

2

Précis to a Practical Unified Theory of Cognition and Action: Some Lessons from EPIC Computational Models of Human Multiple-Task Performance

David E. Meyer and David E. Kieras

ABSTRACT Experimental psychology, cognitive science, and human-factors engineering have progressed sufficiently far that a practical unified theory of cognition and action is now foreseeable. Such a theory soon may yield useful quantitative predictions about rapid human multiple-task performance in applied settings. Toward this end, an executive-process interactive control (EPIC) architecture has been formulated with components whose assumed properties emulate fundamental perceptual, cognitive, and motor processes. On the basis of EPIC, a theorist may construct detailed computational models that characterize multiple-task performance under both laboratory and real-world conditions. For example, EPIC computational models provide good accounts of response latencies and accuracy from the psychological refractory-period procedure, aircraft cockpit operation, and human-computer interaction. As a result, major commonalities in performance across various task domains have been discovered, and useful heuristic principles for designing person-machine interfaces have been identified. The substantive and methodological lessons learned from these advances constitute an instructive précis to further utilitarian theoretical unification.

Like the quest of Indiana Jones, the adventurous anthropologist in *Raiders of the Lost Ark* (Kasdan, Lucas, and Kaufman 1981), our journey to the Holy Land for Attention and Performance XVII has brought us in search of an alluring mystical treasure. The treasure we seek is a unified theory of cognition and action through which human performance can be understood and predicted in a variety of contexts, spanning elementary laboratory paradigms and complex real-world situations. Although not so sacred as the Lost Ark of the Covenant, such a theory would have great value for both applied psychological science and the present volume. By design, this volume concerns the cognitive regulation of human performance, with special emphasis on interactions between theory and practical applications. Ultimately, these interactions and future progress from them will grow best if researchers succeed at constructing a coherent conceptual framework in which scientific knowledge is synthesized about several complementary topics, including goal-directed behavior, top-down supervisory executive processes, performance strategies, attentional control mechanisms, and conscious appraisal of the world. As one kindred seeker enthusiastically proclaimed: "There is nothing more practical than a good theory" (Gopher and Koriat, chap. 1, this volume).¹

Moving onward with such enthusiasm, we start by reviewing prior developments that have paved the way toward our formulating a unified theory of cognition and action. On the basis of these developments, we introduce a functional architecture to emulate fundamental characteristics of the human information-processing system. Next, using this architecture, we construct illustrative computational models to account for observed aspects of multiple-task performance in typical laboratory paradigms and real-world situations. From these accounts, we derive some instructive lessons with respect to how further attempts at theoretical unification should proceed. Although the chapter does not end with a complete, veridical unified theory, it does provide some guidance toward this objective (see also, Kieras and Meyer 1995, forthcoming; Kieras, Wood, and Meyer 1997; Meyer and Kieras 1994, 1997a, 1997b, Meyer et al. 1995).

2.1 HISTORICAL BACKGROUND

Our efforts toward formulating a unified theory of cognition and action have been inspired by several prophets of experimental psychology, cognitive science, and human-factors engineering. Among them, an especially noteworthy sage was the late Allen Newell, who, nearly a quarter century before *Attention and Performance XVII*, published a provocative chapter entitled "You Can't Play 20 Questions with Nature and Win" (Newell 1973a). Newell's thesis was that experimental psychology has neglected to take sufficient stock of the big theoretical picture, thereby hindering cumulative scientific progress in the study of mind and behavior. As Newell saw it, too much effort had been expended on conducting narrow empirical studies to test seductively simple binary hypotheses (e.g., early versus late attentional selection, serial versus parallel memory search, and imaginal versus propositional knowledge). Indeed, such studies have accumulated an impressively large collection of basic facts, on the order of 3,000 "good quantitative regularities" (e.g., see Atkinson et al. 1988; Boff, Kaufman, and Thomas 1986; Meyer and Kornblum 1993), but how they all fit together theoretically remains a "great psychology data puzzle" (Newell 1992).² There are still no general theories that have adequate practical utility across many domains of application. Instead, what psychological theorizing has produced thus far is a set of unrelated micromodels that are relevant only to separate small families of empirical phenomena in limited artificial contexts.

Characteristics of Unified Theories

To go beyond these confines, Newell (1990, 1992) and his colleagues (e.g., Card, Moran, and Newell 1983; Laird, Newell, and Rosenbloom 1987) have advocated the development of unified theories of cognition (UTCs). An ideal UTC would postulate "a single system of mechanisms that operate together to produce the full range of human cognition" (Newell 1990, 1). The moti-

vation for this approach was summarized in Newell's original chapter on the futility of simply playing twenty questions with nature: "Our task in psychology is first to discover the invariant structure of processing mechanisms.... Without such a framework in which to work, the generation ... of new explanations for old phenomena will go on *ad nauseum*" (Newell 1973a, 293, 296).

As envisioned by Newell, the systems on which UTCs are conceptually based should have detailed information-processing architectures with set interconnected components that implement elementary symbolic computational processes for perception, cognition, and action. A complete UTC's architecture ought to be sufficiently powerful that programs executed by it can accurately simulate covert mental processes and overt behavior associated with learning, memory, perceptual-motor skills, language comprehension, decision making, problem solving, and other complex functions. In Newell's opinion (1990, 1992), more than enough empirical data are available now for starting to formulate UTCs—experimental psychology and cognitive science must make theoretical unification be an immediate, principal goal. Some prominent contributors to past Attention and Performance symposia have also called for such unification:

What is urgently needed is ... a computational theory, in the sense outlined by Marr (1982), of the many different functions of attentional selectivity and control ... taking seriously the idea that attentional functions are of many different kinds, serving a great range of different computational purposes (Allport 1993, 205–206).

We need computational theories of interaction between stages. As the number of theoretical entities increases in each area, it becomes increasingly hard to see the implications of combining them. Only computational systems can do this, and they will have the merit of stopping the laxness of definition noted by Allport (Broadbent 1993, 876).

Indeed, UTCs may provide numerous complementary benefits, which include amortization of theoretical constructs, integration of multiple empirical constraints, maximization of process identifiability, solution of irrelevant-specification problems, absolution from Popperian damnation, amplification of scientific progress, and promotion of practical applications (Newell 1990, 18). By combining extant theoretical constructs from diverse sources into one integrated framework, a UTC repays debts owed to past sponsors of experimental psychology. The prospective payoff is enhanced because UTCs account for a wide variety of data whose overall pattern imposes multiple empirical constraints on functional properties that a theory's mechanisms must have. In essence, this maximizes process identifiability and helps clarify what canonical assumptions are most appropriate. Given maximum process identifiability, the theorist can forgo appending "Rube Goldberg kludges" as part of the system, thereby solving irrelevant-specification problems (i.e., haphazard postulation of arbitrary components that contribute to explanations in an unprincipled fashion).

As a result, there may come absolution from Popperian damnation, that is, forgiveness for having proposed simple binary hypotheses with which aberrant bits of data are inconsistent (cf. Popper 1959). Such absolution can facilitate scientific progress; no longer will it be necessary to formulate, test, reject, and reformulate simple theoretical alternatives repeatedly in an endless karmic birth-death cycle of the same old hypotheses.³ Through the progress that UTCs enable, they can be taken into the field for useful practical applications. "A unified theory of cognition is the key to successful applied cognitive science" (Newell 1990, 498). This follows because important real-world tasks engage many aspects of cognition, and effective behavior under such circumstances depends on interactions among many information-processing components; thus, successful applications must treat all of these components in an integrated fashion.⁴

Furthermore, the connection between application and theoretical unification should be taken as a two-way street. While UTCs lead to useful applications, serious concern about practical applications can foster substantial theoretical development and unification. "Applications," Newell (1990, 501–502) tells us, "are critical for the internal conduct of the science. They establish what is worth predicting. They establish what accuracy is sufficient. They establish when a regularity is worth remembering. They establish when a theory should not be discarded.... Applications have a wisdom that the current fashions of theory do not." Mindful of these words and also those spoken by John F. Kennedy, we therefore should ask not just what unified theories of cognition can do for applications, but also what applications can do for cognitive theories (Newell 1990, 500).⁵

Harbingers of Unified Theories

Since Newell's original manifesto (1973a), no complete veridical UTC has been developed. Nevertheless, some promising harbingers of theoretical unification have appeared on the scene (for a comparative review and evaluation, see Newell 1990, 23–36). These include the model human processor of Card, Moran, and Newell (1983), the ACT* system of J. R. Anderson (1983), and the SOAR system of Newell (1990, 1992) and his colleagues (Laird, Newell, and Rosenbloom 1987).⁶

Model human processor The model human processor (MHP) was developed specifically for applications to human-computer interaction (HCI). To predict the speed and accuracy of people's performance during HCI tasks such as text editing, Card, Moran, and Newell (1983) endowed the MHP with a combination of general-purpose memory stores and processing units whose functional characteristics approximated those of the human information-processing system. Among the MHP's memory stores are a short-term working memory and a long-term declarative memory. They involve putative information codes, storage capacities, and durations consistent with

generic data available at the time of the MHP's formulation. Complementing these memory stores, among the MHP's processing units are perceptual, cognitive, and motor processors. The assumed durations of their operations have magnitudes consistent with previously estimated times taken by stages of processing such as stimulus encoding, response selection, and movement production (cf. Sanders 1980; Sternberg 1969). Through integrating the processing units and memory stores of the MHP in a "boxes and arrows" flowchart, acceptably accurate "engineering approximations" of response speed and accuracy across a variety of HCI tasks were obtained by Card, Moran, and Newell (1983).⁷ At the time, their success demonstrated that substantial data from numerous areas of experimental psychology are indeed available for taking significant strides toward utilitarian theoretical unification.

Nevertheless, the scope of the MHP is seriously limited. For example, its perceptual and motor processors do not embody all important characteristics of information processing in the visual, auditory, tactile, ocular, articulatory, and manual modalities. Also, the cognitive processor and working memory of the MHP do not enable computer simulations of human performance with observable outputs under diverse task conditions; that is, Card, Moran, and Newell (1983) did not go much beyond the boxes-and-arrows phase of theory development. As a result, the MHP lacks crucial features that a complete veridical UTC should have.

ACT* In some respects, the ACT* system (Anderson 1983), which evolved from Anderson and Bower's human associative memory (HAM; 1973) model and Anderson's ACT system (1976), is more complete and precise than Card, Moran, and Newell's MHP (1983). For ACT*, Anderson (1983) distinguished explicitly between procedural and declarative knowledge. Pursuing this distinction, he embodied procedural knowledge in ACT* as a formal production system under which various tasks could be performed through sets of production (if condition, then action) rules. The development of ACT* honored Newell's prior suggestion (1973b) that production systems would help construct more complete models of human information processing. ACT*'s production system has a rule interpreter with well-defined properties (e.g., conflict-resolution criteria). Through this interpreter, the conditions of specified production rules are compared to the current contents of working memory, and the rules' actions are executed contingent on the outcomes (matches or mismatches) of these comparisons.

Assumed details of the ACT* rule interpreter, working memory, and other ancillary components enabled Anderson (1983) to account for reaction time (RT) and accuracy data from comprehension and reasoning tasks. Furthermore, with algorithms for compiling and tuning procedural knowledge, phenomena of cognitive-skill acquisition (e.g., the power law of practice; Fitts 1964) are explained by ACT*.⁸ Thus ACT* has more inherent potential than does the MHP to enable computer simulations and to become a bona fide UTC.

Yet ACT* also has significant limitations. No serious treatment of ocular, manual, and articulatory motor control is included thus far as part of the ACT* architecture. Nor have initial applications of ACT* dealt extensively with complex problem solving of the sort addressed previously by Newell and Simon's general problem solver (GPS; 1972). Instead, Anderson (1983) focused mainly on cognitive processes of learning, memory, language comprehension, and inference.

SOAR The gaps left by ACT* and the MHP lead us to the SOAR system of Newell (1990, 1992) and his colleagues (e.g., Laird, Newell, and Rosenbloom 1987). SOAR surpasses these prior harbingers of a UTC by incorporating more detailed assumptions about perceptual-motor and attentional processes in the context of a production-system architecture. Also, extending the approach taken before in GPS (Newell and Simon 1972), an explicit characterization of complex impasse-driven problem solving has been embodied in SOAR's repertoire of cognitive mechanisms. Through an opportunistic "chunking" algorithm, impasses that arise during novel task performance are resolved by SOAR via a heuristic search of problem spaces. The "chunking" algorithm yields new procedural operators that enable automatized perceptual-motor and cognitive skill (cf. Anderson 1983; Schneider and Detweiler 1988; Shiffrin and Schneider 1977). As a result, phenomena at various levels of complexity, including automatic active-memory search, the power law of practice, and means-ends problem solving are treated within SOAR's purview.

Nevertheless, there is still more to the story of human performance than SOAR accommodates. SOAR's components for implementing simulations of ocular, manual, and articulatory motor control remain less well developed than would be desirable. Given this lack of development, interactions between perception and motor control are not yet characterized sufficiently in SOAR. Nor does SOAR have much to say thus far about multiple-task performance. Many questions remain to be answered as part of formulating a unified theory of cognition and action. What executive processes allocate limited perceptual-motor and cognitive resources for scheduling two or more concurrent tasks while satisfying task priorities imposed by prevailing environmental constraints? How do executive processes and skilled performance based on them evolve through transformations of declarative to procedural knowledge? Why might individual performers differ systematically with respect to the types of executive process and degrees of task coordination that they achieve?

2.2 PRESENT OBJECTIVES

The objectives of the present chapter are to foster further theoretical unification in the scientific study of human performance and, concomitantly, to help

answer some of the questions mentioned previously. To do so, we subsequently introduce a functional architecture for emulating basic characteristics of human information processing. Our architecture is called "EPIC," which stands for "executive-process interactive control." Although EPIC does not yet constitute a complete veridical UTC, it supplements prior theories in some significant ways. EPIC enables not only procedural cognition but also motor control and perceptual-motor interactions to be treated explicitly and parsimoniously in conjunction with formal hypotheses about supervisory executive cognitive processes and task-scheduling strategies. Given such treatment, precise computational models can be constructed to explain and predict reaction times (RTs), response accuracy, and other measurable aspects of people's overt behavior across various domains where multiple tasks must be performed concurrently. Our EPIC computational models are applicable to both elementary laboratory paradigms and complex real-world situations. Such applications yield instructive lessons for future work toward theoretical unification in studies of human performance. The transfer of data and theory from laboratory to real world may proceed more quickly by taking these lessons seriously.

Relevance of Multiple-Task Performance

Multiple-task performance under speed stress, a traditional topic of experimental psychology (Meyer and Kornblum 1993), offers an especially relevant venue for developing UTCs. In accord with Newell's terminology (1990, 1992), this topic involves studying *immediate behavior*, that is, responses to stimuli during brief tasks whose performance yields reaction times on the order of $100 \text{ msec} < RT < 1,000 \text{ msec}$. Along the overall time scale of human action, the mental processes that mediate such responding fall in the lower part of what Newell has called the "cognitive band".⁹ Here fundamental symbolic computations are accomplished to access information in various memory stores, make elementary decisions, store intermediate products temporarily, and execute input-output transformations.¹⁰

Centrality of the cognitive band Newell (1990, 1992) has argued that the types of computation done in the cognitive band are fundamental to all intelligent information processing. From the perspective of this argument, RTs for immediate behavior manifest the durations of these computations directly, revealing the nature of the functional architecture that implements them. Any hypothesized architecture must be consistent with available RT data, which impose a strong *real-time constraint* on UTCs. By taking such data thoroughly into account, as focusing on rapid multiple-task performance leads us to do, we may arrive more quickly at detailed specifications for a complete veridical UTC. Thus, during the initial development of SOAR's architecture, Newell (1990, 1992) and his colleagues (e.g., John 1988; John and Newell 1987, 1989; John, Rosenbloom, and Newell 1985) made concerted

efforts to ensure that it was consistent with known facts about stimulus-response compatibility (cf. Duncan 1977; Fitts and Seeger 1953; Kornblum, Hasbroucq, and Osman 1990), active-memory scanning (cf. Sternberg 1969), and transcription typing (cf. Salthouse 1986), which typify systematic empirical phenomena related to immediate behavior. Similar attempts have been made by other progenitors of UTCs (e.g., Anderson, 1983, 1990, 1993), although they do not go as far as EPIC on this score.

Availability of reaction time data In light of these considerations, students of rapid human performance are well situated for contributing to the further development of UTCs. Through techniques of mental chronometry (Donders, 1868; Luce 1986; Meyer et al. 1988; Miller 1988; Roberts and Sternberg 1993; Sanders 1980; Sternberg 1969; Woodworth and Schlosberg 1954), large amounts of RT data that are relevant to immediate behavior have been collected. Much of the available data bear on selective attention (e.g., Eriksen and Yeh 1985; Jonides 1980; Jonides and Yantis 1986; Posner 1980; Treisman 1988; Yantis and Jonides 1986), motor control (e.g., Abrams and Jonides 1988; Fischer and Ramsberger 1984; Ghez, Hening, and Favilla 1990; Meyer and Gordon 1985; Reuter-Lorenz, Hughes, and Fendrich 1991; Rosenbaum 1980; Sternberg et al. 1978), perceptual-motor interaction (e.g., Rosenbaum 1991), and coordination of information-processing operations in multiple-task performance (e.g., Damos 1991; Gopher and Donchin 1986; Meyer and Kieras 1997a,b). This "grist for the mill" is exactly what we need to specify the details of prospective UTCs' functional architectures.

Importance of practical needs With respect to meeting important practical needs, the study of human performance could use more theoretical unification. Multiple tasks must be performed rapidly and accurately in many important real-world situations such as HCI, aircraft cockpit operation, air-traffic control, automobile cellular-telephone communication, power-plant supervision, and so forth. Advances in our understanding of performance under such circumstances are required so that the design of person-machine interfaces, selection of personnel, and training for successful equipment usage may be facilitated. This facilitation would enhance productivity and reduce the frequency of disasters such as those involving Three Mile Island and the naval cruiser *Vincennes*.

Inadequacy of current frameworks Unfortunately, the general-purpose theoretical frameworks being used currently in applications to real-world situations where multiple-task performance plays important roles are less than fully adequate. These frameworks include the SAINT (Chubb 1981) and HOS (Lane et al. 1981; Harris et al. 1987) modeling systems. Although constituting valuable assessment tools, they do not enable precise computer simulations of complex multiple-task performance nor do they have the flexibility and generativity for a wide variety of applications (cf. O'Donnell and Egge-

meier 1986; Sanders 1991; Vreuls and Obermayer 1985). Much room remains for significant contributions to be made by new UTCs in applied settings. Their ultimate benefits may grow dramatically as future technology gives human-factors engineers more options for designing efficient user-friendly person-machine interfaces.

Potential Pitfalls

Of course, one must be vigilant for some potential pitfalls along the way toward a complete veridical UTC; developing such a theory will not be easy.

Necessity of selection In particular, there is the necessity of selection (Newell 1990), which poses the theorist with difficult dilemmas. Among the many extant theoretical concepts about human performance, some must be included and others excluded from any particular UTC, even though most of them have persuasive advocates and potential merit. For example, should a UTC's perceptual processors be equipped with precategorical attentional filters (cf. Allport 1989; Broadbent 1958; Deutsch and Deutsch 1963; Moray 1959; Norman 1976; Treisman 1960, 1964)? Should its cognitive processor have structural decision and response-selection bottlenecks (cf. Allport 1987; Allport, Antonis, and Reynolds 1972; Kahneman 1973; Moray 1967; Navon and Gopher 1979; Neumann 1987; Pashler 1984, 1994a; Welford 1952, 1959, 1967; Wickens 1984)? Should its motor processors use discrete independent movement features (cf. Abrams and Jonides 1988; Ghez, Hening, and Favilla 1990; Goodman and Kelso 1980; Meyer and Gordon 1985; Rosenbaum 1980)? The answers to such questions will not always be self-evident. Depending on what particular assumptions are selected for implementation in various components of an overall architecture, different prospective UTCs each may explain many observable aspects of overt behavior quite well, making it difficult to determine which UTC is most correct.

Turing tar pit Escaping this difficulty is especially problematic because of the so-called Turing tar pit (Newell 1992). In computational modeling, the available programming languages for UTCs are all powerful enough to simulate the same general symbolic transformations and input-output functions (cf. Turing 1937). Thus, at an abstract level, each UTC may have functional capabilities similar to those of its competitors, which can make alternative theories indistinguishable in many respects, thereby trapping theorists in a conceptual morass. Experimentalists therefore should expect to encounter a partial *nonidentifiability problem* when devising empirical tests among competing alternatives.

Degrees-of-freedom problem Closely related to the nonidentifiability problem and Turing tar pit is the degrees-of-freedom problem (Newell 1990, 1992). By construction, a UTC necessarily has many parameters whose

values can change from one context to the next. Within and across particular contexts, the number of "free" parameters may exceed the total degrees of freedom in available data sets. If so, then the theory will be underdetermined with respect to the data; good quantitative accounts provided by the theory will not prove definitively that the theory is apt or informative. In formulating a UTC, care therefore must be taken to impose principled constraints on the theory's potentially free parameters.

Bridges over Troubled Waters

Fortunately, there are a number of supportive heuristic principles for bridging the troubled waters that must be crossed on the way to a complete veridical UTC. Some of these principles are especially helpful in overcoming the necessity-of-selection and degrees-of-freedom problems and the Turing tar pit in designing the theory's functional architecture. Other principles constrain specific models that may be formulated to perform particular tasks on the basis of the architecture.

Maintenance of architectural simplicity and stability To overcome the degrees-of-freedom problem, a UTC's functional architecture should be kept as simple and stable as possible. The theorist must refrain from embellishing the architecture with empirically unsubstantiated, computationally arbitrary, unnecessarily complex, or seductively vague mechanisms. For example, initially postulating an immutable structural response-selection bottleneck (Pashler 1994a; Welford 1959, 1967) or a reservoir of divisible limited capacity (Kahneman 1973) in the architecture's cognitive processor seems inadvisable (Allport 1987; Neumann 1987; Wickens 1991). As Allport (1980, 117–118, 121) forewarned:

Obviously there is a problem of how we know when we are dealing with competition for a single resource.... Once one accepts the idea of general-purpose processing capacity [or central bottlenecks] as a working hypothesis, it becomes temptingly easy to assume, without further ado, that almost any instance of dual-task interference is a result of competition for this same general resource, for "attention."... The theory, at least in its application, appears to be entirely circular.... The result is a strategy of research that can do nothing but chase its own tail.... This [strategy] has been singularly unproductive ... for the discovery of the architectural constraints on concurrent psychological processes.... It merely soothes away curiosity by the appearance of having provided an explanation, even before the data have been obtained.

Given Allport's provisos, the architecture of a UTC ought to include only mechanisms that have firm a priori physical or mental grounds, such as those known to be inherent in either the human body's sensors and effectors or intelligent mind's basic computational needs (Newell 1990). Then fewer degrees of freedom will clutter the theoretical landscape and compromise assessments about goodness of fit to empirical data.

Such benefits likewise may accrue through keeping the numerical parameters of the architecture constant insofar as possible. For example, suppose that in two task contexts, the same stimuli or responses are involved. If so, then setting the parameters of the architecture's perceptual-motor processors to have identical values across both contexts will reduce the degrees-of-freedom problem considerably. Similar beneficial constraints can be imposed on the values of cognitive-processor parameters (Meyer and Kieras 1997a,b).

