Presidential Address, 1980

Surprise! . . . Surprise?

EMANUEL DONCHIN

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

ABSTRACT

The nature of the psychophysiological enterprise is examined as it bears on the study of the endogenous components of event-related brain potentials (ERP). The view is taken that success in Psychophysiology should be measured by the degree to which psychophysiological data can be used in elucidating the processes that underly the behavioral product rather than by the enumeration of psychophysiological "correlates" of behavior. It is proposed that endogenous ERP components are best viewed as *manifestations* of the activities of "subroutines" invoked during the informational transactions of the brain.

A theoretical account of an ERP component consists of the specification of the functional role of the subroutine it manifests. Studies of the P300 components are examined for the contribution they make to the development of such a theory of the P300. Experiments focusing on P300 latency and amplitude are reviewed and it is concluded that P300 may be a manifestation of the processes whereby schemas are revised. The relationship between P300 and the Orienting Reflex is discussed within the framework of this model. It is suggested that P300 amplitude may predict the memorability of events. A preliminary test of this prediction is described.

DESCRIPTORS: Event-related brain potentials, P300, Cognitive task performance, Stimulus probability, Orienting reflex, Memorability of events.

"The electroencephalographic technique of recording brain waves by electrodes on the skull of man has recently been much improved by using computers for electronic summation of otherwise invisible changes of potential . . . In connection with different psychologically defined operations the method is of interest for problems concerning timing, direction, and elementary localization of otherwise inaccessible processes. It can be used also to measure the intensity or degree of a process and is thus a valuable asset for the science of psychophysiology. On the whole it seems likely that this branch of physiology faces a time of expansion." (Granit, 1977, p. 79)

Presidential Address, Society for Psychophysiological Research, Vancouver, British Columbia, October 1980.

The work of the Cognitive Psychophysiology Laboratory would not have been possible without the talents of my collaborators: graduate students, post-doctoral trainees, and colleagues on the faculty of the University of Illinois. The list is long, but it certainly includes Daniel Gopher, who is now with the Technion in Israel, Chris Wickens and Michael Coles, my colleagues in the Psychology Department, Ken Squires, Nancy Squires, John Polich, Leo Towle, and Ted Bashore, all of whom spent some time as post-doctoral associates at the Cognitive Psychophysiology Laboratory, my former students, Ron Herning, Marta Kutas, Connie Duncan-Johnson, Skip Johnson, Dick Horst, Jack Isreal and Greg McCarthy. Then there are those who still toil at Illinois, or who are newly toiling there, notably Demetrios Karis, Earle Heffley, Noel Marshall, Linda Vanasse, and Arthur Kramer. Marlene Calder has provided much aid in the preparation of this report.

The 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. The support of Drs. Craig Fields, Judith Daly, Al Fregley, Don Woodward, Bob O'Donnell, and Dave Otto is gratefully acknowledged.

Address requests for reprints to: Emanuel Donchin, Cognitive Psychophysiology Laboratory, Department of Psychology, University of Illinois, Champaign, Illinois 61820.

0048-5772/81/050493-21$2.10/0

A presidential address provides an opportunity for reviewing the progress of one's research. But, I have already done so before this society at our 1978 meeting in Philadelphia (Donchin, 1979). More studies have been conducted since by our group, and by others (see review by Hillyard & Picton, 1977). Yet, the factual story remains, in all important respects, as I presented it in Philadelphia. We confirmed some of the hypotheses about which I had to be tentative; we added detail in the picture; we have begun to utilize the P300 as a tool in the study of human cognition (Donchin & Isreal, 1980a, 1980b; McCarthy & Donchin, 1978). Nonetheless, if I undertake yet another survey of research on the P300 component[1] of the human event-related potential and its use as a tool in the study of human cognition I will be constructing for you the same edifice, with rococo detail added. I prefer this time to consider the cracks in the edifice.

It is my purpose to emphasize aspects of the data we have obtained that prevent us from making simple, summarizing, statements about our psychophysiological data. I shall conclude from this analysis that further progress in the study of the late endogenous components of the human event-related potential is contingent on the development of a coherent theory of the P300. Many of our assertions that seem to have the "flavor" of a theory are not, in fact, theories. I shall describe the desired features of a theory of P300, and venture to propose a structure, unfortunately rather flimsy, that may be used as a guiding theory. I will conclude with a description of experimental paradigms that are necessary for further progress in the study of P300. These paradigms will be illustrated by a study of the correlation between the amplitude of the P300 elicited by a stimulus and subsequent recall of that stimulus.

The Psychophysiological Enterprise

The theory of P300 I seek should be a *Psychophysiological* theory. By this I mean that it is to be a theory that responds to the fundamental questions asked by Psychophysiology. Of course, Psychophysiologists do not necessarily agree on the nature of the enterprise. For example, it is commonly thought that Psychophysiologists are in the business of seeking the "physiological correlates" of behavior (Andreassi, 1980). I am not sure I agree. Let me examine this proposition. Figure 1a presents a piece of behavior. The person shown in the Figure is a "subject" in an experiment (Karis, Druckman, Lisak, & Donchin, Note 1). As in all our studies, her scalp is adorned with electrodes. The Figure consists of 6 snapshots of the subject's face, taken at successive 500-msec intervals. Clearly, at some point the subject has "behaved": in the frame numbered 01553 she engaged in what an orthodox behaviorist would call "smiling behavior."

This record of "behavior" tells us, presumably, that the subject smiled. The venturesome may even be willing to infer from these pictures that the subject experienced an apparently pleasurable emotion. She may even be considered to have been surprised and pleasantly so (Ekman & Friesen, 1975). This then is the "behavioral" side of our equation.

What are the "correlates" of this behavior? Since we are recording the subject's electroencephalogram and assessing the event-related potentials (ERPs) associated with specific events the subject is performing, we may attempt to find an EEG corre-

Figure 1a. Tracings of 6 photographs of a subject, taken while the subject was "bargaining" with a computer as she was attempting to "purchase a used car" at the lowest possible rate. The subject is facing a Plato terminal. A video camera is placed over the terminal and the subject's facial expressions are recorded continually. The photographs are of a video display taken at 500-msec intervals. The photograph labeled 1552 coincided with the computer display of its "concession." Note the forehead electrode used for recording the EOG and the vertex electrode. Other electrodes are not seen in this photograph.

[1] I intend in this paper to discuss the P300 as though it is a unitary phenomenon. Several investigators have described positive going peaks in the ERP that appear to overlap and perhaps to summate with the P300 in a variety of ways, see reviews by Roth (1979) and Pritchard (1981). Without addressing here the specific issue of the "existence" of these components (see Donchin, Ritter, & McCallum, 1978) suffice it to note that in everyone's gallery of components, no matter how rich, there is always a member whose attributes are identical with what I call here the P300. It is on this parieto-central wave that responds in a specific manner to certain experimental manifestations (Donchin, 1979) that I focus in this report.

late of this smile. In Figure 1b, we show segments of electrophysiological data recorded in the experiment. One of the traces is associated with an event that did elicit a smile; the other, followed an event that did not. These records are clearly different! Are these records then a "physiological correlate" of behavior? Apparently so: different "behaviors" appear to be correlated with different brain activity. Is *this* what psychophysiologists do? Is it enough to observe variance in "behavior" and variance in physiological systems and assess the correlation between the "physiological" and the "behavioral" variables? Suppose we do find that a smile is always accompanied by a typical waveform, have we thereby achieved the goal of psychophysiology? I do not think so.

What, after all, do we learn from the statement that "whenever a subject smiles, a typical waveform appears between two electrodes placed on her head"? What can we learn about the brain, or about smiling? I suggest that not much. My interest is not in the "behaviors" (i.e., overt responses) on which my glance happens to fall but rather on the processes by which the subject encodes and represents the situation in which she finds herself and in the "computations" she performs (not necessarily consciously) as her actions, both overt and covert, are determined by the interaction between these representations and external events. From the perspective of an observer these processes result in overt, measurable manifestations. The subject talks, types, smiles, fidgets, in short—the subject "behaves." We can select for observation one or more of these "behavioral responses" such as the smile, and monitor and measure its various attributes, or record its frequency of occurrence. But, we have, I think, an ulterior motive when we observe this behavior. We are, in point of fact, trying to find out what the subject is "really" doing. The lady in Figure 1 may be surprised. She may be pleased. She may be astonished. She may be worried. She may be hesitant. We can not tell which description of her mental state is correct by merely noting that the subject smiles. But if a smile may indicate any of several underlying processes, do we expect P300 to correlate with the smile or with the underlying process? It seems evident that our interest is focused on the processes. But if so, the interpretation of any observed correlation between P300 and some "behavior" depends, at least to some extent, on the context in which the behavioral responses are recorded. Otherwise, the underlying processes may remain obscure. Thus, correlations between behavior and psychophysiological signals can be interpreted only in the context of the circumstances in which the signals were recorded.

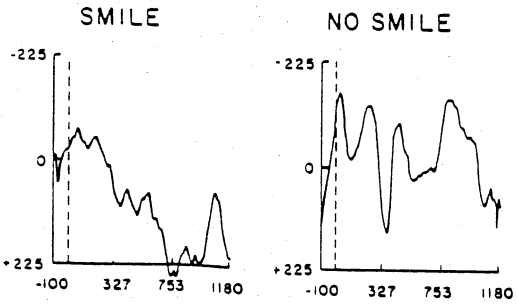

Figure 1b. Two tracings of "raw" EEG taken immediately following a computer concession. The abscissa marks time in milliseconds; the ordinate marks voltage in arbitrary units. The computer concession was displayed at the time indicated by the broken vertical line. One trace was associated with a concession that evoked a smile, the other with a concession that failed to evoke a smile from the subject.

The circumstances in the case considered here are as follows. The subject is playing a game with the computer. It is a game with financial consequences. The subject is trying to "buy a used car" from the computer. It is her task to obtain the car at the lowest possible price. The computer, of course, is trying to sell the car at the highest possible price. In a true oriental bazaar fashion, these two bargainers haggle. Each side concedes a little. Our subject increases her offering from trial to trial. The computer reduces its asking price on every trial. Eventually, the two sides agree and the transaction is completed.

The situation is, of course, somewhat more complex. The computer has a strategy. At any one time, it can be either "mean" or "generous." When the computer is mean, it concedes only 20% of the subject's concession. If the computer is generous, it concedes 80% of the subject's concession. The subject is, of course, interested in keeping the computer generous. But the computer has a habit of changing its strategies. These "switches" in strategy depend on the subject's behavior. The subject's income depends on her ability to detect changes in the computer's strategy. The sooner the subject detects a change in the computer's strategy, the sooner she can do something about it.

