Biological Psychology 26 (1988) 3–67
North-Holland

MODERN MENTAL CHRONOMETRY *

David E. MEYER **

Department of Psychology, University of Michigan, Ann Arbor, MI 48109, U.S.A.

Allen M. OSMAN

Department of Psychology, University of California at San Diego, La Jolla, CA 92093, U.S.A.

David E. IRWIN

Department of Psychology, Michigan State University, East Lansing, MI 48824, U.S.A.

Steven YANTIS

Department of Psychology, Johns Hopkins University, Baltimore, MD 21218, U.S.A.

Mental chronometry, in which conclusions about human information processing are reached through measures of subjects' reaction time, has contributed substantially to studies of cognition and action. During the evolution of the chronometric paradigm, several key issues have emerged. The issues concern (a) the existence of separable processing stages, (b) the degree to which various stages of processing produce partial outputs before they are completed, and (c) the discrete versus continuous form of the outputs. To obtain added temporal resolution, new reaction-time procedures have been developed, including special response-priming and speed-accuracy decomposition techniques that focus on quantitative patterns of reaction-time distributions and error rates. The present article summarizes these developments, starting with a historical review of chronometric research and proceeding to a survey of recent empirical and theoretical innovations. We also discuss the relevance and potential future impact of complementary work by cognitive psychophysiologists on event-related brain potentials and other physiological variables.

1. Introduction

As the scientific study of human information processing has progressed from the middle of the nineteenth century to the present day, experimental

* This article is based on an invited tutorial presentation given at the Fourth International Conference on Cognitive Neuroscience, Dourdan, France, June, 1987. Support for portions of the research reported herein was provided by grant R01 MH 38845 from the National Institute of Mental Health to The University of Michigan, David E. Meyer, Principal Investigator. We thank Sylvan Kornblum, John Kounios, Jeff Miller, and J.E. Keith Smith for their encouragement and many helpful suggestions. The article has also benefited greatly from editorial comments by Michael G.H. Coles, William Gehring, and two anonymous reviewers.

** Correspondence regarding this article and requests for reprints should be addressed to: David E. Meyer, Human Performance Center, Department of Psychology, University of Michigan, 330 Packard Rd., Ann Arbor, MI, 48104, U.S.A.

psychologists have focused considerable attention on the dynamics of cognition and action. In this endeavor, it has been frequently assumed that mental processes are manifested through certain behavioral measures, including subjects' reaction time (RT), response accuracy, and speed-accuracy tradeoffs (Lachman, Lachman, & Butterfield, 1979; Pachella, 1974; Posner & McLeod, 1982; Smith, 1968; Taylor, 1976; Woodworth, 1938). To obtain and analyze such measures, numerous empirical procedures and theoretical models have been developed (e.g., Luce, 1986; McClelland, 1979; Meyer, Irwin, Osman, & Kounios, 1988; Miller, 1982; Sternberg, 1969; Townsend & Ashby, 1983). On the basis of them, extensive inferences have been made concerning the nature of sensation, perception, memory, attention, language, reasoning, problem solving, decision making, and movement control. The overall paradigm is often referred to as *mental chronometry* (Posner, 1978).

Complementing the chronometric paradigm, psychophysiologists have sought additional latent indicators of the mental and physical processes that mediate overt performance (Coles, Donchin, & Porges, 1986). The latter research, which falls under the rubric *cognitive psychophysiology* (Donchin, 1981, 1984), has led to studies involving event-related brain potentials (ERPs), electromyographic (EMG) activity, and other somatic variables. Some results of these studies suggest that the psychophysiological approach can significantly clarify and extend conclusions about information processing reached from reaction-time and response-accuracy data (e.g., Coles, Gratton, Bashore, Eriksen, & Donchin, 1985; Duncan-Johnson & Donchin, 1982; Ford, Roth, Mohs, Hopkins, & Kopell, 1979; Gaillard & Verduin, 1985; Kutas & Hillyard, 1980; Kutas, McCarthy, & Donchin, 1977; Marsh, 1975; McCarthy & Donchin, 1981; Mulder, Gloerich, Brookhuis, Van Dellen, & Mulder, 1984; Renault, Ragot, Lesevre, & Remond, 1982; Ritter, Simson, & Vaughan, 1983; Ritter, Simson, Vaughan, & Macht, 1982). At the same time, lessons learned in mental chronometry may guide the psychophysiological approach, for example, facilitating the interpretation of ERP components. Consequently, a growing wave of enthusiasm has arisen over the wedding of cognitive psychophysiology and the chronometric paradigm (Gaillard & Ritter, 1983; Hillyard & Kutas, 1983; Ritter, Vaughan, & Simson, 1983; Vaughan & Ritter, 1973; Wilkinson, 1967).

The present article provides a concise introduction to the chronometric paradigm and an assessment of the prospects for its marriage with cognitive psychophysiology. It is intended to be a synopsis for a general audience of experimental psychologists, psychophysiologists, and cognitive scientists, including ones who are relatively unfamiliar with technical aspects of mental chronometry and psychophysiological research. Given its brevity and selectivity, this synopsis may seem somewhat biased and oversimplified; the views expressed herein merely reflect our own personal perspectives on the field. Readers who seek a thorough, balanced treatment should also consult the various authoritative reviews cited throughout the remainder of the article.

What follows is divided into four main parts. In the first part, we introduce some basic terms and concepts associated with the paradigm of mental chronometry. Secondly, we summarize the historical evolution of chronometric research on the human mind, establishing a context from which to appreciate the current state of affairs. Following this review, we describe some new techniques developed in our laboratory to supplement classical methods of mental chronometry. These techniques are by no means the only ones now under development around the world, but they do offer instructive examples of recent innovations that a number of investigators have been pursuing. Finally, we speculate a bit about how such innovations may bear on and benefit from the psychophysiological approach to studying cognition and action.

2. Paradigm of mental chronometry

The paradigm of mental chronometry involves experimental procedures in which a subject (i.e., human or other organism) experiences a series of test trials. Each trial starts with a warning signal (e.g., light or tone) followed by a brief foreperiod. The foreperiod serves to maximize the subject's alertness and attention. After the foreperiod, a test stimulus (e.g., visual or auditory pattern) is presented. Given this stimulus, the subject must make an appropriate overt response (e.g., manual movement) quickly and accurately. The subject's reaction time (RT) is measured from the onset of the test stimulus until the response occurs. In a *simple reaction-time procedure*, for example, there would be only one possible stimulus and one possible response. In a *choice reaction-time procedure,* there would be multiple stimuli and multiple responses, with different responses assigned to different stimuli. To encourage good performance, feedback and payoffs are often provided on the basis of the subject's speed and accuracy. For detailed descriptions of these procedures, see Pachella (1974), Woodworth (1938), and Woodworth and Schlosberg (1954).

2.1. Basic stage model

The rationale of the chronometric paradigm can be appreciated more fully in terms of a basic stage model for human information processing (Donders, 1868/1969; Sternberg, 1969). According to this model, performance is mediated by a sequence of time-consuming processes, which include perceptual encoding of a stimulus, retrieving stored information from memory, making decisions based on this information, and preparing an appropriate response. Attempts to discover and analyze such processes have motivated much of the past research involving mental chronometry. The hope has been that through measures of reaction time and response accuracy, elementary processing stages would emerge in clear detail.

2.2. Fundamental questions

To be specific, mental chronometry has concentrated on answering several fundamental questions about the human information-processing system (Pachella, 1974; Posner & McLeod, 1982; Smith, 1968; Townsend & Ashby, 1983). The first question is most basic. Are there, in fact, separate component mental processes that mediate overt responses to stimuli and that take measurable amounts of time to be completed? Assuming that the answer is "yes", one may ask several additional questions about detailed characteristics of these processes. For example, how many different processes are there? What transformations occur during each of them? How much time does a particular process take? Do the processes proceed in a strict sequential fashion with no temporal overlap among them, or do they operate at least somewhat in parallel? Are discrete packets of information or continuous quantities of activation produced as outputs along the way?

At present, we do not have definitive answers to all of these questions. Indeed, one might wonder whether mental chronometry will ever answer them completely. The chronometric paradigm is limited in that conclusions from it are typically based on just two observed dependent variables, namely, reaction time and response accuracy. Mental chronometry does not provide a direct look at the underlying processes or products that mediate subjects' performance. Nevertheless, considerable progress has been made over the past 150 years through reaction-time and accuracy data.

3. History of mental chronometry

For purposes of exposition, mental chronometry's history can be divided roughly into four periods, as summarized in table 1. We refer to these periods as *The Prehistoric Times, The Golden Age, The Dark Age,* and *The Renaissance* of mental chronometry. Detailed discussions of them have appeared in several historical accounts (e.g., Boring, 1950; Lachman et al., 1979; Leonard, 1957; Woodworth, 1938). On the basis of such accounts, table 1 lists some major dates and ideas associated with each period.

3.1. Prehistoric Times

The Prehistoric Times of mental chronometry extended from the dawn of intellectual activity to about the middle of the nineteenth century. Until around 1850 A.D., many scholars believed that human thought was instantaneous and that action was governed by an indivisible mind separate from the body (Boring, 1950, Chapters 1 and 2). Little or no effort was devoted before then to devising serious reaction-time procedures, because there seemed no

Table 1

Historical overview of mental chronometry

Key periods	Dates	Major ideas
Prehistoric Times	Before 1850 A.D.	Belief in indivisible mind
		Belief in instantaneous thought
Golden Age	1850 to 1900 A.D.	Measurement of neural-conduction time
		Development of subtraction method
		Discovery of processing stages
		Introduction of speed-accuracy tradeoff curves
Dark Age	1900 to 1950 A.D.	Criticism of subtraction method
		Decline of reaction-time research
Renaissance	1950 A.D. to present	Development of information-processing concepts
		Introduction of additive-factor method
		Micro-analysis of processing stages
		Concentration on serial-versus-parallel distinction
		Concentration on discrete-versus-continuous distinction

point in trying to measure something infinitely fast and essentially unanalyzable. Neither the classical Greek nor British philosophers appear to have stumbled onto the chronometric paradigm, even though they conceived many other fundamental ideas about cognition and action. This lack of insight even characterized the views of sophisticated nineteenth-century scientists such as Johannes Müller, who is considered by many to have been the father of experimental physiology, but who still believed that the rate of neural conduction had the same order of magnitude as the speed of light (Müller, 1838). As Müller put it, "We shall probably never attain the power of measuring the velocity of nervous action; for we have not the opportunity of comparing its propogation through immense space, as we have in the case of light" (translation cited in Boring, 1950, p. 41).

Given the ideas expressed by Müller (1838), it is ironic that the astronomers of his day were among the first investigators to seek practical techniques for measuring the speed of mental processes. Their effort, which presaged the introduction of the simple reaction-time procedure, was motivated by a desire to assess individual differences in subjective temporal judgements about the movements of stars and other heavenly bodies. Some initial results of this work appeared in a report by Bessel (1823), who formulated the *personal equation*, a measure of the difference between two observers' estimates of the times at which particular stellar events occurred. These differences could, in principle, be taken to suggest the existence of variable time-consuming mental processes whose neurophysiological substrates entail relatively low conduction rates. Yet the full implications of Bessel's (1823) report were not appreciated until quite a while later.

About 1850, however, circumstances started changing dramatically. The transition was due largely to innovations by Hermann von Helmholtz, one of the most influential physicists and neurophysiologists of the nineteenth century. Among his many accomplishments, he not only helped articulate the physical principle for the conservation of energy (Helmholtz, 1847) but also introduced the simple reaction-time procedure as an experimental tool for neurophysiological investigation (Helmholtz, 1850/1853). We may therefore regard him as a major forerunner of present-day cognitive psychophysiology.

Using the simple reaction-time procedure, Helmholtz (1850/1853) discovered that the rate of neural conduction in humans is on the order of just 50 m/s, much less than the speed of sound, let alone the speed of light. Although his discovery may sound trivial from our present perspective, it had a tremendous impact at the time. As Boring (1950) has noted:

> In Helmholtz's experiment lay the preparation for all the later work of experimental psychology on the chronometry of mental acts and reaction times. The most important effect of the experiment and all the research that followed upon it was... that it brought the soul to time, ...(capturing) the essential agent of mind in the toils of natural science. ...it did more than any other single bit of research to advertise the fact that mind is not ineffable but a proper subject for experimental control and observation by him who is wise enough to conceive the necessary means. (pp. 42&45)

The research by Helmholtz is particularly relevant for us because it exemplifies how the chronometric paradigm can be wedded to a neurophysiological framework in reaching new intellectual syntheses such as are sought by cognitive psychophysiology.

3.2. Golden Age

With the initial development and application of the simple reaction-time procedure, the Golden Age of mental chronometry ensued. It continued for the next 50 years or so, up to about 1900 A.D. To illustrate how investigation grew over this and subsequent periods, we have sketched a partial "family tree" of chronometric research, which appears in fig. 1.

This tree diagram is intended as an informal heuristic framework rather than a complete taxonomy of mental chronometry. On the right side of the tree are some noteworthy contributions that, for the most part, have dealt with conventional reaction-time data in efforts to apply or test the basic stage model of information processing outlined earlier. On the left side of the tree are other noteworthy contributions that, by contrast, have emphasized measures of response accuracy as well as speed (e.g., via speed-accuracy tradeoff curves) and have entertained alternative models with salient stochastic or statistical-decision components. For the sake of brevity, we have pruned the chronometric family tree somewhat arbitrarily, so many additional important

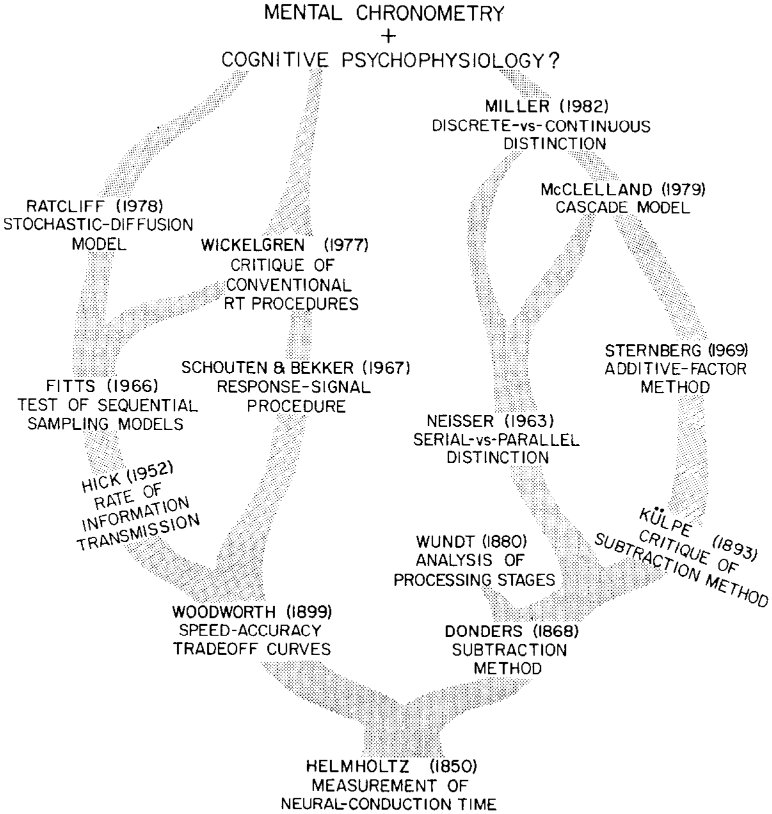

Fig. 1. A partial family tree of mental chronometry. (The right side of the tree contains some noteworthy contributions that have dealt primarily with conventional reaction-time data and direct applications or tests of stage models of information processing. On the left side are other contributions that have emphasized speed-accuracy tradeoff curves and advocated alternative models with salient stochastic or statistical-decision components.)

examples are missing from it. Nevertheless, our metaphorical forestry may help at least a bit in leafing through subsequent material.

3.2.1. Speed-accuracy tradeoff curves

As the left side of fig. 1 indicates, one branch of investigation that stemmed from Helmholtz's (1850/1853) original work proceeded onward to research by Woodworth (1899), who plotted speed-accuracy tradeoff curves in analyzing data on movement control. These curves involve functions of response accuracy (or error rate) versus the speed (or reaction time) with which a subject performs a given task. Congruent with the time-honored adage, "haste makes waste," they typically reveal a tradeoff such that increased speed of performance leads to decreased accuracy. Woodworth (1899) used speed-accuracy tradeoff curves to show how the spatial precision of voluntary movements decreases systematically as movement velocity increases. From his results, he identified several sources of random variability in the human motor system (cf. Meyer, Abrams, Kornblum, Wright, & Smith, 1988; Meyer, Smith, & Wright, 1982; Schmidt, Zelaznik, Hawkins, Frank, & Quinn, 1979). We will not dwell further on the import of Woodworth's (1899) discoveries for the moment, but we will describe later how speed-accuracy tradeoffs are still highly relevant to mental chronometry.

Meanwhile, let us consider the other branch of investigation that stemmed from Helmholtz's original work (fig. 1, right side). This brings us to F.C. Donders, among the most widely cited investigators in the history of mental chronometry. There are two principal reasons for Donders' influence. First, he supplemented the simple reaction-time procedure by developing the choice reaction-time procedure, and second, he devised an analytical technique for estimating the durations of component processing stages (Donders, 1868/1969).

3.2.2. Subtraction method

Donders' technique is known as the subtraction method. It provides a way of analyzing data from three different types of reaction-time procedures, which Donders called *Task A*, *Task B*, and *Task C*. Task A entails a simple reaction-time procedure with a single stimulus and response, as Helmholtz (1850/1853) used. Task B entails a choice reaction-time procedure with multiple stimuli and multiple responses. Donders reasoned that Task B would require a subject to discriminate among the various possible stimuli and to select among the various possible responses on each trial, whereas Task A would not require these processes because there would be nothing to discriminate or select among. So he proposed to estimate the combined durations of the discrimination and selection processes by subtracting the reaction time for Task A from the reaction time for Task B (Donders, 1868/1969).

Donders (1868/1969) also attempted to estimate the duration of each process separately by examining results from his Task C, which entailed a go/no-go reaction-time procedure with multiple stimuli but only a single response. In Task C, subjects had to make a response to one stimulus, but withhold responding to all other stimuli. According to Donders, Task C therefore required stimulus discrimination but not response selection, because subjects would always know beforehand what response to make, if one was needed. By subtracting the reaction time for Task C from the reaction time for Task B, he tried to estimate the selection process's duration. Also, by subtracting the reaction time for Task A from the reaction time for Task C, he tried to estimate the discrimination process's duration.

Of course, Donders' subtraction method requires several strong assumptions (Pachella, 1974; Sternberg, 1969). One assumption is that component mental processes such as stimulus discrimination and response selection are strictly successive stages whose durations combine additively to yield an overall reaction time. Another assumption is that in switching from a simple to choice or go/no-go reaction-time procedure, stages of processing may be inserted or deleted in a pure fashion without changing the time course or outputs of other concomitant processes. If either of these assumptions is violated, then the whole enterprise would tumble like a house of cards.

3.2.3. Discovery of processing stages

Despite the strong, perhaps dubious, assumptions entailed by it, many investigators were enthusiastic about Donders' subtraction method. The reasons for this enthusiasm are obvious. In principle, the method offers the possibility of measuring stage durations as a function of various factors. By applying the subtraction method, one might conceivably discover a whole host of mental processes and analyze their true nature. The prospect was so great that eminent psychologists like Wilhelm Wundt, the founder of the world's first official experimental-psychology laboratory, turned much of their efforts toward exploiting the subtraction method and discovering stages of information processing (e.g., Wundt, 1880).

Wundt's initial applications of the subtraction method proved extremely fruitful. As part of his endeavors, a new task called the *D-reaction* was introduced. Like Donders' (1868/1969) earlier Task C, it involved multiple stimuli and a single response, but subjects had to make this response for each and every stimulus as soon as they thought that they had identified the stimulus correctly. Using Task D along with Tasks A, B, and C, Wundt claimed to have isolated and measured more than half a dozen distinct types of process, including reflexes, voluntary impulses, perception, apperception, cognition, association, and judgment (cited in Boring, 1950, p. 149). Numerous other investigators in Wundt's laboratory and elsewhere also began exploiting the chronometric paradigm (e.g., Cattell, 1886; Exner, 1873; Jastrow, 1890; Lange, 1888; Merkle, 1885).

