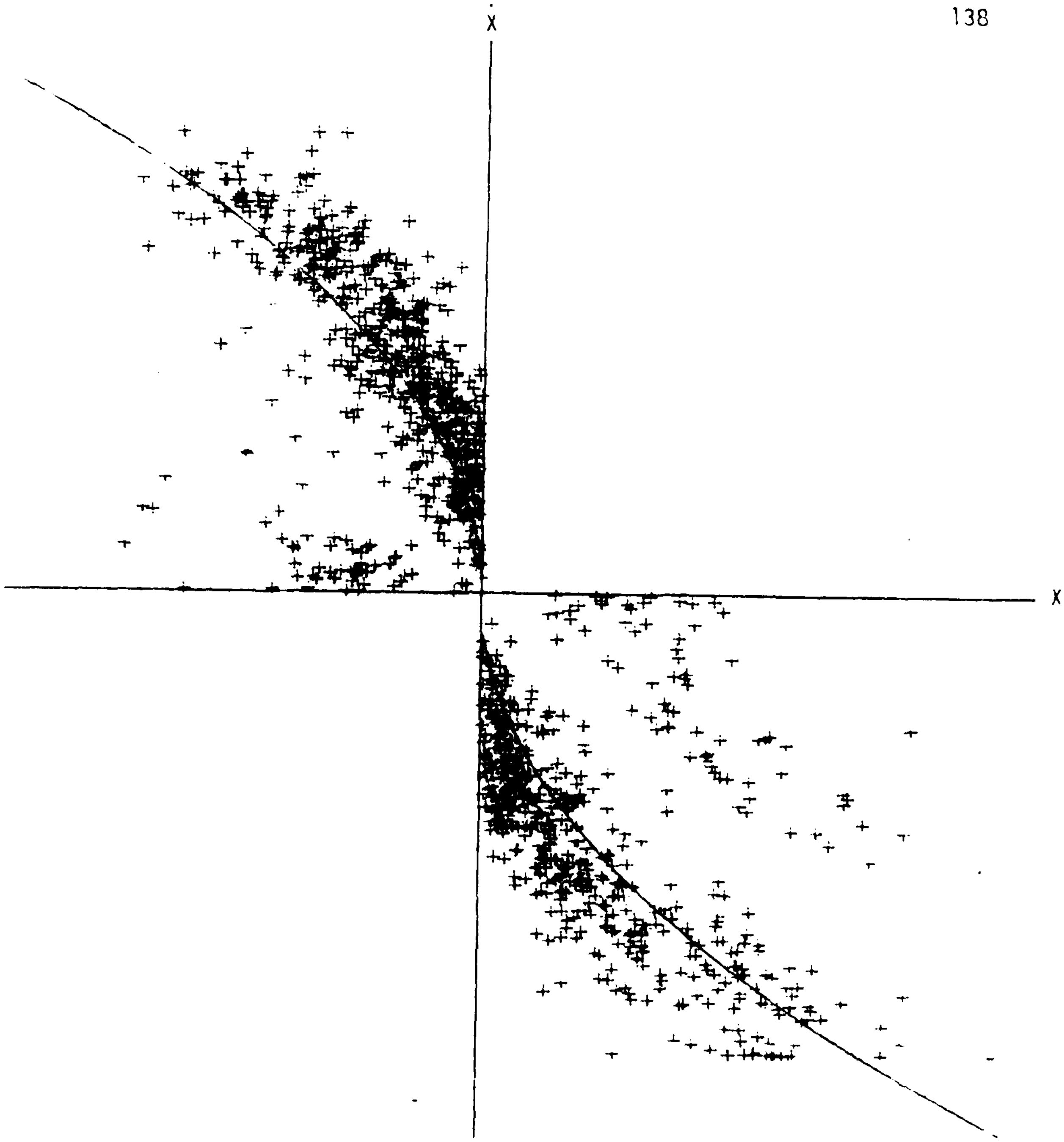


Figure 1 . Pseudo-Quickened Display - Subject #19

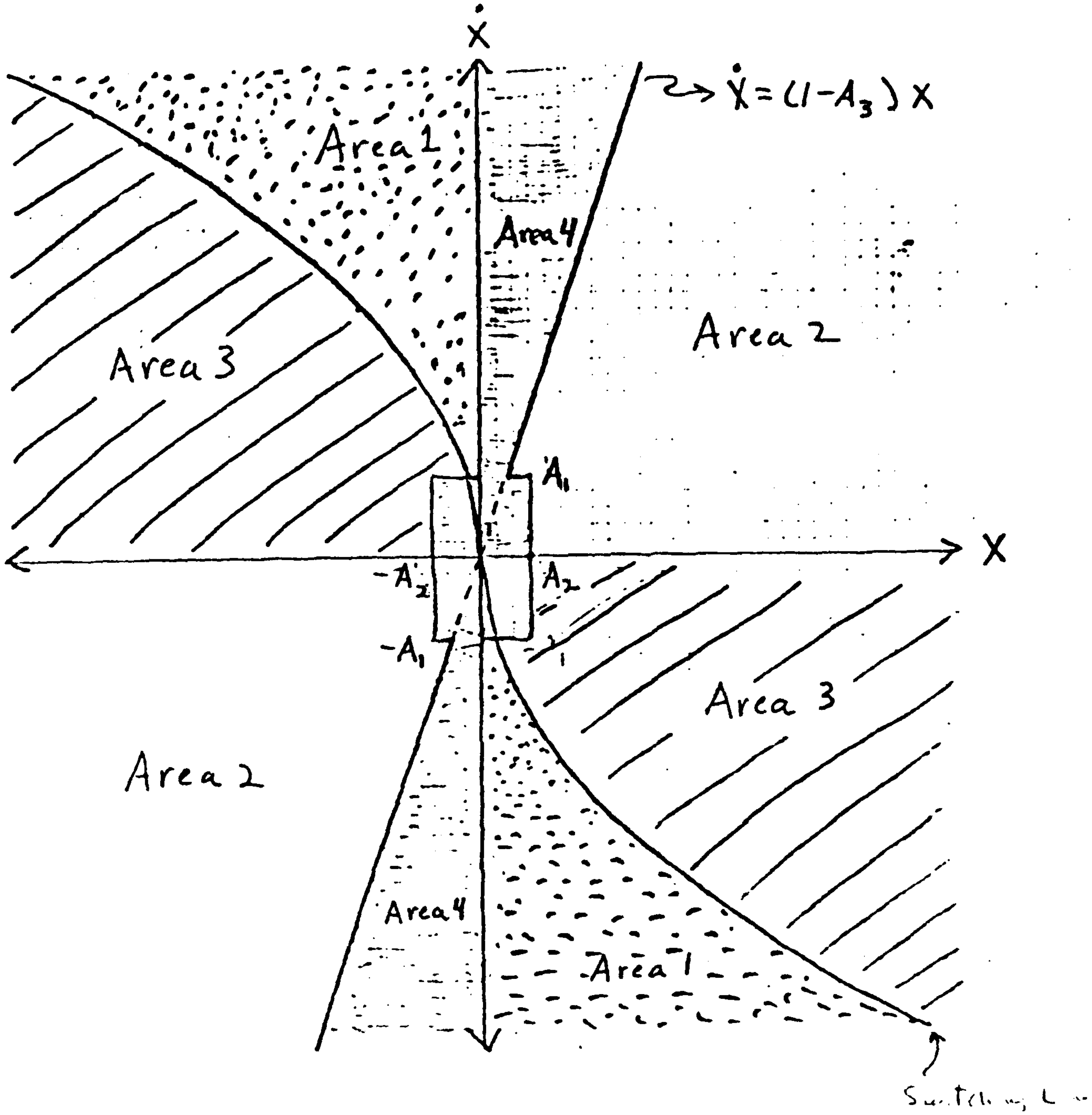


Figure 2

State-space plot of error position x velocity. Ideal switching line is shown. ERPs were averaged and compared when the error state was in the areas shown. Area 1: High demand. Area 4: Low demand.