A sample of the events in the experiment is shown in Figure 2. The experimental session consists of a series of interchanges in which the subject types a number, indicating a concession. This is followed by a computer concession, which is displayed as a two-digit number. Is there a difference between the subject's reactions to computer concessions that reflect a switch in the computer's strategy and reactions to the other concessions? Figure 1a shows one subject's overt reaction, on one occasion, to an ac-

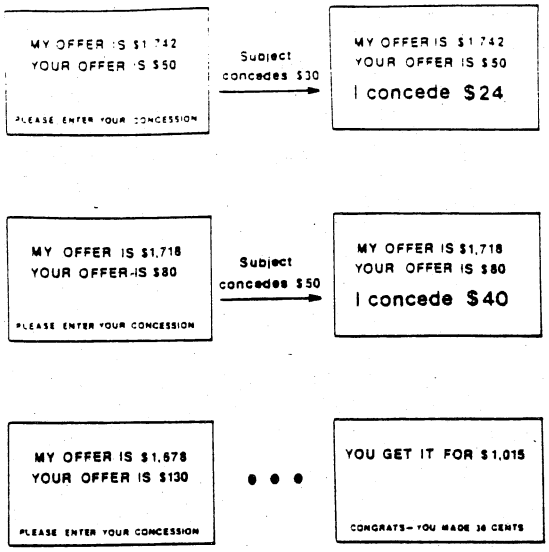

Figure 2. The sequence of events in the bargaining experiment. In the frames appears the material displayed to the subject on the Plato terminal. The computer is the "speaker." Thus "my offer" is the computer's current "asking price." After the first display the subject typed 30, indicating a willingness to pay $80 for the car. The computer (in its generous mode) conceded 80% of the subject's concession (i.e., $24) and displayed both offer and counter offer asking the subject for another offer. This continues until the session is terminated, as shown in the last frame.

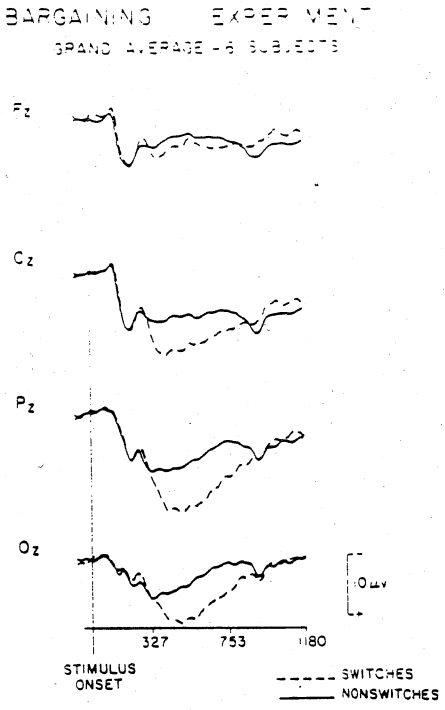

Figure 3. Event Related Potentials, averaged over 6 subjects, recorded from the Frontal (F_z), Central (C_z), Parietal (P_z), and Occipital (O_z) electrodes. The eliciting events were the appearances of the computer's concessions on the screen at the instant indicated by the vertical bar. The dashed line represents the ERP elicited by concessions that constituted a switch in the computer's strategy. The solid line—concessions that were not switches.

tion by the computer. It turns out, however, that such facial responses were quite rare. Most of the time, the subject's face remained impassive.[2] The ERPs, on the other hand, were quite expressive. A systematic study of 6 subjects in each of two experiments showed that the ERPs elicited by computer concessions that reflected a strategy switch were quite different from the ERPs elicited by concessions that did not reflect a change in the computer's strategy.

Figure 3 presents a summary of the data. The ERP elicited by switches is characterized by a large P300 component. Computer concessions that are not switches-in-strategy do not elicit a P300. All subjects respond in this fashion and the response is detectable in the records of the individual trials. It is odd, perhaps, that an overt "behavior" that communicates with utter immediacy (the smile) is elicited rarely. Yet, the physiological response, the ERP, is clear, reliable, and interpretable. These data, then, present a problem for those who assume that "behavior" is reflected by the subject's smile, while the P300 is in some special category of observations that are "correlates" of the behavior. In these terms there appears to be almost no correlation between the psychophysiological (the P300) and the behavioral (the smiles) data. However, perhaps the smile is a poor index of "behavior" in this context. Perhaps the P300 is as much "behavior" as is the smile, it is yet another product of the processes of representation and computation that are at the core of behavior. It may be an incidental product, a "sign" rather than a "code" in Uttal's (1969) terminology. Yet, the processes it manifests are part and parcel of the behavior we seek to understand. These considerations suggest that Psychophysiologists should not seek physiological correlates of "behavior." Rather, our task is to find ways in which our understanding of behavioral processes is enhanced by the study of psychophysiological signals.

Let me repeat, within the framework of this example, the question. What *is* the nature of the psy-

[2]Dr. Dan Druckman, of Math Tech, Inc. who collaborated with us in this project, is now studying the facial features to see if there is a systematic pattern of facial responses to the changes in strategy.

chophysiological enterprise." If it seeks "correlates," what is being correlated with what? We record a physiological response. Something inside the cranium is activated, reliably, whenever the computer switches strategies, and is manifested on the scalp by the P300. The subject is sitting in front of a terminal, typing her concessions, reading the computer's messages, checking them in her memory against the computer's previous concessions. The subject is capable of understanding the situation. After a while, the subject's behavior becomes quite regular and she is in control of the situation. But—the smile, or any other specific overt "behavioral" indicator, does not indicate reliably that this is the case. It is only the rate with which the subject's earnings rise, and the strategies with which she deals with the computer that are reliable indicators of the subject's "behavior." Behavior appears to be, as it always is, more than an arbitrary response selected by an experimenter for generating numbers that can be used in calculations. Rather, behavior is an integrated stream of multiple, parallel processes that interact in myriad ways. These processes have numerous products, and the products can be measured: a key press, a joystick manipulation, a cough, a frown, a tear, a sigh. All these are *products*. We seek the *processes* that underlie these products. *Each product, be it a button press, or an ERP component, is interesting only to the extent that it tells us something about the underlying processes*. The products may or may not be correlated with each other. But this is not really important. If different products provide complementary information about the processes, they are in a very useful sense "correlated." Success in the psychophysiological enterprise requires that the psychophysiological data provide insight concerning the processes, rather than a list of correlations between products. Can the ERP that we, and colleagues in other laboratories, have been generating in the past decade, be used to understand such processes? I shall be examining this question in the remainder of this paper.

The "Oddball" Paradigm

Most investigators of the P300 component, looking at the data in Figure 3, may, upon reflection, recognize the experiment as yet another form of the "classical" oddball paradigm. This "bargaining" experiment is, in some sense, quite similar to studies that date back to Sutton's original experiments (Sutton, Braren, Zubin, & John, 1965) and reported frequently since (see reviews by Donchin et al., 1978; Sutton, 1979; Tueting, 1979; Desmedt & Debecker, 1979).

The oddball paradigm is illustrated in Figure 4. The Figure summarizes a parametric analysis of the

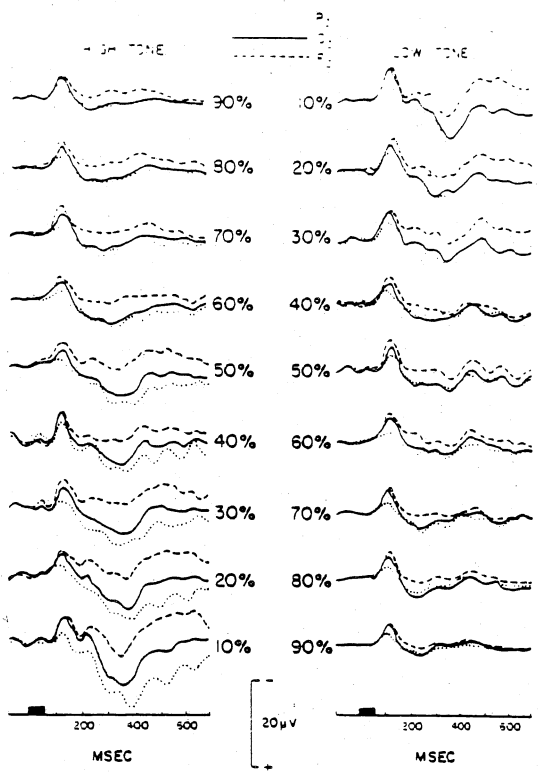

Figure 4. The ERPs elicited by high and low tones presented in Bernoulli series with the probabilities as indicated. Data are shown for experimental conditions in which the subjects were instructed to count the high tones in the series. Each row presents the ERPs recorded in the same series, with the complementary probabilities. The data collected at each of three electrodes are superimposed. (From Duncan-Johnson & Donchin, 1977)

paradigm reported by Duncan-Johnson and Donchin (1977). The experimental design is familiar. One of two distinctive events can occur on each trial. The events occur in a sequence. One or the other is selected randomly with complementary probabilities. The rare event elicits a large P300. The P300 component is largest at the parietal electrode and its amplitude is affected by the probability of the stimulus. We have shown that it is the *subjective probability* of the stimulus, not its prior probability, that determines the amplitude of the P300 (Donchin, 1979; Donchin & Isreal, 1980a; Squires, Donchin, Herning, & McCarthy, 1977; Squires, Wickens, Squires, & Donchin, 1976). The statement that two distinctive events are used in the series must be interpreted rather broadly. Figure 5 presents examples from several studies in which subjects were presented with quite different stimuli that could be placed in one or another category. The data are clear: diverse physical events, if instruc-

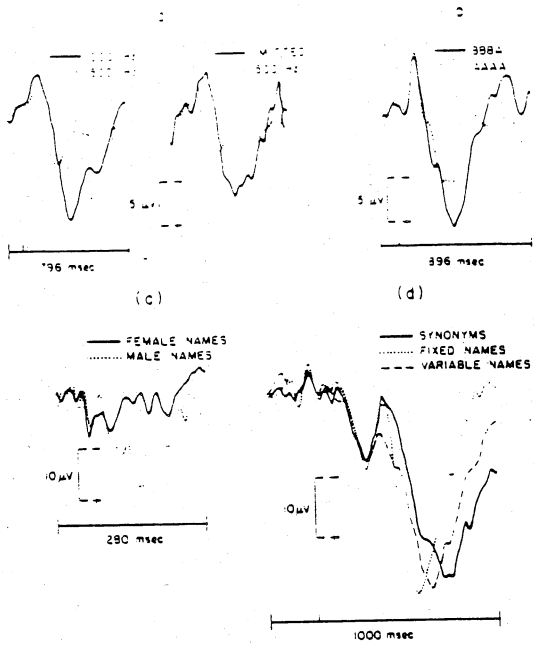

Figure 5. Illustration of different conditions under which the P300 can be elicited. In (a) are shown data obtained in a standard oddball paradigm when two tones are presented in a Bernoulli series (P(low) = .20) and when one of the two tones is omitted. In (b) both traces were elicited by the high tone presented in a Bernoulli series where the two tones were equiprobable. Even though both traces were elicited by the same physical tone they differ in the amplitude of the P300. The solid line represents an ERP preceded by 4 high tones (see Squires et al., 1976). In (c) the Bernoulli series was constructed from a Bernoulli series of names displayed on a computer terminal: 20% of the names of males, the rest names of females. A similar experiment is shown in (d) where the dotted line represents an ERP elicited by the name DAVID that was mixed in a Bernoulli series with a .20 probability with the name NANCY that appeared with a .80 probability. The dashed line was obtained with the same stimuli shown in (c) while the solid line was elicited by synonyms of PROD, that appeared, again with a .20 probability, mixed with other words. The data in (c) and (d) are from McCarthy, Kutas, and Donchin (Note 2), and from Kutas, McCarthy and Donchin (1977) (Copyright 1977 by the American Association for the Advancement of Science. Reproduced with permission.) The figure was prepared by McCarthy (1980).

tions to the subject sort them into categories, will elicit a P300 *jointly* as a category (Johnson & Donchin, 1980; Kutas & Donchin, 1979; Polich, Vanasse, & Donchin, 1981; Courchesne, Hillyard, & Courchesne, 1977).