3.3. Dark Age

Unfortunately, mental chronometry's early successes were rather short lived. In 1893, Oswald Külpe, one of Wundt's most influential students, published a devastating critique of the subtraction method. Külpe (1893/1909) was distressed because the method tended to produce inconsistent results. Estimates of stage durations obtained in various laboratories often differed substantially from each other, depending on the experience and mental set of the subjects being tested. A likely cause of the problem was that the assumption of pure insertion had failed. Switching from one reaction-time procedure to another may not merely insert or delete some processing stage; it may also change the quality of other concomitant stages shared across different tasks (cf. Ach, 1905; Watt, 1905).

So not long after its star had risen, the subtraction method fell into disrepute. The fall was sufficiently precipitous that it left the field of mental chronometry in what might, by contrast, be characterized as a Dark Age. This period lasted through much of the first half of the twentieth century, during which there were relatively few studies that compared performance in simple versus choice or go/no-go reaction-time procedures, and during which only limited theoretical developments occurred concerning the dynamics of information processing. The decline of such chronometric research is apparent in some subsequent major publications, including Stevens' (1951) *Handbook of experimental psychology* and Osgood's (1953) *Method and theory in experimental psychology*. Neither of these famous works mentioned the word "reaction time" except briefly in passing, nor did they cite Donders (1868/1969) and his seminal contributions.

Of course, we do not mean to imply that research involving mental chronometry completely ceased from 1900 to 1950. Some ultimately noteworthy discoveries about the dynamics of cognition and action did appear at the time (Woodworth, 1938; Woodworth & Schlosberg, 1954). These included research by Woodrow (1914) on foreperiod effects, Telford (1931) on the psychological refractory period, Stroop (1935) on perceptual and response competition, and Mowrer (1940; cf., Gibson, 1941) on subjective expectancies. Nevertheless, it took a considerable while before this latter work garnered the full recognition that is deserved.

3.4. Renaissance

Fortunately, around 1950, prospects for the chronometric paradigm started to brighten once again. Several intellectual forces were responsible for the Renaissance of mental chronometry. Part of the impetus was provided by developments in the computer and communication sciences (e.g., Shannon, 1948; Turing, 1950). Through these advances, new theoretical tools became available for building detailed models of human information processing. Also, to test such models, experimental psychologists began thinking harder about how one might collect more powerful reaction-time data. Their efforts have led to further growth in research with the chronometric paradigm (Lachman et al., 1979; Luce, 1986; McGill, 1963; Neisser, 1967; Posner & McLeod, 1982; Smith, 1968; Townsend & Ashby, 1983).

Some places where the growth has occurred are shown among the left branches of mental chronometry's family tree (fig. 1). In 1952, for example, Hick demonstrated how speed-accuracy tradeoff curves may be interpreted on the basis of ideas from mathematical communication theory, yielding measures of subjects' information-transmission rates (cf. Hyman, 1953; Merkle, 1885). Following Hick's (1952) work, mathematical models that treat the tradeoff between speed and accuracy in terms of stochastic processes (e.g., random sequential sampling) were tested by Fitts (1954, 1966) and other investigators (Audley, 1960; LaBerge, 1962; Laming, 1968; Link, 1975; Pike, 1973; Ratcliff, 1978; Stone, 1960). New methods for manipulating and measuring response accuracy as a function of speed were also introduced during this period (e.g., the response-signal procedure of Schouten & Bekker, 1967; cf. Reed, 1976; Corbett & Wickelgren, 1978). We will return again to the topic of speed-accuracy tradeoff curves, but before then, let us direct our attention to the additional branches at the top right of our tree diagram.

The tree's top right branches (fig. 1) stem directly from the tradition established by Donders (1868/1969). Because of problems encountered with his subtraction method, some later investigators have sought alternative ways to study stages of information processing without relying on the assumption of pure insertion. A major advance along these lines occurred through efforts by Sternberg (1969), who developed the additive-factor method of analyzing reaction-time data.

3.4.1. Additive-factor method

Sternberg's (1969) additive-factor method has two major objectives. One is to discover stages of processing. The other is to determine exactly what factors influence these stages. In effect, achieving such goals would help advance Donders' (1868/1969) original theoretical ideas, but because of how it works, the additive-factor method obviates the dubious assumption of pure insertion that undermined the subtraction method.

The logic of the additive-factor method is simple but powerful. Suppose information processing involves a series of strictly successive stages with no temporal overlap among them. Also, suppose there are three separate factors, F_1, F_2, and F_3, such that F_1 influences a relatively early stage of processing (e.g., stimulus encoding or memory retrieval), while F_2 and F_3 influence some later stage (e.g., response selection). Then these factors ought to have a certain pattern of effects on mean reaction times. If two factors like F_1 and F_2 influence different stages, their effects should be additive, because under the serial stage model, reaction time is a sum of the component stage durations. On the other hand, if two factors like F_2 and F_3 influence the same stage, their effects would most likely interact (Sternberg, 1969).

Reversing this logic, Sternberg (1969) proposed two rules of inference for interpreting reaction-time data. One rule states that whenever an investigator finds two or more factors whose effects on mean reaction times are additive, it may be concluded that at least two distinct stages of processing are involved. The other rule states that whenever two or more factors whose effects interact with each other are found, then it may be concluded that they influence at least one stage in common. By manipulating various factors and looking for patterns of additivity and interaction, one may thus try to determine how many different processing stages exist and what the stages do. In studies of short-term memory retrieval, for example, Sternberg (1969) used factors such as the visual quality of stimulus displays, the length of memorized lists, and the probabilities of alternative responses. These factors had additive effects on mean reaction times, suggesting at least three different stages: stimulus encoding, memory retrieval, and response selection.

An appealing feature of the additive-factor method is that it does not require inserting or deleting entire stages of processing. To apply the method, one merely has to identify factors that influence the component stage durations underlying mean reaction times within a given task. The method avoids having to compare results from simple and choice reaction-time procedures, which made Donders' (1868/1969) original approach susceptible to failures in the assumption of pure insertion.

Still, some other assumptions related to those of Donders (1868/1969) are embodied in the additive-factor method. In particular, the method assumes that stages of processing do not have any temporal overlap with each other. It also assumes that the outputs of stages are discrete all-or-none quanta of information whose quality does not depend on the levels of whatever factors are used to manipulate stage durations (Sternberg, 1969). Furthermore, there is an assumption that the factors can selectively influence the durations of different stages. If any of these assumptions happen to be violated, then the logic of the additive-factor method would crumble, just like the subtraction method did.

3.4.2. Cascade model

An illustration of how one might obtain such violations has been described by McClelland (1979), who formulated a cascade model of information processing. Like Sternberg's (1969) discrete stage model, the cascade model assumes that subjects' performance involves processes such as stimulus encoding, memory retrieval, decision making, response selection, and so forth. However, these processes are not assumed to take place in a strictly serial fashion or to produce discrete quantized outputs of information. Instead, McClelland (1979) proposed that several processes may take place in a parallel contingent fashion, with a continuous flow of activation going from one process to the next.

Given this arrangement, which seems plausible in terms of known facts about neural anatomy and physiology, it can be shown that Sternberg's (1969) additive-factor method no longer necessarily applies. Contrary to the method's inference rules, McClelland (1979) demonstrated how the cascade model can yield interactive effects of factors on mean reaction times even if those factors influence functionally different processes. He also demonstrated how the cascade model can yield additive-factor effects despite having component processes that are not strictly serial.

3.4.3. Related issues

A number of other investigators have likewise taken issue with assumptions embodied in the additive-factor method and discrete serial stage model on which it rests. During the 1960s, for example, Neisser and his colleagues reported several well-known studies whose results suggested parallel rather than serial processing (Neisser, 1963; Neisser & Beller, 1965; Neisser, Novick, & Lazar, 1963). They found that subjects performed some visual-search tasks just about as fast for ten target letters as for one or two, albeit with an increased error rate. After this discovery, Townsend (1974) proved mathematically that some special cases of parallel-processing models could mimic apparent serial processing for search tasks in which mean reaction times are not constant but instead increase linearly with the number of items to be searched. Such discoveries have stimulated formal analyses from various alternative theoretical perspectives, including ones with hybrid combinations of serial and parallel processing (e.g., Schweikert, 1978, 1983; Turvey, 1973). The conceptual substrate of McClelland's (1979) cascade model therefore has a rich and lengthy heritage.

At the same time, further doubts have arisen about the potential value of conventional reaction-time procedures (cf. fig. 1, left branches). For any condition of such a procedure, the obtained data may be viewed as constituting a single point on an underlying speed-accuracy tradeoff curve (i.e., the so-called *macro-tradeoff*; Pachella, 1974). From this perspective, one might argue that more complete results are needed than a conventional reaction-time procedure can provide. Perhaps inferences about the human information-processing system necessarily require an entire speed-accuracy tradeoff curve. The force of the latter argument has motivated some concerned critics to strongly favor alternative procedures in which response speed and accuracy are manipulated systematically over a wide range (e.g., Ollman, 1977; Reed, 1976; Wickelgren, 1977). This would seem particularly appropriate if processing involves a gradually increasing output of activation and partial information, as claimed under the cascade model (McClelland, 1979) and other continuous models (e.g., Edwards, 1965; Eriksen & Schultz, 1979; Ratcliff, 1978). With these models, subjects have a broad range of options to trade accuracy for speed, so it may be crucial to determine exactly what strategies they do adopt as a function of speed stress.

3.5. Important theoretical distinctions

In light of controversies surrounding the conventional reaction-time procedures, additive-factor method, and discrete serial stage model, it has become clear that future progress based on the paradigm of mental chronometry will require dealing more thoroughly with two important theoretical distinctions (Meyer, Irwin, et al., 1988; Miller, 1988; Townsend & Ashby, 1983). One distinction concerns the extent to which two (or more) component processes can and do overlap temporally with each other, embodying parallel rather than serial execution. The other distinction concerns the extent to which a given process transforms information in a continuous fashion and transmits it in a steady stream of output, rather than engaging in discrete transformations and transmitting intermittent quantized outputs. These distinctions must be pursued with vigor because their treatment will determine if and when the additive-factor method, or some other inferential framework, is most appropriate for analyzing and interpreting reaction-time data. Consequently, several investigators have begun more detailed chronometric research, taking various approaches to address the serial-versus-parallel and discrete-versus-continuous distinctions in human information processing.

3.5.1. Serial-versus-parallel distinction

With respect to the serial-versus-parallel distinction, chronometric research during the 1980s has looked especially for evidence of parallel contingent processing, in which recipient processes receive partial outputs from some source processes and the recipient processes start before the source processes finish. A representative example of such research appears in the work of Miller (1982, 1983). His approach involves testing whether subjects may become partially prepared for a forthcoming response by processing various dimensions of a presented stimulus (cf. Rosenbaum, 1980). The results of this investigation imply that under some (but not all) circumstances, stimulus evaluation sends partial outputs to response preparation, and the preparation process overlaps in time with the functionally prior evaluation process. These implications, if valid, would support McClelland's (1979) cascade model and other related ones that assume contingent parallel processing (e.g., Eriksen & Schultz, 1979; McClelland & Rumelhart, 1981; Turvey, 1973).

3.5.2. Discrete-versus-continuous distinction

Regardless of whether processes such as stimulus evaluation and response preparation occur serially or in parallel, it still remains to be determined whether they involve discrete or continuous transformation and transmission of information. This fact is illustrated, for example, by differences between the cascade model of McClelland (1979) and another account, the asynchronous discrete-coding model, proposed by Miller (1982). Both of these models assume parallel contingent processing, with the stimulus-evaluation process transmitting partial information to the response-preparation process so that preparation may begin before evaluation has finished. They differ, however, regarding the format of the information transmitted. Under the cascade model, as mentioned already, partial information consists of continuous increasing activation. The asynchronous discrete-coding model does not entail such activation. Instead, outputs by its evaluation process consist of quantized, temporally separate, information packets corresponding to stimulus values on each of a few basic feature dimensions. Given possibilities like this, one must treat the discrete-versus-continuous distinction in its own right, according it equal status with the serial-versus-parallel distinction. [1]

Some representative examples of work on the discrete-versus-continuous distinction appear in our own research. Here we outline two related approaches that we have taken to obtain a close look at the intermediate processes and products of cognition and action. One approach deals with whether response preparation involves discrete or continuous processes (Meyer, Yantis, Osman, & Smith, 1984, 1985; cf. Yantis & Meyer, 1988). The other approach deals with whether stimulus evaluation is discrete or continuous (Meyer & Irwin, 1981; Meyer, Irwin, et al., 1988).

4. Discrete versus continuous response preparation

To pursue the discrete-versus-continuous distinction for response preparation, we have adopted a *varied-priming procedure* (Meyer et al., 1984, 1985). For example, in one version of this procedure (Meyer et al., 1985; Experiment 1), the events were as follows: On each trial, an initial warning signal was first displayed briefly (500 ms). Next a prime stimulus was presented for a variable duration (from 0 to 700 ms). The prime stimulus consisted of either a printed word such as TREE or a nonword such as MAFE, which were sampled randomly from a large pool of letter strings. The subject did not respond overtly to the prime stimulus, but had to evaluate it in preparation for subsequent events. After the prime stimulus, there was a final brief (85 ms) warning signal, which helped equalize the subject's alertness regardless of how long the priming interval lasted. Then we presented a test stimulus consisting of either a right or left arrow. When the test stimulus appeared, the subject had to respond by pressing a corresponding right or left index-finger key quickly and accurately. We measured the subject's reaction time from the onset of the test stimulus until the response occurred. Response accuracy was recorded as well.

Our varied-priming procedure incorporated a strong contingency between the lexical status of the prime stimulus and the spatial orientation of the subsequent test stimulus. When the prime stimulus was a word, we always

[1] For present purposes, we use the term "discrete" in referring to any set of internal states, levels, and so forth that has a finite or countably infinite number of members. We do not necessarily restrict it to sets that have only two possible members, for example, as in all-or-none states of response preparation (cf. Miller, 1988).

followed it with a right arrow that required a right-hand response. When the prime stimulus was a nonword, we always followed it with a left arrow that required a left-hand response. In principle, this contingency allowed the subject to prepare his or her response in advance of the test stimulus, depending on how long the priming interval was.

The design of the procedure included three different conditions that varied the degree to which the subject's responses were actually primed. In one of these, the *completely-primed condition,* the prime stimulus had a relatively long duration (700 ms). This was enough time for the subject to finish evaluating the prime stimulus and fully prepare a response before the test stimulus appeared. As a result, the responses in the completely-primed condition were relatively fast. In another case, the *unprimed condition*, the subject did not receive any information to prepare a response ahead of time. We achieved this by proceeding directly from the initial warning signal to the final warning signal without displaying the prime stimulus at all (0 ms priming interval). [2] As a result, the responses to the test stimuli were relatively slow. Finally, there was a *partially-primed condition.* Here an intermediate level of response preparation was induced by presenting the prime stimulus for a medium duration (around 200 ms), which yielded moderate reaction times. [3] From quantitative comparisons of the reaction-time distributions obtained under the unprimed, partially-primed, and completely-primed conditions, inferences can be made about whether response preparation takes place in a discrete or continuous fashion throughout the priming interval.

4.1. Rationale of varied priming

The rationale of our varied-priming procedure is illustrated more fully in the three panels of fig. 2. This figure shows what representative distributions of reaction times from the unprimed, partially-primed, and completely-primed conditions should look like if response preparation involves a discrete all-or-none process with only two preparatory states. For reasons mentioned already, the unprimed condition has a distribution of relatively slow times, corresponding to a state of no advance preparation (fig. 2, top panel). The completely-

[2] An alternative way of implementing the unprimed condition is to display an uninformative (neutral) prime stimulus such as "XXXX" for some period of time between the initial and final warning signals. This has the advantage that it allows one to assess the efficacy of the final warning signal for maintaining subjects' alertness regardless of the length of the priming interval (Meyer et al., 1985; cf. Yantis & Meyer, 1988).

[3] The exact duration of the prime stimuli in the partially-primed condition was adjusted through a staircase tracking algorithm that yielded a distribution of reaction times located midway between those in the unprimed and completely-primed conditions. Given the flexibility of this algorithm, which takes account of details in individual subjects' performance, we have referred elsewhere to the varied-priming procedure as the *adaptive-priming procedure*. For further details, see Meyer et al. (1985).

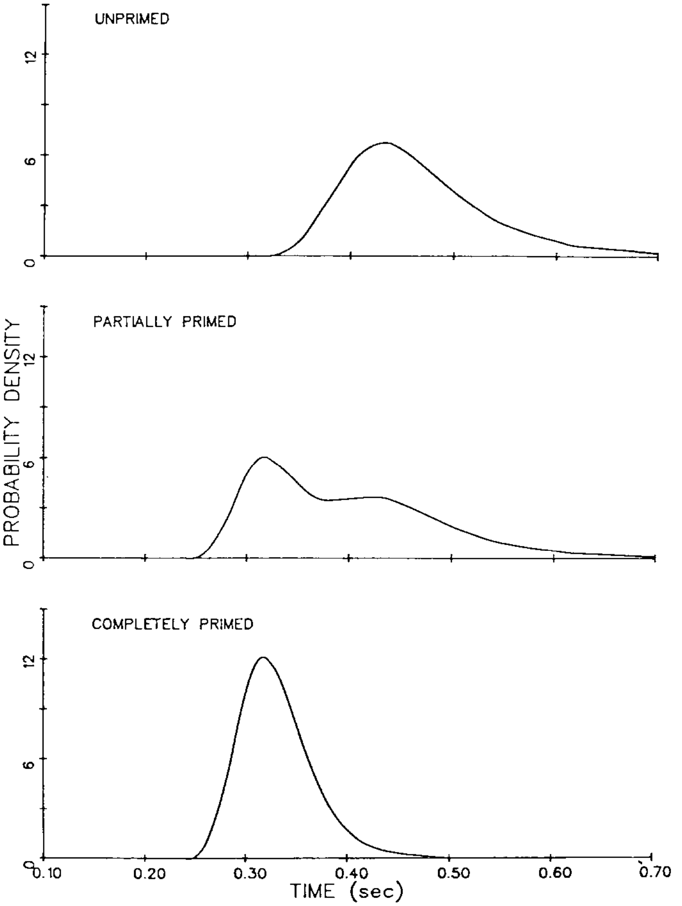

Fig. 2. Hypothetical reaction-time distributions predicted by a discrete all-or-none stage model of response preparation. (The top, middle, and bottom panels show distributions associated with the unprimed, partially-primed, and completely-primed conditions of the varied-priming procedure. From "Temporal properties of human information processing: Tests of discrete versus continuous models," by D.E. Meyer, S. Yantis, A.M. Osman, & J.E.K. Smith (1985). *Cognitive Psychology, 17,* 445–518. Copyright 1985 by Academic Press. Reprinted with permission of authors and publisher.)

primed condition has a distribution of relatively fast times, corresponding to a state of full preparation (fig. 2, bottom panel). However, the most important case concerns the intermediate reaction-time distribution for the partially-primed condition (fig. 2, middle panel).

Given a discrete all-or-none process of response preparation, the partially-primed condition should exhibit a distribution such that part of it comes from the completely-primed condition and part of it comes from the unprimed condition. This is because if the preparation process is all-or-none, then on any particular trial in the partially-primed condition, a subject can be in only one of two preparatory states, namely, unprepared or fully prepared (Meyer et al., 1985). Which state actually obtains after a medium prime duration will vary randomly from trial to trial, because fluctuations may occur in how long the subject takes to change from the unprepared to the fully-prepared state. As a result, the reaction times associated with partial priming should be a perfect mixture of those found under the two extreme priming conditions (i.e., unprimed and completely primed). The effects of such mixing may be seen in the relatively large variance of the partially-primed distribution and in the extensive overlap of its tails with those of the unprimed and completely-primed distributions. [4]

In contrast, a continuous process of response preparation would not yield this sort of mixture. For example, consider fig. 3. Here we have shown what reaction-time distributions in the unprimed, partially-primed, and completely-primed conditions would be like under McClelland's (1979) cascade model, which assumes that response preparation involves a continuous growth of activation rather than a discrete all-or-none transition between two extreme preparatory states. The distribution in the partially-primed condition (fig. 3, middle panel) is not a perfect mixture of those in the unprimed condition (fig. 3, top panel) and completely-primed condition (fig. 3, bottom panel). Instead, its shape is similar to theirs, its tails do not overlap extensively with theirs, and its variance is moderate. This is because if response preparation increases continuously during the priming interval, then in the partially-primed condition, there should be a truly intermediate level of preparation with its own unique pattern of facilitation (McClelland, 1979; Meyer et al., 1984, 1985). The occurrence of partial priming would not stem merely from mixing the unprepared and fully-prepared states on a stochastic basis. So by examining the reaction-time distribution in the partially-primed condition and comparing it with those in the unprimed and completely-primed conditions, we may reach further conclusions about whether response preparation is a discrete or continuous process.