Embodiment of perceptual-motor mechanisms As the preceding discussion also implies, UTCs can gain more power and testability from being physically "embodied." If a UTC makes direct contact with the real world through explicitly represented perceptual and motor mechanisms, then the theory's degrees of freedom at a cognitive level are likely to be reduced, and the Turing tar pit of abstractly conceivable but concretely implausible alternative computational algorithms may be circumvented. Again, Allport (1980, 144, 145, 148) has summarized the basic point quite well:

The constraints of the human body set upper limits on the degrees of freedom of our physical action. A limb cannot be in two positions at once. We cannot shift our gaze simultaneously to right and left, nor vocalize two different syllables at the same time.... Certainly, many of the phenomena attributed hitherto to "attentional" or "general-capacity" limitations can be seen to depend on situations in which separate inputs compete for or share control of the same category of action.... It may be that until we have a better description of what is being done by at least some of the sub-systems, [other] questions about the overall architecture will just be premature.

A prospective UTC therefore should have an array of well-defined perceptual processors for the principal stimulus input modalities, and motor processors for the principal response output modalities should be included, too.

Respect for neurophysiological plausibility To be taken seriously, a prospective UTC must have neurophysiological plausibility as well. Ultimately, whatever architectural components are postulated in the theory ought to accord with mechanisms in the "biological band" of human information processing (Newell 1990, 1992; see also note 10). This requirement further justifies omitting immutable structural decision and response-selection bottlenecks in the cognitive processor of the theory's architecture. Instead, the architecture should enable substantial asynchronous distributed parallel information processing (cf. Rumelhart and McClelland 1986). As Neumann (1987, 362) argued:

[There is no] physiologically established limit on the information that can be picked up at the same time. Neither are there obvious neurophysiological grounds for the assumption that dual-task performance is limited by the hardware properties of the brain. [Instead] there is an immense amount of parallel computation going on simultaneously in the awake brain (see Anderson and Hinton 1981; Creutzfeldt 1983); and there are many sub-systems that integrate information from different sources without an indication of limited capacity.

Sensitivity to task demands Another crucial principle for formulating computational models of multiple-task performance involves being fully sensitive to the logical demands of the tasks at hand. By doing so, the set of plausible models that can perform the tasks will become considerably smaller and more well defined. As a result, precise thorough accounts that have minimal degrees of freedom may quickly emerge for available data.

For example, we (Meyer and Kieras 1997a,b) have found this with respect to the psychological refractory-period (PRP) procedure, which entails performing two discrete RT tasks in rapid succession (Bertelson 1966; Kantowitz 1974; Pashler 1994a; Smith 1967). Under the PRP procedure, the instructions to participants typically demand that one task be "primary" and the other "secondary"; the participants are supposed to produce rapid accurate primary-task responses before making secondary-task responses.¹¹ These requirements strongly constrain which models can account for RT data from the PRP procedure, especially when the models' architecture has no inherent structural cognitive response-selection bottleneck. The PRP procedure mandates that in essence, viable models must emulate such a bottleneck at some point during the course of secondary-task performance even though one would not otherwise be needed.¹² By honoring this mandate, our research has achieved accurate accounts of RT data from not only the PRP procedure but also other related dual-task paradigms.

Application of GOMS analysis Useful guidance for representing the logical consequences of procedural instructions and task demands is provided by GOMS (goals, operators, methods, and selection rules) analysis (Card, Moran, and Newell 1983; John 1990; John and Kieras 1996; John, Vera, and Newell 1994; Kieras 1988; Newell 1990). In this analytical technique, the first step involves identifying and organizing the goals and subgoals for all present tasks, as dictated by the prevailing physical environment, instructions about task priorities, and so forth. Next, procedural methods are formulated so that these goals and subgoals may be achieved. As part of this formulation, sequences of operators (i.e., perceptual, cognitive, and motor transformations) are chosen from an inventory provided by the UTC's functional architecture. The choice of operators is governed by explicit selection rules that tailor the operator sequence to be sufficient and efficient. For example, these rules may invoke the *rationality principle* (Anderson 1990, 1993; Card, Moran, and Newell 1983; Newell 1990), according to which a system's operations should have maximum expected utility. We have found GOMS analysis to be especially useful in modeling performance of some relatively complex practical tasks associated with HCI (Kieras and Meyer forthcoming; Kieras, Wood, and Meyer 1997). In contrast, failure to apply GOMS analysis may leave the theorist trapped in the Turing tar pit, for, as Newell (1973a) originally forewarned, "the same human subject can adopt many radically different methods for the same basic task, depending on goals, background knowledge, and minor details of payoff structure....

To predict a subject you must know: (1) his goals; and (2) the task environment.... [Until these factors are taken into account] we will not be able to bring the problem of specifying subject's methods under control" (Newell 1973a, 293, 299, 301).

Compliance with real-time constraint Finally, the functional architecture of a UTC must comply with the real-time constraint on immediate behavior (Newell 1990, 1992). People perform elementary tasks through relatively simple combinations of operations at the lower end of the "cognitive band," producing reaction times on the order of $100 \text{ msec} < RT < 1,000 \text{ msec}$. Having to describe such rapid performance in terms of mechanisms that also respect known properties of their underlying neural substrates provides more signposts for circumventing the Turing tar pit and degrees-of-freedom problem.

Taxonomy of Lessons

The remainder of this chapter outlines numerous methodological and substantive lessons we have learned from adhering to the aforementioned heuristic principles for developing a unified theory of cognition and action. The substantive lessons highlight specific new empirical facts and theoretical conclusions about the human information-processing system. The methodological lessons highlight inherent nuances of the scientific enterprise whereby UTCs and computational models of multiple-task performance may be formulated.

Lists of the methodological and substantive lessons appear in Tables 2.1 and 2.2, respectively. So that these lessons might be more memorable, we adopt the classical pedagogic practice (e.g., Saunders 1757) of expressing them with brief epigrammatic statements adapted from various familiar sources. Perhaps by learning about our trials and tribulations, other experimental psychologists, cognitive scientists, and human-factors engineers will make faster progress toward understanding human multiple-task performance.

For example, three instructive methodological lessons should be apparent in light of the preceding discourse:

Methodological lesson 1: "Now is the hour" The moment has come to make more progress toward theoretical unification in the scientific study of human performance. Ample data and theoretical concepts are available to support such an advance. Persistent neglect of theoretical unification will waste precious resources and postpone utilitarian transfer of theory and data from the laboratory to practical real-world applications.

Methodological lesson 2: "Reaction time is of the essence" The development of complete veridical UTCs will benefit greatly from RT data that experimental psychology has amassed while studying the performance of

Table 2.1 Methodological Lessons

Lesson number	Source of lesson	Epigrammatic statement
1	Newell 1973, 1990, 1992	"Now is the hour"
2	Newell 1990, 1992	"Reaction time is of the essence"
3	Newell 1973, 1990, 1992	"No pain, no gain"
4	PRP procedure	"Our cup runneth over"
5	PRP procedure	"Expect the unexpected"
6	PRP procedure	"Be careful what you ask for; you might actually get it"
7	AEC models ^a	"Average at your own risk"
8	AEC models ^a	"Seek and you shall find; knock and it shall be opened unto you"
9	NYNEX study of TAOs ^b	"Unification enables application"
10	NYNEX study of TAOs ^b	"Keep your sunny side up"
11	NRL cockpit study ^c	"If you've seen one, you've seen 'em all"
12	NRL cockpit study ^c	"Be thankful for the Second Golden Rule"
13	Feynman 1985	"Psychological science can (and will) be fun"
14	Gopher 1996	"There's nothing more practical than a good theory"

a. See Lauber et al. 1994; Meyer et al. 1995; Meyer and Kieras (1997b); Schumacher et al. (forthcoming).

b. See Kieras, Wood, and Meyer 1995, 1997.

c. See Kieras and Meyer 1995, forthcoming.

Table 2.2 Substantive Lessons

Lesson number	Source of lesson	Epigrammatic Statement
1	SRD model ^a	"Response selection is not like pouring bottled wine"
2	SRD model ^a	"It's difficult to leap before you look"
3	AEC models ^b	"Variety is the spice of life"
4	AEC models ^b	"Wherever there's a will, there are ways"
5	AEC models ^b	"You can teach young performers new tricks"
6	Schumacher et al. 1997	"Dual-task performers can share and share alike"
7	NYNEX study of TAOs ^c	"TAOs know the way"
8	NYNEX study of TAOs ^c	"What goes around comes around"
9	NYNEX study of TAOs ^c	"Hand movements obey the Boy Scout motto"
10	NYNEX study of TAOs ^c	"Practice makes (nearly) perfect"
11	NRL cockpit study ^d	"Daring task scheduling is 'the right stuff'"
12	NRL cockpit study ^d	"The eyes have it"
13	NRL cockpit study ^d	"Covert shifts of visual attention are like poor Yorick"

a. Based on EPIC computational modeling of mean RTs in Hawkins, Rodriguez, and Reicher (1979).

b. Based on EPIC computational modeling of mean RTs from PRP procedure in Lauber et al. 1994.

c. Based on EPIC computational modeling in Kieras, Wood, and Meyer 1995, 1997.

d. Based on EPIC computational modeling in Kieras and Meyer 1995, forthcoming.

various basic tasks. These data, which impose strong real-time constraints on a UTC's functional architecture, are abundant and sorely in need of further integration. *Attention and Performance XVII* together with other volumes of this symposium series therefore can play a crucial role in fostering theoretical unification.

Methodological lesson 3: "No pain, no gain" Future steps toward complete veridical UTCs will not be easy. Considerable misdirection, stumbling, and frustration will occur along the way because of the vast current inventory of potentially apt theoretical constructs, the Turing tar pit, the degrees-of-freedom problem, and other concomitant obstacles. Together, these hinderances may cause the adventurous theorist considerable fatigue and suffering, just as many daunting challenges confronted Indiana Jones in his search for the Lost Ark of the Covenant (Kasdan, Lucas, and Kaufman 1981). Yet the ultimate prize is worth the hardship; both basic research and practical applications will benefit enormously from the elegance, integration, explanatory power, and predictive capability of a complete veridical UTC. Ultimately, UTCs are the only way to achieve such benefits (Newell 1973a, 1990, 1992).

To provide a more detailed framework for further lessons, we next introduce our EPIC information-processing architecture.

2.3 THE EPIC ARCHITECTURE

Figure 2.1 shows a schematic diagram of EPIC. It consists of interconnected LISP software modules for symbolic perceptual, cognitive, and motor information processing. We have designed these modules to emulate basic components of the human information-processing system and to provide a basis for realistic computational models of multiple-task performance. As acknowledged before, EPIC's organization builds on previous work by a number of theorists (e.g., Anderson 1976, 1983, 1990, 1993; Card, Moran, and Newell 1983; Hunt and Lansman 1986; Laird, Newell, and Rosenbloom 1987; Newell 1973a,b, 1990, 1992).

Architectural Components

During computer simulations with EPIC, its perceptual processors receive information from simulated sensors that transduce stimuli presented through input devices (e.g., display screens and headphones) in a virtual task environment. After specified parametric delays, symbolic stimulus codes are sent by the perceptual processors to the declarative working memory of EPIC's cognitive processor. The cognitive processor maintains the contents of working memory, executes procedures for performing various tasks, and instructs the motor processors by sending them symbolic response codes about what actions to take. The motor processors prepare and produce

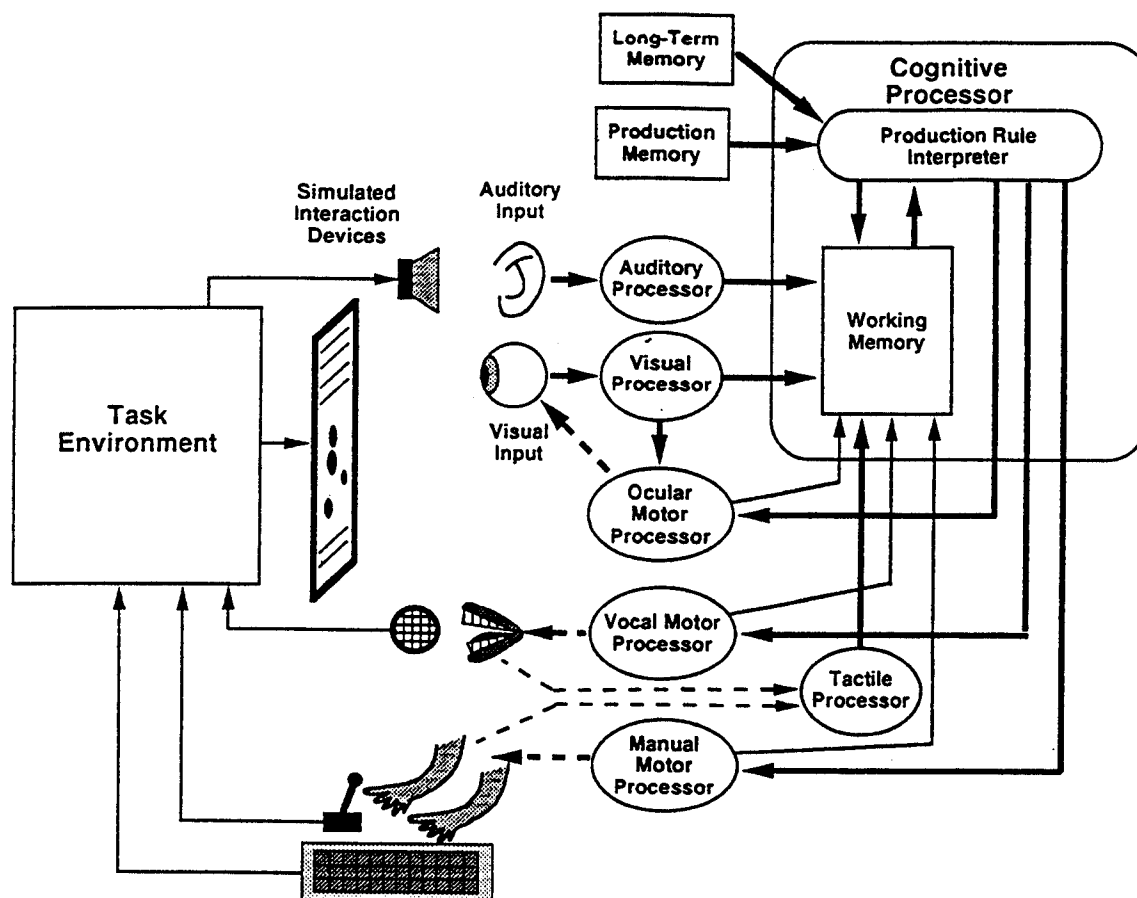


Figure 2.1 Diagram of the executive-process interactive control (EPIC) information-processing architecture.

movements by simulated effectors that operate output devices (e.g., keyboards, joysticks, and microphones) in the virtual task environment.

Together, EPIC and its task environment provide a basis for modeling multiple-task performance in a variety of contexts. The components of the architecture are tailored to be generally applicable and consistent with available empirical data about the nature of the human information-processing system (e.g., Atkinson et al. 1988; Boff, Kaufman, and Thomas 1986; Meyer and Kornblum 1993; Woodworth and Schlosberg 1954). Using the architecture and computational models based on it, a theorist can watch a simulated performer do single or multiple perceptual-motor and cognitive tasks, just as an experimenter observes the performance of a real person.

Perceptual processors In EPIC, the visual, auditory, and tactile sensory modalities each have their own perceptual processor. The inputs to the perceptual processors are stimuli transduced by simulated eyes, ears, and haptic receptors. For example, EPIC's eyes have retinas with foveal, parafoveal, and peripheral regions, so the quality of the inputs to the visual perceptual pro-

cessor depends on the retinal locations of stimuli. The perceptual processors produce symbolic codes for stimulus features such as the locations, shapes, sizes, colors, and identities of visual objects, which are kept in "object files" (cf. Treisman 1988). Stimulus detection and identification times are parameters for our computer simulations (cf. Woodworth and Schlosberg 1954). Their contributions are modulated by the fact that EPIC's perceptual processors function asynchronously and simultaneously with the cognitive processor, whose operations depend on the contents of working memory.

Working memory EPIC's working memory has various subdivisions for storing symbolic information (Kieras et al. forthcoming). Some subdivisions of working memory contain visual, auditory, and tactile perceptual stimulus codes (cf. Baddeley 1986), which may decay over time unless EPIC's cognitive processor refreshes them. Other subdivisions of working memory contain response codes, efference copies of movements, and control codes for the goals, procedural steps, and status of current tasks, which contribute to operations by the cognitive processor.

Cognitive processor EPIC's cognitive processor is programmed with production rules, each of which states that if its specified conditions are true, then its specified actions should be executed. For a rule's conditions to be true, they must match items currently in working memory. When such matches occur, the rule's actions update the contents of working memory and send instructions to EPIC's motor processors. In this way, production rules can be used to perform various perceptual-motor and cognitive tasks.

For example, during a primary auditory-manual choice-reaction task, the following rule might be used by EPIC's cognitive processor to instruct the manual motor processor that it should prepare and produce a keypress with the left index finger in response to an 800 Hz tone:

```
IF
  ((GOAL DO TASK 1)
   (STRATEGY TASK 1 IS IMMEDIATE)
   (AUDITORY TONE 800 ON)
   (STEP DO CHECK FOR TONE 800))
THEN
  ((SEND-TO-MOTOR (MANUAL PERFORM LEFT INDEX))
   (ADD (TASK 1 RESPONSE UNDERWAY))
   (ADD (STEP WAIT FOR TASK 1 RESPONSE COMPLETION))
   (DEL (STEP DO CHECK FOR TONE 800))
   (DEL (AUDITORY TONE 800 ON))).
```

The actions of this rule, which not only instructs the manual motor processor but also adds and deletes items in working memory, would be executed whenever working memory contains all of the items in the rule's conditions. For each current task that EPIC is supposed to perform, there

would be a set of such rules in procedural memory. Also, complementing these task rule sets, procedural memory may contain sets of executive-process rules that help manage the contents of working memory and that coordinate multiple-task performance with respect to prevailing task priorities.¹³

Theoretical RTs for multiple-task performance stem from additional properties of EPIC's cognitive processor. During performance simulations, task and executive production rules are applied by the production-rule interpreter of the cognitive processor, which is based on a parsimonious production system (PPS; Covrigaru and Kieras 1987). Under PPS, the production-rule interpreter operates through a series of processing cycles, whose individual durations vary stochastically around a mean value that typically equals 50 msec.¹⁴ At the start of each cycle, the conditions of all rules currently in procedural memory are tested against the present contents of working memory. At the end of each cycle, for every rule whose conditions completely match the contents of working memory, all of the rule's actions are executed. Given the complexity of representative multiple-task performance, several successive processing cycles may be required to complete each of two or more concurrent tasks, yielding EPIC's theoretical RTs.

At present, however, we assume that there is no limit on how many production rules can have their conditions tested and actions executed during a processing cycle. EPIC's cognitive-processor cycle durations have the same distribution regardless of how many production rules are involved. In this respect, our system architecture therefore differs radically from some past theoretical frameworks; it does not have an inherent structural cognitive response-selection bottleneck or limited reservoir of divisible processing capacity (cf. Kahneman 1973; Moray 1967; Pashler 1984, 1994a; Welford 1959, 1967). On the contrary, with appropriate sets of executive and task production rules, EPIC's cognitive processor may select responses and do other operations simultaneously for concurrent tasks, avoiding between-task interference at this "central" level.

A metatheoretical rationale for such assumptions appears in Meyer and Kieras (1997a). Our reasons conform to the principles of parsimony and neurophysiological plausibility espoused by other proponents of unification in human-performance theory (e.g., Allport 1987; Neumann 1987; Newell 1990, 1992). Some empirical support for the present assumptions about the capacity of EPIC's cognitive processor is provided by studies of multiple-task performance in which virtually perfect time-sharing between two tasks had occurred (e.g., Allport et al. 1972; Greenwald and Shulman 1973; Hirst et al. 1980; Koch 1993, 1994; Schumacher et al. 1997; Shaffer 1975; Wickens 1984).

Motor processors Nevertheless, at a "peripheral" level, EPIC's motor processors do act like bottlenecks similar to ones proposed by some other theorists (e.g., Kantowitz 1974; Keele 1973; Keele and Neill 1978; Reynolds 1964). In our architecture, the ocular, manual, and vocal response modalities

each have their own motor processor. Typically, the inputs to the motor processors are symbolic codes for responses that have been selected by the cognitive processor with its production rules.¹⁵ The outputs by the motor processors are movements of simulated eyes, hands, and mouth, which interact with the virtual task environment. For example, the manual motor processor can produce various styles of hand movement such as pointing, keypressing, typing, and joystick plying. The ocular motor processor can produce eye movements through either voluntary cognitive control or reflexive perceptual control (cf. Fischer and Ramsberger 1984; Rafal, Henik, and Smith 1991; Reuter-Lorenz, Hughes, and Fendrich 1991). Although the ocular, manual, and vocal motor processors may all operate at the same time, each of them individually is a single-channel mechanism that limits the overall rate of overt movements.

To be precise, we assume that after receiving the symbolic code for a selected response, a motor processor converts it to elementary movement features that the response should have overtly. For example, a keypress by the manual motor processor might have features that specify the movement style, hand, and finger to be used (e.g., PRESS, LEFT, INDEX). Consistent with some empirical results (e.g., Abrams and Jonides 1988; Meyer and Gordon 1985; Rosenbaum 1980; Yaniv et al. 1990), the movement features for an overt response are prepared serially, with each feature-preparation step consuming on the order of 50 msec. After all of the movement features for a response have been prepared, the response is produced overtly through a final initiation step that likewise takes on the order of 50 msec. Thus, while symbolic response codes for concurrent tasks may be selected in parallel by EPIC's cognitive processor, the production of distinct overt responses by the same motor processor would have to be temporally staggered, causing potential between-task or "structural" interference (cf. Kahneman 1973).

An especially important case of such interference involves concurrent tasks that each require manual responses. EPIC has only one motor processor devoted to preparing and initiating movements by the two (i.e., right and left) hands. For multiple manual tasks, substantial between-task interference is therefore possible at the peripheral motor level even when the two tasks utilize different hands and different sensory modalities. Effective coping with such interference requires judicious supervisory control. That this control is needed under these circumstances has been demonstrated amply by past studies of manual movement production in multiple-task performance (e.g., Ivry et al. 1994, forthcoming; McLeod 1977).

Contributions by Attention and Performance symposia From the previously cited studies that helped stimulate the development of EPIC, it now should be clear that past Attention and Performance symposia have made major contributions to our thinking. Literally dozens of these studies have been published as part of this series, and many of them have entailed reports of informative RT data. Their prevalence in the present chapter provides

substantial inspiration for our first two methodological lessons (see table 2.1): "Now is the hour" and "Reaction time is of the essence."

Formulation of EPIC Computational Models

We formulate explicit EPIC computational models of multiple-task performance in terms of complementary production-rule sets, which specify the operations of EPIC's cognitive processor. First, for each task at hand, a distinct set of production rules that perform the task with the architecture's various components must be written. The task production rules translate internal stimulus codes to internal response codes and keep other records associated with the individual tasks. Second, a set of production rules for a supervisory executive process must be written. The executive production rules adaptively coordinate progress on the individual tasks so that instructions about task priorities are obeyed and the tasks do not disrupt each other at peripheral perceptual-motor levels. Such coordination is achieved by monitoring the contents of working memory and inserting or deleting task goals together with other control items at appropriate moments along the way. For example, the following executive production rule might be applied to start processing for primary and secondary choice-reaction tasks of a PRP procedure while ensuring that primary-task responses have higher priority than secondary-task responses:

```
IF
  ((GOAL DO DUAL CHOICE RT TASKS)
   (STRATEGY AUDITORY-MANUAL TASK 1)
   (STRATEGY VISUAL-MANUAL TASK 2)
   (VISUAL CENTER EVENT DETECTED ON)
   (NOT (TRIAL UNDERWAY)))
THEN
  ((SEND-TO-MOTOR MANUAL RESET)
   (ADD (TRIAL UNDERWAY))
   (ADD (GOAL DO TASK 1))
   (ADD (GOAL DO TASK 2))
   (ADD (STRATEGY TASK 2 MODE IS DEFERRED))
   (ADD (STRATEGY UNLOCK ON MOTOR-SIGNAL MANUAL
         STARTED LEFT))
   (DEL (VISUAL CENTER EVENT DETECTED ON))
   (ADD (STEP MOVE EYES TO RIGHT))
   (ADD (STEP WAIT-FOR TASK 1 DONE))).
```

The executive production rules for scheduling and coordinating tasks may change, depending on the particular task combinations, priorities, and subjective strategies involved. Our EPIC computational models of multiple-task performance therefore extend previous proposals by theorists who have emphasized the importance of supervisory control in cognition and action

(e.g., Baddeley 1986; Duncan 1986; Logan 1985; Neisser 1967; Norman and Shallice 1986; Shallice 1972).