The bargaining experiment fits well within this paradigm. Two rather elaborately defined classes of events occur in this experiment. *All* events are "computer concessions," all consist of the appearance of a digit on the screen at a given point in time. It is possible, however, to classify some of the concessions as "switches" in the computer's strategy. Other concessions are not switches. Shifts in strategy (switches) are fairly rare in this experiment, occurring only on 10% of the trials. Our prediction that the switches would elicit a P300 is confirmed.

We may have here a glimmer of an interesting correlation between aspects of behavior and a psychophysiological response. Note, however, that the behavior is not "a smile." It is not *any* specific response by the subject. Behavior refers here to the subject's perception of a contextual relationship between the specific stimuli presented, the past history of the sequence, the subjects' strategies, and the experimental instructions. All these determine jointly the meaning of the numbers that appear on the screen. It is this meaning that reflects itself in the amplitude of the P300 (Sutton, 1979). On the basis of data like these we can assert that surprising events elicit a large P300 component. But, the statement that P300 is elicited by surprising events is an assertion about the antecedent conditions of the P300, it tells us nothing about the process, or processes, manifested by the P300. Thus, it does not constitute a theory of P300.

Specifications for a Theory of P300

I want to emphasize the distinction between an enumeration of the antecedent conditions and a process theory of a phenomenon. To observe any phenomenon, certain antecedent conditions must be satisfied. For example, rain cannot be observed without clouds, appropriate atmospheric conditions, specific properties of clouds, and so on. Yet a theory of rain will not be considered adequate if it merely lists the antecedent conditions for rain. For one thing, none of these antecedent conditions *by itself* will produce rain. More important, a theory of rain must specify the process by which precipitation is brought about. Of course, a process theory of a phenomenon must generate as predictions all that is known about its antecedent conditions. The enumeration of the antecedent conditions is thus an important theoretical endeavor because it establishes constraints on the theory; but the theory must present a model of the process.

What shape should a process theory of P300 assume? The label "P300" is assigned to a "wave," a voltage change that can be recorded between two electrodes. There are those who would feel that the ultimate theory should consist of the specification of the intracranial origin of the potentials. The identification of the neuronal assembly, or assemblies whose activation causes this field potential to

appear between the electrodes, would constitute a sufficient theory. Thus, for example, some would consider an adequate theory the assertion that the P300 is produced by the summation of postsynaptic potentials in the entorhinal cortex. I would consider it an important and crucial observation, but not a satisfactory theory, because this assertion will not allow me to derive the antecedent conditions for a P300. It is not that the data on the intracranial sources will be useless. On the contrary, knowing the intracranial source of an ERP component would help in the development of a theory, but it is insufficient.[3] The problem is that specifying the generator source regresses the *theoretical* question from the need to account for the process reflected by the P300 wave to a need to account for the process undertaken by the intracranial structure. Because, in either case, our goal is the specification of the role *in the information transactions of the brain* that is played by the P300 process. What we need, then, is a theory that will explain the data on P300 in terms of the human information processing system.

I start from the assumption that the ERPs we record on the scalp are manifestations of intracranial activity. I assume that in the course of information processing some action will result in a confluence of biophysical activities that have an observable scalp manifestation. It is not necessary that this activity take place in any one nucleus or in a discrete anatomical locus. The electrical potential we observe may well be the summation of the activity of an elaborate complex of anatomical entities (see Goff, Allison, Williamson, & Van Gilder, 1979; Vaughan, 1974). The important assumption is that the neural activity that contributes to appearance of the potential on the scalp is involved in (is activated in connection with) a specific information processing sub-task. We need to specify its function.

A metaphor that expresses the concept is the "subroutine" (Donchin, Kubovy, Kutas, Johnson, & Herning, 1973; Donchin, 1975). Subroutines are units of processing designed to perform a specific task required in a variety of different circumstances. Thus any given subroutine may be involved in quite disparate programs. One may encounter, for example, a need for matrix-inversion in a factor-analysis program, in a meteorological program or in the analysis of elementary particles. The tasks served by the subroutine may be quite diverse but they share a need for the services of the subroutine. To understand the subroutine we need to elucidate the common role it plays within its different contexts. In other words, a process theory of a subroutine is an attempt to abstract the similarity across its seemingly diverse contexts. This, no doubt, requires that we try to enumerate the conditions antecedent to the invocation of the subroutine. Matrix-inversion subroutines, for example, are generally invoked by programs in which matrices are part of the program's data structure. An adequate description of a subroutine will specify its inputs and the operations it performs on these inputs. A full understanding (or a process theory) of the subroutine will require in addition that we understand the role it plays in the program for which it was invoked. Thus, an enumeration of the consequences of the subroutine's invocation is a necessary component of a specification of the role it plays in the primary task.

I think of the process manifested by P300 in much the same way. The process is invoked in a diverse set of circumstances. I assume, however, that these circumstances are similar in that all, at some point, show a need for the services performed by the P300 process. It is our task to try to identify the functional role of the neural activity manifested by P300 in the informational transactions of the brain. We may begin by enumerating the antecedent conditions for the elicitation of a P300. These may lead us to speculate on the nature of the process. Ultimately, any theory we develop about the nature and function of the process should lead us to specific predictions about the consequences of the P300. That is, we should be able to assert what, if anything, will vary in the course of human information processing, and the behavior it controls, as a function of the invocation of the P300 process.

What, then, do we know about the P300? What does the edifice to which I referred earlier look like? It seems to be an arch resting on two columns: subjective probability and task-relevance. The story is reasonably well known; for a very thorough review see Pritchard (1981). Rare events elicit a P300. The rarer the event, the larger the P300. The events eliciting the P300 need to be defined, not in terms of their physical structure, but rather in terms of high-level categories imposed on the external environment by its psychological structure. Furthermore, as I noted above, the evidence is reasonably strong that it is not the objective relative frequency of events that is the crucial variable, but rather it is the perception of this relative frequency, the subjective probability or "expectancy," that is the crucial variable.

This is often summarized by the simple statement that "surprising" events elicit a P300. We have asserted that the amplitude of the P300 can be used to measure that subjective probability of events (Donchin & Isreal, 1980a). But do we then suggest

[3]Of course, for a clinical application, nothing is as important as are the data on the sources generating the components.

that the P300 is itself a manifestation of surprise? Not necessarily.

The ERP as a Tool and the ERP as a Phenomenon

We must distinguish between two statements. To say that surprising events elicit a P300 or that the amplitude of the P300 is proportional to the subjective probability of the eliciting event, is quite different from saying that the P300 is *itself* a manifestation of surprise. Recall that I assume that the electrical activity we record from the scalp is a manifestation of the activity of some intracranial processor or processors. The theory I seek is to be a description of the functional significance of this process. The theory should elucidate the specific processing activities undertaken by the neuronal population whose activity is manifested on the scalp by a component. The assertion that surprise is an antecedent condition for the activation of this processor is not equivalent to the statement that this processor's activity *is* the process of surprise.

The distinction often ignored is between statements and studies that are of value as we attempt to demonstrate the utility of the P300 component as a tool in the study of human information processing and studies and statements that lead to the development of an understanding, a theory, of the P300 phenomenon itself. I will illustrate this distinction by reviewing studies of the latency of the P300 and studies in which the amplitude of P300 is assessed in the context of dual task performance.

The Latency of the P300

There is strong evidence that the latency of P300 is proportional to the time required for stimulus categorization (Kutas et al., 1977). The latency of P300 is relatively independent of the time required for response selection and execution. Many experiments bear on this issue (see review by McCarthy, 1980), a useful illustration is provided by McCarthy and Donchin (1981). The stimuli shown in Figure 6 were used to manipulate two variables known to affect human reaction time. On each trial of the experiment, one of these matrices, or a matrix like it, appeared on a display screen. The words RIGHT or LEFT always appeared somewhere in the matrix. However, on half the trials, this command word was clearly legible because it was surrounded by # signs. In the other half of the trials, the command was hard to detect because it was embedded in a matrix of random characters (the "noise"). Naturally, the subjects responded more slowly to the commands embedded in noise. To complicate matters, each of these stimuli was preceded either by the word "same," or by the word "different." The word "same" indicated to the subject that he must follow

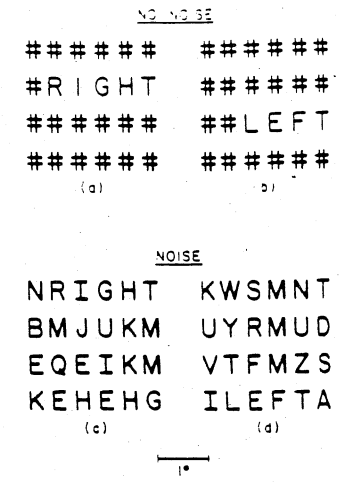

Figure 6. Each of the 4 character matrices was used as a stimulus in an experiment designed to test the hypothesis that P300 latency is affected primarily by stimulus evaluation and is largely independent of response selection and execution processes. (From McCarthy & Donchin, 1981. Copyright 1981 by the American Association for the Advancement of Science. Reprinted with permission.)

the command as shown. That is, respond with the right hand to the word "right" and with the left hand to the word "left." When the word "different" appeared on the screen, the subjects responded to the command "right" with the left hand and the command "left" with the right hand. Subjects respond faster when they are doing exactly what they are told; they are slowed when they have to perform an incompatible response. There was no interaction between the effect of response incompatibility on reaction time and the effect of the noise on reaction time. If we accept traditional "additive factors" logic (Sternberg, 1969), we can infer that the two variables manipulated operate at different stages of processing. There is ample evidence that the addition of noise to a stimulus retards reaction time because it interferes with stimulus encoding and evaluation (Biederman & Kaplan, 1970; Shwartz, Pomerantz, & Egeth, 1977; Frowein & Sanders, 1978). Response incompatibility, on the other hand, retards reaction time because it adversely affects the subject's ability to select and execute a response (Fitts & Deininger, 1953; Broadbent & Gregory, 1962; Schvaneveldt & Chase, 1969).