4.2. *Evidence of all-or-none preparation*

Some results obtained with the varied-priming procedure appear in fig. 4. These data come from a representative subject who participated during several hundred trials of the task just described (Meyer et al., 1985, Experiment 1).

[4] To be precise, let $f_u(t)$, $f_p(t)$, and $f_c(t)$ represent probability-density functions of reaction times in the unprimed, partially-primed, and completely-primed conditions, respectively. Also, let π represent the probability that a subject enters the fully-prepared state under the partially-primed condition, and let $1-\pi$ represent the probability that the subject remains in the unprepared state. Then according to the discrete all-or-none model of response preparation, $f_p(t) = \pi f_c(t) + (1-\pi) f_u(t)$.

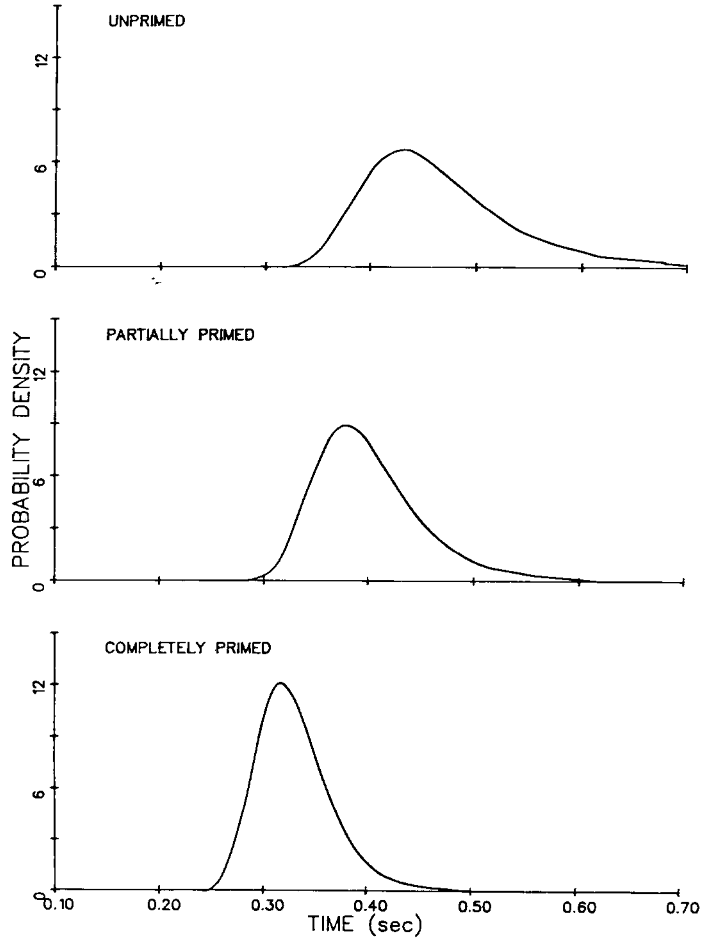

Fig. 3. Hypothetical reaction-time distributions predicted by a continuous cascade model of response preparation. (The top, middle, and bottom panels show distributions associated with the unprimed, partially-primed, and completely-primed conditions of the varied-priming procedure. From "Temporal properties of human information processing: Tests of discrete versus continuous models," by D.E. Meyer, S. Yantis, A.M. Osman, & J.E.K. Smith (1985). *Cognitive Psychology, 17,* 445–518. Copyright 1985 by Academic Press. Reprinted with permission of authors and publisher.)

The three panels of fig. 4 contain solid histograms of observed reaction times for the unprimed, partially-primed, and completely-primed conditions, respectively. As shown there, the relative frequency of fast responses increased as the duration of the prime stimulus increased, reflecting a substantial priming effect.

Next let us consider the dashed curves superimposed on each histogram of fig. 4. These curves represent theoretical estimates of reaction-time distribu-

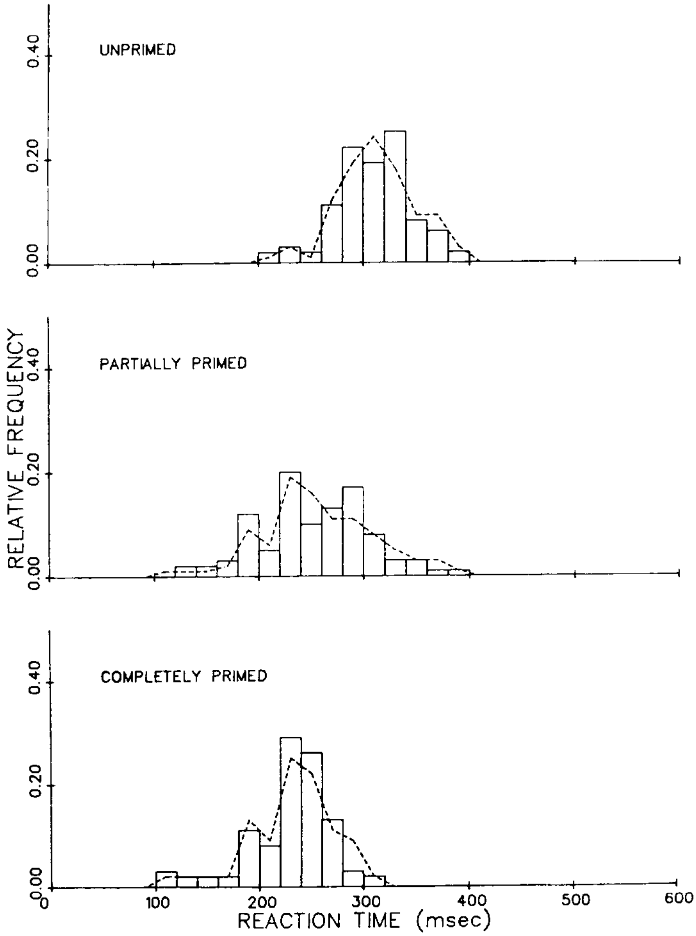

Fig. 4. Distributions of reaction times that support a discrete all-or-none stage model of response preparation. (The top, middle, and bottom panels show data for a representative subject in the unprimed, partially-primed, and completely-primed conditions of the varied-priming procedure. Solid histograms indicate relative frequencies of observed times, and dashed curves indicate best-fitting frequency distributions based on the all-or none model. From "Temporal properties of human information processing: Tests of discrete versus continuous models," by D.E. Meyer, S. Yantis, A.M. Osman, & J.E.K. Smith (1985). *Cognitive Psychology, 17,* 445–518. Copyright 1985 by Academic Press. Reprinted with permission of authors and publisher.)

tions that come closest to the data in terms of maximum likelihood and that form a family in which the estimated distribution for the partially-primed condition is a perfect mixture of the estimated distributions for the unprimed and completely-primed conditions. [5] If response preparation were a discrete all-or-none process with only two preparatory states, namely, unprepared and fully prepared, then the dashed lines would fit the histograms exactly except for random statistical noise. Indeed, the fit is very good, reflecting a nearly perfect mixture ($\chi^2(13) = 16.5$; $p > 0.05$). The partially-primed condition (fig. 4, middle panel) yielded reaction times with the largest variance and tails that extensively overlapped those from the other two conditions (fig. 4, top and bottom panels). Similar results were obtained for a number of other subjects as well (Meyer et al., 1985, Experiments 1 and 2). It therefore appears that under a least some circumstances, response preparation may be an essentially discrete all-or-none process, consistent with Sternberg's (1969) serial stage model.

4.3. Evidence of intermediate preparatory states

Of course, we do not mean to imply that response preparation is always a discrete all-or-none process. The experiment in which we obtained evidence of just an unprepared state and a fully-prepared state was relatively simple. Subjects produced only two different responses there (i.e., keypresses with left and right index fingers). Relationships between the test stimuli (left and right arrows) and responses were highly compatible. Also, the prime stimuli (words and nonwords) were completely informative about the required responses. Once the subjects identified a word or nonword prime, they knew exactly which keypress to make subsequently. However, under more complex circumstances that place greater demands on subjects' processing capacity, the results may be somewhat different.

For example, we have conducted another experiment with the varied-priming procedure to demonstrate the existence of intermediate preparatory states (Meyer et al., 1985, Experiment 3). Rather than including just two pairs of test stimuli and responses, this next experiment included four pairs. Each test stimulus was an upward arrow presented at one of four different locations demarcated by a horizontal array of underlined spaces on a display screen. Each response was a key press made with either the index or middle finger of the right or left hand. The responses were assigned to the test stimuli via a relatively complex mapping. In particular, the left-end arrow ($\overset{\uparrow}{-}\,-\,-\,-$) required a right middle-finger keypress; the left-middle arrow ($-\,\overset{\uparrow}{-}\,-\,-$) required

[5] The estimates were derived through an iterative algorithm that maximizes goodness-of-fit in terms of a quasi chi-square statistic for multinomial mixture distributions. Details regarding this algorithm appear in Smith, Meyer, Yantis, and Osman (1982).

a right index-finger keypress; the right-middle arrow $(- - \overset{\uparrow}{-} -)$ required a left index-finger keypress; and the right-end arrow $(- - - \overset{\uparrow}{-})$ required a left middle-finger keypress.

As in our previous experiments (e.g., Meyer et al., 1985, Experiment 1), the test stimuli on some (but not all) trials were preceded by informative prime stimuli that consisted of various words and nonwords. When the prime stimulus was a word, it indicated that the subsequent test stimulus would require a response with one of the right-hand fingers. A nonword prime stimulus indicated that the subsequent test stimulus would require a response with one of the left-hand fingers.

Three priming conditions were included here, analogous to the unprimed, partially-primed, and completely-primed conditions of our previous experiment (cf. Meyer et al., 1985, Experiment 1). Under the unprimed condition, no informative primes occurred before the test stimuli, allowing subjects no advance preparation for the subsequent responses. Under the partially-primed condition, the informative primes (i.e., words and nonwords) occurred with a medium duration (roughly 200 ms), which did not always permit the subject to finish processing them completely but allowed an intermediate level of response preparation to be achieved. Under the completely-primed condition, the informative primes occurred with a long duration (700 ms), which practically always permitted subjects to finish processing them and to achieve a relatively high level of response preparation. However, the completely-primed condition did not allow as much preparation as in our previous study, because the prime stimuli only cued which hand (i.e., right or left); not which particular finger (i.e., index or middle), to use for making the response. We suspected that the change in this condition, combined with the larger stimulus-response ensemble and more complex mapping of test stimuli onto responses, might yield somewhat different results than before. It seemed possible, in particular, that rather than having just two discrete states (i.e., all or none), the preparation process would pass through some additional intermediate states, because of the increased demands made by the task on subjects' processing capacity and cognitive resources.

Some results consistent with the latter expectations appear in fig. 5. These data come from responses made by a representative subject with the middle finger of his right hand during several hundred trials involving the four alternative stimulus-response pairs (Meyer et al., 1985, Experiment 3). Again we have shown histograms of observed reaction times (solid bars) and best-fitting theoretical distributions (dashed curves) for a discrete all-or-none model of response preparation. Unlike before (cf. fig. 4), the model does not fit very well here ($\chi^2(6) = 30.6$; $p < 0.001$). The model's failure can be seen by looking at the reaction times for the partially-primed condition (fig. 5, middle panel). In this case, the dashed curves do not closely match the solid histograms,

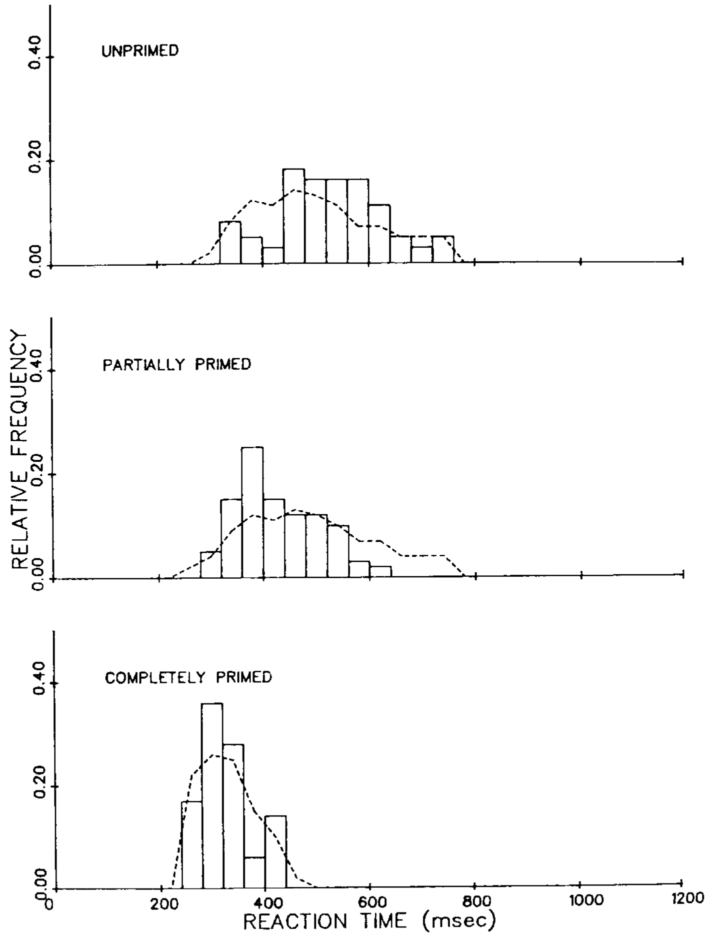

Fig. 5. Distributions of reaction times from the varied-priming procedure with four stimulus-response pairs and a complex stimulus-response mapping. (The top,middle, and bottom panels show data obtained for right middle-finger responses by a representative subject in the unprimed, partially-primed, and completely-primed conditions, respectively. Solid histograms indicate relative frequencies of observed times, and dashed curves indicate best-fitting frequency distributions based on the discrete all-or-none stage model, which is rejected here. From "Temporal properties of human information processing: Tests of discrete versus continuous models," by D.E. Meyer, S. Yantis, A.M. Osman, & J.E.K. Smith (1985). *Cognitive Psychology, 17,* 445–518. Copyright 1985 by Academic Press. Reprinted with permission of authors and publisher.)

indicating that the results do not correspond to an essentially perfect mixture of reaction times from the unprimed and completely-primed conditions. Instead, the medium-duration prime stimuli yielded reaction times that had only moderate variances and short tails compared with those from the other priming conditions (fig. 5, top and bottom panels). [6]

We infer, therefore, that the number of states in response preparation may depend on the size of the stimulus-response ensemble, the complexity of the stimulus-response mapping, and the response effector being used. As ensemble size and mapping complexity increase, there may be finer and finer gradations among preparatory states, reflecting a diffusion caused by greater demands on subjects' processing capacity (Meyer et al., 1985). The varied-priming procedure and analysis of reaction-time mixture distributions provide a way to assess changes in the preparation process as a function of task demands. [7]

5. Discrete versus continuous stimulus evaluation

With these conclusions in mind, we next describe another approach taken in our laboratory to pursue the discrete-versus-continuous distinction (Meyer & Irwin, 1981; Meyer, Irwin, et al., 1988). This approach is called the *speed-accuracy decomposition technique.* It focuses more closely on the processes of stimulus discrimination and evaluation.

5.1. Procedure for speed-accuracy decomposition

The experimental procedure for speed-accuracy decomposition involves two types of test trials, regular and signal, that are interleaved together.

5.1.1. Regular trials

During the regular trials, the events are similar to those in a conventional choice reaction-time procedure. Each regular trial starts with a warning signal (e.g., visual fixation mark) followed by a positive or negative test stimulus (e.g., word or nonword). Subjects must react to the test stimulus quickly but accurately, making either a positive or negative response (e.g., right or left keypress). We instruct subjects to take just enough time in evaluating the test stimulus to ensure that the response is almost always correct. Reaction time is measured from the onset of the test stimulus until the response occurs, and response accuracy is also recorded.

[6] In this experiment, similar results were obtained for middle-finger responses made by other subjects (Meyer et al., 1985, Experiment 3). However, index-finger responses followed the same pattern as found during our previous study with only two test stimuli and responses, in which response preparation was all-or-none (fig. 4). A possible explanation for the dependence of the results on finger type appears in Meyer et al. (1985).

[7] Some additional applications of the varied-priming procedure and mixture-distribution analysis appear in Yantis and Meyer (1985, 1988). For example, we have used our approach to demonstrate that spreading activation between concepts in semantic and episodic memory is not a discrete all-or-none process, but instead has quasi-continuous properties.

5.1.2. Signal trials

Like the regular trials, the signal trials start with a warning signal followed by a positive or negative test stimulus. The subjects are instructed to begin processing the test stimulus in the same way as on the regular trials, aiming toward a correct response that has a reasonably short latency. They may react as soon as they have determined the correct response. However, at some moment after the onset of the test stimulus, a response signal (e.g., auditory tone) is presented. It the subjects have not reacted to the test stimulus yet, then upon detecting the response signal, they must immediately produce their best guess about what the correct response is, regardless of whether or not they have finished evaluating the test stimulus. We vary the lag between the onset of the test stimulus and response signal, allowing the subjects more or less time before being interrupted by the response signal. Reaction time and response accuracy are measured as a function of the signal lag and the identity of the test stimulus, as in other response-signal procedures (e.g., Corbett & Wickelgren, 1978; Reed, 1976; Schouten & Bekker, 1967).

Another important feature of the signal trials is that we mix them randomly with the regular trials. At the start of each trial, the subjects cannot tell whether or not a response signal will be presented subsequently. So they must approach each test stimulus with a consistent mental set, trying to respond as soon as they either finish their stimulus evaluation or detect the response signal. This lets us directly compare results from the different trial types, which in combination provide more power for drawing inferences than does either a conventional reaction-time or speed-accuracy tradeoff experiment alone. Given how speed-accuracy decomposition works, we refer to the combination of regular and signal trials as a *titrated reaction-time procedure* (Meyer, Irwin, et al., 1988).

5.2. Theoretical objective

5.2.1. Information-accumulation functions

The objective of the titrated reaction-time procedure and speed-accuracy decomposition technique can be understood more fully in terms of fig. 6. We assume that when a test stimulus is presented on a regular or signal trial, subjects initiate a process of stimulus evaluation that, if not interrupted, typically yields a correct response. The evaluation process presumably involves accumulating information about what the response should be. Fig. 6 illustrates various forms that this accumulation might have during a representative trial. Using some algebraic equations outlined later, we analyze the information accumulation as a function of time, determining whether it entails discrete or continuous outputs.