Proceedings, 26th Annual Meeting of the Human Factors Society, October, 1982.

PSEUDO-QUICKENING: A NEW DISPLAY TECHNIQUE FOR THE CONTROL OF HIGHER ORDER SYSTEMS

Richard Gill

Wright State University
Dayton, Ohio

Christopher D. Wickens, Emanuel Donchin, & Robert E. Reid

University of Illinois at Urbana-Champaign
Champaign, IL

ABSTRACT

The concept of the pseudo-quickened display is introduced as a technique for aiding the control of higher order systems. In this display the intensity of the cursor is employed as a cue for optimal switching of manual control input. Performance on this display is compared with performance on an unaided, a quickened, and a phase plane display. Some advantages over the conventional display in performance are demonstrated by all three aided displays. When all aided groups transferred to an unaided conventional display, only those trained with the pseudo-quickened display showed benefits of prior training. This group also performed better than those who trained only with the conventional display.

INTRODUCTION

One of the fundamental underlying goals of the manual control research effort has been to develop more effective techniques for the control of complex dynamic systems. A pervasive challenge in this area has been presented by the requirement to control higher order systems: Those in which the number of time integrations between input and output is greater than or equal to two. The response of such systems is characterized as being both sluggish and unstable.

Two important reasons may be cited for the difficulty in controlling higher order systems. The first is related to the concept of optimal control which requires that the control input be a function of the momentary system state. The system state is defined as a minimum set of independent state variables, which, along with the system input, completely defines the system output for all future times. It can be shown that the required number of state variables is equal to the system order (Kirk, 1970). Therefore, as system order increases, the operator must perceive a greater number of state variables in order to optimally control the system.

The second reason for the inherent difficulty in controlling such systems is the required complexity of the operator's internal model. It is assumed that in controlling a dynamic system, the operator develops and continually updates a cognitive model in working memory which relates his control inputs to the system's responses or output (Kelley, 1968; Jagacinski & Miller, 1978). This "internal" model of the system's input-output characteristics is then employed to determine the control input necessary to produce any desired response. Since the optimum control input must be a function of all system states, it is clear that as system order increases, so must the complexity of the internal model.

Several techniques have been developed to aid in the manual control of higher order systems. Two of these techniques investigated in the present research are the phase plane and the quickened displays (Kelley, 1968; Poulton, 1974). In a phase plane display, system position is plotted as a function of system velocity. A theoretically optimal switching line, which represents the system state (combination of position and velocity) at which a control reversal must be implemented to prevent an overshoot, can then be added to the display (Miller, 1969; Platzer, 1955). In a quickened display, system position, velocity, and acceleration are combined via an analytically determined optimum ratio of 1:4:8 (Searle, 1951). Such a combination of state variables allows the operator to observe the results of his control input sooner than normal, thereby helping to prevent overshoot.

In the current research we describe and evaluate a unique display concept for higher order tracking termed pseudo-quickening. It entails the use of a conventional display in which the cursor position is directly proportional to the system error (Figure 1a). Information concerning the higher derivatives of the system state is then coded via the cursor's intensity. When the system state is such that a control reversal should be implemented, as specified by the theoretical switching line, the appropriate side of the cursor will increase in brightness (Figure 1b) indicating the commanded direction of the required control input. In the figure, error position is still left of center, but the operator must apply a hard left correction to stabilize the system at the zero error reference. In short, the algorithm for controlling the cursor intensity is the optimal switching line, as employed in the phase plane display, but with an allowance made for the operator's effective time delay (.5 seconds) between when the cursor intensifies and when the operator actually responds. This format provides compatible command information to the controller (Roscoe, 1968) in a manner that is much more economical in display space than is the phase plane display.