If the process manifested by P300 is invoked when, and only when, a stimulus has been identified as belonging to an unlikely category, then the subject must be able to complete the categorization of the stimulus before the P300 is elicited. On the other hand, there is no particular reason why response selection and execution should affect the P300 la-

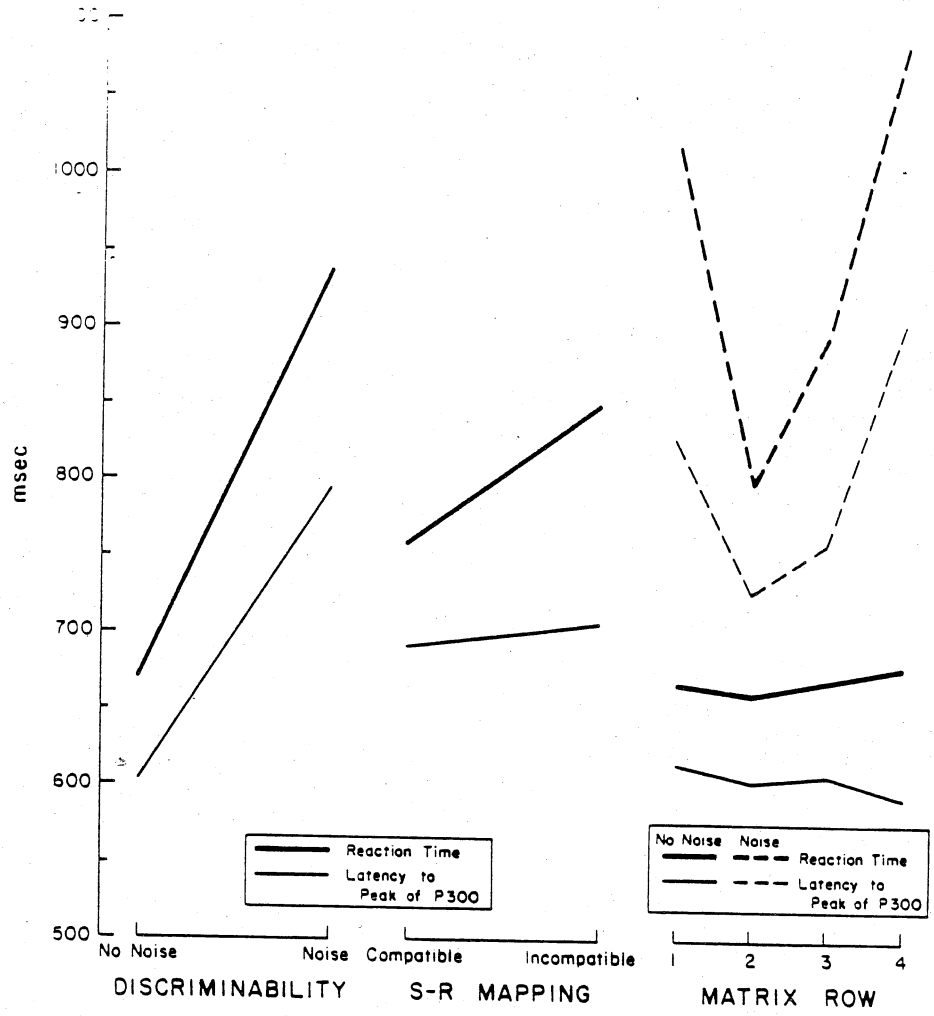

Figure 7. Reaction Times and P300 Latencies recorded in response to the stimuli shown in Figure 6. The data shown are averages over the subjects. Note that while Reaction Times are affected both by the addition of noise-characters to the stimulus and by the need to perform an incompatible response, the latency of P300 is affected solely by the addition of the noise-characters. The rightmost panel shows the effects of the position of the command word in the matrix on Reaction Time and P300 Latency. (From McCarthy & Donchin, 1981. Copyright 1981 by the American Association for the Advancement of Science. Reprinted with permission.)

tency. We predicted therefore that P300 latency, unlike reaction time, would be affected only by the presence of the noise. The need for incompatible responses would not affect P300 latency.

The data confirmed this hypothesis. As shown in Figure 7, P300 latency was strongly affected by the presence of noise and was not significantly affected by the compatibility between the stimulus and the response. This was a very useful experiment. It supports the assertion that P300 can serve as a *dependent* variable for studies of human information processing. The latency allows us to address a range of problems in cognitive psychology that require, for their effective solution, a measure of mental timing uncontaminated by response selection and execution processes.

Duncan-Johnson's (1981) analysis of the Stroop phenomenon is a case in point. Polich and Donchin (Note 3) examined the effects on reaction time of the relative frequency of words in the language and the relative frequency with which words of a given class are used in an experiment. The frequency of words in the language affected both reaction time and P300 latency, while the frequency of words used in an experiment affected reaction time, but had only a minor effect on P300

latency. The data suggest then, that the effect of word frequency on RT is due to encoding and categorization processing while the effect of local frequency on RT is due largely to response selection and execution processes.

While these studies are of utility, the data say relatively little about the *nature* of the P300 process, and do not lead directly to a theory. Their import to a theory of P300 is largely negative. They tell us what must be completed *before* P300 is elicited and they tell us which proceses are largely irrelevant to the determination of the P300 amplitude or latency. We know that the process in which we are interested is invoked after certain activities have been completed. And that, of course, is important. But, what process is invoked as a consequence of the categorization if the category is both important and improbable? That remains to be seen.

Studies of Resource Allocation

Consider another example. Wickens and Donchin, with several associates, primarily Jack Isreal and Art Kramer, have been assessing the utility of the P300 as a measure of "workload." These studies provide yet another illustration of how a study of the P300 that allows us to use P300 as a tool raises, rather than resolves, questions about the nature of the process (Isreal, Wickens, & Donchin, 1979; Isreal, Chesney, Wickens, & Donchin, 1980a, 1980b; Donchin & Isreal, 1980; Wickens, Isreal, & Donchin, 1977; Wickens & Tsang, Note 4). The logic of these experiments derives from a model that assumes that information processing activities require "resources." We assume that at any given instant there is a finite set of resources and if these resources are deployed in the services of one set of tasks, they are not available for serving the purposes of yet another task (Navon & Gopher, 1980).

Our experimental procedures are depicted in Figure 8. The subject is tracking a moving target and, during tracking, a series of tones is presented through a pair of earphones. The subject must count some stimuli and ignore the others. In Figure 9, we show the ERPs elicited during this study. As usual, a large P300 is elicited by the rare tones when the subject is merely counting tones. If we require the subject to track while counting, there is a substantial reduction in the amplitude of the P300. Suppose we make the tracking more difficult by re-

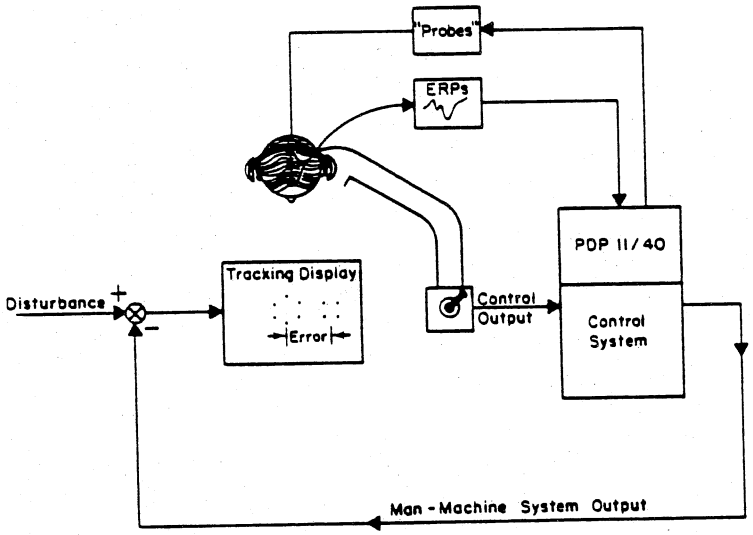

Figure 8. A schematic representation of the arrangement used in studies of the amplitude of P300 elicited by a Bernoulli series of tones during the concurrent performance of a tracking task. The computer is controlling the display of the target's motions as well as the behavior of the cursor that is under the subject's control. The target can be made to move in either one or two dimensions. The computer also generates the tones and records the subject's ERPs. (From Isreal et al., 1980b. Copyright 1980 by the Human Factors Society. Reprinted with permission.)

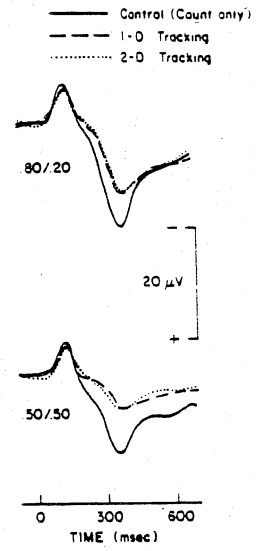

Figure 9. The ERPs elicited by the rare tones presented while the subject was counting tones (solid line) and when the subject was counting tones concurrently with one-dimensional and two-dimensional tracking. Data are shown for two Bernoulli series using different probability mixes. Data are averages over all experimental subjects. (From Isreal et al., 1979. Copyright 1979 by the Human Factors Society. Reprinted with permission.)

quiring the subject to track a target that moves in two rather than in one dimensions. We find that even though the task becomes more difficult (as indicated by increased reaction time and increased error count), the P300 amplitude remains unchanged (Isreal et al., 1979). We have to explain why the P300 amplitude is not affected by increasing task difficulty even though it is severely affected by the introduction of the initial tracking task.

In an extensive series of experiments, we obtained evidence for the assertion that P300 amplitude is specifically sensitive to perceptual load. If the perceptual demands presented by a primary task are increased the amplitude of the P300 will be reduced. On the other hand, if task difficulty is increased by augmenting the load on response processes only a small effect on the amplitude of P300 can be obtained. This again is a finding of considerable utility. There is a need for techniques for measuring, assessing and analyzing human workload. There is a need for measures that are fine-tuned to specific components of the demands placed on subjects by tasks, as well as for the traditional measures, like reaction time, that are affected by a multiplicity of processes (Wickens, 1980). We can, with the help of the P300, proceed to analyze complex situations using P300 amplitude as a dependent variable that allows a definition of the components of performance that are affected by certain experimental manipulations.

Consider, for example, the display shown in Figure 10. It is used in an experiment conducted by Kramer, in collaboration with Wickens and myself. Our subjects are performing a complex task. A large open rectangle appears on the screen and moves from right to left in a random trajectory. The subject controls the trajectory of the small rectangle and, with a joystick in the right hand, has to bring these two rectangles together. When the two rectangles are superimposed, they begin to rotate. It is the subject's task, once the two rectangles are in close proximity, to manipulate a knob with the left hand to match the orientation of the two rectangles. There are two phases, then, in the experiment. In the first phase the subject must "capture" the large rectangle; the other phase requires the subject to align the orientations of the two boxes. The second phase is considerably more difficult than the first. To make life even more difficult, the subject performs the task using either a first-order or a second-order control system. Second-order systems are considerably more difficult to control than first-order systems (Baty, 1971; Fuchs, 1962; McRuen, 1980; Navon & Gopher, 1980; Wickens & Tsang, Note 4).

While the subject is engaged in these maneuvers,

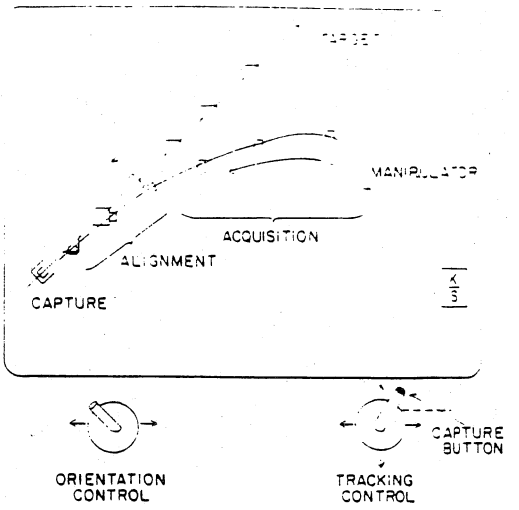

Figure 10. The display used by Kramer, Wickens, and Donchin (Note 5) for assessing the effects of system-order on the amplitude of P300 elicited by tones presented while the subject is attempting to align the "manipulator" under his control with the "target" that is under the computer's control. The manipulator is moved on the screen by the tracking-control stick. As soon as the manipulator acquires (comes near) the target, the target begins to rotate and the subject must use the orientation control to align the manipulator with the target. When he is satisfied he presses the capture button. Performance is graded according to the final position of the target and the manipulator. The function relating the tracking control to the manipulator's movement determines the system's order.

tones or flashes are presented in the basic oddball paradigm. We study variations in the ERPs elicited by these probe stimuli as the difficulty of the primary task is manipulated. Some of the results are shown in Figure 11. The pattern of task difficulty reflects itself in the amplitude of the P300. The amplitude of P300 is smaller during the more demanding alignment phase than it is during the capture phase. Furthermore, within these phases, a smaller P300 is elicited by the probes when the subject is confronted by a second-order system.