For example, consider the three panels of fig. 6. One possibility here is that stimulus evaluation entails a discrete all-or-none accumulation of information

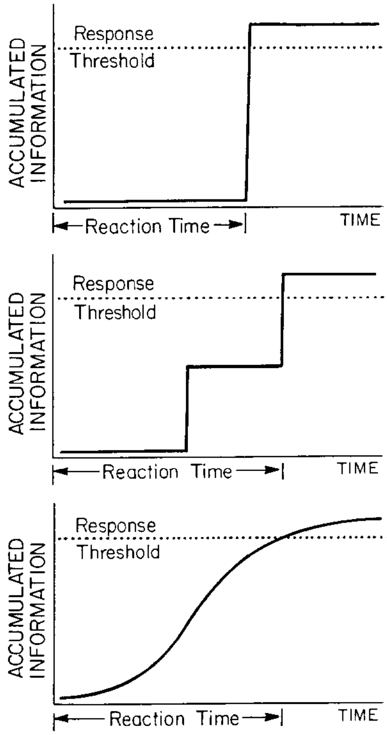

Fig. 6. Alternative information-accumulation functions that may characterize the process of stimulus evaluation. (The top, middle, and bottom panels illustrate discrete all-or-none, discrete multistate, and continuous outputs of accumulated information, respectively. The dotted horizontal lines are information thresholds that, when crossed, trigger the production of overt responses. From "The dynamics of cognition and action: Mental processes inferred from speed-accuracy decomposition," by D.E. Meyer, D.E. Irwin, A.M. Osman, & J. Kounios (1988). *Psychological Review, 95,* 183–237. Copyright 1988 by the American Psychological Association. Reprinted with permission of authors and publisher.)

over time, exhibiting a function with a single step from a low to a relatively high information level (fig. 6, top panel). Under these circumstances, a response would be initiated when the step function crosses the dotted high threshold. A second possibility is that stimulus evaluation entails discrete information accumulation, but that the evaluation process outputs one or more intermediate quanta of partial information before the threshold is crossed, exhibiting a function with multiple steps (fig. 6, middle panel). This multistep function and the above one having a single (all-or-none) step would be consistent with various discrete stage models of information processing (e.g., Miller, 1982; Sternberg, 1969). In contrast, a third possibility is that the evaluation process might produce a steadily increasing flow of partial information until the response threshold is crossed (fig. 6, bottom panel). This would be consistent with McClelland's (1979) cascade model and other deterministic continuous models (e.g., Eriksen & Schultz, 1979; McClelland & Rumelhart, 1981; Turvey, 1973; Wickelgren, 1977). Our goal is to discover which temporal patterns of accumulated partial information actually occur during the performance of various cognitive tasks.

5.2.2. Limitations of speed-accuracy tradeoff curves

Why are both regular and signal trials needed in order to test these alternative possibilities? If we included only signal trials, we could certainly plot observed response accuracy as a function of signal lag, obtaining a speed-accuracy tradeoff curve of the sort studied by previous investigators (e.g., Corbett & Wicklegren, 1978; Fitts, 1966; Pachella & Pew, 1968; Reed, 1976; Schouten & Bekker, 1967). This might provide some insight into how the process of stimulus evaluation accumulates partial information over time.

Unfortunately, there are some serious limitations to conclusions that can be drawn from standard speed-accuracy tradeoff curves. These curves may confound performance achieved on the basis of various information levels (Meyer & Irwin, 1981; Meyer, Irwin, et al., 1988; Schmitt & Scheirer, 1977; Wickelgren, 1977). For example, a smooth tradeoff curve that seems superficially consistent with continuous accumulation of partial information (McClelland, 1979; Wickelgren, 1977) could result from an all-or-none output of information by a discrete two-state evaluation process. In particular, suppose that on each trial there is a single sharp transition from a very low level to a high level of accumulated information, as the top panel of fig. 6 illustrates. Suppose also that the time at which this transition takes place varies randomly from trial to trial. Then when response accuracy is averaged across trials, a smooth tradeoff curve would result because of smearing, even though the underlying evaluation process is really discrete. So standard speed-accuracy tradeoff curves can obscure the true manner in which partial information is accumulated during rapid performance.

Speed-accuracy decomposition is designed to help deal with this problem. By analyzing data from the regular and signal trials in combination, we can measure how much partial information has been accumulated at various moments after the onset of a test stimulus, without the degree of confounding that may contaminate standard speed-accuracy tradeoff curves. The regular trials let us remove the contribution of above-threshold information to subjects' responses on the signal trials, thereby revealing the residual contributions of below-threshold partial information as a function of time. Such analyses would not be possible with data from either signal trials or regular trials alone, but taken together, the two trial types of the titrated reaction -time procedure provide deeper insights into the intermediate products of stimulus evaluation (Meyer, Irwin, et al., 1988).

5.3. Parallel sophisticated-guessing model

Some more details regarding the speed-accuracy decomposition technique appear in fig. 7, which shows a *parallel sophisticated-guessing model* for analyzing results from the titrated reaction-time procedure (Meyer & Irwin, 1981; Meyer, Irwin, et al., 1988). According to this model, a subject begins each regular and signal trial by initiating "normal" processes of stimulus evaluation at the onset of the test stimulus. The normal processes are assumed to be programmed to accumulate sufficient information for producing a typically correct response. They have an underlying distribution of completion times that may depend on the nature of the test stimulus, and when completion is reached, an overt response occurs.

In addition, the model incorporates a set of guessing processes. The guessing processes take place on the signal trials but not the regular trials. They supposedly begin at the onset of each response signal. Their function is to generate a response immediately after the signal is detected, relying on whatever partial information has been accumulated by the normal processes

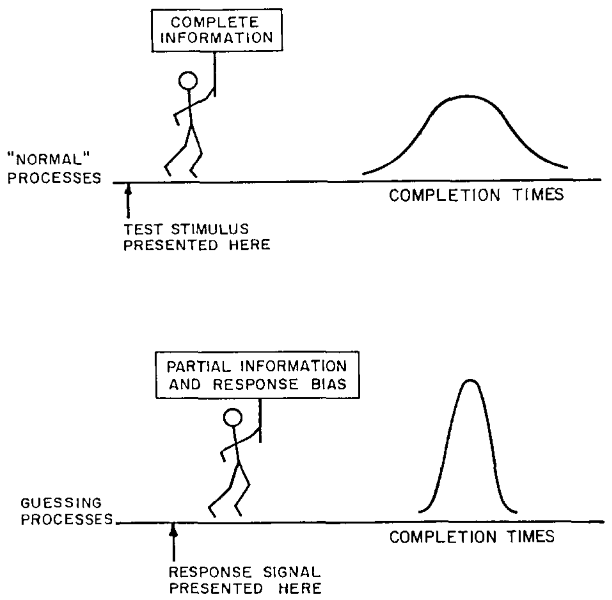

Fig. 7. Parallel sophisticated-guessing model used for analyzing results from regular and signal trials in the speed-accuracy decomposition technique. (From "The dynamics of cognition and action: Mental processes inferred from speed-accuracy decomposition," by D.E. Meyer, D.E. Irwin, A.M. Osman, & J. Kounios (1988). *Psychological Review, 95,* 183–237. Copyright 1988 by the American Psychological Association. Reprinted with permission of authors and publisher.)

up to then. We assume that the guessing processes race with the normal processes and that the winner of the race determines the subject's overt response. This assumption parallels ones made in several previous race models of information processing (e.g., Kornblum, 1973; Logan & Cowan, 1984; Ollman & Billington, 1972; Osman, Kornblum, & Meyer, 1986). The accuracy of the responses produced by the guessing processes is called the *guessing accuracy*. The time at which the guessing processes finish relative to the onset of a test stimulus is called the *guessing-completion time*.[8] We treat the guessing-completion times and the completion times of the normal processes as independent random variables (cf. Kornblum, 1973; Ollman & Billington, 1972).

Given reaction times and accuracies of overt responses from the regular and signal trials, we analyze them in terms of the parallel sophisticated-guessing model. The results can be used to test the model's assumptions about the nature of the race between the normal and guessing processes. On the basis of such tests, it has turned out that these assumptions are reasonably valid (Meyer, Irwin, et al., 1988). More important, the parallel sophisticated-guessing model provides a measure of how much partial information has been accumulated by the normal processes of stimulus evaluation at each moment before they have reached completion. This measurement involves estimating the accuracy and completion times of the guessing processes along with those of the normal processes.

5.3.1. Guessing-completion times

The guessing-completion times are estimated through the following equation:

$$P(t_{gs} \leq C) = \frac{P(t_s \leq C) - P(t_r \leq C)}{1 - P(t_r \leq C)}. \tag{1}$$

Here t_{gs} is a random variable that represents the completion times of the guessing processes on the signal trials, t_s is a random variable that represents the observed reaction times on the signal trials, t_r is a random variable that represents the observed reaction times on the regular trials, and C is an arbitrary non-negative constant. We can determine $P(t_s \leq C)$ and $P(t_r \leq C)$

[8] For reasons discussed elsewhere (Meyer, Irwin, Osman, & Kounios, 1988), the guessing-completion times are not measured relative to the onset of the response signal. This is because we want to plot the guessing accuracy as a function of how much time the normal processes have to accumulate partial information before being beaten by the completion of the guessing processes.

directly from the signal-trial and regular-trial reaction times for any value of C, so we can also determine $P(t_{gs} \leq C)$, using Equation 1. This gives us an estimated cumulative distribution function (i.e., $F_{gs}(C) = P(t_{gs} \leq C)$ for the guessing-completion times. [9]

5.3.2. Guessing accuracy

The accuracy of the guessing processes, which reflects the amount of partial information available to them from the unfinished normal processes at the moment when the subject detects the response signal, are estimated through the following equation:

$$P_g(\text{correct}|t_{gs} < t_n) = \frac{P_s(\text{correct}) - \left[P(t_n \leq t_{gs})P_n(\text{correct}|t_n \leq t_{gs})\right]}{1 - P(t_n \leq t_{gs})}. \qquad (2)$$

Here $P_g(\text{correct}|t_{gs} < t_n)$ represents the guessing accuracy, that is, the probability that the guessing processes produce a correct response when their completion time (t_{gs}) is less than the completion time (t_n) of the normal processes. Using Equation 2, we calculate the guessing accuracy in terms of three other estimable quantities: $P_s(\text{correct})$, $P_n(\text{correct}|t_n \leq t_{gs})$, and $P(t_n \leq t_{gs})$. P_s(correct) represents the probability of correct responses on the signal trials; it is estimated directly from the observed signal-trial response accuracy. $P_n(\text{correct}|t_n \leq t_{gs})$ represents the probability of correct responses by the normal processes when their completion times are less than or equal to those of the guessing processes; it is estimated directly from a combination of the observed regular-trial response accuracy, regular-trial reaction times, and derived distribution of guessing-completion times (Equation 1). $P(t_n \leq t_{gs})$ represents the probability that the normal processes win the race with the guessing processes; it is estimated from a combination of the observed regular-trial reaction times and derived guessing-completion times. The regular trials provide pure measures of both the completion times and the accuracy of the completed normal processes. This is possible because, under the parallel

[9] The derivation of Equation 1 is straightforward (Meyer, Irwin, Osman, & Kounios, 1988). When the normal and guessing processes race with each other, the reaction time on a signal trial (t_s) will exceed C if and only if the completion time of the guessing processes (t_{gs}) and the completion time of the normal processes (t_n) both exceed C. Under the parallel sophisticated-guessing model, these completion times are assumed to be independent random variables, so $P(t_s > C) = P(t_n > C \text{ and } t_{gs} > C) = P(t_n > C)P(t_{gs} > C)$. In turn, this equation implies that $[1 - P(t_s \leq C)] = [1 - P(t_n \leq C)][1 - P(t_{gs} \leq C)]$. The completion times ($t_n$) of the normal processes are directly estimable from the reaction times (t_r) on the regular trials. Substituting t_r for t_n in the latter equation and rearranging terms then yields Equation 1.

sophisticated-guessing model (fig. 7), only the normal processes mediate regular-trial performance. [10]

By plotting the estimated accuracy of the guessing processes versus the guessing-completion times, we may discriminate among different patterns of accumulated partial information such as those illustrated in fig. 6 (Meyer, Irwin, et al., 1988). For example, suppose that the guessing accuracy, which reflects the amount of available below-threshold partial information, remains at a base (chance) level as the lag of the response signal and the guessing-completion time increases. Then this would suggest that the normal processes of stimulus evaluation entail an all-or-none accumulation of information over time (fig. 6, top panel), consistent with a basic discrete stage model (e.g., Sternberg, 1969). If, instead, the guessing accuracy increases rapidly and reaches a stable intermediate plateau after some extended period of chance guessing, then this would suggest an information-accumulation function that has at least one intermediate upward step from its base level (fig. 6, middle panel), consistent with discrete models whose processes produce partial quantized outputs (e.g., Miller, 1982; cf. Ratcliff, 1988). Steady growth of guessing accuracy over time would, by contrast, indicate gradual partial-information accumulation, as expected from some continuous models (e.g., McClelland, 1979).

5.4. Applications of speed-accuracy decomposition

To illustrate the speed-accuracy decomposition technique, we will briefly summarize two experiments in which it has been used for measuring the accumulation of partial information as a function of time during various lexical-decision tasks (Meyer, Irwin, et al., 1988; cf. Meyer & Schvaneveldt, 1971, 1976; Meyer, Schvaneveldt, & Ruddy, 1975). The test stimuli for both experiments were words and nonwords. In one experiment (Meyer, Irwin, et al., 1988, Experiment 5), single words like MAID served as positive stimuli, and single nonwords like BROP served as negative stimuli. Subjects had to determine whether each letter string was a word or nonword and make a corresponding response to indicate their yes-no lexical decision. Six signal lags (viz., 100, 135, 170, 205, 240, and 275 ms on the average) were used during the

[10] The derivation of Equation 2, like the derivation of Equation 1, is straightforward (Meyer, Irwin, Osman, & Kounios, 1988). When the normal and guessing processes race with each other, the observed response on a signal trial will be correct if and only if one of two mutually exclusive outcomes occurs: (a) the normal processes win the race and produce a correct response, or (b) the guessing processes win the race and produce a correct response. The probability of the first outcome is simply $P(t_n \le t_{gs})P_n(\text{correct} | t_n \le t_{gs})$. The probability of the second outcome is $[1 - P(t_n \le t_{gs})]P_g(\text{correct} | t_{gs} < t_n)$. So the probability of a correct signal-trial response, P_s(correct), can be expressed as a sum of the probabilities of these two outcomes. Rearranging the terms of this latter expression yields Equation 2.

signal trials of this experiment. In a second experiment (Meyer, Irwin, et al., 1988, Experiment 3), words and nonwords were paired with each other to form the positive and negative stimuli. The positive stimuli consisted of word pairs like MAID-ROCK and nonword pairs like BROP-MISK, whereas the negative stimuli consisted of pairs with one word and one nonword, such as COAT-WASK or LURT-BARN. This required subjects to decide whether the two letter strings of each pair had the same or different lexical status and then make a corresponding response. Five signal lags (viz., 150, 425, 475, 525, and 575 ms on the average) were used during the signal trials of this experiment.

5.4.1. Results for yes-no lexical decisions

Some results of the experiment involving yes-no lexical decisions about single letter strings appear in fig. 8. The top panel of the figure shows estimated cumulative distribution functions of reaction times (i.e., $F_r(C) = P(t_r \leq C)$) for the positive (word) and negative (nonword) stimuli on regular trials. They represent the completion times of the normal processes under the parallel sophisticated-guessing model. In contrast, the middle panel shows corresponding cumulative distribution functions of reaction times ($F_s(C) = P(t_s \leq C)$) obtained at each of the six signal lags on signal trials, representing completion times associated with the winners of the race between the normal and guessing processes induced by the response signal. At the bottom are estimated cumulative distribution functions of guessing-completion times ($F_{gs}(C) = P(t_{gs} \leq C)$) for each stimulus type and signal lag. These were derived by applying Equation 1 to the results in the top and middle panels, extracting the guessing-completion time distributions from the signal-trial reaction times by partialling out the contributions made by rapidly completed normal processes on signal trials. [11]

As fig. 8 indicates, the reaction times on regular trials were relatively long (top panel). Presenting the response signals on signal trials reduced reaction times significantly, and the reduction grew systematically as the signal lag decreased, presumably because of contributions from the guessing processes induced by the response signals (middle panel). The forms of the estimated guessing-completion time distributions (bottom panel) suggest that decreasing the signal lag shifted these distributions toward the lower end of the time scale but did not otherwise alter their shapes a great deal. [12] This outcome supports

[11] Each distribution in the top and middle panels of fig. 8 was estimated by Vincentizing the specified reaction-time data across individual subjects, including times from both correct and incorrect responses. The reaction times were Vincentized because this provides a way of averaging data while preserving the shapes of individual distribution functions and obtaining unbiased estimates of mean reaction-time quantiles (Thomas & Ross, 1980).

[12] As mentioned earlier, the six signal lags associated with the data in fig. 8 averaged about 100, 135, 170, 205, 240, and 275 ms, respectively. Thus, from the figure, it may be seen that regardless of signal lag, the guessing processes were typically completed within about 200 ms after the onset of a response signal.

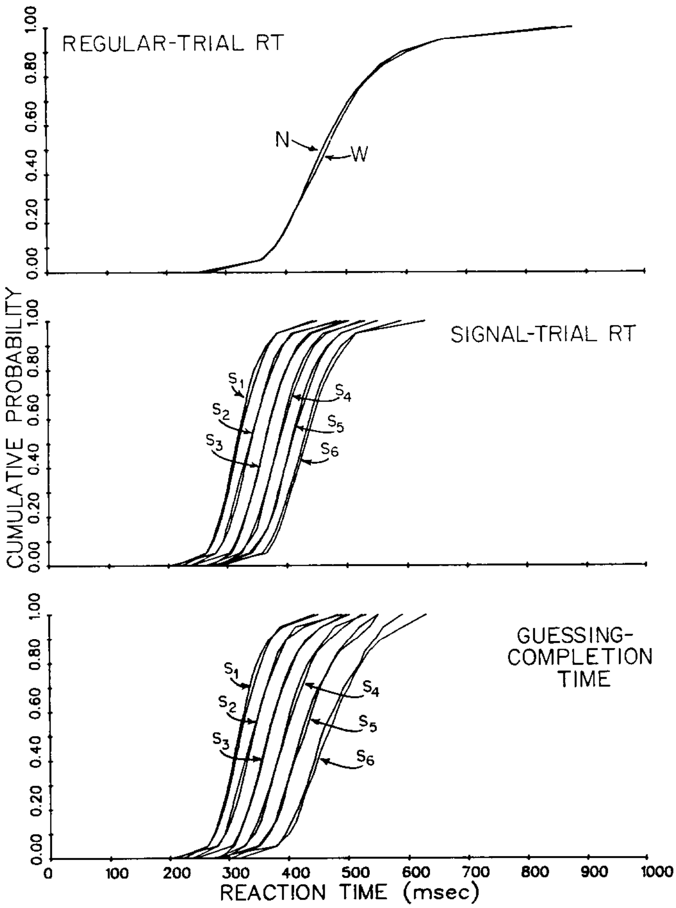

Fig. 8. Vincentized cumulative distribution functions for each stimulus type and signal lag in the experiment with yes-no lexical decisions. (The top, middle, and bottom panels illustrate respectively reaction times on regular trials, reaction times on signal trials, and guessing-completion times derived through Equation 1 of the parallel sophisticated-guessing model. The symbols W and N denote data from word and nonword stimuli, respectively. The symbols s_1 through s_6 indicate which response signals were used to obtain the data, in order of ascending signal lag. From "The dynamics of cognition and action: Mental processes inferred from speed-accuracy decomposition," by D.E. Meyer, D.E. Irwin, A.M. Osman, & J. Kounios (1988). *Psychological Review, 95,* 183–237. Copyright 1988 by the American Psychological Association. Reprinted with permission of authors and publisher.)

the assumptions of the parallel sophisticated-guessing model (Meyer, Irwin, et al., 1988) and provides a basis for using the guessing-completion times in estimating the accuracy of the guessing processes.