Assessment of EPIC Computational Models

We assess our EPIC computational models by simulating multiple-task performance under test conditions that mimic those in which empirical data from human participants are collected. During these assessments, an environment-simulation program and human-simulation program are executed conjointly on a computer workstation. The environment-simulation program provides a sequence of stimulus inputs to the human-simulation program and receives a sequence of response outputs from it, just as an experimenter would test a human participant by presenting real stimuli and observing his or her overt behavior. The human-simulation program consists of the EPIC architecture and production-rule sets in EPIC's cognitive processor, which transform stimulus inputs to response outputs through systematic operations like those outlined before (figure 2.1). Both the environment-simulation program and EPIC's software modules are written in LISP. The sets of executive and task production rules used for the human-simulation program conform to the syntax of the PPS production-rule interpreter (Covrigaru and Kieras 1987). Also, as detailed elsewhere (Meyer and Kieras 1997a,b), execution of the simulation programs entails setting the numerical values of parameters in the task environment and EPIC architecture.

After each simulation run, EPIC's outputs may be compared with observed results from human participants. Insofar as the simulated data match empirical data, this would suggest that our models should be taken as potentially veridical descriptions of how human multiple-task performance is achieved. We have found that for at least some models, good fits between simulated and empirical data (e.g., RTs and error rates) may be obtained with relatively few "free" parameters.

Overview of Applications

The subsequent sections of this chapter describe three representative task domains for which we have formulated and applied some EPIC computational models of multiple-task performance: (1) the PRP procedure, a basic laboratory paradigm that embodies some fundamental aspects of multiple-task performance also found under real-world circumstances; (2) human-computer interaction in a practical context, the servicing of requests by customers to telephone operators for the initiation of collect phone calls; and (3) concurrent visual-manual tracking and tactical decision making in military aircraft operation, another practical context. From focusing on these diverse task domains and our models for them, the potential value of unification in human-performance theory may become clearer, and more lessons relevant to the search for a unified theory of cognition and action may emerge.

2.4 APPLICATION TO THE PSYCHOLOGICAL REFRACTORY-PERIOD PROCEDURE

The PRP procedure is a popular laboratory paradigm for studying human multiple-task performance (Bertelson 1966; Kantowitz 1974; Meyer and Kieras 1997a,b; Pashler 1994a; Smith 1967). Many chapters in *Attention and Performance* volumes have been based on it (see, for example, Kornblum 1973; Koster 1969; Meyer and Kornblum 1993; Sanders 1967, 1970). Such popularity stems from the procedure's simplicity, fecundity, and similarity to important real-world situations in which people must perform perceptual-motor and cognitive tasks concurrently. It is therefore fitting that we have formulated our initial EPIC computational models to explain and predict representative PRP data (Meyer and Kieras 1992, 1994, 1996, 1997a,b; Meyer et al. 1995). Any bona fide UTC should take these data seriously, and as we show later, doing so sets the stage for analyses of multiple-task performance in more complex practical contexts.

Methodology

In a representative experiment with the PRP procedure, there is a series of discrete trials during which two distinct tasks must be performed more or less concurrently. On each trial, a warning signal is followed by a stimulus (e.g., visual letter or auditory tone) for the first task. Given the task 1 stimulus, a participant must make a fast and accurate task 1 response (e.g., press a finger-key or say a word). Soon after the task 1 stimulus, another stimulus is presented for the second task. The perceptual modality and semantic category of the task 2 stimulus may differ from those of the task 1 stimulus. The time between the two stimuli is the *stimulus-onset asynchrony* (SOA), which typically ranges between 0 and 1 sec. Given the task 2 stimulus, the participant must make a fast and accurate task 2 response. The effector for the task 2 response may differ from that for the task 1 response. In most cases, instructions for the PRP procedure require that task 1 has higher priority than task 2; they may also urge participants to make the task 1 response first (see note 11).

The experimenter analyzes the RTs for tasks 1 and 2 to assess how much the two tasks interfere with each other. Specifically, mean RTs for task 2 may be plotted versus the SOA, forming PRP curves, which ordinarily decline as the SOA increases. Depending on various methodological details, this SOA effect—also called the “PRP effect”—can combine either additively or interactively with the effects of other factors (e.g., stimulus discriminability, response-selection difficulty, and movement complexity). Models of multiple-task performance should account for the absolute magnitudes of the RTs and the observed patterns of factor effects on them.

Lessons from the PRP Procedure

Several methodological and substantive lessons (see tables 2.1 and 2.2) may be learned from the PRP procedure and our efforts to formulate EPIC computational models for performance under it.

Methodological lesson 4: "Our cup runneth over" The literature of experimental psychology contains many alternative patterns of PRP curves, manifesting various additivities and interactions among the effects of SOA and other factors on mean RTs for task 2. Even when we confine our attention to factors that presumably affect just one particular processing stage (e.g., response selection) for the secondary task, considerable diversity appears in the PRP-curve patterns obtained through manipulating those factors. Metaphorically, our cup runneth over. The PRP procedure provides a copious fermentation of data to be explained quantitatively and modeled computationally.

Substantive lesson 1: "Response selection is not like pouring bottled wine" Because of the apparent diversity in PRP-curve patterns, a crucial substantive lesson has emerged as well. Results from our computational modeling suggest that after only moderate practice, people do not select responses to stimuli through an immutable structural bottleneck that only can deal with one task at a time. Response selection is not like pouring bottled wine. On the contrary, responses to two or more stimuli apparently can be and are sometimes selected concurrently, thereby enabling substantial temporal overlap between streams of processing for different tasks. Although in accord with our initial assumptions about EPIC's cognitive processor, such overlap patently contradicts some past hypotheses about how multiple-task performance takes place under the PRP procedure. This contradiction has taught us a fifth methodological lesson, which concerns the attitude one should adopt when pursuing theoretical unification.

Methodological lesson 5: "Expect the unexpected" Unexpected conceptual twists and turns are to be expected along the way. For example, before our journey with EPIC began, the best-known explanation of the PRP effect on secondary-task RTs was the response-selection bottleneck (RSB) hypothesis (Welford 1959, 1967; McCann and Johnston 1992; Pashler 1984, 1990, 1993, 1994a; Pashler and Johnston 1989). Past theorists have argued for it based on quasi-additive effects of SOA and task 2 response-selection factors in some PRP experiments. However, EPIC computational models of people's performance under this procedure raise grave doubts about these arguments. Such abrupt changes in theoretical direction are likely to happen repeatedly down the road, because a major benefit of seeking UTCs is the discovery of new and uncharted but promising conceptual territory.

Methodological lesson 6: "Be careful what you ask for; you might actually get it" Also relevant here is a sixth methodological lesson that pertains directly to why past theorists have been misled in their advocacy of the traditional RSB hypothesis. As mentioned before, the instructions to participants under the PRP procedure request that responses for task 1 have higher priority and earlier onsets than responses for task 2. This request explicitly constrains participants to use some bottleneck mechanism that restricts the "flow" of information processing for task 2, so that responses for task 2 do not occur before those for task 1 at short SOAs. For example, in an EPIC computational model, precluding out-of-order task 2 responses requires imposing temporary strategic software bottlenecks through executive control. Theorists who advocate the traditional RSB hypothesis may have been misled because experimenters who adopted the PRP procedure actually got the kind of performance that participants were asked to produce. Consistent with methodological lesson 6, care therefore must be taken to accommodate the role of task instructions in modeling multiple-task performance (cf. Newell 1973a, 1990).

An Instructive PRP Study

We first learned these lessons through formulating EPIC computational models to account for results from an instructive PRP study by Hawkins, Rodriguez, and Reicher (1979).

Procedure In this study, there were four different primary tasks, which involved either auditory stimuli (tones) or visual stimuli (printed letters) and either manual responses (keypresses by left-hand fingers) or vocal responses (spoken words). Participants performed each primary task together with one or the other of two different secondary tasks, which involved either two or eight visual stimuli (digits) and two manual responses (keypresses by right-hand fingers). For each combination of tasks 1 and 2, the SOAs ranged from 0 to 1,200 msec. These manipulations let participants' RTs for the two tasks be measured jointly as a function of the SOA, task 1 perceptual modality, task 1 motor modality, and task 2 response-selection difficulty (Sanders 1980; Sternberg 1969).¹⁶ Thus, this study provides a large set of data with which to test alternative models of basic human multiple-task performance, thereby exemplifying the riches mentioned before in methodological lesson 4.

Results Figure 2.2 shows empirical mean RTs for tasks 1 and 2 reported by Hawkins, Rodriguez, and Reicher (1979). Several salient aspects of these data should be noticed.

First, consider the empirical mean RTs for task 1 (large unfilled symbols on solid curves). Primary-task responses took longer when the task 1 stimuli were auditory rather than visual (figures 2.2A and 2.2C vs. figures 2.2B and

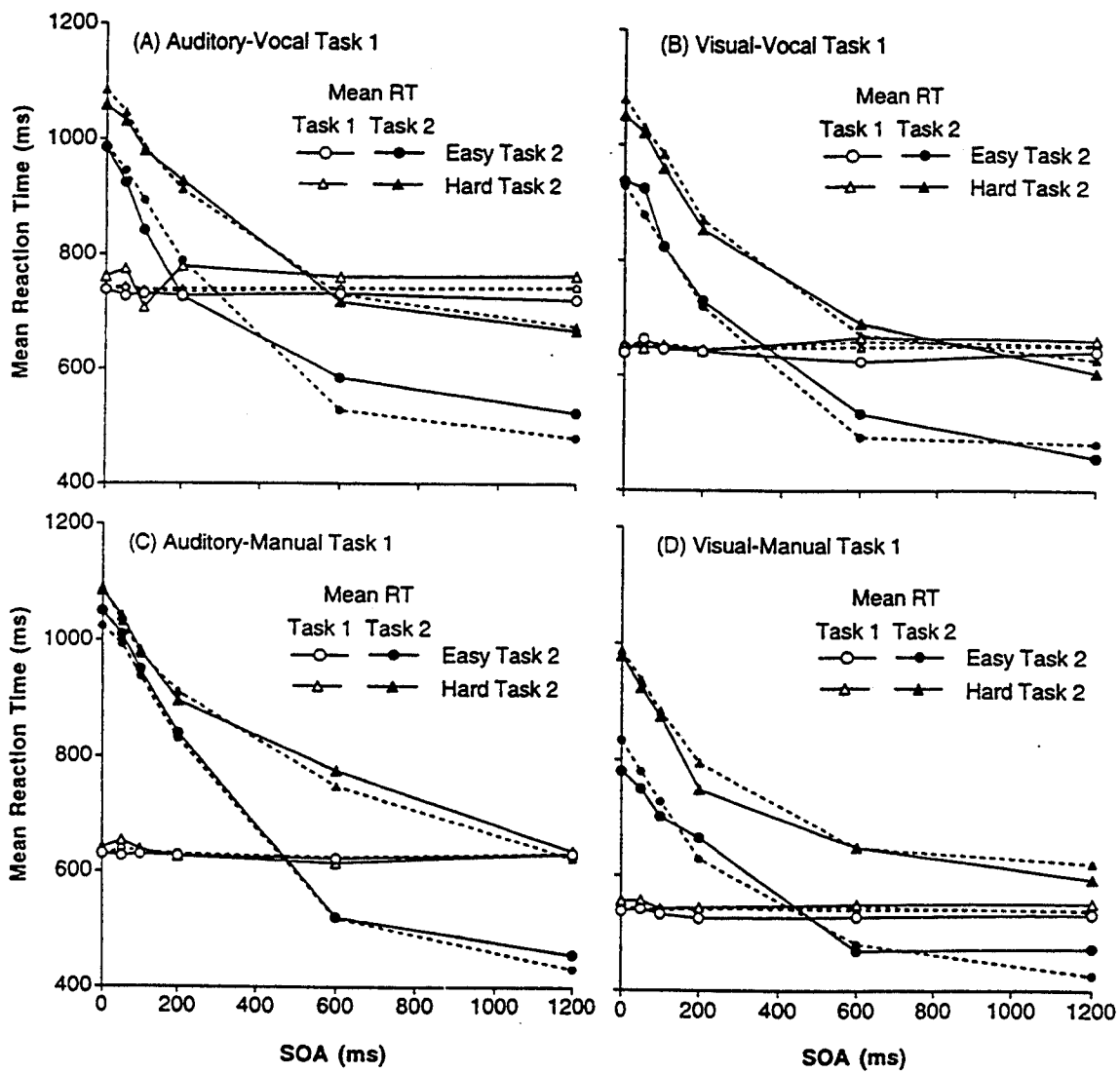


Figure 2.2 Results based on the PRP procedure in Hawkins, Rodriguez, and Reicher (1979). Large symbols on solid curves represent empirical mean RTs; small symbols on dashed curves represent simulated mean RTs produced by the strategic response-deferment (SRD) model. Filled circles and triangles represent mean RTs for task 2 when response selection in task 2 was, respectively, easy and hard; unfilled circles and triangles represent corresponding mean RTs for task 1. A. Simulated versus empirical mean RTs for a combination of auditory-vocal task 1 and visual-manual task 2. B. Simulated versus empirical mean RTs for a combination of visual-vocal task 1 and visual-manual task 2. C. Simulated versus empirical mean RTs for a combination of auditory-manual task 1 and visual-manual task 2. D. Simulated versus empirical mean RTs for a combination of visual-manual task 1 and visual-manual task 2.

2.2D) and when the task 1 responses were vocal rather than manual (figures 2.2A and 2.2B vs. figures 2.2C and 2.2D). These effects of the task 1 stimulus and response modalities were essentially additive. This suggests that stimulus encoding and movement production for task 1 occurred during temporally separate stages of processing (Sternberg 1969). Furthermore, regardless of the task 1 stimulus and response modalities, neither the SOA nor the difficulty of response selection in task 2 affected the mean RTs for task 1 very much. This suggests that consistent with typical instructions in the PRP procedure, participants almost invariably gave task 1 priority over task 2.

Next, let us consider the empirical mean RTs for task 2 (large filled symbols on solid curves). Regardless of which task 1 was involved, the difficulty of task 2 affected mean RTs for task 2 directly (figures 2.2A through 2.2D), as we would expect if task 2 response selection took longer when task 2 involved more S-R pairs. Also, regardless of which task 1 was involved, the SOA affected mean RTs for task 2 inversely. Presumably this happened because after the shorter SOAs, some stage of processing in task 2 had to be postponed temporarily until processing for task 1 progressed enough that task 2 responses would seldom, if ever, precede task 1 responses.

However, the relationship between the effects of SOA and task 2 difficulty on the mean RTs for task 2 differed as a function of which task 1 was involved. For example, when a visual-manual task 1 was involved (figure 2.2D), mean RTs for task 2 were affected almost additively by the SOA and response-selection difficulty in task 2. Such additivity also tended to occur when a visual-vocal task 1 was involved (figure 2.2B). Given locus-of-slack logic (McCann and Johnston 1992; Meyer and Kieras 1997a,b; Pashler 1984), these additivities would be consistent with the traditional RSB hypothesis. Nevertheless, when either an auditory-vocal or auditory-manual task 1 was involved, substantial interactions occurred between the effects of SOA and task 2 response-selection difficulty on mean RTs for task 2 (figures 2.2A and 2.2C); the difficulty effects were considerably less at shorter SOAs than at longer SOAs, yielding marked underadditive interactions. By locus-of-slack logic, such underadditivity is inconsistent with the RSB hypothesis (Karlin and Kestenbaum 1968; Keele 1973; Meyer and Kieras 1997a,b; Schumacher et al. forthcoming; Schvaneveldt 1969). On the contrary, it appears here that response-selection processes for tasks 1 and 2 took place concurrently after shorter SOAs.

Such heterogeneous patterns of additivity and interaction involving manipulations of SOA and other factors (e.g., S-R compatibility) that influence response-selection difficulty for task 2 also have been reported elsewhere (e.g., Schumacher et al. 1996, forthcoming). Frequent, but not universal, temporal overlap of the response-selection processes for tasks 1 and 2 therefore may occur under the PRP procedure. Of course, this is what led us to substantive lesson 1, methodological lesson 5, and methodological lesson 6 at the start of the present section.

Theoretical implications Given these lessons and the discoveries on which they rest, the traditional RSB hypothesis must be abandoned. Instead, a new and more apt account of multiple-task performance that takes what we have learned to heart is needed for the PRP procedure. The next subsection outlines what this new account entails.

Adaptive Executive-Control Models

To account quantitatively for human performance under the PRP procedure, we have formulated a class of adaptive executive-control (AEC) models based on the EPIC architecture (Meyer and Kieras 1996, 1997a,b; Meyer et al. 1995). Our AEC models incorporate executive processes that flexibly control the extent to which secondary-task processes may overlap temporally with primary-task processes. Figure 2.3 outlines how such control is achieved.

According to this view, performance of each task progresses through a sequence of stages, including stimulus identification, response selection, and

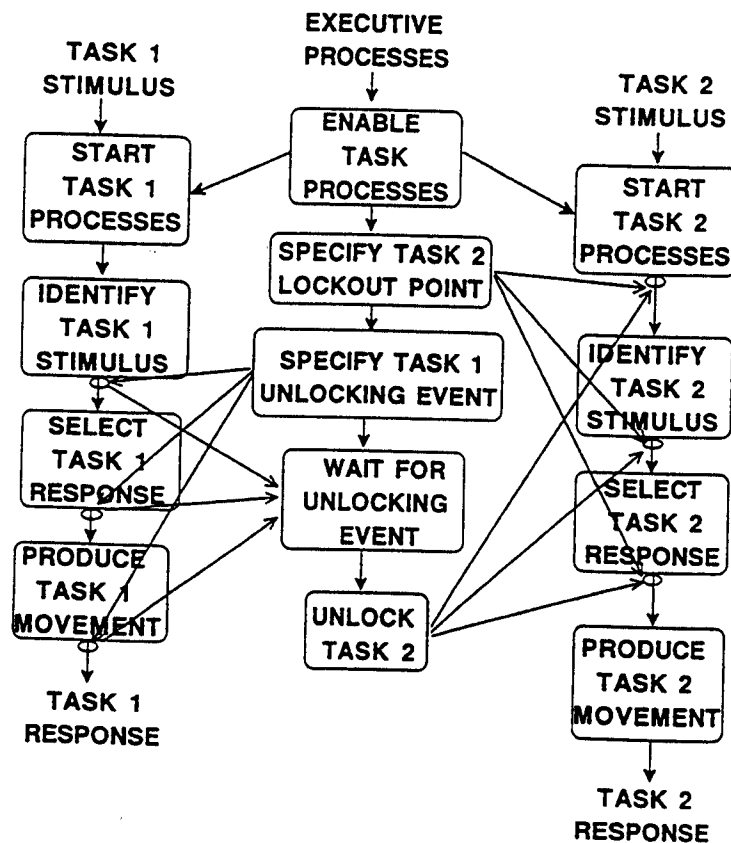


Figure 2.3 Adaptive executive-control (AEC) models of multiple-task performance for the PRP procedure. Diagonal lines with arrows that extend rightward from executive processes to secondary-task processes indicate alternative lockout points for task 2 (right-side ovals). Diagonal lines with arrows that extend leftward from executive processes to primary-task processes indicate alternative unlocking events for task 1 (left-side ovals).

movement production, consistent with discrete stage models (Sternberg 1969; Sanders 1980). An executive process coordinates progress on the primary and secondary tasks by optionally postponing one or more stages of processing for task 2 until task 1 has finished. The supervisory functions of the executive process include (a) enabling the primary-task and secondary-task processes to begin at the start of each trial; (b) specifying a temporary task 2 lockout point; (c) specifying a temporary task 1 unlocking event; (d) waiting for the task 1 unlocking event to occur; and (e) unlocking task 2 processes so that their responses may be completed. Together, these functions ensure that instructions associated with the PRP procedure are satisfied (i.e., task 1 responses receive higher priority and occur before task 2 responses) even though there is enough cognitive processing capacity to perform the two tasks concurrently with little or no interference between tasks. Through the particular combination of task 2 lockout point and task 1 unlocking event that it uses, the executive process can adjust exactly how much priority is given task 1 over task 2.

Lockout points for task 2 By definition, the lockout point for task 2 is a point during the course of task 2 such that when it has been reached, further processing for task 2 stops temporarily until task 1 enters a "done" state. Under our AEC models, there are at least three alternative lockout points for task 2 (figure 2.3, right-side ovals), located before the onsets of, respectively, stimulus identification, response selection, and movement production for task 2. Depending on whether the executive process uses a premovement, preselection, or preidentification lockout point, the processes for task 2 would overlap more or less with those processes for task 1 after short SOAs.

Unlocking events for task 1 The amount of temporal overlap between processes for task 1 and those for task 2 also depends on which unlocking event is used for task 1. By definition, this is an event during the course of task 1 such that when it occurs, task 1 is deemed to be "done," and the executive process permits processing for task 2 to progress beyond the lockout point for task 2. Under our AEC models, there are several alternative unlocking events for task 1 (figure 2.3, left-side ovals); task 1 may be deemed "done" immediately after either its stimulus identification, response selection, or movement production stage finishes. Depending on whether the executive process uses a postidentification, postselection, or postmovement unlocking event, the processes for task 2 would overlap more or less with those for task 1 after short SOAs.

Particular cases Overall, the class of AEC models includes many particular cases. For each possible combination of lockout point for task 2 and unlocking event for task 1, there is a specific set of executive production rules that can implement this combination, achieving a certain preferred amount of temporal overlap between the two tasks. Which executive rule set

is used under what circumstances varies with task instructions, strategic goals, perceptual-motor requirements, prior practice, cognitive style, and personal preference.

From this perspective, the choices of a lockout point and unlocking event are analogous to the choice of a decision criterion (β) in signal-detection theory (Tanner and Swets 1954), which may vary with the relative payoffs and costs assigned to one type of response outcome versus another. For example, some models within the AEC class mimic a response-selection bottleneck by using a preselection lockout point for task 2 and a postselection unlocking event for task 1. Such "cautious" task scheduling could be preferred when prevailing circumstances strongly encourage that task 1 responses always precede task 2 responses. Other models within the AEC class mimic a movement-production bottleneck by using a postselection/preproduction lockout point for task 2 and a post-production unlocking event for task 1 (cf. De Jong 1993; Keele 1973). Such "daring" task scheduling could be preferred instead when circumstances strongly encourage that task 2 responses be produced almost as quickly as task 1 responses at short SOAs.

Strategic Response-Deferment Model

Among models in the AEC class, one with which we have worked extensively is the strategic response-deferment (SRD) model. This model is interesting and apt because as each trial evolves during the PRP procedure, its executive process first uses a post-response-selection lockout point for task 2 but later briefly imposes a pre-response-selection lockout point, depending on how far the processes for task 2 have progressed when the pre-specified unlocking event for task 1 occurs. Given such adaptive executive control, mean RTs for task 2 produced by the SRD model closely match various patterns of empirical PRP curves from previous experiments with the PRP procedure. These patterns and the model's goodness of fit take into account not only the effects of SOA but also effects caused by manipulations in the relative difficulties of the primary and secondary tasks (Meyer and Kieras 1997a,b; Meyer et al. 1995).