From these data we infer that the effects of system order on task difficulty are largely *perceptual*. I am willing to assert this because I have confidence in our interpretation of P300. This analysis is consistent with the suggestion that second-order systems are more difficult to control than first-order systems because a first-order system provides a consistent mapping of stimuli on responses, and requires, therefore, automatic, rather than controlled processing (Schneider & Shiffrin, 1977). The data I just described are useful in a human engineering context. Yet, these data do leave a fundamental gap in our understanding of the *process* manifested by the P300.

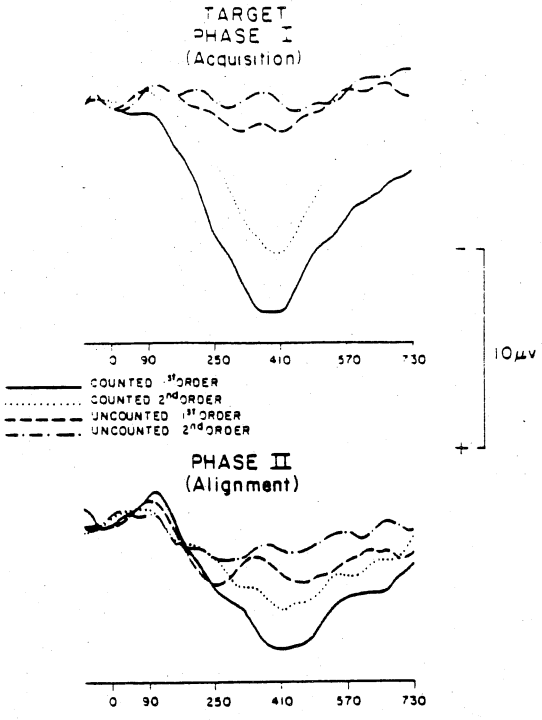

Figure 11. The ERPs, averaged over subjects, elicited by tones presented while the subjects were performing the task shown in Figure 10. Note that the amplitude of P300 is larger during the easier acquisition phase than it is during the more difficult alignment task. Moreover, the amplitude of P300 decreases as the order of the system is increased.

Consider the results. We assert that somewhere in the system there exists a pool of resources that is being utilized as the subjects count. These resources are used as the subject detects and identifies the tones, and updates one or another stimulus count. The process manifested by P300 is available to a lesser degree if the subject is tracking targets on a screen even though the subject continues to count tones. The implication is that the amount of some useful, but not essential, commodity varies in an analog fashion with variations of task difficulty. Thus, the tracking task shares a "resource" with the counting task. As tracking has priority, it seems to consume more of the commodity and less is left over for counting.

But what is that commodity? We know it has something to do with perceptual processing. It may have something to do with switching from automatic to controlled processing (Schneider & Shiffrin, 1977). But that is about as far as we can go. Again, the experiments are fascinating. The data are of considerable utility in human engineering. But the nature of the process we are trying to understand remains obscure. We have more facts about the antecedents of P300 but not a theory. We assert that events that are task relevant and rare elicit a large P300. The larger the probability, the smaller the P300. The more important the event, the larger the P300. If the series of events that elicit the P300 are embedded in a task that competes for the subject's attention with yet another task that places priority demand on stimulus evaluation processes, we are likely to observe a reduction in the amplitude of P300, suggesting a decrease of available resources.

The Probability Effects: Some Problems

But this outline of antecedent conditions is itself problematic. We have, for example, to take the statements about probability and the P300 with caution. There is nothing wrong with the data from which we and others inferred a very strong dependence of P300 amplitude on the subjective probability assigned to the stimulus. However, we know that there are circumstances in which we can vary the subjective probability of a stimulus without a corresponding effect on P300. Consider, for example, the study by Heffley, Wickens, and Donchin (1978) whose paradigm is labeled "continuous display" in Figure 12. Subjects are observing targets that move

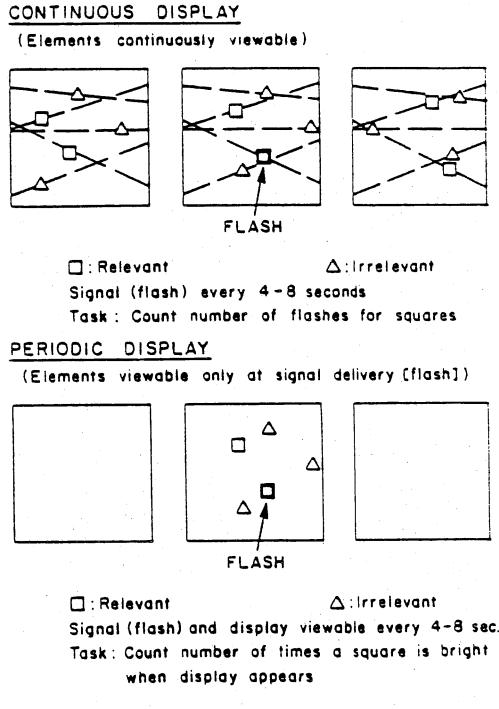

Figure 12. The two display modes used in the monitorin experiments by Heffley (Note 6).

on a screen in random trajectories. Occasionally, one target intensifies. These intensifications are probe events. The targets come in two flavors: triangles and squares. Intensification can occur in either of these classes. The subject is instructed to attend to one brand of targets and to ignore the other. That is, the subject counts either triangles or squares. As is shown in Figure 13 there is a large difference between the P300 amplitudes elicited by the counted and the uncounted targets. The counted probes elicit a large P300. The uncounted probes fail to elicit any P300. These data demonstrate in a striking fashion the sensitivity of the P300 process to task-manipulations. However, these same data open one of the larger cracks in our P300 edifice. One of the independent variables manipulated in this study was the probability with which the target stimuli appeared on the screen. The number of tracks assigned to each of the targets and the total number of tracks were varied systematically. Furthermore, we could vary the probability that one or the other target would intensify. Thus, for all practical purposes, this was very much an "oddball" experiment.

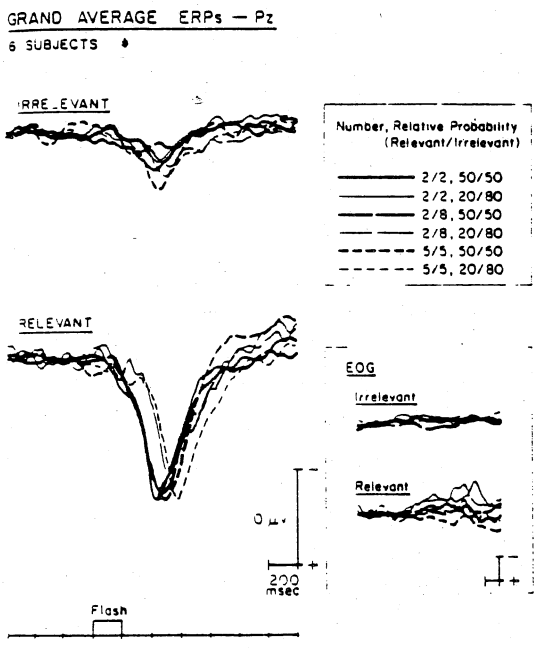

Figure 13. The ERPs elicited by the intensifications of triangular and square targets moving on a screen in random trajectories from left to right. The insert indicates the number of targets on the screen for each trace (targets/non targets) and the probability (×100) that the stimulus will intensify. Note that relevant stimuli elicit a large P300. The amplitude of the P300 is not affected by the probability that the events will occur. However, the latency is clearly affected by the number of items on the screen and by the probability. (After Heffley et al., 1978)

If one ignores the fact that targets are moving on the screen, then the display can be described as a sequence of flashes, some of triangles, some of squares, that occur with a certain probability, much like in the oddball paradigm. Yet, the probability of the targets had no effect on the *amplitude* of the P300, even though the probabilities did affect the latency of the P300, as seen in Figure 13. Very rare events elicited a P300 with a longer latency but with the same amplitude as P300 elicited by frequent targets. How come? Why did probability fail to work in this experiment? Hard as we tried, we could not ignore the problem.

Perhaps the fact that the targets were moving on the screen made the subjects aware of the position of all the targets. The subjects were thus able to ignore the non-targets so that it did not matter how many non-targets there were on the screen and the relative frequency of the targets played no role. But, this explanation was ruled out when Heffley conducted an experiment using the "periodic display" of Figure 12. The subject was presented with two experimental conditions. The "continuous" condition replicated the experiment described above. Triangles and squares were moving on the screen; probability and task relevance were varied. The results duplicated the results of the previous experiment. In the second condition, the trajectories of the stimuli were blanked and stimuli appeared on the screen only when one of the targets was intensified. We reasoned that the subject would not be able, in this condition, to block out the irrelevant events, since it was not possible to predict where they would appear. We predicted that in this condition the probability effect would be observed. This prediction was *not* confirmed. Again, the relevant events elicited a much larger P300 than did the irrelevant events. But the probability of the event did not affect the amplitude of the P300 it elicited.

These results are troublesome for our edifice. We know that if we present triangles and squares in the middle of the screen, the triangles, appearing with a probability of .20, will elicit a large P300. This is the standard "oddball" paradigm. And yet, when the stimuli appear in random loci on the screen the probability effect is diminished. Heffley tried a variety of manipulations in an attempt to determine why the probability effect is diminished. It turns out that the variable that determines whether or not the probability of the stimulus will affect P300 amplitude is the interval between the stimuli (see Heffley & Donchin, Note 7).

The data suggest that probability played no role in the experiment in which targets moved on the

screen because the intervals we allowed between stimuli were as long as 9 sec. The probability effect seems to appear only if we reduce the interval between stimuli to about 1500 or 2000 msec (McCarthy & Donchin, 1976). It is clear that we cannot build our P300 arch on a solid column of subjective probability. The column is cracked in places.

The Antecedents of the P300

Let me then review what we know about the antecedent conditions for a P300. A primitive concept is that of an event. It stands to reason that if we study *event*-related potentials we must have a definition of an event. We can easily provide examples of events. A stimulus is an event. A high tone, a low tone, the appearance of a number on a screen, the movement of a branch as the wind rustles through the tree, all are events. Yet, the situation is more complex. Because what we define as events depends on the level of the description we apply to the environment. Consider, for example, the bargaining experiment described above. We know that we trigger the computer at a given instant. The ERP activity shown in Figure 3 can be defined solely as it relates to the occurrence of a trigger pulse in the computer. But that will not do. We are not interested in trigger pulses. We could describe the event entirely from the perspective of the physicist. The trigger pulse was activated in synchrony with the illumination of a set of plasma dots. We can also say that "*a number*" appeared on the screen. The dots were in a well-defined pattern recognized by the subject as the number "65." If one also takes into consideration what the subject knows about the experiment, we can call the appearance of this number a "*computer concession*." Other numbers appeared on the screen, all physically identical with the "computer concession" but they did not play the same role in the subject's task. Furthermore, if we take into consideration the specific sequence of occurrences in the immediate local neighborhood in which a given number appeared, we must distinguish between a computer concession made within the framework of the strategy of the computer as perceived by the subject or a computer concession that deviates from the subject's conception of the computer's strategy. We then call the P300-eliciting events "*switches*."