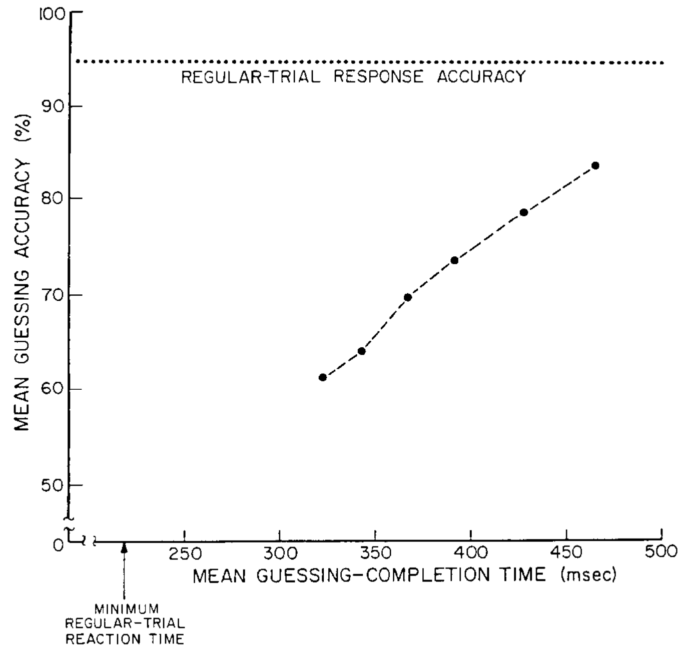

Fig. 9. Mean guessing accuracy (solid points) versus mean guessing-completion times obtained for the six response-signal lags in the experiment with yes-no lexical decisions. (The horizontal dotted line indicates the level of response accuracy on regular trials. The marker on the horizontal axis indicates the minimum regular-trial reaction time for word stimuli, which reflects the completion times of the fastest normal processes. For guesses completed later than the fastest normal processes, the guessing accuracy increases steadily over an extended interval, suggesting a continuous accumulation of partial information during stimulus evaluation. From "The dynamics of cognition and action: Mental processes inferred from speed-accuracy decomposition," by D.E. Meyer, D.E. Irwin, A.M. Osman, & J. Kounios (1988). *Psychological Review, 95,* 183–237. Copyright 1988 by the American Psychological Association. Reprinted with permission of authors and publisher.)

Some results regarding the accuracy of the guessing processes, which we estimated by applying Equation 2 to the response accuracies and reaction times on the regular and signal trials, appear in fig. 9. [13] Here we have plotted the estimated guessing accuracy (solid points) versus the mean guessing-completion time for each of the six response-signal lags. The guessing accuracy reflects how much partial information was accumulated by the normal processes of stimulus evaluation when they did not finish before the response signal was detected and the guessing processes were completed.

[13] Error rates averaged about 5% on the regular trials and, not surprisingly, increased in a monotonic fashion on signal trials as the lag of the response signal decreased.

As fig. 9 shows, the guessing accuracy increased steadily over time. After the longest guessing-completion time, which corresponds to the longest signal lag, the guessing accuracy was nearly 85%. It approached the level achieved by the completed normal processes, which produced approximately 95% correct responses on the regular trials (fig. 9, dotted horizontal line). There was no evidence of any intermediate plateau in the accuracy function. Our findings therefore offer evidence against the hypothesis that stimulus evaluation during yes-no lexical decisions involves a discrete accumulation of information with only one or two sharp transitions (cf. fig. 6, top and middle panels). Instead, it appears that partial information accumulated in a gradual fashion, consistent with assumptions embodied in McClelland's (1979) cascade model and other deterministic continuous models. [14]

However, the present results come from just one case, namely, that in which test stimuli were single letter strings. A different pattern of accumulated partial information could emerge under other circumstances. Indeed, this is what happened in our experiment with same-different lexical decisions about pairs of words and nonwords (Meyer, Irwin, et al., 1988, Experiment 3).

5.4.2. Results for same-different lexical decisions

Some results regarding the same-different lexical decisions appear in fig. 10. This figure shows estimated cumulative distribution functions of regular-trial reaction times (top panel), signal-trial reaction times (middle panel), and guessing-completion times (bottom panel) for each stimulus type and signal lag. We obtained these estimates in the same way as for our previous experiment (cf. fig. 8). There were marked stimulus-type effects on regular trials and signal-lag effects on signal trials. When Equation 1 of the parallel sophisticated-guessing model was applied to the regular-trial and signal-trial reaction times, it yielded relatively well-behaved distributions of guessing-completion times. Consistent with the model's assumptions, the mean guessing-completion times decreased as the signal lag decreased, while the shapes of their distributions were otherwise reasonably invariant as a function of stimulus type and signal lag.

Fig. 11 shows some corresponding estimates of guessing accuracy derived through Equation 2 of the parallel sophisticated-guessing model based on observed signal-trial and regular-trial accuracies. We have plotted the accuracy of the guessing processes as a function of the guessing-completion times after the five response-signal lags used here. The overall pattern is rather different

[14] Another case in which stimulus evaluation may produce a continuous growth of activation over time has been reported by Kounios, Osman, and Meyer (1987). Using speed-accuracy decomposition to analyze results from a sentence-verification task, we found that the accuracy of guessing processes increased steadily as response-signal lag and guessing-completion times increased, just as in our experiment with yes-no lexical decisions (fig. 9).

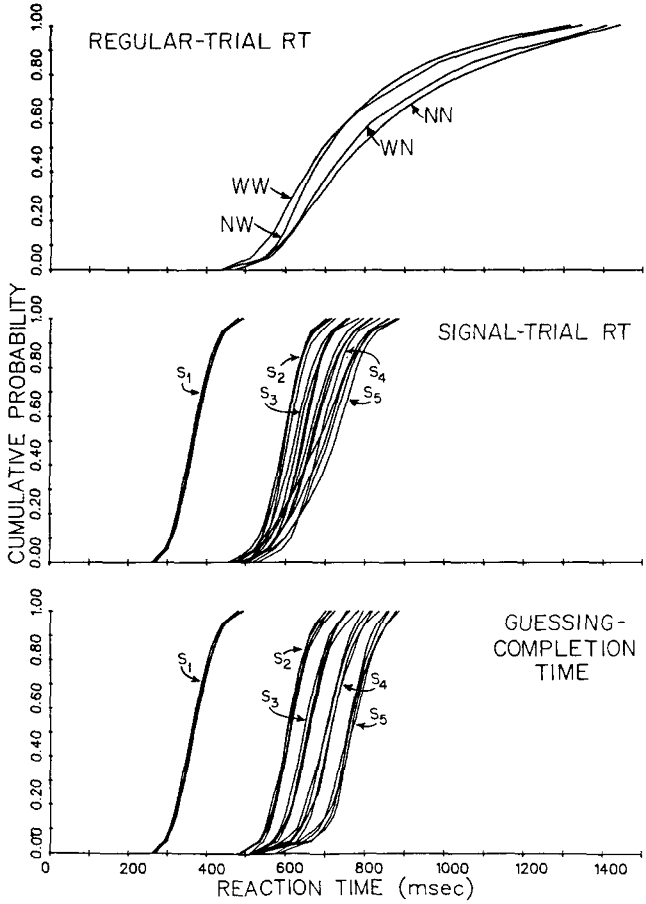

Fig. 10. Vincentized cumulative distribution functions for each stimulus type and signal lag in the experiment with same-different lexical decisions. (The top, middle, and bottom panels illustrate respectively reaction times on regular trials, reaction times on signal trials, and guessing-completion times derived through Equation 1 of the parallel sophisticated-guessing model. The symbols WW, NN, WN, and NW denote data from word-word, nonword-nonword, word-nonword, and nonword-word stimuli, respectively. The symbols s_1 through s_5 indicate which response signals were used to obtain the data, in order of ascending signal lag. From "The dynamics of cognition and action: Mental processes inferred from speed-accuracy decomposition," by D.E. Meyer, D.E. Irwin, A.M. Osman, & J. Kounios (1988). *Psychological Review, 95,* 183–237. Copyright 1988 by the American Psychological Association. Reprinted with permission of authors and publisher.)

than before (cf. Fig. 9). Guessing accuracy did not increase steadily as the signal lag and guessing-completion time increased. Instead, it remained at a virtually chance level for an extended period, until the guessing-completion time nearly equalled the minimum reaction time on regular trials (see fig. 11,

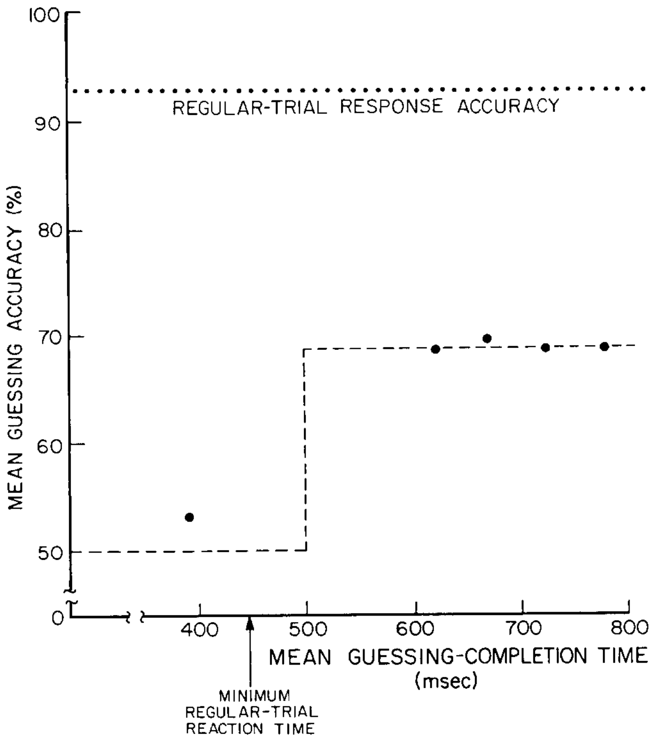

Fig. 11. Mean guessing accuracy (solid points) versus mean guessing-completion times for the five response-signal lags in the experiment with same-different lexical decisions. (The horizontal dotted line indicates the level of response accuracy on regular trials. The marker on the horizontal axis indicates the minimum regular-trial reaction time obtained with word-word stimuli, which reflect the completion time of the fastest normal processes. The dashed step function corresponds to the intermediate output of a hypothetical three-state discrete process of stimulus evaluation, as shown previously in the middle panel of fig. 6. From "The dynamics of cognition and action: Mental processes inferred from speed-accuracy decomposition," by D.E. Meyer, D.E. Irwin, A.M. Osman, & J. Kounios (1988). *Psychological Review, 95,* 183–237. Copyright 1988 by the American Psychological Association. Reprinted with permission of authors and publisher.)

lowest filled circle vs. vertical arrow).[15] Then a flat plateau in the accuracy function emerged soon thereafter (i.e., between 400 and 600 ms after the onset of the test stimuli), as indicated by the dashed step function.

This pattern may have come from a discrete process of stimulus evaluation that produces at least one intermediate quantum of partial information just

[15] The mean guessing-completion time for the response signal that had the shortest lag differed on the order of just 50 ms or so from the minimum regular-trial reaction time. Yet at that lag, the guessing accuracy was still only 52.9%, even though the fastest normal processes (i.e., those associated with the minimum reaction times on regular trials) were highly accurate.

before being completed (fig 6, middle panel). It appears that the information was truly partial. The plateau in guessing accuracy fell at about 68%, substantially below the estimated accuracy of the completed normal processes, which yielded around 93% correct responses on the regular trials (fig. 11, dotted horizontal line). Our findings for same-different lexical decisions are consistent with a discrete stage model of the sort assumed under Sternberg's (1969) additive-factor method. Unlike yes-no lexical decisions about single letter strings, the same-different decisions do not suggest that activation grew in a monotonic deterministic fashion to the level set by normal evaluation processes for initiating correct responses on regular trials (e.g., as in a cascade model, McClelland, 1979; cf. Ratcliff, 1988). [16]

Several factors could explain why a discrete process of stimulus evaluation perhaps occurred in this latter case (Meyer, Irwin, et al., 1988). Subjects had to compare the lexical status of paired words and nonwords. The comparison process was biased toward producing a binary output; two items either had the same lexical status or they did not. Also, lexical status was itself binary; a string of letters was either a word or not. This heavy emphasis on dealing with binary variables may have induced the process of stimulus evaluation to enter a discrete rather than continuous mode of performance. The existence of different performance modes and their variation with task demands are revealed by our speed-accuracy decomposition technique. For more detail regarding these conclusions and their justification, see Meyer, Irwin, et al., (1988).

6. Current status and future prospects

In summary, the current status of mental chronometry can be characterized by considering how its family tree has grown over the years (fig. 1). Emerging from the original work of Helmholtz (1850/1853), the tree's right branches have spread through ideas stimulated by Donders' (1868/1969) subtraction method and the discrete stage model on which it rests. The left branches of the tree have taken a complementary direction, following Woodworth's (1899) development of speed-accuracy tradeoff curves and notions about stochastic variability in human performance. This growth has led ultimately to concern over the serial-versus-parallel and discrete-versus-continuous distinctions, which differentiate alternative models of information processing and guide a

[16] There are some continuous models for which the data from the same-different lexical decisions seem more consistent. Ratcliff (1988) has shown that under certain circumstances, a stochastic diffusion model involving a random drift of response strength can yield an apparent plateau in guessing accuracy after intermediate signal lags. Other aspects of our results may, however, cast doubt on the latter alternative (Meyer, Irwin, Osman, & Kounios, 1988).

quest for more powerful chronometric methodology (Meyer, Irwin, et al., 1988; Miller, 1988). Although we have not included our own research as part of the tree diagram, one may view it as a natural off-shoot of past advances. To be specific, the speed-accuracy decomposition technique embodies a hybrid combination of procedures associated with the tree's right and left branches, perpetuating the traditions of both Donders (1868/1969) and Woodworth (1899). As such, it is intended to address certain basic issues related to the seminal research of these classical investigators.

Of course, many questions about the dynamics of cognition and action still remain. The paradigm of mental chronometry has not yet proven sufficient to isolate and analyze all of the important parts of the human information-processing system. One might thus wonder what can be done henceforth to augment the paradigm and to ensure its continued fertility.

This brings us at last to the proposed marriage between mental chronometry and cognitive psychophysiology (fig. 1, tree top). As mentioned in the introduction to our survey of the chronometric paradigm, some investigators have expressed considerable enthusiasm about supplementing reaction-time and speed-accuracy tradeoff procedures with batteries of psychophysiological indicators (ERPs, EMG recordings, etc.). The wave of enthusiasm has risen partly in response to obvious weaknesses in the chronometric paradigm. Because standard behavioral measures obtained through mental chronometry represent the total duration and final output of many processing stages in combination, they do not offer an especially close look at underlying component processes. However, cognitive psychophysiology can perhaps help overcome this limitation by examining these components more directly.

An eloquent expression of such possibilities has appeared in recent articles by Coles, Gratton, and their colleagues, who have noted:

> The task of describing the information processing system would be considerably easier if we had measures of the activity of particular elements of the system, as well as measures of its output. This is where psychophysiology may help. If psychophysiological measures are sensitive to *particular* information processing activities, then we should be able to use them to understand how these processes interact to produce the behavioural output (Coles & Gratton, 1986, p. 409).
>
> In particular, the (psychophysiological) measures can provide information about the interactions between processes associated with stimulus evaluation and processes that are required for the actual execution of responses (Coles et al., 1985, p. 529).

The latter sentiments are not entirely new. Their heritage dates back to the time of Helmholtz (1850/1853), who introduced the simple reaction-time procedure to supplement his physiological experiments on the rate of neural conduction. It remains, nevertheless, a vibrant hope that cognitive psychophysiology may significantly augment modern mental chronometry and thereby yield deeper insights into the processes and products studied previously through the chronometric paradigm.

It is likewise important to recognize that the marriage between cognitive psychophysiology and mental chronometry need not be an asymmetric partnership wherein the psychophysiological approach bears most of the burden for subsequent advances. The chronometric paradigm brings a rich set of assets to this marriage, from which cognitive psychophysiologists can profit substantially. The assets include: (a) a set of versatile procedures for collecting experimental data, (b) many basic facts about the speed and accuracy of human performance in various cognitive tasks, (c) precise alternative models for characterizing the representation, transformation, and transmission of information through the information-processing system, (d) analytical techniques to help evaluate data and test models, (e) significant questions and issues to guide future research, (f) fundamental lessons, learned the hard way, about the vicissitudes of research on information-processing dynamics. By taking each of these assets into account, the psychophysiological approach may evolve more fruitful applications of its methodology and more meaningful interpretations of its measures (Chase, McCarthy, Squires, & Schvaneveldt, 1984).

To illustrate what prospects lie ahead, the following sections discuss and evaluate advances made so far through the marriage between mental chronometry and cognitive psychophysiology. We will briefly outline the basics of the psychophysiological approach. Next we review interesting results from some representative studies that have taken this approach to expand on several facets of the chronometric paradigm. Then we assess the success of the overall endeavor as it now stands, summarizing its accomplishments and noting some dangerous pitfalls that could impede its further progress.

7. ERPs and information processing

Our discussion here concentrates primarily on psychophysiological research involving event-related brain potentials (ERPs), which are transient voltage fluctuations generated in neural tissue immediately before or after the occurrence of stimulus events. In this research, cognitive psychophysiologists have recorded ERPs from subjects' scalps during the performance of various experimental tasks similar to those used as part of the chronometric paradigm (e.g., Coles, Donchin, & Porges, 1986; Gaillard & Ritter, 1983; Hillyard & Kutas, 1983). The time-locked electrical activity evoked in the brain by presented stimuli has revealed systematic components of the ERP. These components may be measured in terms of peak-to-trough or base-to-peak deflections of the ERP signal after stimulus onset, thereby tapping related phases of information processing.

7.1. Classification of ERP components

There are two general classes of ERP components; *exogenous* and *endogenous* (Hillyard & Kutas, 1983). The exogenous components occur within 100 ms or so after the onset of a stimulus. They may vary with physical parameters of the stimulus, but are essentially obligatory and depend little, if at all, on cognitive demands of the subjects' task. Their properties are believed to reflect relatively peripheral sensory mechanisms. In contrast, the endogenous components occur 100 ms or more after stimulus onset. They typically emerge when the subjects' task entails certain central processes associated with the ERP. Their properties are believed to reflect the processes of perception, attention, memory retrieval, decision, response preparation, and so forth. For present purposes, it is the endogenous components that will concern us most.

Cognitive psychophysiologists have characterized the endogenous components of the ERP more precisely on the basis of several defining criteria: (a) positive or negative polarity, (b) modal latency from the moment of stimulus onset to the moment at which the component's peak occurs, (c) morphology of the component's waveform, (d) spatial distribution, the component's relative amplitude at different recording sites on the scalp, (e) pattern of sensitivity to various experimental factors, (f) underlying source, that is, the neural generator(s) from which the component emanates. With such criteria as a frame of reference, several endogenous ERP components have been identified, including N200, P300, N400, and the RP (readiness potential). Some of these components appear directly in the overall waveform of the ERP, whereas others (e.g., N_A and Nd) emerge when one waveform is subtracted from another (Hillyard & Kutas, 1983). For example, the P300 component represents a positive voltage deflection that, in generic experiments, peaks at a modal time around 300 ms after stimulus onset. It tends to have greatest amplitude at scalp sites located near the parietal regions of the brain, and is sensitive to factors associated with variations in subjects' expectancies (e.g., stimulus probability; Duncan-Johnson & Donchin, 1982). [17]

Because of background noise in the brain's on-going electrical activity, detailed analyses of the endogenous ERP components require signal-processing and pattern-recognition techniques (Coles, Gratton, Kramer, & Miller, 1986). One may filter the ERP digitally and/or average it across trials to attenuate

[17] There is still some residual ambiguity about whether certain ERP components belong to the exogenous or endogenous class. This ambiguity applies, for example, to the N100 component, which some authors (e.g., Renault, 1983) have treated as exogenous but others (e.g., Donchin, McCarthy, Kutas, & Ritter, 1983) have treated as endogenous. The N100 component may exhibit exogenous properties but appear, in part, to be endogenous (i.e., sensitive to cognitive task demands) because another endogenous component, the Nd wave, which depends on subjects' attention, overlaps with it (Hansen & Hillyard, 1980; Hillyard & Kutas, 1983; Näätänen & Michie, 1979).

contributions of the noise. Techniques such as linear discriminant analysis (Donchin, 1969), template matching through cross correlation (Kutas et al., 1977), and simple peak picking provide possible ways to estimate parameters of particular components. In applying these techniques, ancillary assumptions must of course be made. For example, parameter estimation in terms of discriminant analysis assumes that each underlying component's morphology and temporal locus are invariant across trials.

7.2. *Inference rules*

After extracting the endogenous components of the ERP and estimating their parameters, cognitive psychophysiologists have used several related inference rules to interpret obtained results theoretically. These rules are not officially codified in any single place, but one may educe them from examining a variety of representative studies. They concern the implications of joint factor effects on behavioral measures (e.g., reaction time) and parameters of ERP components (e.g., amplitude and latency). An especially important parameter in this respect is *peak latency*, that is, the length of time between stimulus onset and the occurrence of a component's greatest amplitude. In particular, the peak latency has often been treated as an indicator of mental processes thought to precede and mediate an individual component. Underlying this interpretation are several implicit assumptions, which we discuss more fully later, a key one being that the termination of certain processes determines the moment when the peak amplitude of a component occurs.