Details of executive process Figure 2.4 outlines the executive process of the SRD model in more detail. At the start of each trial during the PRP procedure, the executive process puts task 1 in an *immediate response-transmission mode* and task 2 in a *deferred response-transmission mode*. While task 2 is in deferred mode, the symbolic identities of its responses may be selected and sent to declarative working memory, but overt response movements for task 2 are not produced by EPIC's motor processors. This constraint is imposed by adding a special control note to working memory, which specifies a post-selection/preproduction lockout point for task 2. Putting task 1 in immediate mode lets its responses be selected and sent to their motor processor as

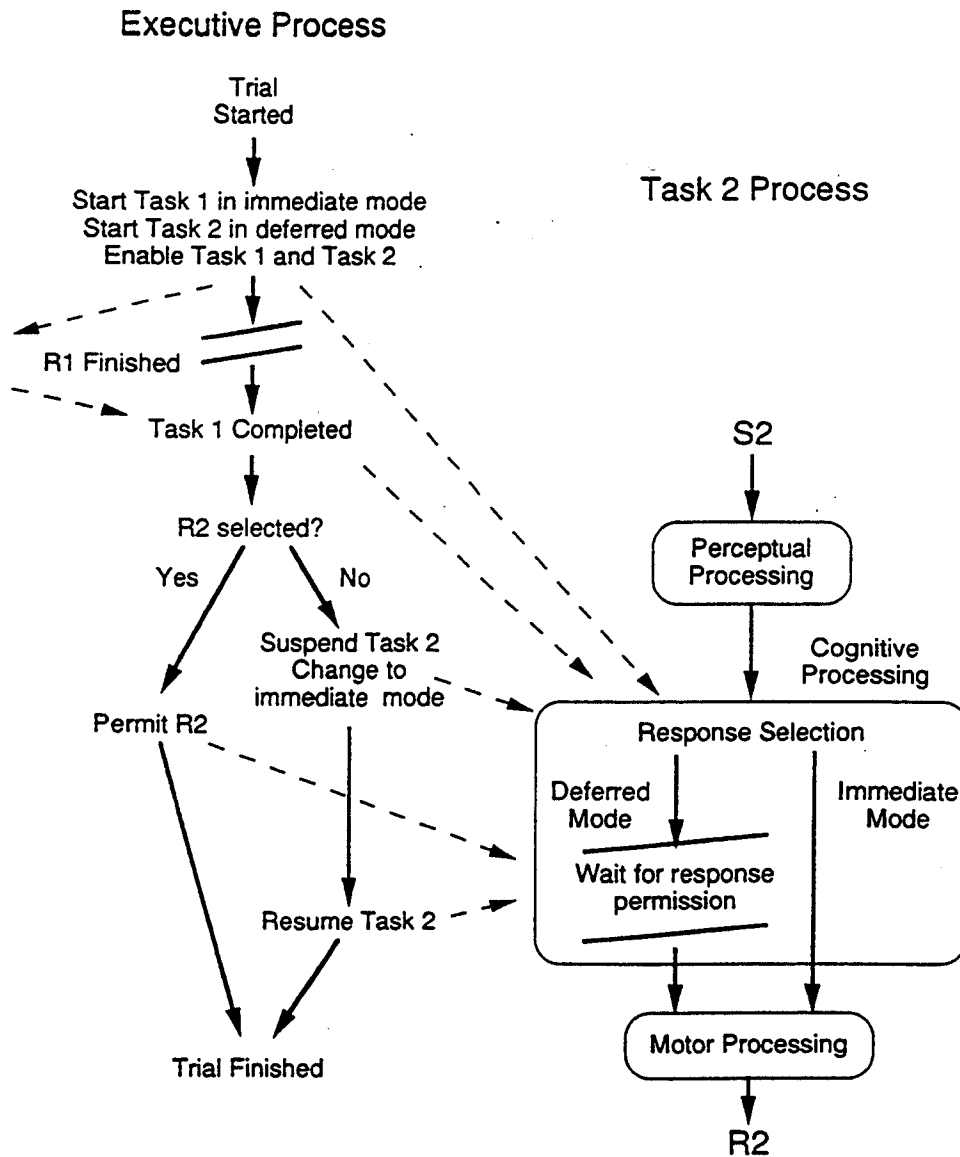


Figure 2.4 Steps taken by the executive process of the SRD model to unlock processes in task 2 for the PRP procedure after task 1 has been declared "done." Breaks in the vertical time lines shown by diagonal hash marks represent variable time intervals whose durations depend on the SOA and temporal properties of prior processes.

quickly as possible for movement production. This freedom is enabled by adding another control note to working memory. When the task 1 unlocking event later occurs (e.g., the overt response movement for task 1 is initiated), the executive process temporarily suspends task 2 and shifts it to immediate mode, after which task 2 is resumed. Following this transition, the identities of previously selected responses for task 2 may be transferred from working memory to their motor processor for movement production. If response selection has not yet finished for task 2 before it is shifted to immediate mode, then subsequently the production rules for task 2 will both select and send the identities of its responses directly to their motor processor.¹⁷

Alternative paths of information processing and reaction time equations for task 2 Because of how its executive process works, five alternative paths of information processing (different sequences of operations) may lead from stimuli to response movements for task 2 in the SRD model (Meyer and Kieras 1997a, figures 10 through 13). Which path is taken during a particular trial of the PRP procedure depends on the SOA and the relative difficulty of task 1 versus task 2. Associated with each path is a distinct equation that defines the RTs for task 2 in terms of the model's parameters and the SOA (Meyer and Kieras 1997a, table 3). Under some experimental conditions, all five paths and equations contribute to the RTs for task 2 over the interval of positive SOAs. Under other experimental conditions, the RTs for task 2 stem from only a subset of these paths and equations. Consequently, the SRD model implies that the SOA and other factors (e.g., response-selection difficulty for task 2) can affect mean RTs for task 2 either additively or interactively, depending on prevailing experimental conditions (Meyer and Kieras 1997a, figure 15).

Account of mean reaction times in Hawkins et al. Given its theoretical implications, we have applied the SRD model successfully in accounting for the mean RT data of Hawkins, Rodriguez, and Reicher 1979. A summary of the obtained account appears in figure 2.2, which shows simulated mean RTs from the SRD model versus empirical mean RTs as a function of SOA for various combinations of primary and secondary tasks. With respect to both tasks 1 and 2, the simulated RTs fit the empirical RTs fairly well regardless of which perceptual and motor modalities were involved during task 1. For task 2, the simulated RTs accurately approximate the interactive and additive effects of SOA and response-selection difficulty on the empirical RTs. The SRD model required relatively few context-dependent parameters to achieve its goodness of fit; the number of parameter values used here was markedly less than the number of reliable one-degree-of-freedom contrasts in the empirical mean RT data. Similarly, the SRD model (Meyer and Kieras 1997a,b) has provided accurate parsimonious quantitative accounts of mean RTs from a variety of other studies with the PRP procedure (e.g., Karlin and Kestenbaum 1968; McCann and Johnston 1992; Pashler 1990). This led us to substantive lesson 1.

Further Lessons

Our SRD model and other related members of the AEC class have also taught us other substantive and methodological lessons (see tables 2.1 and 2.2).

Substantive lesson 2: "It's difficult to leap before you look" A second substantive lesson concerns the role of eye movements during concurrent choice-reaction tasks. As figure 2.2 indicates, large underadditive interactions between the effects of SOA and response-selection difficulty on mean RTs

for task 2, which manifest temporal overlap between concurrent response-selection processes, occurred in Hawkins, Rodriguez, and Reicher 1979 when task 1 involved auditory stimuli and task 2 involved visual stimuli. However, when both tasks involved visual stimuli, these SOA and difficulty effects were approximately additive, suggesting that the selection of responses for task 2 usually took place after the selection of responses for task 1 had finished. Why and how did this happen? Intervening eye movements provide the answer.

When both tasks in Hawkins, Rodriguez, and Reicher 1979 involved visual stimuli, a large (>5 degrees) visual angle separated them. Thus, on each trial there, participants first had to look at the stimulus for task 1 and later had to make a time-consuming saccadic eye movement to look at the stimulus for task 2. This intervening saccade, which was not needed when task 1 involved auditory stimuli, presumably prevented the response-selection processes for tasks 1 and 2 from temporally overlapping. By when the stimulus for task 2 had been identified and response selection for task 2 had begun after the saccade, response selection for task 1 already would have finished even if the SOA was very short. The consequences of this constraint are embodied in the SRD model's simulated RTs for task 2 under these conditions (figures 2.2B and 2.2D). We therefore learn here that although procedural cognitive processes may have the capacity to select responses concurrently for multiple tasks, such capacity will not be apparent if a peripheral perceptual-motor bottleneck (also referred to as "structural interference"; Kahneman 1973) precludes its benefits.

Substantive lesson 2 is relevant to interpreting results not only from Hawkins, Rodriguez, and Reicher 1979 but also from other studies as well. For example, McCann and Johnston (1992, experiment 2) reported a study in which the SOA and difficulty of response selection for task 2 affected the mean RTs for task 2 almost additively. Again this additivity may have stemmed from mandatory eye movements intervening between the onsets of stimuli for task 1 and task 2 (Meyer and Kieras 1997b). If so, then McCann and Johnston's results would not support the traditional RSB hypothesis *per se*; instead, a peripheral perceptual-motor bottleneck could account for them. Given such considerations, viable UTCs must incorporate realistic treatments of the contributions and constraints associated with eye movements during human multiple-task performance.

Substantive lesson 3: "Variety is the spice of life" A third substantive lesson from our work with the PRP procedure is that there are interesting individual differences in how people coordinate multiple-task performance. According to our AEC models (see figure 2.3), people may have two alternative types of strategy for scheduling performance of the primary and secondary tasks in the PRP procedure (Meyer and Kieras 1996, 1997b). One type of strategy is cautious. *Cautious scheduling strategies* use relatively early (e.g., preselection) lockout points for task 2 and relatively late (e.g., post-

production) unlocking events for task 1, as if there were a cognitive response-selection bottleneck. This allows little temporal overlap between processes for task 1 and those for task 2, increasing RTs for task 2 after short SOAs in order to minimize the likelihood that overt responses for task 2 might precede those for task 1 and thereby violate the PRP procedure's requirements. By contrast, a second type of scheduling strategy is daring. *Daring scheduling strategies* use relatively late (e.g., postselection) lockout points for task 2 and relatively early (e.g., preproduction) unlocking events for task 1, consistent with there being no response-selection bottleneck.¹⁸ This allows greater temporal overlap between processes for task 1 and those for task 2, decreasing RTs for task 2 at short SOAs but increasing the likelihood that overt responses for task 2 might precede those for task 1 and violate the PRP procedure's requirements. Which scheduling strategy is adopted under what circumstances presumably depends on factors such as the subjective difficulties of tasks 1 and 2, people's prior experience with multiple-task situations, and their personal preferences for conservative or aggressive task performance.

Given these considerations, we predicted that under the PRP procedure, some participants' RT data would manifest cautious task scheduling and others' would manifest daring task scheduling. Indeed, this prediction has been confirmed already in a study by Lauber et al. (1994, experiment 2; see also Meyer et al. 1995). Their study replicated the one of Hawkins, Rodriguez, and Reicher (1979) with an auditory-manual task 1 and visual-manual task 2, except that Lauber et al.'s task 1 was more challenging.¹⁹ This change encouraged more participants to use cautious rather than daring task scheduling (Meyer and Kieras 1997b).²⁰ As a result, diverse patterns of PRP curves appeared in the participants' RT data.

Figures 2.5A and 2.5B illustrate this diversity clearly. In figure 2.5A are the PRP curves of a participant for whom the effect of response-selection difficulty on mean RTs for task 2 increased reliably as the SOA decreased. Here the negative SOA-by-difficulty interaction suggests a cautious scheduling strategy that used a hybrid combination of preselection and postselection lockout points for task 2 (Meyer and Kieras 1997b). By contrast, figure 2.5B shows the PRP curves of a different participant, for whom the effect of response-selection difficulty on mean RTs for task 2 decreased reliably as the SOA decreased. The latter positive SOA-by-difficulty interaction is exactly opposite to the former one (cf. figure 2.5A), suggesting a daring scheduling strategy that consistently used a postselection lockout point for task 2 (Meyer and Kieras 1997b). The difference between the task-scheduling strategies of these two participants occurred even though their mean RTs for task 1 were similar.²¹ This outcome, consistent with substantive lesson 3, again demonstrates how striving toward a complete veridical UTC can uncover instructive new phenomena that otherwise might go unnoticed from the perspective of older theoretical viewpoints such as the traditional RSB hypothesis.²²

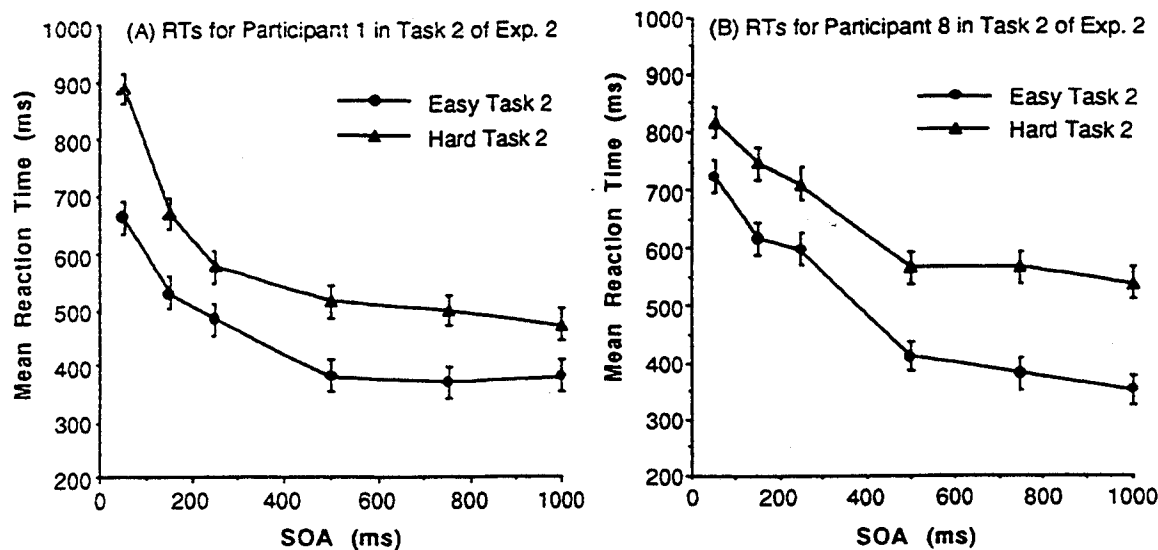


Figure 2.5 Results from using the PRP procedure in Lauber et al. 1994, experiment 2. A. Mean RTs for task 2 as function of SOA and response-selection difficulty in task 2 for participant 1. B. Mean RTs for task 2 as function of SOA and response-selection difficulty in task 2 for participant 8.

Substantive lesson 4: "Wherever there's a will, there are ways"
 Related to the preceding analysis is a fourth substantive lesson; executive cognitive processes and various strategies of task performance are everywhere. Apparently, they play crucial roles not only in complex cognitive domains like memory (Reitman 1970) and problem solving (Newell and Simon 1972) but also in seemingly simpler domains like selective attention and elementary multiple-task performance (cf. Moray 1979). Of course, this generality is not too surprising in light of some past precedents.

For example, consider sensory psychophysics (Green and Swets 1996; Krantz 1969; Luce 1963; Tanner and Swets 1954). At first, the dominant model in sensory psychophysics was high-threshold theory (HTT). Analogous to the traditional perceptual and response-selection bottleneck hypotheses (Broadbent 1958; Pashler 1994a; Welford 1959, 1967), HTT assumed that human observers detect simple sensory stimuli (e.g., light flashes and tone bursts) through a discrete all-or-none threshold mechanism, wherein the subjective stimulus intensity must exceed some constant absolute level to be detected. Because of this threshold's putative rigidity, it provided little room for observers' decision criteria and judgment strategies. As a result, many problematic psychophysical data went unexplained.

Ultimately, however, signal-detection theory (SDT) emerged on the scene, reconciling phenomena that had previously bedeviled HTT. Unlike in HTT, no discrete absolute high threshold is assumed in SDT. Instead, SDT characterizes observers' detection performance in terms of stochastic processes that involve a continuum of sensory states and adjustable decision criteria. According to this characterization, observers set their decision criteria (beta values)

strategically to achieve various preferred frequencies of "hits" for stimulus signals and "correct rejections" for noise, depending on prevailing reward schemes. These strategic adjustments account well for the forms of receiver operating-characteristic (ROC) curves and the different points that observers adopt on them.

Similar to this account is our use of the SRD and other AEC models for interpreting the forms of PRP curves (Meyer et al. 1995; Meyer and Kieras 1997b). In essence, the lockout points for task 2 and the unlocking events for task 1 in the AEC models play much the same conceptual role that SDT's decision criteria play. If the detection of simple sensory signals involves sophisticated supervisory executive processes, it is only natural that multiple-task performance does too. The moral of our story is that a complete veridical UTC must acknowledge such processes and provide a basis for computationally modeling the diverse performance strategies associated with them.

Methodological lesson 7: "Average at your own risk" That human multiple-task performance is mediated by various task-scheduling strategies has other methodological implications as well. Results from paradigms like the PRP procedure should not be averaged across participants without first checking for systematic individual differences among them. As mentioned elsewhere (e.g., Estes 1956; Meyer et al. 1988; Siegler 1987), forgoing such checks can lead to seriously erroneous conclusions.

For example, consider figure 2.6A. Here we have plotted RTs for task 2 obtained by averaging over a whole group of participants who were in the same experiment from which figure 2.5 came. After such averaging, the mean RT data for task 2 appear to form "parallel" (vertically equidistant) PRP curves that embody nearly additive effects of SOA and response-selection difficulty. If only these averaged data had been considered, then one might conclude that the participants all performed according to the traditional RSB hypothesis (cf. McCann and Johnston 1992; Pashler 1984, 1994a; Pashler and Johnston 1989). However, this conclusion would not be correct with respect to the individual participants' data patterns, as indicated by figures 2.5A, 2.5B, and 2.6B.

In particular, figure 2.6B shows interactions between the effects of SOA and response-selection difficulty on mean task 2 RTs for each individual participant who contributed to figure 2.6A. Here the distribution of interactions is rather diffuse; one participant had an approximately null interaction (i.e., additive effects of SOA and task 2 difficulty), but three participants had markedly negative interactions (e.g., see figure 2.6B, participant 1), and four others had various magnitudes of positive interaction including some that were quite large (e.g., see figure 2.6B, participant 8). This is not what would happen if a traditional response-selection bottleneck mediated every participant's performance. To the contrary, some participants apparently used task-scheduling strategies that were cautious (i.e., lacked overlapping response-selection processes), but others used strategies that were more

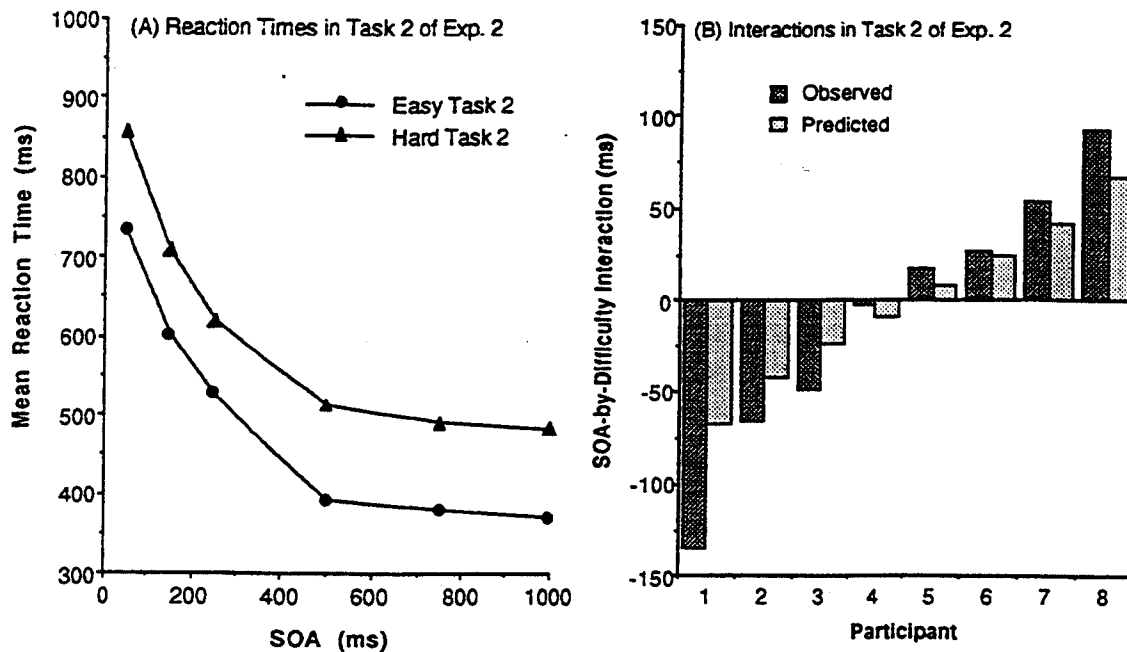


Figure 2.6 Further results from using the PRP procedure in Lauber et al. 1994, experiment 2 (cf. figure 2.5). *A.* Mean RTs for task 2 as function of SOA and response-selection difficulty in task 2 obtained by averaging over a group of eight participants. *B.* Magnitudes of interaction between the effects of SOA and response-selection difficulty on mean RTs in task 2 for the individual participants who contributed to figure 2.6A. Dark vertical bars represent the participants' observed interactions. Light vertical bars represent the distribution of predicted interactions that should have occurred if all participants used the same cautious strategy of task scheduling and the observed interactions differed only because of between-trial RT variance.

daring (i.e., had overlapping response-selection processes).²³ Averaging the RT data for task 2 across these two subgroups of participants obscures the crucial differences between them, creating the illusion that the traditional RSB hypothesis has been supported. Consequently, one might wonder how often such illusions have occurred on past occasions when experimenters advocated the RSB hypothesis after finding seemingly additive SOA and difficulty effects for task 2 in averaged group PRP curves.²⁴

Substantive lesson 5: "You can teach young performers new tricks" Despite the preceding concerns, we have learned a fifth substantive lesson; under at least some circumstances, most if not all participants in a group can be trained to adopt the same daring scheduling strategy.²⁵ This follows because the lockout points for task 2 and the unlocking events for task 1 in our AEC models (see figure 2.3) are presumably flexible and open to change through appropriate practice or instructional manipulations. These models therefore predict that even participants who initially prefer cautious scheduling strategies when task 1 is difficult may come eventually to prefer daring task scheduling.