The definition of an event appears to depend on a rather high level of abstraction. We must take into account the subject's strategies, the instructions given the subject, the course of the experiment, the specific expectations the subject has with respect to the immediately following stimuli, the subject's recall, and the income she has thus far obtained, as well as her interest in the experiment. Thus, the very definition of "an event" in this context calls for a large number of assumptions. It, in fact, holds within itself the germs of a theory. Much of the work we have done in the past decade, was, in fact, an attempt to show that the definition of the events to which the P300 is related must include a description of the psychological rather than the physical situation.

Let us assume that we have agreed on a definition of the "events" and that we have a series of such events. Let us further assume that we can agree which events are crucial. We can say that of the class of all events that can occur at a given instant in a given situation, some are more relevant to the task the subject is performing. I define task relevance here in terms of the ability of the stimulus to resolve subject uncertainties, uncertainties that must be resolved if the subject is to perform the assigned tasks correctly and efficiently; this, of course, depends, as Sutton (1969) has pointed out, more on the subject than on the experimenter. The stimulus is relevant to the extent that the subject is extracting information from the stimulus (Johnson & Donchin, 1978). Information is, by definition, a function of the subject's uncertainties. The subject's uncertainties depend, among other things, on what it is that the subject is attempting to do at any given time, or on how much prior information the subject has. The data show that the stimulus must be task relevant if P300 is to be elicited. Secondly, the data are quite strong in supporting the assertion that for any given level of task relevance, if stimulus probability is varied, the amplitude of P300 will also vary. The rarer the stimulus, the larger the P300. This, however, is not always the case. On occasion, highly probable stimuli will elicit a P300 and, on occasion, fairly irrelevant stimuli of extreme rarity will elicit a large P300.

The P300 and the Orienting Reflex

It is time to return, then, to the challenge I accepted earlier in this paper, to describe the subroutine of human information processing that is likely to be undertaken by the P300 process. A statement that has often been made about the P300 holds promise. It has not escaped people's notice (Ritter, Vaughan, & Costa, 1968; Sutton et al., 1965) that the attributes of the P300 are remarkably similar to the attributes that one often finds enumerated for a phenomenon that has been quite dear to the hearts of psychophysiologists. I refer, of course, to the orienting reflex (OR) (Kimmel, Van Olst, & Orlebeke, 1979; Graham, 1979; Naatanen, 1979) Events defined with the complexity that is very similar to the complexity required for defining P300

eliciting events, events that are task relevant and improbable, tend to elicit the OR. I will not define the orienting reflex or review here the literature on the OR or elaborate on the history of the study of this phenomenon. I think, however, that there is general agreement that the attributes characterizing the OR are remarkably similar to the corresponding attributes of the P300. There is only one attribute in which the two are somewhat discrepant. This has, in fact, led many to believe there can be no relationship between the P300 and the OR. I refer to the rapid habituation presumed to take place in the OR; habituation that is not commonly observed for the P300. Many investigators argue that it is a characteristic property of ORs that they habituate. After a few presentations of the critical stimuli the response disappears. In fact, the rate of habituation is one of the most studied attributes of the OR. On the other hand, P300 can be elicited repeatedly with no obvious habituation.

Is this an insurmountable obstacle to the incorporation of the P300 results within the same framework that accounts for the OR? I think not. Several investigators have noted (see reviews in Kimmel et al., 1979) that the concept of the habituating OR is, to an extent, an artifact of the experiment that has been most frequently done with ORs. More often than not, in studies of the OR, the eliciting stimuli had no significance. They were events that carried no information to the subject. There was little the subject had to do about the stimuli. Under this set of circumstances, P300 will also habituate very rapidly. It is common laboratory folklore that the very first stimuli presented in any series elicit a P300, whether they are frequent, rare, relevant or irrelevant. What is even more striking, however, is the fact that the rapid habituation may not be a necessary attribute of the OR. Luria and Homskaya (1970) have shown that stimuli that have what they call "signal value," that is, stimuli that I would call "task relevant," continue to elicit an OR without much habituation. Their experiment was very similar in structure to the oddball paradigm. So the correspondence between the conditions antecedent to a P300 and the conditions antecedent to an OR are quite similar in all respects.

Does this mean that P300 is a "component" of the orienting reflex (Öhman, 1979)? Yes and no. It depends on the meaning of the phrase "a component of." It has been suggested that, like the OR, the P300 is an output of a mismatch detector, a detector that identifies the occurrence of mismatches between a neuronal model and a physical stimulus eliciting the ERP (Bauer & Squires, 1972; Thatcher; 1977). This position, especially in its early incarnations, tended to emphasize the *physical* mismatch between stimuli and ignored the rather elaborate complexity of the events required to elicit a P300. But, a mismatch between a computer concession that *is* a "switch" and a computer concession that is not a "switch" in our bargaining experiment is clearly not a mismatch between constellations of physical attributes. It depends on a complex integration of the information that exists in the subject's head about the situation. The phrase "the information that exists in the subject's head about the ongoing environment" is often replaced by phrases like "neuronal model," or "world model," "schema," "map" or "representations." The idea of a schema, or a neuronal model, is, of course, central to the concept of the OR (Sokolov, 1969). The OR is thought to result from a mismatch between external events and the schema. Once such a mismatch is detected, an OR may, or may not, be elicited. Whether or not the OR will be activated depends, of course, on the importance attributed to the mismatch. Some mismatches are more violent than others.

The OR has many "components." What is common to OR components is that they are what I called elsewhere "tactical responses" (Donchin et al., 1978). Upon detection of a mismatch, action must be taken. The subject must be ready to react to the external events. Heart rate changes, galvanic skin response can be recorded, pupils dilate, a host of autonomic activities are activated (Van Olst, Heemsha, & Ten Kortenaar, 1979). All are action-oriented, *tactical* responses. Is the P300 one of those? I think not. I think not because it seems clear that the P300 is not related to the actions the subject is taking in response to the stimulus eliciting the P300. I will not review here the evidence for this assertion, as I have discussed it in detail elsewhere (Donchin, 1975; Donchin, 1979; Donchin & Isreal, 1980a, 1980b; Donchin et al., 1978). For the present, I shall assume that the process manifested by the P300 is not elicited for the purpose of tactically responding to a given stimulus in a given trial, but rather to what I called strategic information processing. This is the information processing that will affect the manner in which we respond to *future* stimuli. It is information processing that is undertaken for the purpose of evaluating expectancies, shifting strategies, etc. All of these can, of course, be subsumed under one rubric. *These are activities that affect our schema rather than our actions.*

A curious lacuna in the literature on the OR is that while it very often assumes the existence of the neuronal model, it generally does not explore how the neuronal model is built and what are the effects on the neuronal model of the detection of mismatches. Of course, the habituation process modu-

lates the neuronal model. As the relevant stimuli are repeated a few times, something is adjusted in the system that determines that no further reaction is required. Pribram and McGuiness (1975) referred to this process as "context updating."

A reasonable interpretation of the literature on P300 is that the process manifested by P300 is associated with the updating of the schema. The schema may be conceptualized as a large and complex map representing all available data about the environment. It is the reservoir of information that is necessary for performing whatever tasks require active processing at any time. Some of the information has just been delivered and may still reside in various stages of dynamic memory. Other information resides in longer term memory and may need to be made available, on a temporary basis, for integration into the overall schema. The system is quite fluid. There must be some priority weighting system that is associated with the schema and determines which of its aspects is related to which tasks. Representations decay because of misuse or because of shifting strategies and tasks. New information is brought in. Choices are made in the process of using this schema. When there is a need, the model is revised by building novel representations through the incorporation of incoming data into schema based on long-term memory data. It is likely that it is this updating process that we see manifested by the P300. This view is consistent with the strong effect of probability on P300. As it is indeed the unexpected (but relevant) that dictates revisions of the neuronal model. Also comprehensible is the fact that probable but highly relevant events do sometimes elicit a P300, as well as the failure of irrelevant events to do so. It is the degree to which the event requires a revision of the model, not its inherent attributes, that is the crucial determinant of P300.

Summary and Conclusions: The P300 and Memory

Does this concept of P300 qualify as a theory? Not in the degree of detail needed. Much remains to be elaborated, considered, and concretized. The suggestion, however, is valuable in pointing to a class of experiments that will allow us to consider not only the antecedent conditions for a P300 but also its consequences. The model I propose assumes that the P300 is intimately involved with the process of memory modification or, if you will, learning. This is, of course, quite consistent with the Zeitgeist. In the literature on the "Classical Conditioning" or the traditional literature on animal learning (Medin, Roberts, & Davis, 1976), the concept of surprise has virtually replaced the concept of reinforcement. Things appear to be learned if, and only if, they are surprising. In the neurophysiological literature, we find increasing emphasis on the role of the norepinephrine system in the incorporation of memories (Gold & Van Buskirk, 1975, 1978; Gold & Sternberg, 1978; Sternberg & Gold, 1980). Things are apparently learned if, and only if, they activate this system. Contemporary views of learning place considerably more emphasis on learning as the reorganization of storage processes rather than on learning as a process of strengthening of associations. As much else in modern cognitive psychology, we find that the concepts have been elaborated with elegance and thoughtfulness by McDougal (1920). In the chapter on Memory, in his book on "Body and Mind," McDougal criticizes the tradition began by Ebbinghaus in the study of memory. McDougal does not begrudge Ebbinghaus the importance of his studies. He feels, however, that the study of the learning of nonsense syllables is concerned with "habit" not with memory. To McDougal, habit is the strengthening of neural associations. Memory he considers to be something quite different. He gives, among others, the following example:

> "A number, say ten, points of light are thrown simultaneously, for a small fraction of a second, upon a screen, and I am required to draw a map of the points. If the spots are irregularly distributed I find this quite impossible to achieve; and perhaps it is necessary to repeat the flash now thirty to fifty times before I can succeed in constructing a tolerably correct map." On the other hand "My eye rests for a moment on a photograph or drawing of a striking face that is unknown to me. The drawing consists of a great number of points, lines, and areas, arranged in extremely complex fashion, yet after a brief glance I am able to picture the face with considerable accuracy, perhaps even after the lapse of a day or a month." (pp. 337–338)

For years to come, McDougal points out, he may recall in vivid detail impressions ingrained in a fleeting instant. It is this representation that McDougal calls memory. This view of memory has been stated vividly by the anonymous author of the treatise on memory *Ad Herenium* published in 82 BC. This author is quoted by Yates (1969) as saying:

> "When we see in every day life things that are petty, ordinary, and banal, we generally fail to remember them, because the mind is not being stirred by anything novel or marvelous. But if we see or hear something exceptionally base, dishonourable, unusual, great, unbelievable, or ridiculous, that we are likely to remember for a long time. Accordingly, things immediate to our eye or ear we commonly forget; incidents of our childhood we often remember best. Nor could this be so for any other reason than that ordinary things easily slip from the memory while the striking and the novel stay longer in the mind."

There is a clear distinction between the memorable and the mundane. But what is the distinction?

What is it about events that makes some remarkable and others occur without leaving a trace? It is not unreasonable to suggest that events are remembered if they require, upon their occurrence, a restructuring of our mental models (which is presumably what happens when we are surprised). It is conceivable that these are the processes that are manifested by the P300.