We will consider six important inference rules here. Some of these provide a basis for assessing the functional significance of ERP components and associating them with particular mental processes. Others are intended to demonstrate the existence of separate processing stages or to determine the locus of factor effects in such stages (cf. Sternberg, 1969). We do not claim that the set of rules outlined below is foolproof or exhaustive, but it does serve to illustrate the style of reasoning often adopted by cognitive psychophysiologists.

Rule 1: Functional significance of ERP components. If an experimental factor is believed to influence a particular mental process, and if the parameters of an ERP component depend on this factor, then infer that the component is a manifestation either of this process or of some other subsequent related process. With this rule, one can assess a component's functional significance and narrow the range of processes that the component may conceivably manifest.

Rule 2: Locus of factor effects. If an ERP component is believed to manifest a particular mental process, and if an experimental factor has equal effects on the component's mean peak latency and mean reaction time, then infer (a) that the factor only influences this process or some other preceding one(s), and (b) that the process(es) in which the effect takes place mediate overt responses to presented stimuli. With this rule, the effects of experimental factors may be localized in a subset of processing stages that form part of the stimulus-response pathway. Rule 2 also provides a link between components of the ERP and observed behavior.

Rule 3: Locus of factor effects. If an ERP component is believed to manifest a particular mental process, and if an experimental factor has a greater effect on mean reaction time than on the component's mean peak latency, then infer that the factor influences some additional (e.g., subsequent) process(es) whose operation mediates overt responses to presented stimuli but not the generation of the component. Moreover, a corollary of this rule is that if the factor in question has any effect on the mean peak latency of the component, then infer that its effect also occurs partly during some prior process(es) whose operation does mediate both the component and overt responses to presented stimuli. The latter inference complements the ones drawn from Rule 2. In certain respects, however, Rule 3 is more powerful than Rule 2, because it not only helps determine the locus of factor effects but also demonstrates the existence of other related processes. Further ways of achieving such demonstrations are embodied in the next two rules.

Rule 4: Existence of processing stages. If an ERP component is believed to manifest some mental process, and if two experimental factors have additive effects on the component's mean peak latency, then infer that these factors influence two distinct processing stages, including (a) either a stage associated with this component or some preceding stage, and (b) another even earlier stage. In addition, a corollary of this rule is that if two factors have interactive effects on a component's mean peak latency, then infer that these factors both influence either the process manifested by the component or some preceding process. Rule 4 therefore constitutes a direct extension of Sternberg's (1969) additive-factor method to ERP component latencies. To apply it, one must manipulate two factors orthogonally and measure their joint effects on a single component's latency. Interestingly, with ERPs, there is also another complementary way of testing for the existence of processing stages. As stated in Rule 5, this alternative involves manipulating a single factor and measuring its respective effects on the mean peak latencies of two successive ERP components.

Rule 5: Existence of processing stages. If an early ERP component and a later ERP component are believed to manifest two distinct mental processes, and if an experimental factor has equal effects on the mean peak latencies of the two components, then infer that the processes constitute nonoverlapping stages and that this factor influences the first of these stages or some even earlier stage. Rule 5, like Rule 4, may be used both to demonstrate the existence of processing stages and to localize the effects of factors in them. No direct analog of it is available for reaction-time data, because they do not provide estimates of the termination for each of two (or more) individual stages that operate in tandem.

Rule 6: Locus of factor effects. If an ERP component is believed to manifest a particular processing stage, and if the component's peak latency but not its onset time (i.e., the time between stimulus presentation and the start of the component) depends on an experimental factor, then infer that the factor influences this stage but not any earlier stages. This rule further constrains the locus of effects that a factor has in selected stages of processing (cf. Rule 2).

Rationale

The rationale for the six inference rules should be apparent already. It rests on the basic stage model of information processing formulated originally in mental chronometry (Donders, 1868/1969; Sternberg, 1969). As was the case earlier, the working assumption here is that processing stages are strictly successive (i.e., have no temporal overlap). It is also assumed that the peak latency of an ERP component includes the summed durations only of those stages up to and including the one putatively manifested by the component (plus a possible residual delay). Given the latter assumption, the effect of a factor on the duration of an early stage should propagate through the system to affect later components' peak latencies by the same amount (Rule 5). Furthermore, two factors that influence different stages whose durations both contribute to a component's peak latency should affect this latency additively, just as they do overt reaction time (Rule 4). However, because ERP latencies embody selected subsets of stage durations, they may have greater diagnosticity than do reaction times, which combine the durations of all stages up to and including response execution.

For example, fig. 12 shows one way that these possibilities might be realized. Here we have depicted a hypothetical version of the stage model in which individual components of the ERP are tentatively linked to successive mental processes. According to this view, a series of stages (S_i; $i = 1, 2, \ldots, n$) leads from stimulus input to response output, with various factors (viz., F_1, F_2, and F_3) affecting the durations of different stages. Associated with each stage is an ERP component (C_i; $i = 1, 2, \ldots, n$). Under the indicated arrangement, the six inference rules for interpreting component latencies and reaction times would apply.

Note that the respective components do not necessarily have to emerge in the ERP waveform at exactly the same times as the corresponding stages take place. Some residual delay could intervene between the execution of a stage

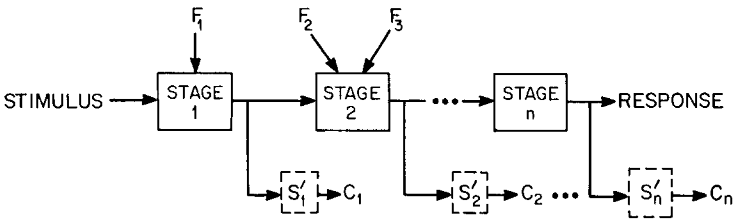

Fig. 12. Stage model in which individual components ($C_1, C_2, \ldots, C_n$) of the ERP are tentatively linked to successive processing stages ($S_1, S_2, \ldots, S_n$). (The solid boxes indicate stages in the main stream of processing that leads from stimulus input to response output. The dashed boxes indicate other ancillary stages ($S'_1, S'_2, \ldots, S'_n$) that are not in the main stream but that may also be precursors of the components. F_1, F_2, and F_3 denote three factors that influence various stages.)

and the occurrence of its associated ERP component. Other ancillary stages (S'_i; $i = 1, 2, \ldots, n$) outside the main stream of processing (fig. 12, dashed boxes) could also be immediate precursors of the components and lengthen their latencies. Factor effects on the durations of mainstream processing stages (fig. 12, solid boxes) would, nevertheless, appear fully in the peak latencies of selected components, assuming that the peak latencies are indeed a function of stage-termination times. More precisely, the key assumption is that $t_{ci} = t_{s1} + t_{s2} + \ldots + t_{si} + t_{di}$, where t_{ci} is the peak latency of the ith component, t_{si} is the duration of the ith main stage, and t_{di} is extra residual delay ($i = 1, 2, \ldots, n$) due to ancillary processing events outside the main stream.

7.3. *Illustrative studies*

To illustrate some applications of the preceding inference rules for interpreting reaction times and ERP component latencies theoretically, we will consider some findings from four representative studies. These studies concern the mental processes of stimulus discrimination and identification, memory retrieval, and response preparation manifested through the N_A, N200, P300, and RP components, as well as EMG activity and reaction-time data. Together they exemplify the kinds of accomplishments that have resulted from wedding the chronometric paradigm with cognitive psychophysiology.

7.3.1. *McCarthy and Donchin (1981)*

Our first example comes from a study by McCarthy and Donchin (1981, 1983). Here subjects made responses to the printed words RIGHT and LEFT by pressing buttons with their right and left hands. Two independent variables were manipulated orthogonally. One was stimulus discriminability; either the words appeared against a background array of pound (#) symbols (high-discriminability condition), or they appeared against a background of other random letters (low-discriminability condition). The other variable was stimulus-response compatibility; either the word RIGHT required a right-hand button press and the word LEFT required a left-hand button press (high-compatibility condition), or the mapping between the stimulus words and button presses was reversed (low-compatibility condition). On each trial of these conditions, three dependent variables were measured: Peak latency of the P300 component, overt reaction time, and response accuracy.

Replicating results from a related experiment by Sternberg (1969), McCarthy and Donchin (1981, 1983) found that stimulus discriminability and stimulus-response compatibility had additive effects on mean reaction time. This outcome supports Sternberg's (1969) original suggestion that the discriminability and compatibility factors may influence two different stages of processing, to wit, stimulus evaluation (e.g., encoding and identification) and response selection, respectively. Furthermore, the mean of the P300 component's peak latency varied nearly as much with stimulus discriminability as did mean reaction time, but this latency depended hardly at all on stimulus-response compatibility.

Given these results, McCarthy and Donchin (1981, 1983) concluded that the latency of the P300 component reflects the duration of stimulus-evaluation processes, and that the effects of stimulus discriminability are localized in these processes (cf. Duncan-Johnson & Donchin, 1982). They also concluded that whatever processing stage is influenced by stimulus-response compatibility occurs after stimulus evaluation has been completed. Their conclusions follow from Rules 1 through 3 outlined earlier (see *Inference Rules*), illustrating how the approach of cognitive psychophysiology can simultaneously reinforce hypotheses developed previously through mental chronometry and how behavioral measures can help determine the functional significance of ERP components.

7.3.2. Ford et al. (1979)

Our second example comes from a study by Ford et al. (1979). Here subjects had to perform a memory-scanning task similar to one that Sternberg (1966, 1969) has used. Two groups of subjects were included; young and elderly adults. Their task involved a series of trials, on each of which a short list of items (alphanumeric characters) was retained in memory and a speeded decision was made about whether or not a presented test stimulus belonged to the list. The peak latency of the P300 component, overt reaction time, and response accuracy were measured as a function of the list length (i.e., memory load) and subjects' age.

Consistent with results reported originally by Sternberg (1966), Ford et al. (1979) found that mean reaction time increased linearly as list length increased. The linear list-length effect on mean reaction time may reflect a process in which short-term memory is searched serially for each test stimulus (Sternberg, 1966, 1969). In addition, there was a linear effect of list length on the peak latency of the P300 component (Ford et al., 1979; cf. Marsh, 1975). This latter effect did not depend significantly on the age of the subjects. Although the P300 component's peak latency was a bit longer for the elderly subjects, their latency function had about the same slope as that of the young subjects.

In terms of Rule 4, it therefore appears that list length and age influence two separate stages of processing, one or both of which take place before the process manifested by the P300 component. For example, list length may influence a serial memory-search stage, whereas age influences a prior encoding process. The P300 component's peak latencies suggest that the rate of memory search, as quantified by the list-length effect on them, was about the same regardless of the subjects' age.

However, list length had a significantly greater effect on mean reaction times than on peak latencies of the P300 component. An especially big difference between the list-length effects for these two measures occurred in the elderly subjects. Assuming reaction times manifest both stimulus evaluation and other processes thereafter, whereas the P300 component is a relatively pure manifestation of stimulus evaluation, Ford et al. (1979) inferred that list length influences not only a memory-search stage but also one or more subsequent stages in which responses are mobilized for execution. The young and elderly subjects may differ considerably from each other during the latter stage(s), with the elderly subjects taking more time there than the young do. This inference, which goes beyond conclusions reached by Sternberg (1966, 1969) about encoding, memory search, and response execution, illustrates another application of Rule 3 for demonstrating the existence of processing stages and determining the locus of factor effects.

7.3.3. Coles et al. (1985)

Further insights regarding the process of response execution and its relation to stimulus evaluation have been obtained by Coles, Gratton, and their colleagues (Coles & Gratton, 1986; Coles et al., 1985) who again used Rules 1 and 3 for making theoretical inferences. Here subjects produced rapid squeeze responses with their right and left hands for two different target letters, H and S, respectively. The stimuli were presented visually at a central fixation point and were surrounded by a flanking array of "noise" characters. In one case, the compatible-noise condition, the noise characters had the same identity (e.g., H) as the target stimulus (e.g., H). In another case, the incompatible-noise condition, the noise characters had the same identity as the other unpresented target stimulus (e.g., S). Several dependent measures of subjects' performance were recorded under each condition, including response accuracy, reaction time, peak latency of the P300 component, latency of onset in EMG activity, and subcriterion squeeze activity (i.e., squeezing that did not exceed a threshold set for overt responses). Paralleling the results of Ford et al. (1979) and some other investigators (e.g., Duncan-Johnson & Donchin, 1982), Coles et al. (1985) found that a selected factor, namely, noise incompatibility, increased the P300 component's mean peak latency but had an even greater effect on mean reaction time. It was concluded, therefore, that noise incompatibility influenced both the process of stimulus evaluation and a subsequent process associated with response execution, as Rule 3 dictates.

To analyze response execution more closely, Coles et al. (1985) examined EMG and subcriterion squeeze activity as a function of noise incompatibility. When there was incompatible noise, subjects often tended toward partial incorrect covert responses before eventually producing complete correct overt responses. On the basis of these and other related results, a continuous-flow model of stimulus evaluation and response execution was advocated (cf. Eriksen & Schultz, 1979). According to this model, parallel activation of an incorrect response increases temporarily over time as the evaluation process transmits partial information about the identities of the incompatible-noise characters, thereby creating response competition before the central target stimulus has been evaluated completely and used to initiate a correct response.

Some additional evidence that activation in the response system grows continuously has come from subsequent analyses by Coles and Gratton (1986). They plotted the results of Coles et al. (1985) in terms of a normalized speed-accuracy tradeoff curve. This normalization entailed graphing subjects' response accuracy (percentage of correct responses) versus a ratio computed by dividing the peak latency of the P300 component into the overt reaction time on a trial-by-trial basis. For each of several small intervals over the range of this ratio ($0.2 \leq RT/P300 \leq 1.0$), the accuracy of responses whose parameters fell in that interval was determined. [18] The results revealed a smooth tradeoff curve, which first tended toward below-chance accuracy because of response competition and then reversed direction, gradually increasing toward perfect accuracy.

According to Coles and Gratton (1986), this outcome implies a continuous process of response preparation before overt physical movement takes place. By dividing the P300 latency into the observed reaction time on each trial, it is possible that they at least partially removed contributions in the speed-accuracy tradeoff curve from a discrete all-or-none stimulus-evaluation process with stochastic transition times. The normalized tradeoff curve may manifest the nature of residual response-preparation processes more clearly, helping to overcome our previous criticism of standard (average) speed-accuracy tradeoff curves (Meyer & Irwin, 1981; Meyer, Irwin, et al., 1988).

[18] In technical terms, this curve constitutes a form of "micro-tradeoff" between speed and accuracy (Pachella, 1974), representing accuracy as a function of trial-by-trial changes in processing speed.

In seeking more evidence to test the continuous-flow model and response-competition hypothesis, Coles and Gratton (1986) also examined another aspect of the ERPs, which they have termed the *lateralized readiness potential* (LRP; cf. Coles, Gratton, & Donchin, 1988). The LRP was measured by recording activity with electrodes on the scalp over the motor cortices of the two cerebral hemispheres and then subtracting the activity for one hemisphere from the activity for the other hemisphere. This measure yielded a number of interesting results. Before relatively fast correct and incorrect responses, the LRP tended to increase such that it anticipated the side of the impending response, even when a test stimulus had not appeared. The observed tendency suggested a form of "aspecific response priming." After stimuli were presented on trials involving the incompatible-noise condition, the LRP often dipped briefly toward the side of the incorrect response, even when a correct overt response ultimately occurred. The latter dip helps confirm the response-competition hypothesis. Finally, EMG activity in the muscles used to produce the overt responses typically began at a moment when the LRP reached a set level, regardless of what the stimulus condition was and how long the LRP took to reach this level. It therefore appears that response execution may entail crossing a fixed threshold of activation, as postulated in the cascade model (McClelland, 1979) and other continuous models (e.g., Ratcliff, 1978).

7.3.4. Ritter, Simson, and Vaughan (1983)

Complementing these detailed findings about response preparation and execution, Ritter, Simson, and Vaughan (1983) have studied more closely the processes associated with pattern recognition and stimulus categorization, applying Rules 5 and 6 to draw theoretical inferences. Their study included both a simple reaction-time task (one stimulus-response combination) and a go/no-go reaction-time task (multiple stimuli, only one type of which required a response; cf. Donders, 1868/1969). On each trial of the simple-RT task, subjects lifted their right index fingers as quickly as possible when a prespecified visual test stimulus (e.g., the pattern < >) appeared. On each trial of the go/no-go RT task, there were two different possible classes of test stimuli (e.g., the patterns < > and > <, respectively). These stimulus classes had unequal probabilities of occurrence, namely, 0.8 and 0.2. For the stimuli with a probability of 0.2, subjects lifted their right index fingers as quickly as possible, yielding the "go" trials. No overt responses were supposed to be produced for the stimuli with a probability of 0.8, yielding the "no-go" trials. The go/no-go RT task therefore required stimulus evaluation on all trials and response selection on some trials, whereas the simple-RT task required little evaluation or selection.

As part of this task manipulation, another factor was also varied. Either the test stimuli were relatively easy to discriminate (e.g., the pattern < > vs. the pattern > <), or they were harder to discriminate (e.g., the pattern > > > >

vs. the pattern < > > > >). ERPs and overt reaction times were recorded as a function of the discriminability factor and task type (simple RT vs. go/no-go RT).

For these measures, Ritter, Simson, and Vaughan (1983) focused specifically on two ERP components. First, the average ERP obtained in the simple-RT task was subtracted from the average ERP obtained in the choice-RT task with the stimuli that had a 0.8 probability of presentation. This yielded a residual negative ERP component, termed $\mathrm{N_A}$, whose peak latency occurred on the order of 250 ms after stimulus onset. The $\mathrm{N_A}$ component's peak latency was hypothesized to manifest completion of a pattern-recognition stage. Although the 0.8-probability stimuli required no overt responses, pattern recognition may have been needed for them because they had to be discriminated from the 0.2-probability stimuli, whereas no such discriminations were needed in the simple-RT task.

Second, the average ERP obtained in the choice-RT task with the 0.8-probability stimuli was subtracted from the average ERP obtained with the 0.2-probability stimuli. This yielded another negative ERP component, termed N2, whose peak latency occurred on the order of 300 ms after stimulus onset. The N2 component's peak latency was hypothesized to manifest completion of a subsequent stimulus-categorization stage.

To support these hypotheses regarding the $\mathrm{N_A}$ and N2 components, Ritter, Simson, and Vaughan (1983) examined the magnitudes of the peak latencies and their variation with stimulus discriminability. The N2 component's peak latency in the average ERP exceeded the $\mathrm{N_A}$ component's peak latency, and the discriminability factor had almost identical effects on both of them, difficult discriminations yielding greater latencies in each case. Consistent with Rule 5, this outcome led Ritter, Simson, and Vaughan (1983) to infer that two distinct stages of processing (viz., pattern recognition and stimulus categorization) were associated with $\mathrm{N_A}$ and N2, respectively, and that stimulus discriminability influenced the first of these stages.

That the discriminability factor might influence some processing stage earlier than pattern recognition was rejected on other grounds. Ritter, Simson, and Vaughan (1983) found that although the $\mathrm{N_A}$ component's peak latency depended on stimulus discriminability, its onset latency did not. Paralleling Rule 6, this result suggests that the discriminability effect must have occurred in the recognition stage, because if the effect had taken place any earlier, then it should have propagated ahead to the $\mathrm{N_A}$ onset latency, not just the $\mathrm{N_A}$ peak latency (see *Rationale, Inference Rules*).