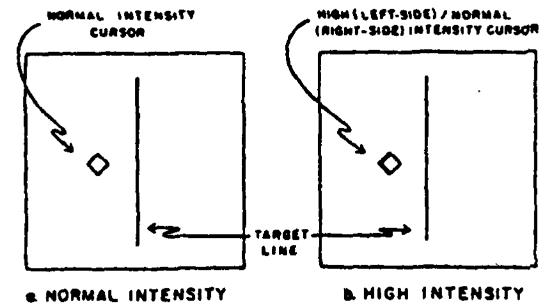


Figure 1: The pseudo-quickened display.

THE EXPERIMENT

The purpose of the experiment was to compare the tracking performance and the transfer of training characteristics of the three augmented displays (phase plane, quickened, and pseudo-quickened).

All 20 subjects (right-handed males) performed a single axis compensatory tracking task with second order dynamics attempting to minimize RMS error in the face of a .32 Hz random disturbance input. Control was exercised by manipulating a spring-loaded joystick in the left-right direction. The experiment was divided into two main phases. At the outset of Phase I (training), five subjects were randomly assigned to each of the four display groups: quickened, phase plane, pseudo-quickened, and conventional. Subjects in the first three groups, those with the aided displays, then tracked for 50 - 2 minute blocks. For the final 20 blocks (transfer phase), these subjects transferred to the unaided display. Subjects in the unaided, control condition tracked for all 70 blocks with the conventional display.

During the final 40 blocks (i.e., blocks 30-70), all subjects performed a secondary tone-counting task to assess workload via electrophysiological techniques. The results of this aspect of the data will not be reported here, except as they pertain to performance on the primary tracking task.

RESULTS

Figure 2 is a plot of RMS error, averaged across subjects, for each display type vs. block number 31-70. Three separate analyses were per-

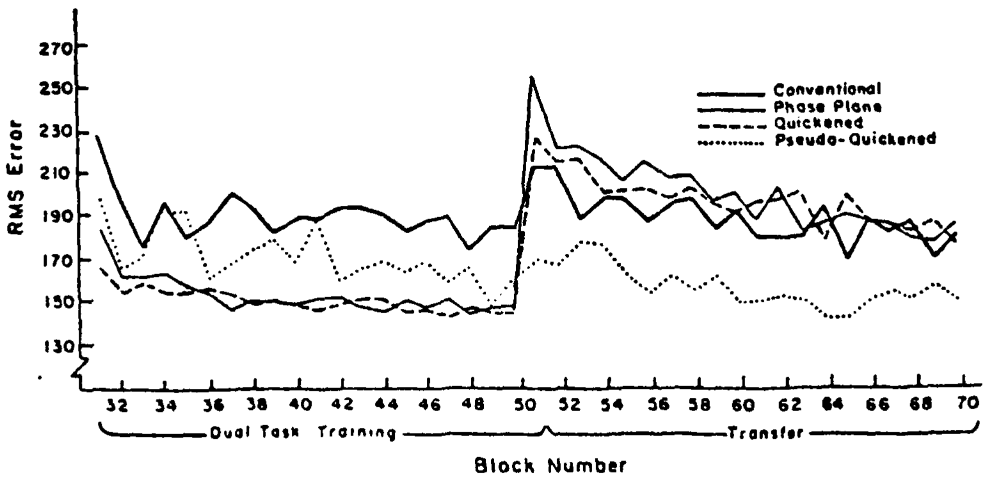


Figure 2: RMS error for all groups during training and transfer. All data are collected during performance of the secondary task.

formed on the data. One for training trials without the secondary task (1-30), one for training with the secondary task (32-50), and the final during the transfer period.