One can make, therefore, very specific predictions about the *consequences* of a P300. It is that events that elicit a P300 are remembered better than events that do not elicit a P300. A strong version of this hypothesis is that the memorability of events will be proportional to the amplitude of the P300 they elicit. Ted Bashore, Demetrios Karis, Monica Fabiani, and I are now testing this prediction. Our experimental paradigm is shown in Figure 14. The subject is presented, during a test phase, with a series of words displayed, one at a time, on a computer terminal. Half the words are displayed at a somewhat dimmer intensity than the other half. The subject is told that he has to memorize the dim words (or, for other subjects, the bright words). After the presentation is completed the subject, after a brief rest period, is tested again with a "test" series of words. Half of the words appearing in the test phase were presented in a previous list as words to-be-remembered. Half of the remaining words were to-be-ignored in the training phase. The rest were brand new words. Of course, all words appear with the same intensity in the test phase. The subject's task is to identify those words he was instructed to memorize. After this recognition test we test the subject's recall of the words, by asking for a list of all words he recalls. We predicted that the larger the amplitude of the P300 elicited by a word *in the training phase*, the larger the probability that the word would be recognized in the test phase. Data from both the training phase and the test are shown in Figure 15. The solid line represents the ERPs elicited by words that the subject was supposed to memorize. The dotted line shows the ERPs elicited by words that the subject did not have to memorize. The probability was equal that a word would be either bright or dim. As there is almost no difference between the ERP elicited by the bright and that elicited by the dim words, we could not test our prediction using these data. However, note that in

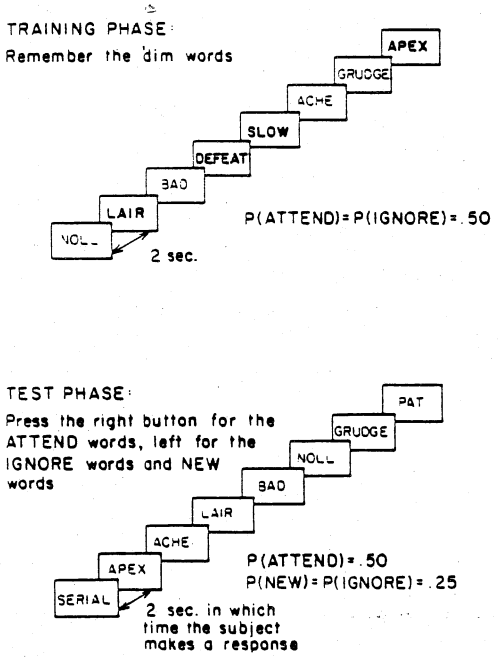

Figure 14. Design of the Memory study. In the test phase the subject is presented with a series of words, either bright or dim, and is instructed to memorize the bright or the dim words. In the test phase the to-be-remembered words are presented with "new" words and some of the to-be-ignored words. Following these sessions the subject's recall for the words is tested.

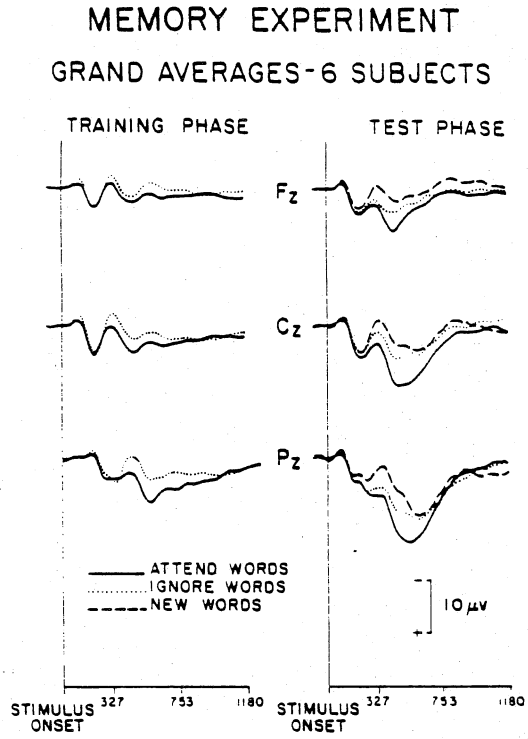

Figure 15. The ERPs, averaged over subjects, in the training and test phases of the Memory study. Note absence of difference between "attend" and "ignore" words in the training phase and the larger P300 elicited by the attend words in the test phase.

the test phase there were large differences between the ERPs elicited by the words that were dim in the training phase and words that were bright in the training phase. This effect cannot be attributed to intensity as, in this phase, all words were presented with equal intensity and with equal probability. A partial test of our hypothesis could, therefore, be attempted. As the to-be-remembered words did elicit a large P300 in the test phase, we could evaluate the extent to which the amplitude of the P300 elicited by an item in the test phase was correlated with the subsequent recall of that item. We sorted trials in the test phase depending on whether or not these items were actually recalled in the recall test. The results of this analysis averaged over the 6 subjects are shown in Figure 16. The solid line is the ERP elicited during the *test phase* by words that were recalled in a post experimental test debriefing of the subject. The dashed line is the ERP elicited by words that were not recalled. Note that all words used to elicit the ERPs displayed in Figure 16 were presented as "to-be-remembered" in the training phase, and elicited a substantial P300 during the test phase. However, the amplitude of this P300 varies, and this variation appears to be related to the items' subsequent recall.

Of course, these data are but preliminary and the results are inconclusive. We are encouraged by the pattern of the data, but cannot take these as confirming the hypothesis that P300 amplitude predicts recall, or recognition. In current studies we are attempting with a number of experimental manipulations to assure that a P300 is elicited by the to-be-remembered items in the training phase. We expect to focus on this work for the next few years. The point at this time is that studies such as these are required. If we are to understand the functional significance of P300, we need to ascertain the *consequences* of the P300.

Elucidating the functional significance of the P300 is one of the important tasks for Cognitive Psychophysiology. The P300 is a manifestation of brain processes elicited reliably with a narrow and known set of antecedent conditions. The time, I think, is past when work that merely adds features to the story, that repeats additional descriptions of antecedent conditions, is sufficient. We must, I think, accept the challenge to develop a strong, sound, and solid theory of the P300 and develop the empirical evidence that can support or reject such a theory. The emphasis must be, therefore, on elucidating the consequences of the P300. Those who will be able to provide answers to this question will have made an important contribution to psychophysiology.

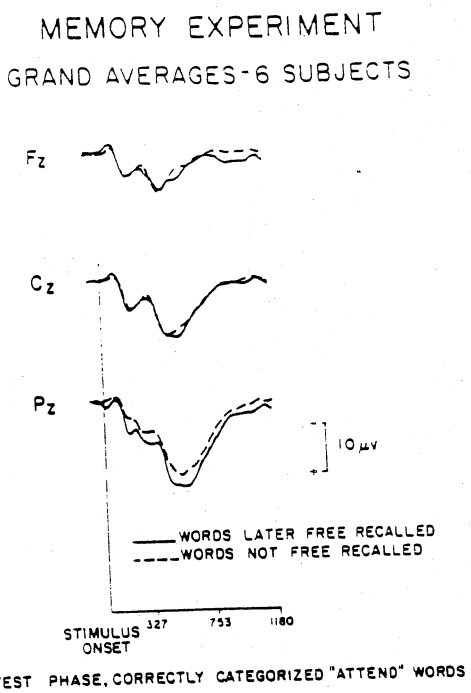

Figure 16. A reaveraging of the data shown as solid lines (attend words) in Figure 15 for the test phase. The data were sorted prior to reaveraging according to the words that were subsequently recalled in the recall test.

REFERENCES

Andreassi, J. L. *Psychophysiology: Human behavior and physiological response*. New York: Oxford University Press, 1980.

Baty, D. Human transformation rates in one to four axis tracking. *7th Annual Conference on Manual Control*. NASA S. P. No. A281, 1971.

Bauer, J. W., & Squires, K. C. Computer signal detection by monitoring auditory evoked potentials. *Perception and Psychophysics*, 1972, *11*, 301–308.

Biederman, I., & Kaplan, R. Stimulus discriminability and SR compatibility: Evidence for independent effects on choice reaction time. *Journal of Experimental Psychology*, 1970, *86*, 434–439.

Broadbent, D. E., & Gregory, M. Donder's B and C reactions and SR compatibility. *Journal of Experimental Psychology*, 1962, *63*, 575–578.

Courchesne, E., Hillyard, S. A., & Courchesne, R. Y. P3 waves to the discrimination of targets in homogeneous and heterogeneous stimulus sequences. *Psychophysiology*, 1977, *14*, 590–597.

Desmedt, J. E., & Debecker, J. Waveform and neural mechanisms of the decision P350 elicited without prestimulus

CNV or readiness potential in random sequences of near-threshold auditory clicks and finger stimuli. *Electroencephalography & Clinical Neurophysiology*, 1979, *47*, 648–670.

Donchin, E. Brain electrical correlates of pattern recognition. In G. F. Inbar (Ed.), *Signal analysis and pattern recognition in biomedical engineering*. New York: John Wiley, 1975. Pp. 199–218.

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., & Isreal, J. B. Event-related brain potentials: Approaches to cognitive psychology. In R. E. Snow, P. A. Federico, & E. W. E. Montague (Eds.), *Aptitude, learning, and instruction: Cognitive process analyses* (Vol. 2). Hillsdale, NJ: Lawrence Erlbaum Associates, 1980. Pp. 47–82. (a)

Donchin, E., & Isreal, J. B. Event-related potentials and psychological theory. In H. H. Kornhuber & L. Deecke (Eds.), *Motivation, motor and sensory processes of the brain: Electrical potentials, behavior, and clinical use*. Progress in Brain Research. Amsterdam: Elsevier-North Holland, 1980. Pp. 697–715. (b)

Donchin, E., Kubovy, M., Kutas, M., Johnson, R. Jr., & Herning, R. I. Graded changes in evoked response (P300) amplitude as a function of cognitive activity. *Perception and Psychophysics*, 1973, *14*, 319–324.

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

Duncan-Johnson, C. C. P300 latency: A new metric of information processing. *Psychophysiology*, 1981, *18*, 207–215.

Duncan-Johnson, C. C., & Donchin, E. On quantifying surprise: The variation in event-related potentials with subjective probability. *Psychophysiology*, 1977, *14*, 456–467.

Ekman, P., & Friesen, W. V. *Unmasking the face: A guide to recognizing emotions from facial clues*. Englewood Cliffs, NJ: Prentice Hall, 1975.

Fitts, P. M., & Deininger, R. L. SR compatibility: Correspondence among paired elements within stimulus and response codes. *Journal of Experimental Psychology*, 1953, *46*, 199–210.

Frowein, H. W., & Sanders, A. F. Effects of visual stimulus degradation, SR compatibility and foreperiod duration on choice reaction time and movement time. *Bulletin of the Psychonomic Society*, 1978, *12*, 106–108.

Fuchs, A. The progression-regression hypothesis in perceptual-motor skill learning. *Journal of Experimental Psychology*, 1962, *63*, 177–182.

Goff, W. R., Allison, T., Williamson, P. D., & Van Gilder, J. C. Scalp topography in the localization of intracranial evoked potential sources. In D. Otto (Ed.), *Multidisciplinary perspectives in event-related brain potential research*. EPA-60019-77-043, Washington, D. C.: U. S. Government Printing Office, 1979. Pp. 526–532.