7.3.5. Summary

Taken as a whole, the four ERP-RT studies that we have reviewed here offer an impressive illustration of how cognitive psychophysiology has both enhanced mental chronometry and benefited from it. Through judicious application of Rules 1 through 6 for theoretical inferences based on psychophysiological variables, one may make strides toward assessing the functional significance of ERP components, demonstrating the existence of information-processing stages, and determining the locus of factor effects in those stages. The range of mental processes spanned by this approach extends from relatively early stages of pattern recognition (e.g., Ritter, Simson, & Vaughan, 1983) to intermediate stages such as short-term memory search (e.g., Ford et al., 1979) and later stages involving response selection, motor programming, and execution (Coles & Gratton, 1986; Coles et al., 1985; McCarthy & Donchin, 1981, 1983). A number of other examples could have been mentioned as well (e.g., see Gaillard & Ritter, 1983; Hillyard & Kutas, 1983). In particular, studies that use combinations of RT and ERP measures to focus on the interface between stimulus evaluation and response execution seem especially suited to pursuing the serial-versus-parallel and discrete-versus-continuous distinctions outlined previously under the chronometric paradigm (cf. Meyer, Irwin, et al., 1988; Miller, 1988).

8. Limitations of cognitive psychophysiology

Still, despite the many accomplishments by studies like those above, cognitive psychophysiology has some serious limitations. There are significant weaknesses in the theoretical and methodological foundations of the psychophysiological approach as it currently exists. These weaknesses could undermine conclusions reached through the preceding inference rules, creating frustrations that might precipitate an acrimonious ERP-RT divorce. Unless new groundwork takes place to improve the present situation, the marriage of mental chronometry and cognitive psychophysiology will face a rough road ahead.

8.1. Theoretical weaknesses

8.1.1. Relation between ERP latencies and stage-termination times

One critical weakness in the theoretical foundation of cognitive psychophysiology involves the putative relation between the latencies of ERP components and key events during associated mental processes. As we mentioned already, many psychophysiologists have assumed that the peak latencies of selected components (e.g., N200, P300, etc.) correspond closely to the times at which particular stages of processing (e.g., stimulus evaluation) terminate. A component's peak latency is not necessarily believed to equal a stage's termination time exactly, but the peak latency is often interpreted as if it and the termination time of a stage differ only by a constant amount due to some residual delay of the component in the ERP waveform (e.g., see Coles &

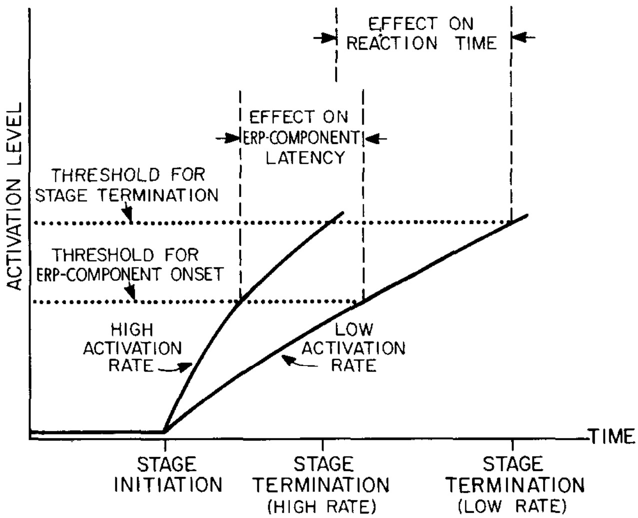

Fig. 13. Activation mechanism for a stage of processing in which factor effects on the rate of activation may produce larger changes in mean reaction time than in mean peak latency of an ERP component associated with the stage. (The solid curves indicate different rates of activation induced by different levels of a given factor. The dotted horizontal lines indicate intermediate and high thresholds that respectively trigger the release of the component and terminate the stage when activation crosses them.)

Gratton, 1986; Coles et al., 1985; Ford et al., 1979; McCarthy & Donchin, 1981; Ritter, Simson, & Vaughan, 1983). Up to now, this interpretation has received little real justification. The primary basis for focusing on peak latencies appears to be that they are easier to measure than onset latencies, offset latencies, and other ERP time points.

It is possible instead that the onset latency, offset latency, or some other nondescript time point of an ERP component actually corresponds more closely to the termination of a processing stage than does the peak latency. Perhaps the component's peak latency manifests events during the middle of the stage's execution rather than an event at its end. If such were the case, then many of the previous inference rules for interpreting ERP and reaction-time measures would have to be reformulated, and the conclusions derived from them would, in at least some cases, need reassessment (for a similar critique, see Miller, 1988).

For example, consider the following scenario, as shown in fig. 13. Suppose that a particular stage of processing entails a continuous activation mechanism with two thresholds, one set at an intermediate level and the other at a somewhat higher level. Furthermore, suppose that as activation increases gradually during this stage, it triggers the subsequent onset of a corresponding ERP component when the intermediate threshold is crossed, whereas the stage terminates and transmits its output to later response-related processes when the high threshold is crossed. Then for factors that affect the duration of the first stage, the peak latency of the component would not necessarily exhibit the same magnitudes of effects as do overt reaction times. To be specific, if these factors alter the rate at which activation rises toward threshold, their effects on mean reaction time would exceed their effects on mean peak latency, even though the factors do not influence any other later (or earlier) processes. Consequently, troubling violations of Rule 3 and of conclusions based on it could arise.

To see how this might cause problems, let us review the study by Ford et al. (1979). Some results obtained there, which we summarized previously (*Illustrative Studies*), revealed a larger effect of list length on mean reaction time than on the P300 component's peak latency in a memory-scanning task. Applying Rule 3 of ERP interpretation, Ford et al. (1979) took these results to imply that list length influences one or more response-related stages after stimulus evaluation has been completed. This inference rests on the assumption that, regardless of list length, the mean peak latency of the P300 component manifests the end of stimulus evaluation, differing by no more than a constant amount from the average time at which the evaluation process terminates relative to stimulus onset. Another possibility, however, is that some intermediate event during stimulus evaluation triggers the P300 component before the evaluation process reaches completion and transmits its output for subsequent decisions and responses. The differential effects of list length on mean peak latencies and mean reaction times could result from the sort of multiple-threshold activation mechanism suggested above (Ratcliff, 1978). If the system works like this, then contrary to the conclusions of Ford et al. (1979), one should not infer that list length affects the duration of any later stage (e.g., decision, response selection, and execution) after stimulus evaluation has terminated.

A similar account would perhaps explain the results of other studies whose experimental factors (e.g., stimulus probability, noise compatibility, etc.) have yielded larger effects on mean reaction times than on peak latencies of P300 and related ERP components (e.g., Duncan-Johnson & Donchin, 1982). We do not know for sure that any selected component necessarily manifests all of the activity in some stage of processing associated with it (Miller, 1988). So before reaching definitive conclusions from differences between the magnitudes of factor effects on reaction times and component latencies, better validated models of correspondences between intrastage processing events and temporal loci of ERP components are needed. [19]

8.1.2. Identification of neural generators

As part of this modeling effort, it may prove helpful to identify more precisely the underlying neural generators of ERP components (Allison, Wood, & McCarthy, 1986). At present, cognitive psychophysiologists have not yet determined exactly what sections of neural tissue in the brain's substrates are sources of particular components such as N200, P300, and so forth. These components could conceivably result from a composite of more or less asynchronous activity at multiple intracranial sites whose functions differ considerably (Allison et al., 1986; Wood & Allison, 1981; Wood et al., 1984). If so, then such complexities must be accommodated by detailed models of "stage-wave relations" between mental processing events and ERP-component time points. The models should, in essence, specify joint correspondences involving three related types of entities: functional stages of processing, ERP components, and neural generators.

8.1.3. Treatment of stochastic variability

On the basis of our experience with information-processing models developed through past efforts in mental chronometry, it seems clear that future generations of models in cognitive psychophysiology must incorporate some other important features as well. They should deal directly with the inherent stochastic variability of processing stages. Such stages do not always start or stop at the same time on each trial, nor do the same outputs always emerge from them, even if stimulus inputs remain the same. This variability is unlikely to be well accommodated by attributing it simply to background noise separate from the basic processes at hand; it is not the same as ancillary noise of an electrical sort superimposed on an otherwise deterministic ERP signal. Rather, stochastic elements must play an integral role in the mental and physical processes postulated by theorists for viable models of cognition and action (Audley, 1960; Laming, 1968; Link, 1975; Meyer, Abrams, et al., 1988; Pike, 1973; Ratcliff, 1978; Schmidt et al., 1979; Stone, 1960).

[19] The present concern is not disarmed by reports of frequent cases in which the peak latencies of certain ERP components (viz., P300) have exceeded mean reaction times (e.g., Coles et al., 1985; Duncan-Johnson & Donchin, 1982; Kutas et al., 1977; Ritter, Simson, & Vaughan, 1983). Such cases could arise even though the triggering event for a component takes place during the middle, rather than at the termination, of a processing stage that leads ultimately to overt responses. If the component's peak latency includes a residual delay after the triggering event, because some other ancillary process not required for responding intervenes between the triggering event and the onset of the component, then the peak latency might easily exceed the observed reaction time (fig. 12), but our above argument would still hold.

8.2. Methodological weaknesses

Accompanying the aforementioned theoretical weaknesses in cognitive psychophysiology, there are also significant methodological weaknesses. In particular, current techniques for analyzing psychophysiological data tend to be somewhat inadequate, given the substantial complexity and variability inherent in patterns of ERPs and reaction times. Applications of these techniques may run the risk of reaching erroneous conclusions, and sufficiently powerful alternatives to them do not yet exist. Further work will be required to cure the problem.

8.2.1. Measurement of component latencies

Let us consider, for example, the measurement of latencies associated with underlying components in the ERP waveform. If one's objective is to measure the peak latency for a component such as P300 on a trial-by-trial basis, then there are some promising methods (Coles et al., 1986). These methods entail preliminary enhancement of the ERP signal on each trial (e.g., via vector filtering; Gratton, Coles, & Donchin, 1983) followed by template matching (e.g., Kutas et al., 1977) or, under some circumstances, simple peak picking (e.g., Coles et al., 1985). At present, however, none of them have been perfected; they do not deal fully with difficulties caused by the stochastic fluctuation of an ERP component's morphology and temporal locus over trials, nor do they provide a complete account of contributions due to noise from background brain activity. Some investigators have therefore chosen instead to stick with the conventional technique of measuring latencies for ERP components by first averaging the records of ERPs across trials involving the same experimental condition (e.g., Ritter, Simson, & Vaughan, 1983). In this case, a component's onset latency, peak latency, and so forth may then be derived from the less noisy average ERPs. However, the averaging technique can introduce biases of its own that vitiate subsequent inferences about the dynamics of cognition and action.

To be specific, suppose that we conduct an experiment concerning the effects of two factors, F_1 and F_2, on two different ERP components, C_1 and C_2, one occurring later than the other (e.g., N200 and P300). The experiment might address whether these components respectively manifest two separate stages of processing, which can be tested via the inference rules outlined earlier (viz., Rules 4 and 5). Also, suppose that in testing the stage hypothesis, we average the ERPs within each condition defined by a particular combination of factor levels, and we measure the peak latencies of the components from these averages. Then mistaken conclusions could easily result.

One such mistake might involve rejecting the stage hypothesis inappropriately. Under this hypothesis, the two factors should affect a late component's peak latency additively on each trial, assuming that these latencies correspond to stage-termination times. So if the hypothesis were valid, the means of the peak latencies from individual trials should exhibit additive factor effects as well. Be that as it may, an interaction could still appear in the peak latencies of a relatively late component like C_2 when they are derived from the average ERPs, because the peak latency of a component in an average ERP does not necessarily equal the mean of that component's peak latencies derived trial-by-trial (Callaway, Halliday, Naylor, & Thouvenin, 1984). The two measures may differ substantially even when the peak latencies on individual trials are obtained with perfect accuracy. Given this incipient inequality, a component's peak latency in an average ERP constitutes a potentially biased estimate of the mean termination time for any associated processing stage. The magnitude of the bias may vary in subtle ways, depending on the morphology of the ERP component in question. This could lead the stage hypothesis to be rejected via Rule 4, even though there are underlying successive stages at the level of individual trials.

A similar error could also result through the application of Rule 5 to peak latencies of components derived from average ERPs. For the stage hypothesis to hold under Rule 5, the peak latency of a relatively late component must exhibit exactly the same factor effect as the peak latency of an earlier component does, if the factor effect is localized in a stage manifested by the earlier component. This requirement might hold with peak latencies measured on individual trials but not otherwise. Depending on the component's morphology, averaging ERPs across trials before measuring peak latencies might increase (or decrease) the apparent effect of a factor on late versus early components. The implication is that, when applying Rules 1 through 6, investigators who want to assess stage durations and concomitant factor effects should first measure peak latencies and then average, not the reverse. [20]

Of course, this is not to say that averaging ERPs before measuring their components' peak latencies always produces misleading results. It seems less likely that the latter method would yield spurious additivity than that it would yield spurious interactions among factor effects. So if investigators measure peak latencies after averaging ERPs and find evidence supporting a discrete stage model, then perhaps one can still have some confidence in their conclusions. For example, confidence may still be warranted in some of the conclusions reached by Ford et al. (1979) and Ritter, Simson, and Vaughan (1983).

We suspect, nevertheless, that there are some cases in which averaging ERPs across trials before measuring component latencies may have yielded mistaken conclusions. Consider, in particular, other results from the study by

[20] For reasons analogous to those outlined above, additive-factor analyses of mean peak amplitudes should be done by first measuring a component's peak amplitude on each trial and then averaging, rather than averaging ERPs across trials and then measuring the peak amplitude of the component in the composite waveform.

Ritter, Simson, and Vaughan (1983). Although they reported equal effects of stimulus discriminability on peak latencies of the N_A and N2 components derived from ERP averages, certain other details of their data regarding N_A and N2 suggested that the processes manifested by these components are not entirely separate in time. The peak latency of the N_A component appeared somewhat larger than the onset latency of the N2 component (i.e., the time at which N2 started to rise above base level). Ritter, Simson, and Vaughan (1983) took the latter outcome as evidence that the stimulus-categorization process associated with N2 may start before the pattern-recognition process associated with N_A has finished, consistent with McClelland's (1979) cascade model and Miller's (1982) asynchronous discrete-coding model.

Yet this is not the only possible interpretation. On each trial of Ritter, Simson, and Vaughan's (1983) study, the onset latency of N2 may have exceeded the peak latency of N_A, as expected under a strict serial stage model. If the latencies of the components varied randomly across trials, however, then averaging the ERPs before measuring them could have made the N2 onset latency appear less than the N_A peak latency. The reason is that ERP averages yield components whose onset latencies tend to equal the minimum of the onset latencies from individual trials. So the onset latency of a component in the average ERP will usually be less than the average of the component's onset latencies measured on each trial separately. Also, the latter bias is likely to be greater than the corresponding one for peak latencies, which we discussed earlier. Averaging ERPs before measuring onset and peak latencies would, therefore, yield exactly the pattern of results that Ritter, Simson, and Vaughan (1983) reported, even if processing were strictly serial.

8.2.2. Other difficulties with averaging

When the latencies of ERP components are measured accurately on individual trials and only averaged thereafter, cognitive psychophysiology may still suffer from other difficulties associated with aggregating reaction-time and accuracy data. These difficulties are pervasive and difficult to escape, as mental chronometry has repeatedly discovered. Supplementing chronometric measures with additional results from ERP records will not always suffice to overcome the problem.

To illustrate the degree of difficulty here, let us review the analyses of speed-accuracy tradeoff curves done by Coles and Gratton (1986), whose work was summarized earlier (see *Illustrative Studies*). They sought a way of attenuating the contamination in tradeoff curves caused by averaging accuracy data across trials, which may produce smearing and obscure the presence of discrete (viz., all-or-none) processes. Their approach entailed measuring the peak latency of the P300 component, an indicator of stimulus-evaluation time, on each trial and dividing it into the corresponding overt reaction time for that trial, deriving a resultant RT/P300 ratio. Response accuracy was then averaged across trials as a function of the ratio. This averaging yielded mean accuracy scores at each of several RT/P300 values. The hope was that the obtained results would be controlled for stochastic fluctuations in the completion times of stimulus evaluation, thereby revealing a pure step-function in response accuracy over time, if the evaluation processes were really discrete (i.e., all-or-none). However, a smooth undulating tradeoff curve still emerged. This outcome, like the gradual monotonic increase of accuracy over time in standard speed-accuracy tradeoff curves, might tempt one to conclude that stimulus evaluation and response execution are continuous overlapping processes, as postulated by continuous-flow models (Eriksen & Schultz, 1979; McClelland, 1979).

The problem with such a conclusion is that the approach taken to reach it still does not deal fully with stochastic fluctuations of the times at which stimulus-evaluation and response-execution processes are completed. When the RT/P300 ratio from individual trials is used as a predictor variable in plotting response accuracy, there are at least two significant residual sources of variance that could smear the resultant tradeoff curve. First, the measured P300 component's peak latency may be a somewhat unreliable indicator of stimulus-evaluation time. If only a moderate correlation exists between the completion time of the evaluation process on a trial and the peak latency of P300, then averaging the accuracy of responses over trials that have the same RT/P300 ratio could obscure an underlying step-wise information-accumulation function with stochastic transition times (Meyer & Irwin, 1981; Meyer, Irwin, et al., 1988). Secondly, response preparation and execution may themselves entail an all-or-none process with stochastic transition times, as we observed previously (fig. 4; cf. Meyer et al., 1985). As a result, this would cause additional smearing in the average speed-accuracy tradeoff function, even when the mean response accuracy is plotted versus RT/P300. The RT/P300 ratio does not control for such smearing, because the RT measure contains a contribution from response processes on each trial, which are at least partially uncorrelated with P300 latencies.

A potential cure for the latter problem might be to use psychophysiological measures in combination with our speed-accuracy decomposition technique, analysis of reaction-time mixture distributions, and appropriate mathematical models of information-processing dynamics (Meyer, Irwin, et al., 1988; Meyer et al., 1984, 1985). By relating ERP data obtained on individual trials to the forms of underlying latency distributions, perhaps one can better identify the internal processing states that mediate a given response. In turn, this may help to purify derived speed-accuracy tradeoff curves more fully.

8.2.3. The subtraction problem

Finally, a comment is in order about the practice adopted by some investigators of subtracting data from different conditions to isolate ERP components and to estimate their latencies (e.g., Hansen & Hillyard, 1980; Näätänen & Michie, 1979; Ritter, Simson, & Vaughan, 1983). As mental chronometry's past has revealed, subtraction methodology sometimes leads to rather deceptive outcomes (e.g., Wundt, 1880). When one uses this methodology for analyzing data from conditions that involve different experimental tasks, it can yield embarrassing artifacts caused by uncontrolled variations of subjective strategies or violations in the assumptions of pure insertion and selective influence (Külpe, 1893/1909; cf. Sternberg, 1969). With ERPs especially, such artifacts might produce a spurious proliferation of illusory components, much like what happened when Wundt (1880) was seduced by Donders' (1868/1969) subtraction method and, as a result, obtained evidence for an implausibly large number of distinct processing stages. Careful attention must therefore be given to documenting the functional significance of ERP components that are derived through subtraction methodology before letting them become bed partners in the marriage with mental chronometry. Cognitive psychophysiologists should test the assumptions of pure insertion and selective influence rigorously, checking for ancillary context-dependent changes in subjective strategies, lest the psychophysiological approach enter another Dark Age of the sort that previously befell the chronometric paradigm (table 1).

8.3. Fundamental lessons from mental chronometry

For the marriage between mental chronometry and cognitive psychophysiology to flourish henceforth, we urge in conclusion that psychophysiologists remember several fundamental lessons learned through the trials and tribulations of the chronometric paradigm. These lessons include the following:

There's no free lunch. A strict empiricist orientation will not suffice in cognitive psychophysiology. The field needs precise quantitative models for evaluating chronometric and psychophysiological measures. One cannot interpret these measures fruitfully without making some specific theoretical assumptions about the temporal properties and products of mental processes. Such assumptions also have a cost associated with them; they must be tested as best possible along the way.

Variability is a way of life. Models of human information processing and its psychophysiological substrates must accomodate the fact that reaction-time and ERP data are inherently variable. This variability results, at least in part, from stochastic processing components that play an integral role in task performance. These components have to be characterized thoroughly in any successful modeling effort.

Average at your own risk. Because chronometric and psychophysiological measures may contain both systematic variability and noise, one cannot necessarily handle them simply by averaging data across trials. Such averaging is a two-edged sword. In some cases, it will attenuate the contributions of the noise, but in others, it will also introduce undesirable biases as well. A partial solution for the problem is to examine the detailed forms of distributions associated with ERP component latencies and overt reaction times.