A two-way mixed model ANOVA was performed on these data for each phase with the major factors being display type (four) and block number; subjects were nested within display type. During Phase I (whose data are not displayed here) the main effects of display type and block were found to be significant ($F = 14.0$, $p < .01$; $F = 20.3$, $p < .01$, respectively). That is, there was a significant difference in the attained tracking performance among the display types and there was significant learning as indicated by the main effect of blocks. The insignificance of the display by block interaction ($F = .67$, $p > .1$) suggests that the rate of learning was the same for each display. Since the main effect of display type was found to be significant, separate comparisons were made between pairs of displays. These analyses revealed the phase plane, quickened, and pseudo-quickened displays all to be superior to the unaided display ($F = 34.1$, $p < .01$, $F = 46.4$, $p < .01$, $F = 17.8$, $p < .01$, respectively), while revealing no significant difference between themselves ($p > .1$ in all cases).

During Phase II the ANOVA, excluding trial 31, revealed reliable effects of both block ($F = 2.21$, $p < .01$) and display type ($F = 11.09$, $p < .01$). Planned contrasts in this case revealed the quickened and phase plane displays to be still superior to the conventional; however, the pseudo-quickened display was no longer reliably better than the conventional display ($p = 0.10$).

During the transfer trials (blocks 51 through 70) subjects who had been using the augmented displays were transferred to the conventional display while subjects who were using the conventional display continued to do so. While all groups thus tracked with a conventional display, in the following discussion groups will be defined by the type of display on which the subjects had been previously <u>trained</u>. Since the initial blocks for the augmented displays exhibited both a comparatively higher variability and rate of learning, the first five blocks were deleted prior to any of the following analyses.

In the ANOVA performed on the transfer trials, the main effect of display was found to be significant ($F = 4.8$, $p = .01$), thereby indicating a difference between displays. As suggested by the data in Figure 2, the planned comparisons revealed that this effect was due to the superiority of the pseudo-quickened display. That is, training on this display produced significantly better tracking performance ($p < .05$ in all cases) than any other display, while there were no other significant differences ($p > .1$) observed between the other displays. The lack of significance of the display by block interaction ($F = .99$, $p > .1$) suggests that the superiority of the pseudo-quickened display is likely to be a long term advantage. Finally, the main effect of block was also found to be significant ($F = 3.4$, $p < .01$), which reflects the small but consistent improvement with increased practices which can be observed in Figure 2.

For an additional analyses, the following equation was used to compute the initial percent transfer of training for each subject.

$$TOT = [(NN - TN)/(NN - FO)] \times 100\%$$

where: NN = average of blocks 1 through 5 for each conventional display subject.

TN = average of blocks 51 through 55 for each augmented display subject.

FO = average of blocks 46 through 50 for each augmented display subject.

The resultant percent transfer was average across subjects to give the average percent transfer for each display; a student-t test was then employed to assess the differences between displays. The pseudo-quickened display was found to produce reliably better transfer, with 87.7%, than either the phase plane display (28.8%; $t = 11.3$; $p < .01$) or the quickened display (39.7%; $t + 4.87$; $p < .01$). No significant differences were observed between the phase plane and quickened displays ($t = 1.04$; $p > .1$). In addition, the final percent transfer was computed by employing the same procedure except that "TN" was now averaged across the last five blocks of the transfer session. The results were basically the same; the pseudo-quickened display with 108.9% transfer was reliably better than either the phase plane display (69.4% transfer, $t = 5.31$, $p < .01$) or the quickened display (67.5% transfer, $t = 4.68$, $p < .01$). Once again, there was no difference between the phase plane or quickened displays ($t = .205$, $p > .1$). These results for the final percent transfer are important in that they indicate that the transfer benefits of the pseudo-quickened display are not short-lived, but should continue to produce superior performance.