Gold, P. E., & Sternberg, D. B. Retrograde amnesia produced by several treatments: Evidence for a common neurobiological mechanism. *Science*, 1978, *201*, 367–368.

Gold, P. E., & Van Buskirk, R. Facilitation of time-dependent memory processes with posttrial epinephrine injections. *Behavioral Biology*, 1975, *13*, 145–153.

Gold, P. E., Van Buskirk, R. Posttraining brain norepinephrine concentrations: Correlation with retention performance of avoidance training and with peripheral modulation of memory processing. *Behavioral Biology*, 1975, *13*, 509–520.

Graham, F. K. Distinguishing among orienting, defense, and startle reflexes. In H. D. Kimmel, E. H. Van Olst, & J. F. Orlebeke (Eds.), *The orienting reflex in humans*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1979. Pp. 137–168.

Granit, R. *The purposive brain*. Cambridge, MA: MIT Press, 1977.

Heffley, E., Wickens, C., & Donchin, E. Intramodality selective attention and P300—Reexamination in a visual monitoring task. *Psychophysiology*, 1978, *15*, 269–270. (Abstract)

Hillyard, S. A., & Picton, T. W. Event-related brain potentials and selective information processing in man. In J. Desmedt (Ed.), *Cerebral-evoked potentials in man: The Brussels Symposium*. London: Oxford University Press, 1977. Pp. 1–52.

Isreal, J. B., Chesney, G. L., Wickens, C. D., & Donchin, E. P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. *Psychophysiology*, 1980, *17*, 259–273. (a)

Isreal, J. B., Wickens, C. D., Chesney, G. L., & Donchin, E. The event-related brain potential as an index of display-monitoring workload. *Human Factors*, 1980, *22*, 212–224. (b)

Isreal, J. B., Wickens, C. D., & Donchin, E. The event-related brain potential as a selective index of display load. In C. K. Bensel (Ed.), *Proceedings of the Human Factors Society 23rd Annual Meeting*. Santa Monica, CA: Human Factors Society, 1979. Pp. 558–562.

Johnson, R. E. Jr., & Donchin, E. On how P300 amplitude varies with the utility of the eliciting stimuli. *Electroencephalography & Clinical Neurophysiology*, 1978, *44*, 424–437.

Johnson, R. E. Jr., & Donchin, E. P300 and stimulus categorization: Two plus one is not so different from one plus one. *Psychophysiology*, 1980, *17*, 167–178.

Kimmel, H. D., Van Olst, E. H., & Orlebeke, J. F. *The orienting reflex in humans*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1979. Pp. 425–442.

Kutas, M., & Donchin, E. Variations in the latency of P300 as a function of variations in semantic categorization. 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. 198–201.

Kutas, M., McCarthy, G., & Donchin, E. Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science*, 1977, *197*, 792–795.

Luria, A. R., & Homsakaya, E. D. Frontal lobes and the regulation of arousal processes. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis*. New York: Appleton-Century Crofts, 1970. Pp. 303–330.

McCarthy, G. *The P300 and stages of human information*

processing: An additive factors study. Dissertation thesis, University of Illinois at Urbana-Champaign, 1980.

McCarthy, D., & Donchin, E. The effects of temporal and event uncertainty in determining the waveforms of the auditory event-related potential (ERP). *Psychophysiology*, 1976, *13*, 581–590.

McCarthy, G., & Donchin, E. Brain potentials associated with structural and functional visual matching. *Neuropsychologia*, 1978, *16*, 571–586.

McCarthy, G., & Donchin, E. A metric for thought: A comparison of P300 latency and reaction time. *Science*, 1981, *211*, 77–80.

McDougal, W. *Body and mind, a history and defense of animism.* (5th ed.) London: Methuen and Co., 1920.

McRuen, D. Human dynamics in man-machine systems. *Automatica*, 1980, *16*, 237–253.

Medin, D. L., Roberts, W. A., & Davis, R. T. (Eds.) *Processes of animal memory.* Hillsdale, NJ: Lawrence Erlbaum, 1976.

Näätänen, R. Orienting and evoked potentials. In H.D. Kimmel, E. H. Van Olst, & J. F. Orlebeke (Eds.), *The orienting reflex in humans.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1979. Pp. 61–76.

Navon, D., & Gopher, D. Interpretations of task difficulty in terms of resources: Efficiency, load, demand, and cost composition. In R. Nickerson and R. Pew (Eds.), *Attention and performance VIII.* Englewood Cliffs, NJ: Lawrence Erlbaum Associates, 1980. Pp. 297–318.

Öhman, A. The orienting response, attention and learning: An information processing perspective. In H. D. Kimmel, E. H. Van Olst, & J. F. Orlebeke (Eds.), *The orienting reflex in humans.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1979. Pp. 443–471.

Polich, J. M., Vanasse, L., & Donchin, E. Category expectancy and the N200. *Psychophysiology*, 1981, *18*, 142. (Abstract)

Pribram, K. H., & McGuiness, D. Arousal, activation, and effort in the control of attention. *Psychological Review*, 1975, *82*, 116–149.

Pritchard, W. The psychophysiology of P300. *Psychological Bulletin*, 1981, *89*, 506–540.

Ritter, W., Vaughan, H. G. Jr., & Costa, L. P. Orienting and habituation to auditory stimuli: A study of short-term changes in average evoked responses. *Electroencephalography & Clinical Neurophysiology*, 1968, *25*, 550–556.

Roth, W. T. How many late positive waves are there? 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. 170–172.

Schneider, W., & Shiffrin, R. M. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 1977, *84*, 1–66.

Schvaneveldt, R. W., & Chase, W. G. Sequential effects in a choice reaction time task. *Journal of Experimental Psychology*, 1969, *80*, 1–8.

Shwartz, S. P., Pomerantz, J. R., & Egeth, H. E. State and process limitations in information processing: An additive factors analysis. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, *3*, 402–410.

Sokolov, E. N. The modeling properties of the nervous system. In I. Maltzman & K. Cole (Eds.), *Handbook of contemporary soviet psychology.* New York: Basic Books, 1969. Pp. 671–704.

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 & Clinical Neurophysiology*, 1977, *42*, 1–14.

Squires, K. C., Wickens, C., Squires, N. K., & Donchin, E. The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science*, 1976, *193*, 1142–1146.

Sternberg, D. B., & Gold, P. E. Effects of L and B adrenergic receptor antagonists on retrograde amnesia produced by frontal cortex stimulation. *Behavioral and Neural Biology*, 1980, *29*, 289–302.

Sternberg, S. The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, 1969, *30*, 276–315.

Sutton, S. The specification of psychological variables in an average evoked potential experiment. In E. Donchin & D. B. Lindsley (Eds.), *Average evoked potentials: Methods, results, and evaluations.* NASA SP-191. Washington, D. C.: U. S. Government Printing Office, 1969. Pp. 237–297.

Sutton, S. P300—Thirteen years later. In H. Begleiter (Ed.), *Evoked brain potentials and behavior.* New York: Plenum Press, 1979. Pp. 107–126.

Sutton, S. Braren, M., Zubin, J., & John, E. R. Evoked potential correlates of stimulus uncertainty. *Science*, 1965, *150*, 1187–1188.

Thatcher, R. W. Evoked potential correlates of hemispheric lateralization during semantic information-processing. In S. Harnad, R. W. Doty, L. Goldstein, J. Jaynes, & G. Krauthammer. *Lateralization in the nervous system.* New York: Academic Press, 1977. Pp. 429–449.

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-600/9-77-043. Washington, D. C.: U. S. Government Printing Office, 1979. Pp. 159–169.

Uttal, W. R. Emerging principles of sensory coding. *Perspectives in Biology and Medicine*, 1969, *12*, 344–368.

Van Olst, E. H., Heemsha, M. L., & Ten Kortenaar, T. Stimulus significance and the orienting reaction. In H. D. Kimmel, E. H. Van Olst, & J. F. Orlebeke (Eds.), *The orienting reflex in humans.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1979. Pp. 521–547.

Vaughan, H. G. Jr. The analysis of scalp-recorded brain potentials. In R. F. Thompson & M. M. Patterson (Eds.), *Bioelectrical recording techniques. Part B. Electroencephalography and human brain potentials.* New York: Academic Press, 1974. Pp. 157–207.

Wickens, C. D. The structure of processing resources. In R. Nickerson (Ed.), *Attention and performance VIII.* Englewood Cliffs, NJ: Lawrence Erlbaum Associates, 1980. Pp. 239–258.

Wickens, C. D., Isreal, J., & Donchin, E. The event-related cortical potential as an index of task workload. In A. S. Neal & R. F. Palasek (Eds.), *Proceedings of the Human Factors Society 21st Annual Meeting.* San Francisco: Human Factors Society, 1977. Pp. 282–287.

REFERENCE NOTES

1. Karis, D., Druckman, D., Lissak, R., & Donchin, E. *A psychophysiological analysis of bargaining: ERPs and facial expressions*. Manuscript in preparation, 1981.
2. McCarthy, D., Kutas, M., & Donchin, E. *The latency of P300 in choice reaction time studies*. Manuscript in preparation, 1981.
3. Polich, J. M., & Donchin, E. *The word frequency effect: A psychophysiological analysis*. Manuscript in preparation, 1981.
4. Wickens, C. D., & Tsang, P. *Attention allocation in dynamic environments* (Tech. Rep. EPL-79-3/AFOSR-79-3). Champaign, IL: University of Illinois, Engineering-Psychology Technical Report, June 1979.
5. Kramer, A., Wickens, C. D., & Donchin, E. *The event-related potential as an index of processing demands of a complex task*. Manuscript in preparation, 1981.
6. Heffley, E. *Elements of a neural theory of human information processing derived from analysis of cognitive ERPs*. Manuscript in preparation, 1981.
7. Heffley, E., Wickens, C. D., & Donchin, E. *Psychophysiological analyses of display monitoring*. Manuscript in preparation, 1981.

Announcements

Twenty-First Annual Meeting Society for Psychophysiological Research

From October 29th through November 1st, 1981, the Twenty-First Annual Meeting of the Society for Psychophysiological Research will be held at the Sheraton-Washington Hotel, in Washington, D.C.

Information regarding submission of papers and the deadline for abstracts may be obtained from: Dr. Michael Dawson, Camarillo/UCLA Research Center, Box A, Camarillo, CA 93010.

Registration information may be obtained from: Dr. Robert J. Gatchel, Department of Medical Psychology, U. S. University of the Health Sciences, 4301 Jones Bridge Road, Bethesda, MD 20014.

Ninth Annual Scientific Meeting Psychophysiology Society (London)

From December 16th through 18th, 1981, the Ninth Annual Scientific Meeting of the Psychophysiology Society will be held at St. George's Hospital Medical School, London. The Conference will consist of formal paper sessions, workshops (event-related potentials and information processing, sleep, psychophysiology of relaxation training, measures of blood pressure and myocardial contraction), symposia (psychopharmacology, clinical applications of evoked potentials, behavioural medicine), and science fairs.

For information on the submission of papers and deadlines, contact: Dr. D. Papakostopoulos, Burden Neurological Institute, Stoke Lane, Stapleton, Bristol, BS16 IQT. For details of conference registration and accommodation, contact: Dr. A. Steptoe, Department of Psychology, St. George's Hospital Medical School, Cramner Terrace, Jenner Wing, Tooting, London, SW17.