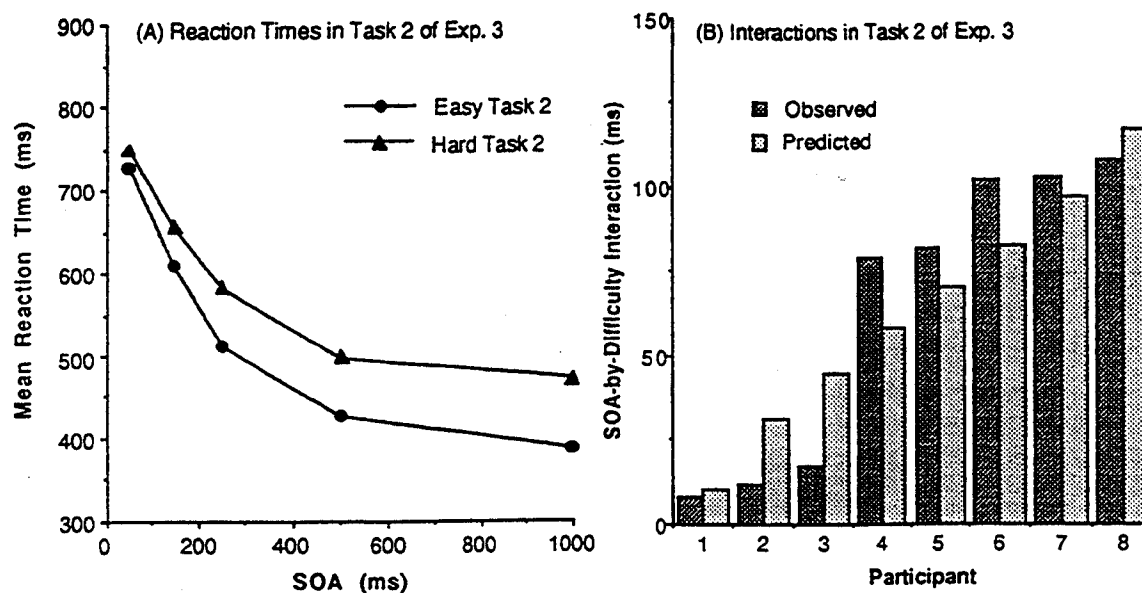


Figure 2.7 Results from using a standard PRP procedure in Lauber et al. 1994, experiment 3, with a hard auditory-manual task 1 and easy or hard visual-manual task 2 after participants had three days of initial variable-priority training. A. Mean RTs for task 2 as function of SOA and response-selection difficulty in task 2 averaged over eight participants. B. Magnitudes of interaction between the effects of SOA and response-selection difficulty on mean RTs in task 2 for individual participants. Dark vertical bars represent the participants' observed interactions. Light vertical bars represent the distribution of predicted interactions that should have occurred if all participants used the same daring strategy of task scheduling and the observed interactions differed only because of between-trial RT variance.

Such predictions are supported by more results from the study by Lauber et al. (1994, experiment 3). Eight new participants were given an initial three-day phase of variable-priority training, after which they were tested in a one-day assessment phase with the standard PRP procedure. The training phase followed Gopher's suggestions (1993) about how to enhance the efficiency of dual-task performance. It required that concurrent auditory-manual and visual-manual tasks be performed with equal priority and no constraints on the serial order of stimuli and responses. The relative difficulties of the auditory-manual and visual-manual tasks also varied orthogonally across trial blocks. Because of instructions given before the training phase started, participants were strongly encouraged to overlap their response-selection processes for the two tasks regardless of their difficulty. After the training phase ended, participants entered the subsequent assessment phase, which involved the same PRP procedure that had yielded the results shown in figure 2.6, where a hard auditory-manual primary task was combined with easy or hard visual-manual secondary tasks (Lauber et al. 1994, experiment 2).²⁶

Some results from this assessment appear in figure 2.7A. Here, although the same relatively hard task 1 is involved as before, the mean RTs for task 2 do not look like those previously obtained by Lauber et al. (1994, experiment 2; cf. figure 2.6A). After variable-priority training, participants who

transferred to the standard PRP procedure had average group PRP curves that embodied a strong positive interaction between the effects of SOA and response-selection difficulty in task 2.²⁷ It therefore appears that these participants often selected their responses for tasks 1 and 2 concurrently after short SOAs, as would happen through a daring scheduling strategy of the type assumed in the SRD model (see figure 2.4). Indeed, the present post-training RT pattern looks much like what Hawkins, Rodriguez, and Reicher (1979) obtained previously with the standard PRP procedure when task 1 was easier (cf. figure 2.2C).

Moreover, it appears that after variable-priority training, most if not all participants performed in about the same fashion. Figure 2.7B shows interactions between the effects of SOA and response-selection difficulty on mean task 2 RTs for each participant who received such training (Lauber et al. 1994, experiment 3). These participants' interactions, unlike previous ones (cf. figure 2.6B), are uniformly positive. The distribution of observed interactions in figure 2.7B is very similar to what should have occurred if every participant used a daring scheduling strategy through which responses for tasks 1 and 2 were selected concurrently. This outcome proves clearly that people can be trained to perform concurrent tasks without an immutable structural response-selection bottleneck.²⁸ These findings reinforce those obtained by Gopher (1993).

Substantive lesson 6: "Dual-task performers can share and share alike"

Our sixth substantive lesson follows from an even more extreme demonstration of the response-selection bottleneck's ephemeral nature. As our assumptions about the capacities of EPIC's cognitive processor predict, we have found that through judicious instruction and training, people can perform two generic choice RT tasks with essentially no interference between tasks (Schumacher et al. 1997). Such virtually perfect time-sharing occurs when five prerequisite conditions prevail in combination: (1) participants are encouraged to give the tasks equal priority; (2) participants are expected to perform each task quickly; (3) there are no constraints on temporal relations or serial order among responses; (4) different tasks use different perceptual and motor processors; and (5) participants receive enough practice to compile complete production-rule sets for performing each task. Results obtained under these conditions confirm and extend previous claims about the existence of perfect time-sharing in multiple-task performance (e.g., Allport, Antonis, and Reynolds 1972; Greenwald and Shulman 1973; Hirst et al. 1980; Koch 1993, 1994; Shaffer 1975; Wickens 1984).

Specifically, consider figure 2.8, which shows mean RTs from the fifth session of an experiment that satisfied the aforementioned conditions (Schumacher et al. 1997). In this experiment, there were an auditory-vocal RT task and a visual-manual RT task.²⁹ The experiment included three trial types: (1) dual task; (2) heterogeneous single task; and (3) homogeneous single

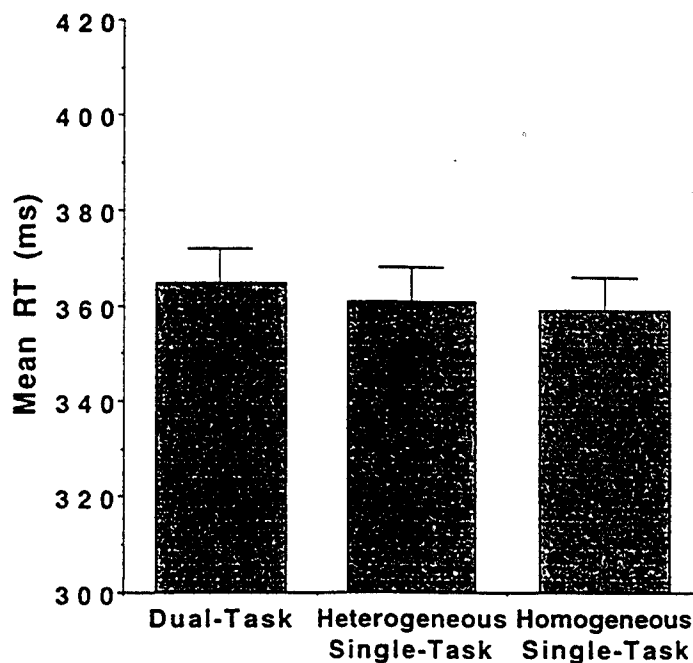


Figure 2.8 Mean RTs from three trial types in an experiment that demonstrates virtually perfect time-sharing under certain prerequisite conditions (Schumacher et al., 1997).

task. On *dual-task* trials, participants performed both tasks simultaneously, and the stimuli for them had a zero SOA. On *heterogeneous single-task* trials, only one stimulus (either auditory or visual) was presented, and only one task was performed. However, the heterogeneous single-task and dual-task trials were interleaved randomly within trial blocks, so before each of these trials, participants were uncertain about which tasks would come next. In contrast, the *homogeneous single-task* trials were arranged such that the same one task had to be performed on each trial throughout a block. Some blocks of homogeneous single-task trials involved the auditory-vocal task, while others involved the visual-manual task. Although the various trial types differed substantially in their nominal demands on participants' information-processing resources, the mean RTs and error rates that resulted from them were nearly equal (see figure 2.8). Furthermore, the individual RTs for the auditory-vocal and visual-manual tasks had values that were essentially independent of each other within the dual-task trials, as perfect time-sharing would entail. This outcome casts grave doubt on the traditional RSB hypothesis but strongly affirms our theoretical claim that skilled performers can test the conditions and execute the actions for two sets of task production rules concurrently.

The results in figure 2.8 also raise another intriguing question. Why have past attempts of some other experimenters (e.g., Pashler 1994b; Ruthruff, Pashler, and Klassen 1995) who tried to "uncork" the putative response-selection bottleneck been unsuccessful? Perhaps the answer is that their experiments did not satisfy one or more of the prerequisite conditions for perfect time-sharing. This seems likely because Pashler's experiment (1994b)

required participants to perform two manual tasks concurrently, and Ruthruff, Pashler, and Klaasen's experiment (1995) required participants to produce their vocal and manual responses in grouped fashion, which imposed strong constraints on the temporal relations among responses. Thus we would expect that these experiments might produce seductively misleading evidence of central response-selection or peripheral perceptual-motor bottlenecks. The repeated fulfillment of such expectations reinforces our prior methodological lesson 6: "Be careful what you ask for; you might actually get it."³⁰

Methodological lesson 8: "Seek and you shall find; knock and it shall be opened unto you" Equally relevant at this juncture is an eighth methodological lesson; along the path to a complete veridical UTC, interesting new predictions may be derived theoretically and confirmed empirically. For example, our discoveries of systematic individual differences in dual-task performance under the PRP procedure (substantive lesson 3), beneficial effects of training on participants' task-scheduling strategies (substantive lesson 5), and virtually perfect time-sharing (substantive lesson 6) all were stimulated by predictions based on the present AEC models and their EPIC architecture. One might therefore be optimistic that further steps toward theoretical unification will yield more successful predictions. Indeed, such optimism is justified by results from subsequent applications of EPIC to modeling multiple-task performance in other domains beyond the PRP procedure.

2.5 APPLICATION TO HUMAN-COMPUTER INTERACTION

Beyond the PRP procedure, one practical domain to which we have applied the EPIC architecture for modeling multiple-task performance is human-computer interaction (Kieras and Meyer forthcoming; Kieras, Wood, and Meyer 1995, 1997). Financial support of this application has come from NYNEX, a regional telephone company in the northeastern United States; our research on HCI concerns performance by telephone operators who provide on-line service to customers making phone calls.

EPIC computational models are especially relevant in this regard. The operators' performance requires processing auditory and visual information from headphones and computer monitors, respectively. On the basis of such information, operators must make decisions about incoming calls while speaking to customers and typing on a keyboard. The interactions between operators and customers are controlled by computer workstations. Dealing with typical calls through a workstation involves several tasks that exercise a variety of perceptual-motor and cognitive resources. The speed and accuracy of an operator in completing each call depends both on the capacities of these resources and on the workstation's design. By applying EPIC computational models to describe and predict this performance, we may help

improve the design of workstations, the selection of operators, and the regimens of training that they receive.³¹

The prospects for such help are good. Empirical studies of operators' performance have revealed patterns of overt movements and response latencies consistent with EPIC's assumptions. From these patterns, it appears that considerable temporal overlap occurs among concurrent perceptual, cognitive, and motor processes in skilled telephone operators (Gray, John, and Atwood 1993; John 1990). Thus other substantive and methodological lessons may emerge through modeling these processes in greater detail (see tables 2.1 and 2.2).

The NYNEX Study

An initial example of our research in this domain focuses on performance by toll assistance operators (TAOs), who serve customers for telephone calls charged to third-party billing numbers.

Procedure During such calls, NYNEX investigators have made audiovisual tapes of representative on-line exchanges between skilled TAOs and customers. These tapes contain records of equipment signals, customers' vocal inputs, TAOs' vocal outputs, and sequences of manual keypresses produced by TAOs while they were using standard computer workstations. From the tapes, the identities and latencies of relevant environmental and behavioral events can be transcribed for data analysis and theoretical interpretation.³²

Each of the taped calls required a TAO to perform several steps: (1) detect the onset of a tone over a pair of headphones, which signaled that a call from a prospective customer had arrived at the workstation; (2) look at a computer monitor for alphanumeric information that specified the call's category; (3) greet the customer making the call; (4) get the billing number to be charged for the call; (5) enter this number and other relevant information into the workstation by typing a series of keystrokes on the computer's keyboard; (6) look at the monitor to confirm whether the typing had been done correctly; (7) complete the connection for the call by pressing a call initiation key; and (8) bid the customer good-bye. The steps performed by the TAOs were analogous to ones that participants perform repeatedly during PRP and serial choice-reaction procedures. It therefore might be expected that in at least some respects, the TAOs' response latencies would resemble those found under laboratory conditions.

Results For example, figure 2.9 shows results from one representative exchange between a TAO and customer. Here observed response latencies of the TAO's manual keystrokes (large filled circles on solid curve) are plotted against their serial positions throughout the exchange. Several features of these data, which typify third-party billing calls, should be noticed. At the start of the keystroke sequence (first serial position), the observed response latencies rise to a maximum and then gradually decrease as the serial positions

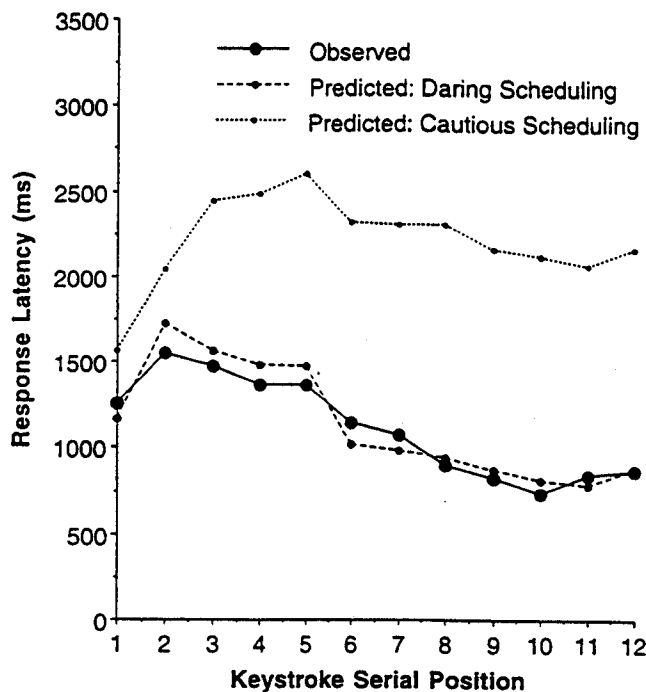


Figure 2.9 Response latencies as a function of keystroke serial position for a sequence of keystrokes typed by a TAO during a representative exchange with a customer. Large filled circles on the solid curve denote observed response latencies. Small filled circles on the dashed curve denote simulated latencies from an EPIC computational model that used a daring scheduling strategy. The upper dotted curve denotes simulated latencies from another model that had an artificial response-selection bottleneck and used a cautious scheduling strategy.

of the keystrokes increase. Consequently, there is a downward latency trend that looks much like the RT curves found during past studies with the PRP procedure (e.g., figure 2.2, mean RTs for task 2), suggesting some postponement of pending processes in order for current processes to be completed. This postponement presumably was necessary so that the TAO's keystrokes occurred in correct serial order, just as instructions for the PRP procedure require that responses for task 2 occur after those for task 1.

The pattern of observed response latencies in figure 2.9 raises intriguing theoretical questions. What type of task-scheduling strategy is used by TAOs to coordinate their performance? Are their strategies consistent with our previous findings for laboratory studies that involved the PRP procedure? How might the modeling of TAOs' performance promote the search for complete veridical UTCs? Can workstation designs be improved on the basis of insights from such modeling?

EPIC Computational Models of TAOs' Performance

Some progress toward answering these questions can be made by formulating EPIC computational models that account more or less satisfactorily for data like those shown in figure 2.9.

Model with cautious scheduling strategy Among the specific models we have tested thus far is one that uses a cautious scheduling strategy (CSS). According to this CSS model, which has an artificial response-selection bottleneck and incorporates the traditional RSB hypothesis, there is no temporal overlap in the response-selection processes whereby TAOs choose their successive overt actions. Nor does the CSS model permit temporal overlap between the overt execution of current eye or hand movements and the covert preparation of subsequent eye or hand movements. Rather, the CSS model assumes that serially ordered movements are mediated by strictly sequential cognitive and motor processes. As a result, the simulated response latencies produced by the CSS model fit the observed response latencies poorly.

For example, the upper dotted curve in figure 2.9 comes from the CSS model, whose simulated response latencies greatly exceed the TAOs' actual response latencies ($R^2 = 0.04$; RMSE = 1,150 msec), especially for key-strokes in early serial positions. The CSS model fails because, unlike real TAOs, it does not temporally overlap response-selection or movement-production processes for successive ocular, manual, and vocal actions. This failure happened although, in other respects, we formulated the CSS model to approximate the TAOs' actual response latencies as best possible despite its cognitive response-selection bottleneck and cautious scheduling strategy.

Model with daring scheduling strategy Another model we have tested in this context is one that uses a daring scheduling strategy (DSS). According to this DSS model, which generalizes our previous SRD model from the PRP procedure (see figure 2.4), the TAOs' successive overt actions are mediated by response-selection and movement-production processes that have substantial temporal overlap. As a result, the simulated response latencies that are produced by the DSS model fit the observed response latencies fairly well.

For example, the lower dashed curve in figure 2.9 comes from the DSS model, whose simulated response latencies approximate the TAOs' actual response latencies closely at almost every serial position ($R^2 = 0.93$; RMSE = 95 msec).³³ Of course, the success of the DSS model would be expected from our previous findings for the PRP procedure, where temporally overlapping stimulus-identification, response-selection, and movement-production processes were demonstrated through the SRD model's goodness of fit (e.g., figure 2.2). In the present context, such temporal overlap may occur again because TAOs have substantial job experience and want to complete their telephone calls quickly so that customers are satisfied and business costs stay low.

Implications for interface design It is beyond the scope of this chapter to discuss at length how designers can improve current computer workstations on the basis of results like those shown in figure 2.9. Nevertheless, a

few general comments that bear on workstation design may be offered here (see also Gray, John, and Atwood 1993; Kieras, Wood, and Meyer 1997).

From figure 2.9, it appears that TAOs can work on multiple tasks concurrently at procedural cognitive and perceptual-motor levels. Workstation interfaces therefore should be designed to exploit this capacity for daring task scheduling. For example, such scheduling may be facilitated by designing interfaces that interleave stimuli and responses in visual-manual tasks with those in auditory-vocal tasks. Also beneficial may be displays that present visual stimuli for successive tasks at predictable adjacent spatial locations. Our findings show these heuristic design principles to be prerequisites for concurrent performance of multiple tasks.

Another related point concerns the prospective role of computational models in evaluating workstation interfaces. Our results (e.g., figure 2.9) show that models with a systematic functional architecture for the human information-processing system can precisely predict major facets of interface users' performance (e.g., response speed and accuracy). In the future, as such models evolve further, a priori predictions may be made with respect to performance achieved both through different interfaces and by different personal styles of operation. Thus interface and personnel testing could become both automated and theoretically motivated in ways that save much time and effort compared to current engineering practices.

Lessons from Modeling of TAOs' Performance

Given our theoretical research on the performance of TAOs, several more substantive and methodological lessons may be deduced (see tables 2.1 and 2.2).

Substantive lesson 7: "TAOs know the way" As demonstrated by our previous research with the PRP procedure, the way to perform multiple tasks efficiently is by daring task scheduling. Such scheduling can occur because there is no structural cognitive bottleneck through which individual responses must be selected sequentially. On the contrary, the human cognitive processor—like EPIC's—has the capacity to select responses for multiple tasks concurrently. Nor is there a single bottleneck mechanism through which every selected response must be produced overtly. Rather, each response modality has its own motor processor; while the physical execution of various movements is under way, these motor processors may prepare the movement features to anticipate subsequent overt responses. The success of the present DSS model in characterizing performance during HCI shows that TAOs also know the way to use such capabilities.

Substantive lesson 8: "What goes around comes around" Complementing substantive lesson 7 is an eighth, correlative one. Through our modeling of multiple-task performance under both simple laboratory and

complex real-world conditions, it has become apparent that theoretical conclusions reached in one context often may be generalized to other seemingly quite different contexts. For example, conclusions reached for the PRP procedure also apply to realistic human-computer interaction, and vice versa. Such generality should be especially reassuring to seekers of a complete veridical UTC.

Substantive lesson 9: "Hand movements obey the Boy Scout motto"

The motto of the Boy Scouts is "Be prepared." Our modeling of TAOs' response latencies (see figure 2.9) suggests that consistent with this motto, TAOs are prepared to anticipate their impending hand and eye movements while current overt responses are under way. It is this anticipatory preparation that contributes in part to the daring scheduling strategies with which TAOs achieve efficient multiple-task performance (Kieras, Wood, and Meyer 1997). For example, such preparation helps increase the rates of manual typing and visual search (cf. Rumelhart and Norman 1982).

Substantive lesson 10: "Practice makes (nearly) perfect"

Consistent with our prior findings about practice effects in the PRP procedure (see figure 2.7), the efficient performance of TAOs may stem from their extensive experience at fielding various types of telephone calls under conditions that encourage daring task scheduling. Given this experience, adaptive executive processes can be evolved to enable substantial flexibility and temporal overlap among response selection, movement production, and other operations for concurrent tasks (cf. Gopher 1993). Persistent dual-task interference, which sometimes lasts for thousands of trials in the standard PRP procedure, and which has misled some advocates of the traditional RSB hypothesis (e.g., Gottsdanker and Stelmach 1971; Pashler 1993), is merely a deceptive stereotypy due to the constant differential task priorities imposed by the PRP procedure's instructions.

Methodological lesson 9: "Unification enables application"

For researchers and practitioners still uncertain about what approach ultimately may lead to the greatest practical payoffs in experimental psychology, cognitive science, and human-factors engineering, the extensions of our EPIC computational models from the PRP procedure to the HCI tasks of TAOs offer a ninth methodological lesson. Efforts toward theoretical unification and the development of UTCs can facilitate the transfer of data and hypotheses from the laboratory to applied settings. Unification enables application. Of course, Newell (1990) anticipated this in his previous endorsements for UTCs, and he was correct, unlike some other less optimistic scientific leaders in the field.

Methodological lesson 10: "Keep your sunny side up" Unfortunately, past pessimism has been rather pervasive about applying laboratory data and

hypotheses to complex realistic situations. For example, consider what the founder of the Attention and Performance symposium series had to say with respect to these matters:

Generalizations from the simple to the complex are sometimes straightforwardly wrong. As an example, one cannot build a predictive model about, say, how people type or play the piano on the basis of the results from research on responses to signals that are presented in rapid succession (the psychological refractory period; Welford 1967)... the "small" paradigm ... is merely concerned with artificially created situations, irrelevant to real-life performance and, hence, to system design.... the general message [therefore] is that we should aim at delineating the range of application of [micro] behavioural models rather than searching for the one ultimate [macro] model. (Sanders 1991, 1006–1008; see also Sanders 1984)

In our opinion, however, such assessments and prescriptions are too pessimistic and restrictive. On the contrary, applications of EPIC computational models jointly to the PRP procedure and HCI demonstrate clearly that data and theoretical concepts from the former "artificial" laboratory domain can transfer successfully to the latter "natural" real-world domain. Although Sanders (1991) merits high respect, he was wrong about the impossibility of this transfer and about the inadvisability of seeking a practical predictive (macromodel) UTC.

Which brings us to our tenth methodological lesson: "Keep your sunny side up." Future investigators should be optimistic about both the scientific and practical prospects for UTCs. There is considerable reason to expect that further strides toward theoretical unification of basic psychological laboratory data and conceptual hypotheses will promote applied human-factors engineering significantly. In fact, more reason for such optimism is provided by our next application of EPIC computational models to a second real-world task domain.