Subtraction methodology opens Pandora's Box. Given that human performance is highly context dependent, ancillary mental processes may change significantly across task conditions, making it impossible to obtain exact estimates of individual stage durations by comparing results from one condition versus another. Failure to heed this warning will breed a host of bedeviling results.

No pain, no gain. Even with thoughtful experimentation and incisive theoretical analyses, some frustrations and setbacks are inevitable. Studying the dynamics of cognition and action is not "a piece of cake," as mental chronometry has already proven. Yet out of the struggle, meaningful progress can emerge through persistent research efforts. So, like mental chronometry and Timex watches, cognitive psychophysiology should gird itself to take a licking and keep on ticking.

References

Ach, N. (1905). *Über die Willenstätigkeit und das Denken.* Göttingen: Vardenhoeck.

Allison, T., Wood, C.C., & McCarthy, G. (1986). The central nervous system. In M.G.H. Coles, E. Donchin, & S.W. Porges (Eds.), *Psychophysiology: Systems, processes, and applications* (pp. 5–21). New York: Guilford Press.

Audley, R.J. (1960). A stochastic model for individual choice behavior. *Psychological Review, 67,* 1–15.

Bessel, F.W. (1823). *Astronomische Beobachtungen in Königsberg, 8,* iii–viii.

Boring, E.G. (1950). *A history of experimental psychology.* New York: Appleton-Century-Crofts.

Callaway, E., Halliday, R., Naylor, H., & Thouvenin, D. (1984). The latency of the average is not the average of the latencies. *Psychophysiology, 21,* 571.

Cattell, J.McK. (1886). The time taken by cerebral operations. *Mind, 11,* 20–242.

Chase, W.G., McCarthy, G., Squires, K.C., & Schvaneveldt, R.W. (1984). Report of Panel IV: Mental chronometry. In E. Donchin (Ed.), *Cognitive psychophysiology: Event-related potentials and the study of cognition. The Carmel Conferences* (Vol. 1, pp. 249–301). Hillsdale, NJ: Lawrence Erlbaum.

Coles, M.G.H., Donchin, E., & Porges, S.W. (1986). *Psychophysiology: Systems, processes, and applications.* New York: Guilford Press.

Coles, M.G.H., & Gratton, G. (1986). Cognitive psychophysiology and the study of states and processes. In G.R.J. Hockey, A.W.K. Gaillard, & M.G.H. Coles (Eds.), *Energetics and human information processing* (pp. 409–424). Dordrecht, The Netherlands: Nijhoff.

Coles, M.G.H., Gratton, G., Bashore, T.R., Eriksen, C.W., & Donchin, E. (1985) A psychophysiological investigation of the continuous flow model of information processing. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 529–553.

Coles, M.G.H., Gratton, G., & Donchin, E. (1988). Detecting early communication: Using measures of movement-related potentials to illuminate human information processing. *Biological Psychology, 26,* 69–89, this issue.

Coles, M.G.H., Gratton, G., Kramer, A.F., & Miller, G.A. (1986). Principles of signal acquisition and analysis. In M.G.H. Coles, E. Donchin, & S.W. Porges (Eds.), *Psychophysiology: Systems, processes, and applications* (pp. 183–221). New York: Guilford Press.

Corbett, A.T., & Wickelgren, W.A. (1978). Semantic memory retrieval: Analysis by speed accuracy tradeoff functions. *Quarterly Journal of Experimental Psychology, 30,* 1–15.

Donchin, E. (1969). Discriminant analysis in average evoked response studies: The study of single trial data. *Electroencephalography and Clinical Neurophysiology, 27,* 311–314.

Donchin, E. (1979). Event-related brain potentials: A tool in the study of human information processing. In H. Begleiter (Ed.), *Evoked potentials and behaviour* (pp. 13–75). New York: Plenum Publishing Corp.

Donchin, E. (1981). Presidential address, 1980. Surprise!...Surprise? *Psychophysiology, 18,* 493–513.

Donchin, E. (Ed.) (1984). *Cognitive psychophysiology: Event related potentials and the study of cognition.* Hillsdale, NJ: Lawrence Erlbaum.

Donchin, E., McCarthy, G., Kutas, M., & Ritter, W. (1983). Event-related brain potentials in the study of consciousness. In R.J. Davidson, G.E., Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (Vol. 3, pp. 81–121). New York: Plenum Publishing Corp.

Donders, F.C. (1868/1969). On the speed of mental processes. Translation by W.G. Koster in W.G. Koster (Ed.), *Attention and performance II* (pp. 412–431). Amsterdam: North-Holland Publishing Co. Original work published in *Onderzoekingen gedaan in het Physiologisch Laboratorium der Utrechtsche Hoogeschool,* Tweede reeks, 1868–1869, *II*, 92–120.

Duncan-Johnson, C.C., & Donchin, E. (1982). The P300 component of event-related brain potential as an index of information processing. *Biological Psychology, 14,* 1–52.

Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology, 2,* 312–329.

Eriksen, C.W., & Schultz, D. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics, 25,* 249–263.

Exner, S. (1873). Experimentelle Untersuchung der einfachsten psychischen Processe. *Pflügers Archiv für gesamte Physiologie, 7,* 601–660.

Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47,* 381–391.

Fitts, P.M. (1966). Cognitive aspects of information processing: III. Set for speed versus accuracy. *Journal of Experimental Psychology, 71,* 849–857.

Ford, J.M., Roth, W.T., Mohs, R.C., Hopkins, W.F., & Kopell, B.S. (1979). Event-related potentials recorded from young and old adults during a memory retrieval task. *Electroencephalography and Clinical Neurophysiology, 45,* 319–330.

Gaillard, A.W.K., & Ritter, W. (Eds.) (1983). *Tutorials in ERP research: Endogenous components.* Amsterdam: North-Holland.

Gaillard, A.W.K., & Verduin, C.J. (1985). Comparisons across paradigms: An ERP study. In M.I. Posner & O.S.M. Marin (Eds.), *Attention and performance XI* (pp. 375–388). Hillsdale, NJ: Lawrence Erlbaum.

Gibson, J.J. (1941). A critical review of the concept of set in contemporary experimental psychology. *Psychological Bulletin, 38,* 781–817.

Gratton, G., Coles, M.G.H., & Donchin, E. (1983). Filtering for spatial distribution: A new approach (Vector Filter). *Psychophysiology, 20,* 443–444 (Abstract).

Hansen, J.C., & Hillyard, S.A. (1980). Endogenous brain potentials associated with selective auditory attention. *Electroencephalography and Clinical Neurophysiology, 49,* 277–290.

Helmholtz, H.L.F. von (1847). Über die Erhaltung der Kraft. *Wissenschaftliche Abhandlungen, I,* 12–75.

Helmholtz, H.L.F. von (1850). Über die Methoden, kleinste Zeittheile zu messen, und ihre Anwendung für physiologische Zwecke. Original work translated in *Philosophical Magazine,* 1853, *6* (Section 4), 313–325.

Hick, W.E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology, 4,* 11–26.

Hillyard, S.A., & Kutas, M. (1983). Electrophysiology of cognitive processing. *Annual Review of Psychology, 34,* 33–61.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology, 45,* 188–196.

Jastrow, J. (1890). *The time relations of mental phenomena.* New York: Hodges.

Kornblum, S. (1973). Simple reaction time as a race between signal detection and time estimation: A paradigm and a model. *Perception & Psychophysics, 13,* 108–112.

Kounios, J., Osman, A.M., & Meyer, D.E. (1987). Structure and process in semantic memory: New evidence based on speed-accuracy decomposition. *Journal of Experimental Psychology: General, 116,* 3–25.

Külpe, O. (1909). *Outlines of psychology: Based upon the results of experimental investigation* (Third Edition). New York: Macmillan. Translation of original work published in 1893.

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

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

LaBerge, D. (1962). A recruitment theory of simple behavior. *Pschometrika, 27,* 375–396.

Lachman, R., Lachman, J.L., & Butterfield, E.C. (1979). *Cognitive psychology and information processing: An introduction.* Hillsdale, NJ: Lawrence Erlbaum.

Laming, D. (1968). *Information theory of choice reaction time.* New York: Wiley.

Lange, L. (1888). Neue Experimente über den Vorgang der einfachen Reaction auf Sinneseindrücke. *Philosophische Studien, 4,* 479–510.

Leonard, J.A. (1957). *A reconsideration of the early history of reaction time studies.* Unpublished manuscript, Applied Psychology Research Unit, Cambridge, England. (Available upon request from the present authors.)

Link, S.W. (1975). The relative judgment theory of two choice reaction time. *Journal of Mathematical Psychology, 12,* 114–136.

Logan, G.D., & Cowan, W.B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91,* 295–327.

Luce, R.D. (1986). *Response times: Their role in inferring elementary mental organization.* New York: Oxford University Press.

Marsh, G.R. (1975). Age differences in evoked potential correlates of a memory scanning process. *Experimental Aging Research, 1,* 3–16.

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

McCarthy, G., & Donchin, E. (1983). Chronometric analysis of human information processing. In A.W.K. Gaillard & W. Ritter (Eds.), *Tutorials in ERP research – Endogenous components* (pp. 251–268). Amsterdam: North Holland.

McClelland, J.L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review, 86,* 287–330.

McClelland, J.L., & Rumelhart, D.E. (1981). An interactive-activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review, 88,* 375–407.

McGill, W.J. (1963). Stochastic latency mechanisms. In R.D. Luce, R.R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. I, pp. 309–360). New York: Wiley.

Merkle, J. (1885). Die zeitlichen Verhältnisse der Willenstätigkeit. *Philosophische Studien, 2,* 73–127.

Meyer, D.E., Abrams, R.A., Kornblum, S., Wright, C.E., & Smith, J.E.K. (1988). Optimality in human motor performance: Ideal control of rapid aimed movements. *Psychological Review, 95.* 340–370.

Meyer, D.E., & Irwin, D.E. (1981, November). *On the time-course of rapid information processing.* Paper presented at the meeting of the Psychonomic Society, Philadelphia. Available as Technical Report No. 43, Michigan Cognitive Science Program, June, 1982.

Meyer, D.E., Irwin, D.E., Osman, A.M., & Kounios, J. (1988). The dynamics of cognition and action: Mental processes inferred from speed-accuracy decomposition. *Psychological Review, 95,* 183–237.

Meyer, D.E., & Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227–234.

Meyer, D.E., & Schvaneveldt, R.W. (1976) Meaning, memory structure, and mental processes. *Science, 192,* 27–33.

Meyer, D.E., Schvaneveldt, R.W., & Ruddy, M.G. (1975). Loci of contextual effects on visual word recognition. In P.M.A. Rabbit & S. Dornic (Eds.), *Attention and performance V* (pp. 98–118). London: Academic Press.

Meyer, D.E., Smith, J.E.K., & Wright, C.E. (1982). Models for the speed and accuracy of aimed movements. *Psychological Review, 89,* 449–482.

Meyer, D.E., Yantis, S., Osman, A.M., & Smith, J.E.K. (1984). Discrete versus continuous models of response preparation: A reaction-time analysis. In S. Kornblum & J. Requin (Eds.), *Preparatory states and processes* (pp. 69–94). Hillsdale, NJ: Lawrence Erlbaum.

Meyer, D.E., Yantis, S., Osman, A.M., & Smith, J.E.K. (1985). Temporal properties of human information processing: Tests of discrete versus continuous models. *Cognitive Psychology, 17,* 445–518.

Miller, J. (1982). Discrete versus continuous stage models of human information processing: In search of partial output. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 273–296.

Miller, J. (1983). Can response preparation begin before stimulus recognition finishes? *Journal of Experimental Psychology: Human Perception and Performance, 9,* 161–182.

Miller, J. (1988). Discrete and continuous models of human information processing: Theoretical distinctions and empirical results. *Acta Psychologica,* in press.

Mowrer, O.H. (1940). Preparatory set (expectancy) – some methods of measurement. *Psychological Review Monograph, 52* (Whole No. 233).

Mulder, G., Loerich, A.B.M., Brookhuis, K.A., Van Dellen, H.J., & Mulder, L.J.M. (1984). Stage analysis of the reaction process using brain-evoked potentials and reaction time. *Psychological Research, 46,* 15–32.

Müller, J. (1838). *Handbuch der Physiologie des Menschen,* Vol. I.

Näätänen, R., & Michie, P.T. (1979). Early selective attention effects on the evoked potential. A critical review and reinterpretation. *Biological Psychology, 8,* 81–136.

Neisser, U. (1963). Decision-time without reaction-time: Experiments in visual scanning. *American Journal of Psychology, 76,* 376–385.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Neisser, U., & Beller, H.K. (1965). Searching through word lists. *British Journal of Psychology, 56,* 349–358.

Neisser, U., Novick, R., & Lazar, R. (1963). Searching for ten targets simultaneously. *Perceptual and Motor Skills, 17,* 955–961.

Ollman, R. (1966). Fast guesses in choice reaction time. *Psychonomic Science, 6,* 155–156.

Ollman, R. (1977). Choice reaction time and the problem of distinguishing task effects from strategy effects. In S. Dornic (Ed.), *Attention and performance VI* (pp. 99–113). Hillsdale, NJ: Lawrence Erlbaum.

Ollman, R.T., & Billington, M.J. (1972). The deadline model for simple reaction times. *Cognitive Psychology, 3,* 311–336.

Osgood, C.E. (1953). *Method and theory in experimental psychology.* New York: Oxford University Press.

Osman, A.M., Kornblum, S., & Meyer, D.E. (1986). The point-of-no-return in choice reaction time: Controlled and ballistic stages of response preparation. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 243–258.

Pachella, R.G. (1974). The interpretation of reaction time in information processing research. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41–82). New York: Halstead Press.

Pachella, R.G., & Pew, R.W. (1968). Speed-accuracy tradeoff in reaction time: Effect of discrete criterion times. *Journal of Experimental Psychology, 76,* 19–24.

Pike, R. (1973). Response latency models for signal detection. *Psychological Review, 80,* 53–68.

Posner, M.I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Lawrence Erlbaum.

Posner, M.I., & McCleod, P. (1982). Information processing models – in search of elementary operations. *Annual Review of Psychology, 33,* 477–514.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Ratcliff, R. (1988). Continuous versus discrete information processing: Modeling accumulation of partial information. *Psychological Review, 95,* 238–255.

Reed, A.V. (1976). List length and the time course of recognition in immediate memory. *Memory & Cognition, 4,* 16–30.

Renault, B. (1983). The visual emitted potentials: Clues for information processing. In A.W.K. Gaillard & W. Ritter (Eds.), *Tutorials in ERP research: Endogenous components* (pp. 159–175). Amsterdam: North-Holland.

Renault, B., Ragot, R., Lesevre, N., & Remond, A. (1982). Onset and offset of brain events as indices of mental chronometry. *Science, 215,* 1413–1415.

Ritter, W., Simson, R., & Vaughan, H.G., Jr. (1983). Event-related potential correlates of two stages of information processing in physical and semantic discrimination tasks. *Psychophysiology, 20,* 168–179.

Ritter, W., Simson, R., Vaughan, H.G., Jr., & Macht, M. (1982). Manipulation of event-related potential manifestations of information processing stages. *Science, 218,* 909–911.

Ritter, W., Vaughan, H.G., Jr., & Simson, R. (1983). On relating event-related potential components to stages of information processing. In A.W.K. Gaillard & W. Ritter (Eds.), *Tutorials in event-related potential research: Endogenous components* (pp. 143–158). Amsterdam: North-Holland.

Rosenbaum, D.A. (1980). Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General, 109,* 444–474.

Schmidt, R.A., Zelaznik, H., Hawkins, B., Frank, J.S., & Quinn, J.T., Jr. (1979). Motor-output variability: A theory for the accuracy of rapid motor acts. *Psychological Review, 86,* 415–451.

Schmitt, J.C., & Scheirer, C.J. (1977). Empirical approaches to information-processing: Speed-accuracy tradeoff functions or reaction time. *Acta Psychologica, 41,* 321–325.

Schouten, J.F., & Bekker, J.A.M. (1967). Reaction time and accuracy. *Acta Psychologica, 27,* 143–153.

Schweikert, R. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology, 18,* 105–139.

Schweikert, R. (1983). Latent network theory: Scheduling of processes in sentence verification and the Stroop effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 353–383.

Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27,* 379–423, 623–656.

Smith, E.E. (1968). Choice reaction time: An analysis of the major theoretical positions. *Psychological Bulletin, 69,* 77–110.

Smith, J.E.K., Meyer, D.E., Yantis, S., & Osman, A.M. (1982, August). *Finite-state models of*

reaction time: Estimation of latency distributions. Paper presented at the meeting of the Society for Mathematical Psychology, Princeton, NJ.
Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153,* 652–654.
Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W.G. Koster (Ed.), *Attention and performance II* (pp. 276–315). Amsterdam: North-Holland.
Stevens, S.S., (Ed.) (1951). *Handbook of experimental psychology.* New York: Wiley.
Stone, M. (1960). Models for choice-reaction time. *Psychometrika, 25,* 251–260.
Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.
Taylor, D.A. (1976). Stage analysis of reaction time. *Psychological Bulletin, 83,* 161–191.
Telford, C.W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology, 14,* 1–36.
Thomas, E.A.C., & Ross, B.H. (1980). On appropriate procedures for combining probability distributions within the same family. *Journal of Mathematical Psychology, 21,* 136–152.
Townsend, J.T. (1974). Issues and models concerning the processing of a finite number of inputs. In B.H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 133–185). Hillsdale, NJ: Lawrence Erlbaum.
Townsend, J.T., & Ashby, F.G. (1983). *Stochastic modeling of elementary psychological processes.* Cambridge, England: Cambridge University Press.
Turing, A.M. (1950). Computing machinery and intelligence. *Mind, 59,* 433–450.
Turvey, M.T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review, 80,* 1–52.
Vaughan, H.G., Jr., & Ritter, W. (1973). Physiologic approaches to the analysis of attention and performance: Tutorial review. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 129–154). New York: Academic Press.
Watt, H.J. (1905). Experimentelle Beiträge zur einer Theorie des Denkens. *Archiv für Gesamte Psychologie, 4,* 289–436.
Wickelgren, W.A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41,* 67–85.
Wilkinson, R.T. (1967). Evoked response and reaction time. In A.F. Sanders (Ed.), *Attention and performance* (pp. 235–245). Amsterdam: North-Holland Publishing Co.
Wood, C.C., & Allison, T. (1981). Interpretation of evoked potentials: A neurophysiological perspective. *Canadian Journal of Psychology, 35,* 113–135.
Wood, C.C., McCarthy, G., Squires, N.K., Vaughan, H.G., Jr., Woods, D.L., & McCallum, W.C. (1984). Anatomical and physiological substrates of event-related potentials. In R. Karrer, J. Cohen, & P. Tueting (Eds.), *Brain and information: Event-related potentials* (pp. 681–721). New York: The New York Academy of Sciences.
Woodrow, H. (1914). The measurement of attention. *Psychological Monographs, 17* (Whole No. 76).
Woodworth R.S. (1899). The accuracy of voluntary movement. *Psychological Review, 3* (2, Whole No. 13).
Woodworth, R.S. (1938). *Experimental psychology.* New York: Holt.
Woodworth, R.S., & Schlosberg, H. (1954). *Experimental psychology.* New York: Holt, Rinehart and Winston.
Wundt, W. (1880). *Grundzüge der physiologischen Psychologie* (2nd Ed.). Leipzig: W. Engelmann.
Yantis, S., & Meyer, D.E. (1985). *Dynamics of spreading activation in information processing.* Paper presented at the meeting of the Psychonomic Society, Boston, MA, November.
Yantis, S., & Meyer, D.E. (1988). Dynamics of activation in semantic and episodic memory. *Journal of Experimental Psychology: General, 117,* 130–147.
Yellott, J.I. (1967). Correction for guessing in choice reaction time. *Psychonomic Science, 8,* 321–322.
Yellott, J.I. (1971). Correction for fast guessing and the speed-accuracy tradeoff in choice reaction time. *Journal of Mathematical Psychology, 8,* 159–199.