DISCUSSION AND CONCLUSIONS

The most dramatic effect in the present results was the pronounced benefit to performance on a conventional display realized by prior training on the pseudo-quickened display. Performance of those subjects was far better than those who had practiced exclusively with either the conventional display or the other two augmented versions. We assume that this advantage is a consequence of the discrete signal information of the pseudo-quickened display, which allowed subjects to acquire a more precise and accurate internal model of the second order system. This conclusion is substantiated by an analysis of the locus of the actual control reversals in the state space, as described by Gill and Wickens (1982). Training with the quickened display was not served in this regard because a somewhat distorted picture of the true second order dynamics is presented with a quickened display (Poulton, 1974). The reason why training with the phase plane display failed to transfer effectively is less clear. It is likely that the continuous availability of the optimal switching line on the display allowed subjects to track effectively without having to "internalize" it, -- a strategy that proved to their detriment when the aid was removed.

While the benefits of the pseudo-quickened display over the conventional display were also realized in single task performance (Phase I), they were somewhat attenuated when the secondary task was required (Phase II). It is possible that this attenuation in performance related to the greater investment in learning the internal model--a benefit that, as noted above, was definitely realized in transfer, and also appears from Figure 2 to be emerging during the final eight trials of training. In any case, whatever the possible costs of the pseudo-quickened display in Phase II, relative to the other augmented displays, these must be balanced against two very real benefits; (1) unlike the quickened display, it presents an undistorted picture of the true state of the system, quite useful in transfer. (2) Unlike the phase plane display, it requires only one physical dimension to capture the state variable information. It is therefore efficient in display economy and quite feasible as a multi-axis display. It is reasonable to assume that these further benefits will be realized under more complex test conditions.

ACKNOWLEDGMENTS

This research was supported by Contract #49620-79-C-0233 from the Air Force Office of Scientific Research. Life Sciences Directorate, Dr. Alfred Fregly, was the technical monitor.

REFERENCES

Gill, R. & Wickens, C.D. A state space analysis of second order tracking with aided displays. Proceedings, 18th Annual Conference on Manual Control. Wright Patterson Air Force Base, June, 1982.

Jagacinski, R.J. & Miller, R.A. Describing the human operator's inter model as a dynamic system. Human Factors, 1978, 20(4), 425-435.

Kelley, C.R. Manual and Automatic Control. New York: Wiley, 1968.

Kirk, D. Optimal Control Theory: An Introduction. Prentice-Hall, Inc., 1970.

Miller, D.C. Behavioral sources of suboptimal human performance in discrete control tasks. MIT Engineering Projects Laboratory Report NSG-107-61, Cambridge, Mass., 1969.

Platzer, H.L. A non-linear approach to human tracking. Franklin Institute, Laboratories for Research and Development, Interim Technical Report 2490-1, Philadelphia, PA, 1955.

Poulton, E.C. Tracking Skill and Manual Control. Academic Press, 1974.

Roscoe, S.N. Airborne displays for flight navigation. Human Factors, 1968, 10, 321-332.

Searle, L.V. Psychological studies of tracking behavior: Part IV - The intermittency hypothesis as a basis for predicting optimum aided-tracking time constants. U.S. Naval Research Laboratory Report 3872, Washington, D.C., 1951.

ABSTRACT

An investigation of redundant auditory tracking.

Amir M. Mane, Christopher D. Wickens
(Manuscript to be submitted to Human Factors)

Four subjects performed a compensatory tracking and a complex visual detection task. Each of these were performed separately and concurrently. The tracking signal was presented in one of three display modes: an intermittent visual cursor on a CRT; auditory signal, varying in pitch and apparent spatial location; redundant visual and auditory combined presentation. When tracking was performed alone, the redundant presentation was marginally advantageous. When performed concurrently with the detection task, detection error increased to the degree that subject had to rely on visual cues in tracking. The redundant presentation was also superior when the combined score on both tasks was considered. The use of analog auditory display proved to be advantageous, particularly when visual attention must be allocated away from a visual tracking display.

3-8

DT