2.6 APPLICATION TO AIRCRAFT COCKPIT OPERATION

A second practical application of EPIC computational models deals with multiple-task performance during cockpit operations in military aircraft (Kieras and Meyer 1995, forthcoming; Meyer and Kieras 1996, 1997b). Our work there has been conducted collaboratively with scientists at the Naval Research Laboratory (NRL) in Washington, D.C. This collaboration extends previous laboratory studies of visual-manual tracking and serial choice reactions (e.g., Brickner and Gopher 1981; Gopher 1993; Gopher, Brickner, and Navon 1982; McLeod 1977; North 1977; Wickens 1976). We have modeled the performance of personnel who must do visual-manual tracking and tactical decision making concurrently under realistic cockpit conditions. As in other related contexts (e.g., HCI by the TAOs of NYNEX), such performance relies on executive cognitive processes to coordinate ocular and manual motor processes with visual and auditory perceptual processes. By

applying EPIC computational models to characterize these processes, we may help improve the design of cockpit control panels, the selection of personnel, and the regimens of training that they receive. The resulting benefits could be relevant to performance in other real-world situations as well (e.g., concurrent automobile driving and cellular telephoning; Gugerty 1997).

The Naval Research Laboratory Study

The empirical data on which we focus for now were collected at the NRL by Ballas, Heitmeyer, and Perez (1992).

Procedure During the Ballas, Heitmeyer, and Perez (1992) study, participants—including some trained pilots—viewed and responded to a computerized visual display similar to those used in military aircraft cockpits. A diagram of this display, which provided information for performing two tasks concurrently, appears in figure 2.10.

On the right side of the display was a window for a visual-manual tracking task. In this window were a cursor and a target (iconic airplane) that moved evasively through space. When participants performed the tracking task, they had to keep the cursor on target by moving a right-hand joystick that controlled the cursor's spatial position. The tracking error (distance between cursor and target) was measured as a function of the tracking task's difficulty.

Meanwhile, on the left side of the display was a window for a tactical decision task. In this window were iconic blips that appeared sequentially

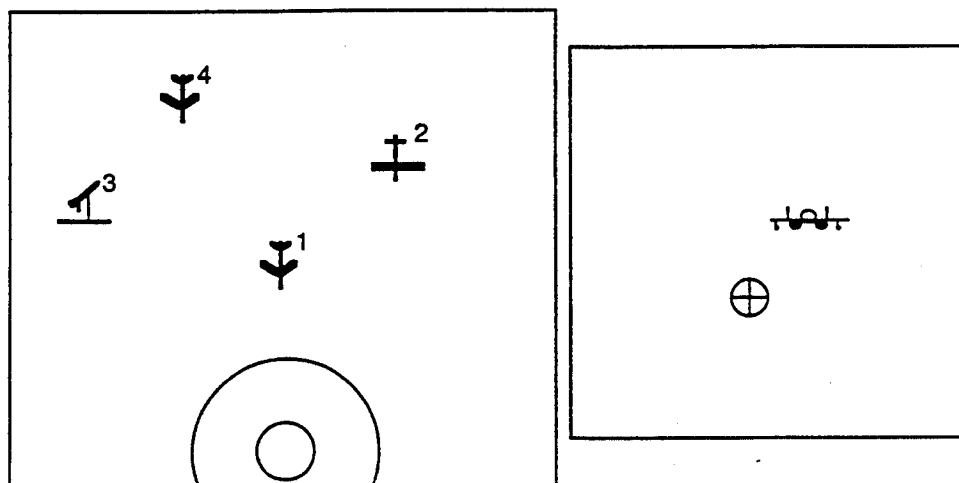


Figure 2.10 Diagram of the visual display used in the Ballas, Heitmeyer, and Perez (1992) study of realistic dual-task performance during aircraft cockpit operation. On the left is the window for a tactical decision task, where the iconic blips represent potentially dangerous objects that had to be classified as "hostile" or "neutral." On the right is the window for a visual-manual tracking task, where the crosshairs of a cursor had to be superimposed on a moving target plane by manipulating a right-hand joystick.

at unpredictable times and locations, depicting potentially hazardous objects (e.g., jet fighters, bombers, and missile sites) whose locations changed gradually over time. When participants performed the tactical decision task, they looked at these blips one after another and indicated which were "hostile" and which were "neutral" by typing on a keyboard with their left hand. Response latencies of the keypresses were measured individually for the blips as a function of their serial positions in the sequence of tactical decisions.³⁴

Performance of the tactical decision and visual-manual tracking tasks occurred during two types of epoch: (1) single task; and (2) dual task. During single-task epochs, participants had to perform a relatively hard version of the tracking task for 2 minutes while disregarding the tactical decision task. During dual-task epochs, the tactical decision task and a relatively easy version of the tracking task both had to be performed concurrently for 2 minutes, with emphasis on making the tactical decisions quickly. Participants alternated back and forth between successive single-task and dual-task epochs. Each transition between epochs was signaled by a brief auditory tone. The sequence of alternating epochs included six single-task and six dual-task epochs per participant.

Through measurements of participants' performance after each transition from a single-task to dual-task epoch, Ballas, Heitmeyer, and Perez (1992) studied the effects of adaptive automation on aircraft cockpit operations. By definition, *adaptive automation* involves two principal elements. First, a computer system takes temporary responsibility for performing one or more tasks (e.g., tactical decision making), allowing a human operator to concentrate on another task (e.g., visual-manual tracking) that has become especially difficult at the time. Second, when the task on which the operator has been concentrating becomes easier subsequently, the operator continues performing it and, in addition, resumes performing the one or more tasks for which the computer took temporary responsibility. Human-factors engineers hypothesize that adaptive automation can ease the operator's mental workload during stressful time periods while maximizing overall performance across a variety of environmental conditions.

However, adaptive automation also can have short-term detrimental effects. For a while after the transition from an epoch of single-task performance to an epoch of dual-task performance, responses to stimuli in the recently resumed task (e.g., tactical decision making) may be relatively slow or inaccurate, manifesting an *automation deficit* until the operator "gets back into the swing of things." The sources of automation deficits are not well understood; they perhaps involve temporary losses of situation awareness (Garner 1997; Gugerty 1997), transient PRP effects of cautious task scheduling, disruptive phasic changes in operators' level of arousal, or a combination of these. Given the design of the Ballas, Heitmeyer, and Perez (1992) study, we can examine the contributions that some of these sources make. Specifically, by applying EPIC computational models to account for data from this study, it is possible to test whether an automation deficit stems partly or

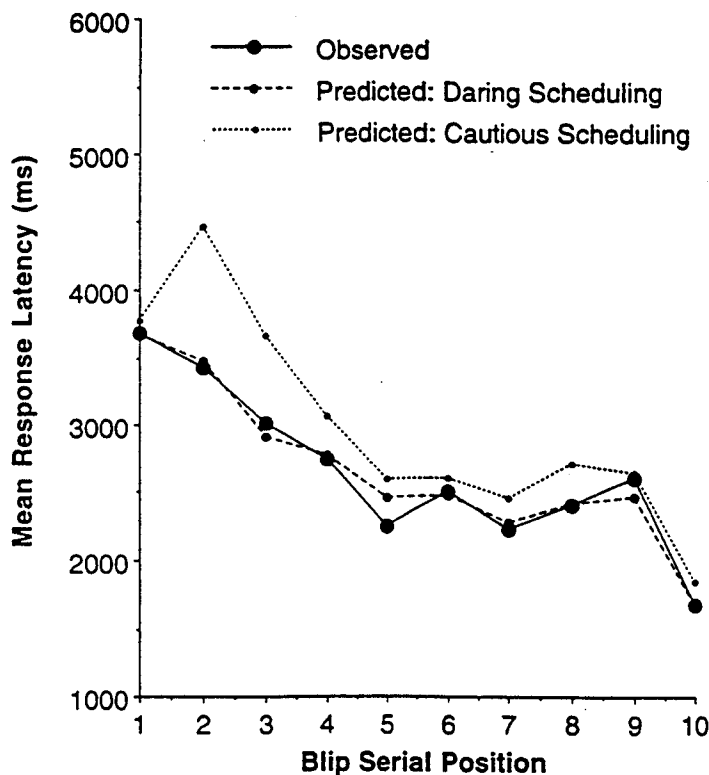


Figure 2.11 Mean response latencies as a function of blip serial position for the tactical-decision task in Ballas, Heitmeyer, and Perez 1992. Large filled circles on the solid curve denote observed latencies. Smaller filled circles on the lower dashed curve denote simulated latencies from an EPIC computational model that used a daring scheduling strategy. The upper dotted curve denotes simulated latencies from another model that had an artificial response-selection bottleneck and used a cautious scheduling strategy. All three latency curves come from a sequence of tactical decisions made immediately after a period during which only the visual-manual tracking task had been performed.

wholly from cautious scheduling strategies in the aftermath of transitions between single-task and dual-task epochs.

Results Some empirical results for making such tests appear in figure 2.11.³⁵ Here observed mean response latencies for the tactical decision task are shown versus the serial positions in which iconic blips were classified following the transitions from single-task to dual-task epochs. The observed latencies tended to be longer at the start of the blip sequence, manifesting an initial automation deficit, after which they decreased as the blip serial position increased. This latency decrease over serial positions formed a downward curve that embodied a PRP effect, reminiscent of mean RTs for task 2 in the standard PRP procedure (cf. figure 2.2).

EPIC Computational Models of Aircraft Cockpit Operation

With respect to the observed response latencies in figure 2.11, we have tested two alternative EPIC computational models: one with a cautious

scheduling strategy (CSS), and one with a daring scheduling strategy (DSS). Task scheduling in each model was constrained because a large visual angle (on the order of 20 degrees) separated the centers of the display windows for the tactical decision and visual-manual tracking tasks (see figure 2.10). This separation precluded participants from performing more than one task at any particular moment during the dual-task epochs because their eyes had to fixate on one window or the other to acquire the detailed visual information needed to perform a particular task. Thus, under both the CSS and DSS models, our simulations during dual-task epochs alternated back and forth between performing the tactical decision task and performing the visual-manual tracking task, mediated by top-level executive processes that permitted the gaze of EPIC's eyes to be shifted between the left and right display windows.

However, the two models differed considerably in their assumptions about the processes that performed the tactical decision task while it was under way. Under the CSS model, there was an intermediate-level executive process that used a cautious scheduling strategy with little temporal overlap among the processes for making tactical decisions about the individual iconic blips in this task's display window. By contrast, under the DSS model, the intermediate-level executive process for the tactical decision task used a daring scheduling strategy, whereby more temporal overlap occurred among the processes for making tactical decisions about individual blips.

Cautious scheduling strategy model Specifically, we programmed the CSS model with an artificial bottleneck that precluded manual keypress responses for multiple iconic blips from being selected concurrently while performance of the tactical decision task was under way. Also, under the CSS model, processes that chose the next blip to be classified and that shifted EPIC's eyes to fixate on it for stimulus identification could not begin until after manual responses for the currently attended blip had been selected and produced. As a result, the upper dotted curve of simulated mean response latencies in figure 2.11 was produced for the tactical decision task during dual-task epochs.

The CSS model's simulated latencies substantially exceeded the observed latencies ($R^2 = 0.869$; RMSE = 439 msec), especially for blips in early serial positions. This overshoot happened even though subject to ancillary architectural constraints, the CSS model was programmed to make the best possible approximation of the observed latencies. Thus our simulations suggest that there is no plausible way in which real participants could have performed the tactical decision task as well as they did if their performance had been limited by a structural response-selection bottleneck or a cautious scheduling strategy.

Daring scheduling strategy model In contrast, the DSS model and its daring scheduling strategy were programmed such that no response-selection

bottlenecks precluded manual keypress responses for multiple iconic blips from being selected concurrently while performance of the tactical decision task was under way. Also, under the DSS model, processes that chose the next blip to be classified and that shifted EPIC's eyes to fixate on it for stimulus identification began before movement production for the previously attended blip had finished. As a result, the lower dashed curve of simulated mean response latencies in figure 2.11 was produced for the tactical decision task. Here the simulated and observed response latencies are fairly close at almost every serial position ($R^2 = 0.975$; RMSE = 90 msec). The DSS model's goodness of fit in figure 2.11 is similar to what we obtained before for the PRP procedure (see figure 2.2) and for TAOs' performance in HCI (see figure 2.9). Furthermore, this model accounts fairly well for participants' degrees of error in the visual-manual tracking task during both single-task and dual-task epochs. It therefore appears that EPIC may be applied for characterizing and predicting multiple-task performance across various realistic task domains.

Implications for design of cockpit control panels Results from the present application of our EPIC computational models suggest that automation deficits in aircraft cockpit operations are not caused simply by cautious task scheduling after transitions from single-task to dual-task epochs. Automation deficits may instead occur because the visual displays for suspended tasks do not sufficiently sustain situation awareness during single-task epochs; operators may need extra time at the start of dual-task epochs to create new mental representations of the current environment for formerly suspended tasks.

If so, our results have some obvious implications for designing cockpit control panels:

- Distances between the display windows of current and temporarily suspended tasks should be minimized, while providing adequate discriminability among visual stimuli, as in advanced "head up" displays;
- Supplementary perceptual aids that help sustain situation awareness for suspended tasks should be provided;
- At the start of dual-task epochs, special visual cues should be available to orient operators' attention automatically toward the most important stimuli for resuming formerly suspended tasks.³⁶

Perhaps implementing such guidelines in combination will enhance the utility of semiautomated aircraft cockpits.

Lessons from Modeling Aircraft Cockpit Operation

More generally, theoretical and practical benefits may also accrue from other substantive and methodological lessons that modeling of aircraft cockpit operation has taught us.

Substantive Lesson 11: "Daring task scheduling is 'the right stuff'"
Our eleventh substantive lesson concerns the nature of "the right stuff" (Wolfe 1979). According to aviation lore, "the right stuff" is what test pilots, "top gun" aviators, and astronauts must have to survive hazardous flight operations. The success of the present DSS model suggests that at least part of "the right stuff" involves an ability to use daring scheduling strategies for concurrent performance of perceptual-motor and cognitive tasks under duress. Without such task scheduling, test pilots, top guns, and astronauts could not survive the crises they encounter, because it would be impossible for them to complete all of their crucial tasks before fatal deadlines have passed. That pilots like the legendary Chuck Yeager have survived to a ripe old age weighs heavily against the traditional RSB hypothesis (cf. Yeager and Janos 1985).

Substantive lesson 12: "The eyes have it" Due to the complex tasks performed in rapid aircraft cockpit operations, the daring scheduling strategies used for them presumably involve temporal overlap among a panoply of mental and physical processes, including ocular response selection, saccadic eye-movement execution, visual stimulus encoding, manual response selection, and finger keypressing. Specifically, our successful DSS model implies that ocular response selection and saccadic eye-movement execution must overlap these other processes in order to account for observed response latencies; otherwise, simulated latencies would far exceed observed latencies. This implication leads directly to our twelfth substantive lesson: "The eyes have it," where "it" is "the right stuff" associated with daring task scheduling. Contrary to claims by some advocates of the traditional RSB hypothesis (e.g., Pashler, Carrier, and Hoffman 1993), we have found no evidence that a structural response-selection bottleneck constrains ocular response selection for voluntary saccadic eye movements.

Methodological lesson 11: "If you've seen one, you've seen 'em all"
Given the similarity between what we have found in modeling aircraft cockpit operation and other related cases of multiple-task performance (e.g., HCI and PRP procedure), an eleventh methodological lesson now becomes apparent. Having seen how multiple-task performance is achieved under one set of circumstances may provide deep insights into how it is achieved under many other circumstances. Methodological lesson 11 therefore constitutes a welcome generalization of our substantive lesson 8 ("What goes around, comes around") and substantive lesson 12 ("The eyes have it"). Seekers of theoretical unification should find this modicum of generality encouraging because it is an important prerequisite to formulating complete veridical UTCs.

Substantive lesson 13: "Covert visual attention shifts are like poor Yorick" Nevertheless, there may be a few exceptions to some of the pre-

vious lessons. In particular, our modeling of aircraft cockpit operation suggests that perhaps substantive lesson 8 and methodological lesson 11 do not hold with respect to covert shifts of visual attention. By definition, covert visual attention shifts supposedly take place in the "mind's eye" (Jonides 1980, 1981); they entail mentally redirecting the focus of attention from one spatial location to another without concomitant overt eye movements. Such attention shifting has been studied extensively in artificial laboratory experiments, and substantial evidence suggests that it may influence visual information processing there (Johnston and Dark 1986; Mangun, Hillyard, and Luck 1993; Posner 1980). However, in the DSS model for concurrent tactical decision making and visual-manual tracking, no covert visual attention shifts contributed to its daring scheduling strategy. This model assumed simply that shifts of visual attention took place through overt saccadic eye movements from one relevant stimulus location to the next. Accounting for observed response latencies and tracking errors did not require covert attention shifts.

The lesson from this is that under some—perhaps even most—real-world circumstances, covert visual attention shifts are like poor Yorick, the deceased king's jester in *Hamlet*. Although experimental psychologists have known them well, such attention shifts may play little if any role for present practical purposes. Instead, perhaps what matter more are the ocular-motor processes whereby overt saccadic eye movements take place to shift the eyes' foveas and gaze between locations.³⁷

Methodological lesson 12: "Be thankful for the Second Golden Rule" Mandates from organizations such as NYNEX and the U.S. Office of Naval Research (ONR), the financial sponsors of our work, also have taught us that researchers should be thankful for the "Second Golden Rule," namely, "Those who have the gold make the rules." For us, this has meant having to focus on solving our sponsors' practical problems. As a result, our research has yielded instructive insights about major empirical phenomena and theoretical processes in both "artificial" laboratory and "natural" real-world contexts that would otherwise have gone unnoticed. We therefore concur with Newell (1990) that the press of practical applications can have substantial benefits to basic science.

2.7 CONCLUSION

Reviewing the lessons that our computational modeling of human multiple-task performance has taught us, we see that they fall into several subcategories. Among the methodological lessons (see table 2.1), some concern attitudes and intellectual orientations that theorists should adopt while seeking a practical unified theory of cognition and action (e.g., methodological lessons 1, 2, 3, 5, 10, and 12). Others are warnings about perilous pitfalls along the way to a complete veridical UTC (e.g., methodological lessons 6

and 7). In compensation for heeding such warnings and maintaining proper attitude, additional methodological lessons offer promissory notes about major benefits that may accrue (e.g., methodological lessons 4, 8, 9, and 11). We hope that together, these lessons will both encourage and guide researchers toward successful unification in human-performance theory.

Also relevant to theoretical unification are the concomitant substantive lessons (see table 2.2). Among them, some describe basic characteristics of the human information-processing system that have become salient through our work and that must be incorporated in the functional architecture of a complete veridical UTC (e.g., substantive lessons 1, 2, 6, and 12). Others summarize important facts about people's preferred task-scheduling strategies (e.g., substantive lessons 3, 4, 7, 8, 9, 11, and 13). Given such facts, it may be possible to improve multiple-task performance considerably with systematic training (substantive lessons 5 and 10). The combined thrust of these substantive lessons reiterates some of Newell's original points (1973a, 1990): practical predictive theories must take into account "hardware properties" of human information processing, environmental task demands, personal goals based on these demands, and preferred strategies for goal attainment. By doing so, future computational models of multiple-task performance may realize the benefits anticipated in our methodological lessons.

Meanwhile, important questions remain with respect to the present project. For example, how close do our EPIC architecture and computational models currently come to satisfying established criteria for what a successful UTC should be? In which directions should future research go to improve EPIC and to make it a more complete veridical theory? Available space does not permit extensive answers here, but we may offer a few partial ones.

Evaluation of EPIC

Admittedly, EPIC lacks features essential to a complete veridical UTC; these omissions are apparent from evaluating EPIC with respect to requisite criteria proposed by Newell (1990, 1992; cf. Seifert and Shafto 1994). According to Newell, a complete veridical UTC must ultimately yield accurate predictive computational models for all of the following functions: (1) comprehending and producing natural language; (2) storing and using large amounts of knowledge; (3) learning from experience; (4) dealing with emotion and motivation; (5) solving problems; (6) behaving creatively; and (7) interacting socially. If so, EPIC and the models associated with it have a long way to go—they satisfy none of these criteria yet.

They do, however, satisfy some of Newell's other requisite criteria (1990, 1992), namely: (8) having a precise and stable architecture; (9) embodying detailed perceptual and motor mechanisms; (10) producing rational, adaptive goal-directed behavior; (11) finishing tasks in real time; and (12) being potentially realizable as a neural system. For example, among EPIC's note-

worthy features are its motor processors for the manual and ocular response modalities, and its enablement of efficient executive cognitive processes for task scheduling. Moreover, many of EPIC's basic assumptions (e.g., those about distributed asynchronous parallel information processing) are consistent with human brain anatomy and neurophysiology (Kieras and Meyer forthcoming; Meyer and Kieras 1997b; cf. Newell 1990). Granted these virtues, we believe EPIC merits future research.

Directions for Future Research

The preceding evaluation suggests several attractive directions for future research with EPIC and new computational models based on it.

Specification of subdivisions in working memory In particular, our future research may specify more precisely the capacities, durations, and interconnections for perceptual and motor subdivisions of EPIC's working memory (Kieras et al., forthcoming). Such specifications can build on previous studies of how working memory contributes to various cognitive processes that underlie general intelligence (e.g., Baddeley 1986; Carpenter and Just 1989; Carpenter, Just, and Shell 1990; Chase and Ericsson 1983; Daneman and Carpenter 1980; Gilhooly et al. 1993; Kyllonen and Christal 1990). This advance will help EPIC support computational models of reasoning, problem solving, and language processes, and will also enable us to formulate more accurate measures of mental workload, whose current practical indicators leave much to be desired (cf. Gopher and Donchin 1986; Moray 1979; O'Donnel and Eggemeier 1986; Wierwille and Connor 1983; Willeges and Wierwille 1979).

Elaboration of declarative long-term memory Another productive extension will involve elaborating EPIC's long-term memory for declarative knowledge. Thus far, EPIC computational models have used only procedural and declarative knowledge in working memory. However, EPIC's long-term memory could store large amounts of organized propositional declarative knowledge as well. Realizing this potential would move EPIC further toward supporting detailed models of language processing; it also would help set the stage for a treatment of skill acquisition, which entails compiling procedural knowledge on the basis of task instructions stored as propositional declarative knowledge in long-term memory (Anderson 1982; Bovair and Kieras 1991; Bovair, Kieras, and Polson 1990; Fitts 1964).

Treatment of skill acquisition We anticipate treating skill acquisition thoroughly in the context of EPIC. If previous conceptions about perceptual-motor and cognitive skill are correct, then people may pass through several distinct acquisition phases as they become expert performers. For example,

Anderson (1982) has proposed an initial declarative stage of skill acquisition followed by several later procedural substages (cf. Fitts 1964). During the declarative stage, performance is presumably mediated by propositional knowledge about how a task should be performed. Using such knowledge apparently requires slow controlled verbal interpretive processes that lead indirectly to overt action. Nevertheless, through practice, propositional knowledge about proper task performance can be converted to executable procedures whereby tasks are performed directly with sets of production rules. The compilation and gradual refinement of these rules characterize successive substages of procedural learning. Because EPIC has both a long-term memory for declarative knowledge and a production-system formalism for procedural knowledge, it provides natural bases with which to characterize various stages of skill acquisition. Learning algorithms such as those proposed by Anderson (1982) and others (e.g., Bovair and Kieras 1991; Bovair, Kieras, and Polson 1990) may be used by EPIC's cognitive processor, enabling the compilation and refinement of production rules for performing multiple as well as single tasks.

As part of this attractive prospect, an important objective will entail describing and predicting how flexible strategies of task scheduling are acquired and incorporated into evolving executive processes. That such acquisition occurs and that it markedly influences eventual performance levels has been demonstrated already (e.g., Gopher 1993; Lauber et al. 1994; Meyer et al. 1995; Schumacher et al. 1996, forthcoming). We know specifically that the efficacy and rate of learning depend on what types of intermediate training are provided. Thus an important next step will involve specifying the learning algorithms through which various training protocols promote both optimized temporal overlap among task processes and efficient allocation of limited perceptual-motor resources.

Incorporation of energetic mechanisms To make EPIC computational models more realistic, we also eventually must endow them with "energetic" mechanisms. Drugs, sleep deprivation, emotional arousal, and other psychophysiological variables can significantly affect rapid human performance (see, for example, Beatty 1982; Broadbent 1971; Frowein 1981). Some ideas that are relevant to how these influences would fit within EPIC have been suggested by Gopher (1986), Sanders (1983), and others (Hockey, Gaillard, and Coles 1986). Different factors related to mental and physical "energetics" might selectively modulate the estimated magnitudes of EPIC parameters such as the perceptual processors' stimulus detection and identification times, the cognitive processor's cycle duration, the working memory's information decay times, and the motor processors' movement-production times. It will be exciting to test whether "energetic" effects are interpretable and predictable from the same perspective that has let us account for multiple-task performance in the PRP procedure, HCI, and aircraft cockpit operation.

Final Lessons

With these future research directions before us, we leave behind two final lessons that bear on the search for a complete veridical UTC.

Methodological lesson 13: "Psychological science can (and will) be fun" As the intellectual approach taken by the Nobel laureate physicist, Richard Feynman, has amply demonstrated, striving toward theoretical unification need not entail "all work and no play" (Feynman 1985; Seifert et al. 1994). Instead, psychological science can (and will) be fun, because prospective UTCs inspire curiosity and excitement while leading researchers through surprising twists and turns along the path of discovery. If the lessons we have outlined in this chapter (see tables 2.1 and 2.2) convey some of these benefits, then our principal objective will have been accomplished.

Methodological lesson 14: "There's nothing more practical than a good theory" This is the keynote on which we began the present chapter and on which we end for now.

NOTES

Funding for preparation of this chapter was provided through grant N00014-92-J-1173 from the U.S. Office of Naval Research (ONR) to the University of Michigan, David E. Kieras and David E. Meyer, principal investigators. The authors thank Terry Allard, Susan Chipman, Harold Hawkins, and Willard Vaughan of the ONR for their continuing encouragement and support. Additional funding was provided by the research laboratory of NYNEX. Research assistance and scientific contributions by past and present members of the Brain, Cognition, and Action Laboratory at the University of Michigan—Dana Apfelblat, David Fencsik, Jennifer Glass, Leon Gmeindl, Cerita Jones, Ryan Kettler, Erick Lauber, Shane Mueller, Eric Schumacher, Molly Schweppe, Travis Seymour, B. J. Woodside, and Eileen Zurbriggen—are gratefully acknowledged. Comments by Eric Schumacher on an initial draft of this chapter are also appreciated.

1. The original form of this proclamation is attributable to Lewin (1951).
2. One impressive illustration of the abundant quantitative regularities now available from experimental psychology has been provided by Salthouse (1986), who tabulated twenty-nine systematic phenomena discovered about transcription typing. Presumably, these stem from interactive properties of skilled typists' perception, attention, memory, and motor control processes. Yet the exact nature of these interactions and the processes that mediate them remains to be elucidated (cf. John 1988; Rumelhart and Norman 1982). If there are on the order of thirty known regularities relevant to transcription typing alone, then it is easy to imagine that research on human performance already has found 3,000 or more regularities overall.
3. Newell's teachings (1973a, 1990, 1992) have many interesting parallels with those of the Buddha. For example, according to the Buddha, the way to escape perpetual mental suffering is to contemplate and unify various basic inner psychological and outer environmental realms (Bukkyo Dendo Kyokai 1985). Similarly, according to Newell, release from intellectual turmoil in experimental psychology and cognitive science will come through deep theoretical contemplation and unification.

4. Applied cognitive science must find a way to sail successfully between what Sanders (1991) called the Scylla of experimental psychology and the Charybdis of human-factors engineering. In his view, experimental psychology has formed a massive Scylla of relatively simple laboratory phenomena that are easy to measure but that also are separate from the "richness of reality," whereas human-factors engineering finds itself caught in the treacherous Charybdis of complex real-life phenomena that are difficult to analyze (Sanders 1991, 1997). The development of UTCs can help us chart a course through the gap that separates these two perils.

5. For example, consider circumstances that involve Fitts' law (1954), under which the mean duration of rapid aimed movements to a target region is a logarithmic function of the target distance divided by the target width. In numerous applied settings where aimed movements must be made, Fitts' law prevails (Meyer et al. 1988). This prevalence suggests that Fitts' law is a fundamental regularity whose underlying mechanisms should play a role as part of any complete veridical UTC (cf. Newell 1990, 3-6).

6. Although we do not discuss them here, there are also other major harbingers of UTCs (e.g., Just and Carpenter 1987; Norman and Shallice 1986; Schneider and Detweiler 1987). More consideration of these may be found elsewhere (Meyer and Kieras 1997a; Newell 1990).

7. By informal convention, an acceptably accurate engineering approximation is one in which the theoretically calculated values deviate from the observed values by no more than 10% of the observed values' magnitudes (Card, Moran, and Newell 1983; Newell 1990).

8. According to the power law of practice, the reaction time (RT) to complete one trial of a task is a power function whose domain is the number (N) of prior practice trials on the task and whose exponent is a negative constant. As a result, $\log RT$ would be a negatively sloped linear function of $\log N$ (Anderson 1983; Fitts 1964; Newell 1990).

9. The lower part of the cognitive band includes two sublevels of information processing: *deliberate acts*, with durations on the order of 100 msec; and *simple operations* (short sequences of deliberate acts used to perform simple tasks), with total durations on the order of 1 sec (Newell 1990). For example, under our EPIC computational models, "firing" a production rule during a typical choice-reaction task constitutes a deliberate act, and using a sequence of such rules to perform the whole task constitutes a simple operation. The part of the cognitive band directly above these two sublevels consists of *composed operations* (sequences of simple operations used to perform complex tasks) with total durations on the order of 10 sec. For example, skilled playing of rapid-transit chess presumably entails composed operations.

10. On the overall time scale of human action, there are also other activity bands. Below the cognitive band is the *biological band*, a physical substrate for the functional architecture of people's information-processing system. The biological band has at least three sublevels: organelle, neuron, and neural circuit. At the neural-circuit sublevel, activities take on the order of 10 msec to complete, which in turn yields the aforementioned approximate 100 msec durations of deliberate acts (Newell 1990; see also note 9). Connectionist network models provide abstract characterizations of the neural-circuit sublevel (Rumelhart and McClelland 1986). Above the cognitive band are the *rational* and *social bands*. The rational band involves complex problem-solving activities with durations on the order of minutes or more. SOAR is especially tailored to characterize information processing in the rational band. Farther upscale, the social band involves the long-term pursuit of people's life goals. No prospective UTC yet deals seriously with the social band. Instead, on the basis of Newell's analyses and arguments (1990), the cognitive band is most important for present purposes.

11. For example, in Pashler (1984, experiment 1), "the subject was instructed to respond as quickly as possible to both tasks in the two-task blocks, with the restriction that the first stimulus must be responded to before the second" (p. 365). Similarly, in Pashler and Johnston (1989, 30), subjects were told that they "should respond as rapidly as possible to the first stimulus," and "the experimenter emphasized the importance of making the first response as promptly as possible."

12. The emulation of a bottleneck is required because without one, secondary-task responses might be selected and produced before primary-task responses, violating the instructions of the PRP procedure (Meyer and Kieras 1997a). According to this logic, the function of the emulated bottleneck is to delay secondary-responses enough that out-of-order responding never happens.
13. At present, the executive and task production rules for EPIC computational models are placed in procedural memory by the theorists who use this simulation system (e.g., Meyer and Kieras 1997a,b). Unlike some other architectures (cf. Anderson 1983; Laird, Newell, and Rosenbloom 1987), EPIC does not yet have the capability to acquire perceptual-motor and cognitive skills through on-line procedural-learning algorithms, although we may augment EPIC with such capabilities in future research.
14. The mean cognitive-processor cycle duration is one of EPIC's temporal parameters that remains the same across different simulations. We set it to 50 msec on both theoretical and empirical grounds. Newell (1990) has argued that theoretically, the time taken for testing the conditions and executing the actions of generic production rules like those used here should be on the order of 50 msec per rule, given known temporal constraints on the neural-network circuits whereby such operations occur at the biological level (see also note 10). Furthermore, Kristofferson (1967) has reported empirical results about perceptual successiveness judgments and choice RTs that suggest a mean cognitive-processor cycle duration of about 50 msec. This value is approximately the same as the alpha rhythm's mean zero-crossing interval in EEG records of brainwave activity (Callaway and Yeager 1960; Kristofferson 1967; Ray 1990).
15. Under some circumstances, responses also may be produced by a motor processor on the basis of sensory information sent directly to it by a perceptual processor via a pathway that bypasses the cognitive processor and involves an automatic "privileged loop" (cf. McLeod and Posner 1984).
16. Reliable effects on response selection for task 2 presumably were caused by Hawkins, Rodriguez, and Reicher's manipulation of S-R numerosity. To infer that such effects occurred there is highly plausible from results of previous research with S-R numerosity manipulations (Schumacher et al. forthcoming). For example, Sternberg (1969) found large reliable interactions between the effects on mean RTs of S-R numerosity and S-R compatibility, a prototypical factor that is believed to affect response selection. By contrast, Sternberg (1969) found relatively little interaction between S-R numerosity and stimulus-legibility effects, suggesting that S-R numerosity affects stimulus encoding hardly at all. These and additional related data imply that most, if not all, of the S-R numerosity effect on mean RTs occurs during response selection rather than other stages of processing (Brainard et al. 1962; Broadbent and Gregory 1965; Gottsdanker 1969; Sanders 1980; Theios 1973).
17. In some respects, the SRD model resembles the hybrid structural-bottleneck model of De Jong (1993), who proposed that both response-selection and movement-production bottlenecks mediate multiple-task performance, integrating hypotheses advocated by Kantowitz (1974), Keele (1973), Pashler (1984, 1994a), Welford (1967, 1980), and others. Similarly, to coordinate progress on tasks 1 and 2 of the PRP procedure, the executive process of the SRD model uses both post-response selection and pre-response selection lockout points for task 2. However, these lockout points are optional, flexible, and adaptively controlled, whereas the bottlenecks of De Jong's hybrid model (1993) are immutable and insensitive to changing task requirements.
18. Thus, by definition, the task-scheduling strategy used in the SRD model (figure 2.4) is daring.
19. Here task 1 involved four rather than two alternative S-R pairs (cf. Hawkins, Rodriguez, and Reicher 1979).
20. Because Lauber et al. (1994, experiment 2) increased the difficulty of task 1, the chances that responses for task 2 would occur prematurely before those for task 1 were potentially greater unless their participants used more cautious task scheduling than did those in Hawkins, Rodriguez, and Reicher (1979).

21. For each participant in figure 2.5, the mean RTs for task 1 averaged about 500 msec and were not affected much by either SOA or task 2 difficulty.
22. Insofar as we know, few advocates of the traditional RSB hypothesis have checked for systematic individual differences in task-scheduling strategies. Presumably, this omission has occurred because, according to the RSB hypothesis, all participants have an immutable structural response-selection bottleneck, and so their patterns of PRP curves would not be expected to differ much from each other.
23. To support this conclusion further, one may compare the light vertical bars (predicted interactions) and dark vertical bars (observed interactions) in figure 2.6B. Derivations described elsewhere (Meyer et al. 1995; Meyer and Kieras 1997b) yielded the light vertical bars, which assume that every participant had a response-selection bottleneck. Given this assumption, the light vertical bars should approximate the dark bars closely, but they do not. Seven-eighths of the dark vertical bars in figure 2.6B are longer than the light vertical bars paired with them, embodying consistently more extreme interactions than the traditional RSB hypothesis predicts. Specifically, it appears that figure 2.6B contains at least two distinct subgroups of participants, some of whom (e.g., participant 1) produced significantly negative interactions between the effects of SOA and response-selection difficulty on mean RTs for task 2, and others of whom (e.g., participant 8) produced significantly positive interactions. This supports our claim that task scheduling involves adaptive executive control and that, because of systematic personal preferences, some but not all participants may adopt daring scheduling strategies even when task 1 is relatively hard.
24. Standard statistical tests demonstrating additive factor effects in averaged group PRP curves (e.g., McCann and Johnston 1992; Pashler 1984; Pashler and Johnston 1989) do not surmount the problems caused by averaging across participants. The power of these tests is greatly weakened by systematic individual differences in patterns of PRP curves because these differences inflate the subject-by-treatment interactions that serve as error variance in the denominators of accompanying *t* and *F* statistics (Hays 1963). Such inflation may explain why some researchers have found seemingly unreliable interactions between the effects of SOA and other factors that influence response selection for task 2 of the PRP procedure. For example, McCann and Johnston (1992, experiment 1) reported $F(3, 69) = 1.94$ in the case of a positive interaction between SOA and S-R compatibility effects on mean RTs for task 2. This interaction was marginal ($.05 < p < .10$ for a unidirectional hypothesis test), but might have been much more reliable in the absence of underlying systematic individual differences, thereby further disconfirming the traditional RSB hypothesis.
25. The participants here were college students. Whether substantive lesson 5 likewise applies to older adults remains an important question for future work. Some initial research suggests that fortunately the answer may be yes (Glass et al. 1997; Kramer et al., chap. 22, this volume).
26. The results shown in figure 2.6 came from participants' third session of testing with a hard primary task and easy or hard secondary task in the standard PRP procedure.
27. The corresponding mean RTs for task 1 did not vary much as a function of SOA or response-selection difficulty in task 2; on average, they equaled 485 msec and 487 msec when task 2 was easy and hard, respectively. These values are close to those found by Lauber et al. (1994, experiment 2) during their previous experiment with a hard primary task in the PRP procedure. Administration of initial variable-priority training seems not to have influenced participants' performance of task 1 much at all.
28. In contrast, some advocates of the traditional RSB hypothesis have based their support on the persistence of PRP effects (dual-task interference) throughout many thousands of practice trials under the standard PRP procedure (Gottsdanker and Stelmach 1971; Pashler 1993). Such persistence is not, however, antithetical to our present conclusions. As mentioned already, the instructions about task priorities for the standard PRP procedure typically require that task 1 re-

ceive more emphasis than task 2 and that responses for task 1 be produced before those for task 2. Given such instructions, which hold throughout practice, some PRP effect always must occur after short SOAs no matter how practice changes participants' task scheduling strategies in other respects. Thus, as before, methodological lesson 6 ("Be careful what you ask for; you might actually get it") applies here as well.

29. For the auditory-vocal task, each stimulus was either a 220, 880, or 3,520 Hz tone, to which participants responded by saying either the word "one," "two," or "three," respectively. For the visual-manual task, each stimulus was the capital letter "O" displayed directly over the left, middle, or right dash in a horizontal row, to which participants responded by pressing a key with either the index, middle, or ring finger of the right hand, respectively. These tasks were similar to those used by Pashler (1990) and others (e.g., McCann and Johnston 1992) in testing the traditional RSB hypothesis.

30. Methodological lesson 6 also is reinforced by another result from the experiment of Schumacher et al. (1997). After our evidence of virtually perfect time-sharing (see figure 2.8) had been collected, we tested these same participants under the standard PRP procedure, asking them to treat the auditory-vocal task as primary and the visual-manual task as secondary. This yielded mean RTs for task 1 that remained short and constant regardless of the SOA, but mean RTs for task 2 increased when the SOA was short, substantially exceeding those shown in figure 2.8. The latter increase would be expected from methodological lesson 6 because the PRP procedure encourages participants to produce their responses for task 2 after those for task 1, which requires temporary postponement of progress on task 2 after short SOAs.

31. The benefits from the application of EPIC computational models could be substantial to NYNEX. For example, an improvement in the design of an interface that reduces the average completion time per call by 1 sec may decrease equipment and personnel costs on the order of a million dollars or more per year.

32. We thank Michael Atwood of the NYNEX Science and Technology Center for providing us with audiovisual tapes of TAOs' on-line performance. Helpful comments by Bonnie John and Rory Stuart about our analyses and modeling of this performance also are gratefully acknowledged.

33. The grand mean of the observed response latencies in figure 2.9 is about 1,100 msec. Relative to this baseline, the RMSE of 95 msec constitutes an 8.5% error of prediction by the DSS model. When working in realistic domains, engineers typically consider theoretical models practically useful when they predict observed numerical values with less than 10% margins of error (Card, Moran, and Newell 1983). The DSS model satisfies this criterion, whereas the CSS model does not. Furthermore, the fit of the DSS model seems satisfactory because the response latencies in figure 2.9 come from single keystrokes. The goodness of fit here is about what one might expect if the model were correct, but each response latency also contained an approximately 10% contribution from perceptual and motor "noise," which would be typical of practiced performers such as TAOs.

34. The response latency for a blip equaled the amount of time between (1) when the color of the blip changed from black to red, blue, or amber; and (2) when a key was pressed to indicate the blip's hostility status. Red blips had to be classified as "hostile"; blue blips had to be classified as "neutral"; amber blips had to be classified as either "hostile" or "neutral" in terms of their position, direction, or speed of movement on the display screen.

35. We thank James Ballas and his colleagues at the Naval Research Laboratory for generously providing us with their data and other helpful information.

36. For example, in the display of the Ballas, Heitmeyer, and Perez (1992) tactical decision task, the iconic blip that should get classified first after the start of each dual-task epoch could be blinked rapidly on and off, thereby orienting the operator's attention to it and helping the operator make a more rapid sequence of tactical decisions.

37. Experimental psychologists who are familiar with *Hamlet* as well as attention theory and practical applications to human-factors engineering have assured us that substantive lesson 14 is indeed apt (N. Moray, personal communication, July 12, 1996).

REFERENCES

- Abrams, R. A., and Jonides, J. (1988). Programming saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 428–443.
- Allport, D. A. (1980). Attention and performance. In G. L. Claxton (Ed.), *Cognitive psychology: New directions*, pp. 112–153. London: Routledge and Kegan Paul.
- Allport, D. A. (1987). Selection-for-action: Some behavioral and neurophysiological considerations of attention and action. In H. Heuer and A. F. Sanders (Eds.), *Perspectives on perception and action*, pp. 395–419. Hillsdale, NJ: Erlbaum.
- Allport, D. A. (1989). Visual attention. In M. I. Posner (Ed.), *Foundations of cognitive science*, pp. 631–682. Cambridge, MA: MIT Press.
- Allport, D. A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of 25 years. In D. E. Meyer and S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*, pp. 183–218. Cambridge, MA: MIT Press.
- Allport, D. A., Antonis, B., and Reynolds, P. (1972). On the division of attention: A disproof of the single-channel hypothesis. *Quarterly Journal of Experimental Psychology*, 24, 225–235.
- Anderson, J. A., and Hinton, G. E. (1981). Models of information processing in the brain. In G. E. Hinton and J. A. Anderson (Eds.), *Parallel models of associative memory*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 369–406.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R., and Bower, G. H. (1973). *Human associative memory*. New York: Winston.
- Atkinson, R. C., Herrnstein, R. J., Lindzey, G., & Luce, R. D., Eds. (1988). *Stevens' handbook of experimental psychology*. 2nd ed. New York: Wiley.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- Ballas, J. A., Heitmeyer, C. L., and Perez, M. A. (1992). *Direct manipulation and intermittent automation in advanced cockpits*. Technical report NRL/FR/5534-92-9375. Washington, DC: Naval Research Laboratory.
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91, 276–292.
- Bertelson, P. (1966). Central intermittency 20 years later. *Quarterly Journal of Experimental Psychology*, 18, 153–164.
- Boff, K. R., Kaufman, L., and Thomas, J. P., Eds. (1986). *Handbook of perception and human performance*. New York: Wiley.
- Bovair, S., and Kieras, D. E. (1991). Toward a model of acquiring procedures from text. In R. Barr, M. L. Kamil, P. Mosenthal, and P. D. Pearson (Eds.), *Handbook of reading research*, vol. 2, pp. 206–229. White Plains, NY: Longman.

- Bovair, S., Kieras, D. E., and Polson, P. G. (1990). The acquisition and performance of text editing skill: A cognitive complexity analysis. *Human-Computer Interaction*, 5, 1-48.
- Brainard, R. W., Irby, T. S., Fitts, P. M., and Alluisi, E. A. (1962). Some variables influencing the rate of gain of information. *Journal of Experimental Psychology*, 63, 105-110.
- Brickner, M., and Gopher, D. (1981). *Improving time-sharing performance by enhancing voluntary control on processing resources*. Technical report AFOSR-77-3131C. Haifa, Israel: Technion—Israel Institute of Technology.
- Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon Press.
- Broadbent, D. E. (1971). *Decision and stress*. London: Academic Press.
- Broadbent, D. E. (1993). A word before leaving. In D. E. Meyer and S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*, pp. 863-879. Cambridge, MA: MIT Press.
- Broadbent, D. E., and Gregory, M. (1965). On the interaction of S-R compatibility with other variables affecting reaction time. *British Journal of Psychology*, 56, 61-67.
- Bukkyo Dendo Kyokai. (1985). *The teaching of Buddha*. Tokyo: Kosaido.
- Callaway, E., and Yeager, C. L. (1960). Relationship between reaction time and electroencephalographic alpha base. *Science*, 132, 1765-1766.
- Card, S. K., Moran, T. P., and Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Erlbaum.
- Carpenter, P. A., and Just, M. A. (1989). The role of working memory in language comprehension. In D. Klahr and K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon*, pp. 31-68. Hillsdale, NJ: Erlbaum.
- Carpenter, P. A., Just, M. A., and Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97, 404-431.
- Chase, W. G., and Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation*, vol. 16, pp. 1-58. New York: Academic Press.
- Chubb, G. P. (1981). SAINT: A digital simulation language for the study of manned systems. In J. Moraal and K. F. Kraas (Eds.), *Manned system design*, pp. 153-179. New York: Plenum Press.
- Covrigaru, A., and Kieras, D. E. (1987). *PPS: A parsimonious production system*. Technical report TR-87/ONR-26. Ann Arbor: University of Michigan, Technical Communication Program.
- Creutzfeldt, O. D. (1983). *Cortex cerebri*. Berlin: Springer.
- Damos, D. L. (1991). *Multiple-task performance*. London: Taylor and Francis.
- Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
- De Jong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 965-980.
- Deutsch, J. A., and Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, 70, 80-90.
- Donders, F. C. (1868). On the speed of mental processes. Translation by W. G. Koster in Koster (Ed.), *Attention and performance II*, pp. 412-431. Amsterdam: Elsevier, 1969. Original work published in *Onderzoekingen gedaan in het Physiologisch Laboratorium der Utrechtsche Hoogeschool*, Tweede reeks, 1868-1869, 2, 92-120.

- Duncan, J. (1977). Response selection rules in spatial choice reaction tasks. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, NJ: Erlbaum.
- Duncan, J. (1986). Disorganization of behaviour after frontal lobe damage. *Cognitive Neuropsychology*, 3, 271-290.
- Eriksen, C. W., and Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 583-597.
- Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin*, 53, 134-140.
- Feynman, R. (1985). *Surely you're joking, Mr. Feynman! Adventures of a curious character*. New York: Norton.
- Fischer, B., and Ramsberger, E. (1984). Human express saccades: Extremely short reaction times of goal directed eye movements. *Experimental Brain Research*, 57, 191-195.
- 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. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning*, pp. 243-285. New York: Academic Press.
- Fitts, P. M., and Seeger, C. M. (1953). S-R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, 46, 199-210.
- Frowein, H. W. (1981). Selective effects of barbiturate and amphetamine on information processing and response execution. *Acta Psychologica*, 47, 105-115.
- Garner, K. (1997). Situational awareness in the tactical air environment. *CSERIAC Gateway*, 7, 12-13.
- Ghez, C., Hening, W., and Favilla, M. (1990). Parallel interacting channels in the initiation and specification of motor response features. In M. Jeannerod (Ed.), *Attention and performance XIII: Motor representation and control*, pp. 265-293. Hillsdale, NJ: Erlbaum.
- Gilhooly, K. J., Logie, R. H., Wetherick, N. E., and Wynn, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. *Memory and Cognition*, 21, 115-124.
- Glass, J. M. B., Lauber, E. J., Schumacher, E. H., Kieras, D. E., and Meyer, D. E. (1997). Aging and multiple-task performance: Task-coordination strategies in young and old adults. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago, May.
- Goodman, D., and Kelso, J. A. (1980). Are movements prepared in parts? Not under compatible (naturalized) conditions. *Journal of Experimental Psychology: General*, 109, 475-495.
- Gopher, D. (1986). In defence of resources: On structures, energies, pools and the allocation of attention. In G. R. J. Hockey, A. W. K. Gaillard, and M. G. H. Coles (Eds.), *Energetics and human information processing*, pp. 353-371. Dordrecht: Nijhoff.
- Gopher, D. (1993). Attentional control: Acquisition and execution of attentional strategies. In D. E. Meyer and S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*, pp. 299-322. Cambridge, MA: MIT Press.
- Gopher, D., Brickner, M., and Navon, D. (1982). Different difficulty manipulations interact differently with task emphasis: Evidence for multiple resources. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 146-157.
- Gopher, D., and Donchin, E. (1986). Workload: An examination of the concept. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of perception and human performance*. Vol. 2, *Cognitive processes and performance*, pp. 41.1-41.49. New York: Wiley.

- Gottsdanker, R. (1969). Choice reaction time and the nature of the choice response. *Psychonomic Science*, 14, 257-258.
- Gottsdanker, R., and Stelmach, G. E. (1971). The persistence of psychological refractoriness. *Journal of Motor Behavior*, 3, 301-312.
- Gray, W. D., John, B. E., and Atwood, M. E. (1993). Project Ernestine: Validating a GOMS analysis for predicting and explaining real-world task performance. *Human-Computer Interaction*, 8, 237-309.
- Green, D. M., and Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Greenwald, A. G., and Shulman, H. (1973). On doing two things at once: 2. Elimination of the psychological refractory period. *Journal of Experimental Psychology*, 101, 70-76.
- Gugerty, L. J. (1997). Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied*, 3, 42-66.
- Harris, R. M., Iavecchia, H. P., Ross, L. V., and Shaffer, S. C. (1987). Microcomputer human operator simulator (HOS-IV). Paper presented at the annual meeting of the Human Factors Society, Santa Monica, CA.
- Hays, W. L. (1963). *Statistics for psychologists*. New York: Holt, Rinehart, and Winston.
- Hawkins, H. L., Rodriguez, E., and Reicher, G. M. (1979). *Is time-sharing a general ability?* ONR technical report 3. Eugene: University of Oregon.
- Hirst, W., Spelke, E., Reaves, C., Caharack, G., and Neisser, U. (1980). Dividing attention without alternation or automaticity. *Journal of Experimental Psychology: General*, 109, 98-117.
- Hockey, G. R. J., Gaillard, A. W. K., and Coles, M. G. H. (Eds.). (1986). *Energetics and human information processing*. Dordrecht: Nijhoff.
- Hunt, E., and Lansman, M. (1986). A unified model of attention and problem solving. *Psychological Review*, 93, 446-461.
- Ivry, R. B., Franz, E. A., Kingstone, A., and Johnston, J. C. (1994). The PRP effect in a split-brain patient: Response uncoupling despite normal interference. Paper presented at the annual meeting of the Psychonomic Society, St. Louis, November.
- Ivry, R. B., Franz, E. A., Kingstone, A., and Johnston, J. C. (forthcoming). The PRP effect following callosotomy: Residual interference despite uncoupling of lateralized response codes. *Journal of Experimental Psychology: Human Perception and Performance*.
- John, B. E. (1988). Contributions to engineering models of human-computer interaction. Ph.D. diss. Carnegie-Mellon University.
- John, B. E. (1990). Extensions of GOMS analyses to expert performance requiring perception of dynamic visual and auditory information. In *Proceedings of the CHI 1990 Conference on Human Factors in Computing Systems*, pp. 107-115. New York: Association of Computing Machinery.
- John, B. E., and Kieras, D. E. (1996). The GOMS family of user interface analysis techniques: Comparison and contrast. *ACM Transactions on Computer-Human Interaction*, 3, 320-351.
- John, B. E., and Newell, A. (1987). Predicting the time to recall computer command abbreviations. In *Proceedings of the CHI 1987 Conference on Human Factors in Computing Systems*. New York: Association of Computing Machinery.
- John, B. E., and Newell, A. (1989). Cumulating the science of HCI: From S-R compatibility to transcription typing. In *Proceedings of the CHI 1989 Conference on Human Factors in Computing Systems*, New York: Association of Computing Machinery.

- John, B. E., Rosenbloom, P. S., and Newell, A. (1985). A theory of stimulus-response compatibility applied to human-computer interaction. In *Proceedings of the CHI 1985 Conference on Human Factors in Computing Systems*. New York: Association of Computing Machinery.
- John, B. E., Vera, A. H., and Newell, A. (1994). Toward real-time GOMS: A model of expert behavior in a highly interactive task. *Behavior and Information Technology*, 13, 255-267.
- Johnston, W. A., and Dark, V. J. (1986). Selective attention. *Annual Review of Psychology*, 37, 43-75.
- Jonides, J. (1980). Towards a model of the mind's eye's movement. *Canadian Journal of Psychology*, 34, 103-112.
- Jonides, J. (1981). Voluntary versus automatic control of the mind's eye's movement. In J. Long and A. Baddeley (Eds.), *Attention and performance IX*, pp. 187-203. Hillsdale, NJ: Erlbaum.
- Jonides, J., and Yantis, S. (1986). Uniqueness of abrupt visual onset in capturing attention. *Perception and Psychophysics*, 43, 346-354.
- Just, M. A., and Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Boston: Allyn and Bacon.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
- Kantowitz, B. H. (1974). Double stimulation. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*, pp. 83-131. Hillsdale, NJ: Erlbaum.
- Karlin, L., and Kestenbaum, R. (1968). Effects of number of alternatives on the psychological refractory period. *Quarterly Journal of Experimental Psychology*, 20, 67-178.
- Kasdan, L., Lucas, G., and Kaufman, P. (1981). *Raiders of the Lost Ark*. Hollywood, CA: Lucasfilm.
- Keele, S. W. (1973). *Attention and human performance*. Pacific Palisades, CA: Goodyear.
- Keele, S. W., and Neill, W. T. (1978). Mechanisms of attention. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of perception*, pp. 3-47. London: Academic Press.
- Kieras, D. E. (1988). Towards a practical GOMS model methodology for user interface design. In M. Helander (Ed.), *Handbook of human-computer interaction*, pp. 135-158. Amsterdam: Elsevier.
- Kieras, D. E., and Meyer, D. E. (1995). Predicting human performance in dual-task tracking and decision making with computational models using the EPIC architecture. In D. S. Alberts, D. Buede, T. Clark, R. Hayes, J. Hofmann, W. Round, S. Starr, and W. Vaughan (Eds.), *Proceedings of the International Symposium on Command and Control Research and Technology*, pp. 314-325. Washington, DC: National Defense University.
- Kieras, D. E., and Meyer, D. E. (forthcoming). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*.
- Kieras, D. E., Meyer, D. E., Mueller, S., and Seymour, T. (forthcoming). Insights into working memory from the perspective of the EPIC architecture for modeling skilled perceptual-motor and cognitive human performance. In A. Miyake and P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.
- Kieras, D. E., Wood, S. D., and Meyer, D. E. (1995). Predictive engineering models using the EPIC architecture for a high-performance task. In I. R. Katz, R. Mack, L. Marks, M. B. Rossan, and J. Nielsen (Eds.), *Proceedings of the CHI 1995 Conference on Human Factors in Computing Systems*, pp. 11-18. New York: Association of Computing Machinery.
- Kieras, D. E., Wood, S. D., and Meyer, D. E. (1997). Predictive engineering models based on the EPIC architecture for a multimodal high-performance human-computer interaction task. *ACM Transactions on Computer-Human Interaction*, 4, 230-275.

- Koch, R. (1993). Die psychologische Refraktärperiode. Ph.D. diss., University of Bielefeld, Bielefeld, Germany.
- Koch, R. (1994). Hick's Law and the psychological refractory period. Paper presented at the KNAW Symposium on Discrete versus Continuous Information Processing, Amsterdam, December.
- Kornblum, S., Ed. (1973). *Attention and performance IV*. New York: Academic Press.
- Kornblum, S., Hasbroucq, T., and Osman, A. (1990). Dimensional overlap: Cognitive basis of stimulus-response compatibility—A model and taxonomy. *Psychological Review*, 97, 253–270.
- Koster, W. G., Ed. (1969). *Attention and performance II*. Amsterdam: North-Holland Publishing Co.
- Krantz, D. H. (1969). Threshold theories of signal detection. *Psychological Review*, 76, 308–324.
- Kristofferson, A. B. (1967). Attention and psychophysical time. In A. F. Sanders (Ed.), *Attention and performance*, pp. 93–100. Amsterdam: North-Holland Publishing Co.
- Kyllonen, P. C., and Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, 14, 389–433.
- Laird, J. E., Newell, A., and Rosenbloom, P. S. (1987). SOAR: An architecture for general intelligence. *Artificial Intelligence*, 33, 1–64.
- Lane, N. E., Strieb, M. I., Glenn, F. A., and Wherry, R. J. (1981). The human operator simulator: An overview. In J. Moraal and K. F. Kraas (Eds.), *Manned system design*, pp. 121–152. New York: Plenum Press.
- Lauber, E. J., Schumacher, E. H., Glass, J., Zurbriggen, E., Kieras, D. E., and Meyer, D. E. (1994). Adaptive PRP effects: Evidence of flexible attention to action. Paper presented at the annual meeting of the Psychonomic Society, St. Louis, November.
- Lewin, K. (1951). *Field theory in social science*. New York: Harper.
- Logan, G. (1985). Executive control of thought and action. *Acta Psychologica*, 60, 193–210.
- Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), *Handbook of mathematical psychology*, vol. 1, pp. 103–189. New York: Wiley.
- Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.
- Mangun, G. R., Hillyard, S. A., and Luck, S. J. (1993). Electrocortical substrates of visual selective attention. In D. E. Meyer and S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*, pp. 219–243. Cambridge, MA: MIT Press.
- Marr, D. (1982). *Vision*. San Francisco: Freeman.
- McCann, R. S., and Johnston, J. C. (1992). Locus of the single-channel bottleneck in dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 471–484.
- McLeod, P. D. (1977). A dual task response modality effect: Support for multiprocessor models of attention. *Quarterly Journal of Experimental Psychology*, 29, 651–667.
- McLeod, P. D., and Posner, M. I. (1984). Privileged loops: From percept to act. In H. Bouma and D. G. Bouwhuis (Eds.), *Attention and Performance X: Control of language processes*, pp. 55–66. Hillsdale, NJ: Erlbaum.
- Meyer, D. E., Abrams, R. A., Kornblum, S., Wright, C. E., and Smith, J. E. K. (1988). Optimality in human motor performance: Ideal control of rapid aimed movements. *Psychological Review*, 95, 340–370.

- Meyer, D. E., and Gordon, P. C. (1985). Speech production: Motor programming of phonetic features. *Journal of Memory and Language*, 24, 3–26.
- Meyer, D. E., and Kieras, D. E. (1992). The PRP effect: Central bottleneck, perceptual-motor limitations, or task strategies? Paper presented at the annual meeting of the Psychonomic Society, St. Louis, November.
- Meyer, D. E., and Kieras, D. E. (1994). *EPIC computational models of psychological refractory-period effects in human multiple-task performance*. EPIC report 2 (TR-94/ONR-EPIC-2). Ann Arbor: University of Michigan.
- Meyer, D. E., and Kieras, D. E. (1996). EPIC: Adaptive executive control of human multiple-task performance. Paper presented at the annual meeting of the Psychonomic Society, Chicago, November.
- Meyer, D. E., and Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review*, 104, 3–65.
- Meyer, D. E., and Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of psychological refractory-period phenomena. *Psychological Review*, 104, 749–791.
- Meyer, D. E., Kieras, D. E., Lauber, E., Schumacher, E., Glass, J., Zurbriggen, E., Gmeindl, L., and Apfelblat, D. (1995). Adaptive executive control: Flexible multiple-task performance without pervasive immutable response-selection bottlenecks. *Acta Psychologica*, 90, 163–190.
- Meyer, D. E., and Kornblum, S., Eds. (1993). *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*. Cambridge, MA: MIT Press.
- Meyer, D. E., Osman, A. M., Irwin, D. E., and Yantis, S. (1988). Modern mental chronometry. *Biological Psychology*, 26, 3–67.
- Miller, J. (1988). Discrete and continuous models of human information processing: Theoretical distinctions and empirical results. *Acta Psychologica*, 67, 191–257.
- Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, 11, 56–60.
- Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologica*, 27, 84–92.
- Moray, N. (1979). The strategic control of information processing. In G. Underwood (Ed.), *Strategies of information processing*, pp. 301–328. London: Academic Press.
- Moray, N. (1979). *Mental workload: Its theory and application*. New York: Plenum Press.
- Navon, D. and Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, 86, 214–255.
- Neisser, U. (1967). *Cognitive psychology*. Englewood Cliffs, NJ: Prentice Hall.
- Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer and A. F. Sanders (Eds.), *Perspectives on perception and action*, pp. 361–394. Hillsdale, NJ: Erlbaum.
- Newell, A. (1973a). You can't play 20 questions with nature and win. In W. G. Chase (Ed.), *Visual information processing*, pp. 283–308. New York: Academic Press.
- Newell, A. (1973b). Production systems: Models of control structures. In W. G. Chase (Ed.), *Visual information processing*, pp. 463–526. New York: Academic Press.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Newell, A. (1992). Unified theories of cognition and the role of SOAR. In J. A. Michon and A. Akyürek (Eds.), *SOAR: A cognitive architecture in perspective*, pp. 25–79. Amsterdam: Kluwer.

- Newell, A., and Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.
- Norman, D. A. (1976). *Memory and attention: An introduction to human information processing*. 2nd ed. New York: Wiley.
- Norman, D. A., and Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, and D. Shapiro (Eds.), *Consciousness and self-regulation*, vol. 4, pp. 1–18. New York: Plenum Press.
- North, R. (1977). *Task components and demands as factors in dual-task performance*. Report no. ARL-77-2/AFOSE-77-2. Urbana: University of Illinois, January.
- O'Donnell, R. D., and Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of perception and human performance*. Vol. 2, *Cognitive processes and performance*, pp. 42.1–42.49. New York: Wiley.
- Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 358–377.
- Pashler, H. (1990). Do response modality effects support multiprocessor models of divided attention? *Journal of Experimental Psychology: Human Perception and Performance*, 16, 826–842.
- Pashler, H. (1993). Dual-task interference and elementary mental mechanisms. In D. E. Meyer and S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*, pp. 245–264. Cambridge, MA: MIT Press.
- Pashler, H. (1994a). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, 116, 220–244.
- Pashler, H. (1994b). Graded capacity-sharing in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 330–342.
- Pashler, H., Carrier, M., and Hoffman, J. E. (1993). Saccadic eye movements and dual-task interference. *Quarterly Journal of Experimental Psychology*, 46A, 51–82.
- Pashler, H., and Johnston, J. C. (1989). Chronometric evidence of central postponement in temporally overlapping tasks. *Quarterly Journal of Experimental Psychology*, 41A, 19–45.
- Popper, K. (1959). *The logic of scientific discovery*. New York: Basic Books.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3–25.
- Rafal, R., Henik, A., and Smith, J. (1991). Extrageniculate contributions to reflexive visual orienting in normal humans: A temporal hemifield advantage. *Journal of Cognitive Neuroscience*, 3, 323–329.
- Ray, W. J. (1990). The electrocortical system. In J. T. Cacioppo and L. G. Tassinary (Eds.), *Principles of psychophysiology*, pp. 385–412. Cambridge: Cambridge University Press.
- Reitman, W. (1970). What does it take to remember? In D. A. Norman (Ed.), *Models of memory*, pp. 470–509. New York: Academic Press.
- Reuter-Lorenz, P. A., Hughes, H. C., and Fendrich, R. (1991). The reduction of saccadic latency by prior offset of the fixation point: An analysis of the gap effect. *Perception and Psychophysics*, 49, 167–175.
- Reynolds, D. (1964). Effects of double stimulation: Temporary inhibition of response. *Psychological Bulletin*, 62, 333–347.
- Roberts, S., and Sternberg, S. (1993). The meaning of additive reaction-time effects: Tests of three alternatives. In D. E. Meyer and S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*, pp. 611–653. Cambridge, MA: MIT Press.

- Rosenbaum, D. A. (1980). Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General*, 109, 475-495.
- Rosenbaum, D. A. (1991). Human motor control. San Diego, CA: Academic Press.
- Rumelhart, D. E., and McClelland, J. L., Eds. (1986). *Parallel distributed processing*. Cambridge, MA: MIT Press.
- Rumelhart, D. E., and Norman, D. A. (1982). *Simulating a skilled typist: A study of skilled cognitive-motor performance*. *Cognitive Science*, 6, 1-36.
- Ruthruff, E., Pashler, H., and Klaasen, A. (1995). Preparation and strategy in dual-task interference. Poster presented at the meeting of the Psychonomic Society, Los Angeles, November.
- Salthouse, T. (1986). Perceptual, cognitive, and motoric aspects of transcription typing. *Psychological Bulletin*, 99, 303-319.
- Sanders, A. F. (1980). Stage analysis of reaction processes. In G. E. Stelmach and J. Requin (Eds.), *Tutorials in motor behavior*, pp. 331-354. Amsterdam: Elsevier.
- Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica*, 53, 61-97.
- Sanders, A. F. (1984). Ten symposia on attention and performance: Some issues and trends. In H. Bouma and D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes*, pp. 3-13. Hillsdale, NJ: Erlbaum.
- Sanders, A. F. (1991). Simulation as a tool in the measurement of human performance. *Ergonomics*, 34, 995-1025.
- Sanders, A. F., Ed. (1967). *Attention and performance*. Amsterdam: Elsevier.
- Sanders, A. F., Ed. (1970). *Attention and performance III*. Amsterdam: Elsevier.
- Saunders, R. (1757). *Poor Richard's almanac*. Philadelphia: B. Franklin's New Printing Office.
- Schneider, W., and Detweiler, M. (1987). A connectionist/control architecture for working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation*, vol. 21, pp. 53-119. Orlando, FL: Academic Press.
- Schneider, W., and Detweiler, M. (1988). The role of practice in dual-task performance: Toward workload modeling in a connectionist/control architecture. *Human Factors*, 30, 539-566.
- Schumacher, E. H., Glass, J., Lauber, E. J., Gmeindl, L., Woodside, B. J., Kieras, D. E., and Meyer, D. E. (1996). Concurrent response-selection processes in multiple-task performance. Poster presented at the annual meeting of the Psychonomic Society, Chicago, November.
- Schumacher, E. H., Lauber, E. J., Glass, J. M. B., Zurbriggen, E. L., Gmeindl, L., Kieras, D. E., and Meyer, D. E. (forthcoming). Concurrent response-selection processes in dual-task performance: Evidence for adaptive executive control of task scheduling. *Journal of Experimental Psychology: Human Perception and Performance*.
- Schumacher, E. H., Seymour, T. S., Glass, J. M., Lauber, E. J., Kieras, D. E., and Meyer, D. E. (1997). *Virtually perfect time sharing in dual-task performance*. Paper presented at the annual meeting of the Psychonomic Society, Philadelphia, PA, November.
- Schvaneveldt, R. W. (1969). Effects of complexity in simultaneous reaction time tasks. *Journal of Experimental Psychology*, 81, 289-296.
- Seifert, C. M., Meyer, D. E., Davidson, N., Patalano, A. J., and Yaniv, I. (1994). Demystification of cognitive insight: The prepared-mind perspective. In R. J. Sternberg and J. Davidson (Eds.), *The nature of insight*, pp. 65-124. Cambridge, MA: MIT Press.

- Seifert, C. M., and Shafto, M. G. (1994) Computational models of cognition. In F. Boller and J. Grafman (Eds.), *Handbook of neuropsychology*, vol. 9, pp. 409–425. Amsterdam: Elsevier.
- Shaffer, L. H. (1975). Multiple attention in continuous verbal tasks. In P. M. A. Rabbitt and S. Dornic (Eds.), *Attention and performance V*, pp. 157–167. New York: Academic Press.
- Shallice, T. (1972). Dual functions of consciousness. *Psychological Review*, 79, 383–393.
- Shiffrin, R. M., and Schneider, W. (1977). Controlled and automatic human information processing: 2. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127–190.
- Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, 116, 250–264.
- Smith, M. C. (1967). Theories of the psychological refractory period. *Psychological Bulletin*, 67, 202–213.
- Sternberg, S. (1969). On the discovery of processing stages: Some extensions of Donders' method. *Acta Psychologica*, 30, 276–315.
- Sternberg, S., Monsell, S., Knoll, R. L., and Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. Stelmach (Ed.), *Information processing in motor control and learning*, pp. 117–152. New York: Academic Press.
- Tanner, W. P., Jr., and Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, 61, 401–409.
- Theios, J. (1973). Reaction time measurements in the study of memory processes: Theory and data. In G. H. Bower (Ed.), *The psychology of learning and motivation*, vol. 7, pp. 43–85. New York: Academic Press.
- Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 12, 242–248.
- Treisman, A. M. (1964). Verbal cues: Language and meaning in selective attention. *American Journal of Psychology*, 77, 206–219.
- Treisman, A. M. (1988). Features and objects: The Fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, 40A, 201–237.
- Turing, A. (1937). On computable numbers, with an application to the Entscheidungs-problem. *Proceedings of the London Mathematical Society*, 42.
- Vreuls, D., and Obermayer, R. W. (1985). Human systems performance measurement in training simulators. *Human Factors*, 27, 241–250.
- Welford, A. T. (1952). The "psychological refractory period" and the timing of high speed performance: A review and a theory. *British Journal of Psychology*, 42, 2–19.
- Welford, A. T. (1959). Evidence of a single-channel decision mechanism limiting performance in a serial reaction task. *Quarterly Journal of Experimental Psychology*, 2, 193–210.
- Welford, A. T. (1967). Single channel operation in the brain. *Acta Psychologica*, 27, 5–22.
- Welford, A. T. (1980). The single-channel hypothesis. In A. T. Welford (Ed.), *Reaction time*, pp. 215–252. San Diego, CA: Academic Press.
- Wickens, C. D. (1976). The effect of divided attention on information processing in manual tracking. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 1–13.
- Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman, J. Beatty, and R. Davies (Eds.), *Varieties of attention*, pp. 63–101. New York: Wiley.

- Wickens, C. D. (1991). Processing resources and attention. In D. L. Damos (Ed.), *Multiple-task performance*, pp. 3-34. London: Taylor and Francis.
- Wierwille, W. W., and Connor, S. A. (1983). Evaluation of 20 workload measures using a psychomotor task in a moving-base aircraft simulator. *Human Factors*, 25, 1-16.
- Williges, R. C., and Wierwille, W. W. (1979). Behavioral measures of aircrew mental workload. *Human Factors*, 21, 549-574.
- Wolfe, T. (1979). *The right stuff*. New York: Farrar, Strauss, and Giroux.
- Woodworth, R. S., and Schlosberg, H. (1954). *Experimental psychology*. New York: Holt.
- Yaniv, I., Meyer, D. E., Gordon, P. C., Huff, C. A., and Sevald, C. A. (1990). Vowel similarity, connectionist models, and syllable structure in motor programming of speech. *Journal of Memory and Language*, 29, 1-26.
- Yantis, S. G., and Jonides, J. (1986). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 121-134.
- Yeager, C., and Janos, L. (1985). *Yeager: An autobiography*. New York: Bantam.