Yunfeng Zhang

4/23/12

Using Computational Cognitive Modeling to Validate and Advance Multitasking Theories

Abstract

Humans routinely perform multiple tasks simultaneously. Understanding the capabilities and limits of human multitasking not only helps design efficient devices to enhance human performance, but also helps uncover the nature of human cognition. Research on multitasking is difficult because multitasking performance is inevitably influenced by many factors such as a person's perceptual, cognitive and motor capabilities, as well as the strategies adopted for managing the conflicts among tasks. Decades of experimental-psychology research has identified many of the invariable factors that influence multitasking, but it is cognitive modeling that shows the potential in integrating the factors and making quantitative predictions about the performance.

This paper reviews the results of past research on multitasking performance, and stresses the increasingly important role that cognitive modeling has played in this research endeavor. The paper discusses multiple resource theory in detail, which is the current predominant theoretical framework in psychology that incorporates

various sensorimotor and cognitive factors. Cognitive architectures (primarily EPIC and ACT-R) are shown here to appropriately implement the majority of such factors in the form of computational simulation. The benefits of this integrated, computational approach is revealed in modeling studies of PRP (psychological refractory period) tasks.

The paper also discusses the influence of task strategies on multitasking performance. This type of exploration is uniquely enabled by cognitive modeling, because the production system adopted by most cognitive architectures provides a means to formally express task strategies. The research on task strategies has shed light on many cognitive functions such as executive processing and human adaptation to a task environment.

Studies on driving are discussed because driving is a major application domain of multitasking. Several driving models are reviewed and their advantages and drawbacks are analyzed. The paper concludes by summarizing the advantages of cognitive modeling over traditional experimental approach and suggesting future research directions for cognitive modeling of multitask performance.

Table of Contents

Introduction

Today, enabled by powerful digital devices, people often engage in multiple tasks at the same time in a variety of scenarios. Typical everyday scenarios include texting while walking or taking notes on a phone while talking with someone. Although these scenarios do not require high efficiency, they can still disrupt people's daily activities if the interaction with the devices is too involved. Other multitasking scenarios, however, demand high efficiency because they often have an impact on personal and public safety. For example, manually operating a GPS while driving is a very dangerous multitasking activity, and a poorly designed user interface can exacerbate the problem. Similarly, devices in emergency vehicles are complex and difficult to use, and yet the drivers often have to operate them while driving due to the demands of their work (Richtel, 2010). Multitasking scenarios are not only a concern of the mobile industry, but also an issue that may occur in the traditional desktop environment. For example, operators in control centers, such as emergency response centers and air traffic control centers, are often presented with a tremendous amount of information that needs constantly monitoring and responding. Perhaps even a small increase in efficiency can help the control centers improve public safety and potentially save millions of losses per year.

Despite the increasing prevalence of multitasking, there are not many agreed-upon theories to guide user interface design to better support multitasking. At the level of theoretical psychology, there is still debate regarding whether people can truly respond to multiple tasks at the same time (Meyer & Kieras, 1997a; Pashler, 1989). At the level of applied psychology, there are no definite conclusions regarding how to effectively reduce the interferences between two tasks.

For example, in driving research, many studies show that visual displays can distract drivers (e.g., Wierwille, 1993), but some other studies suggest that speech-based user interfaces for certain tasks may bring more interferences (e.g., Lee, Caven, Haake, & Brown, 2001). It is not surprising to see such diverse results given that the multitasking domain often involves very complex and highly dynamic tasks. Even in simple multitasking scenarios, the interplay among sensorimotor and cognitive factors can become very complex. It is thus very difficult to study the effect of each factor independently. To better support multitasking, a better understanding of the fundamental capabilities of human information processing is needed, as well as a framework that allows researchers to examine the effects of integration of various factors.

To construct an integrated theory on multitasking, one effective method is to build computational cognitive models. In computational cognitive modeling, different stages and aspects of human information processing, such as visual perception and manual motor control, are simulated with different software modules. The implementation of such a software module is often a computational realization of a psychological theory. The integration of these software modules leads to cognitive architectures, within which a variety of psychological theories can be integrated and tested. Cognitive architectures substantially reduce the need of repetitively building modules of well-studied human information processes and enable modeling at a higher level of abstraction such as task strategies. Because most computational architectures also make assumptions about the time needed to complete various cognitive, perceptual, and motor processes, they can make predictions about task completion time. Perhaps because of this unique feature, cognitive modeling has been shown to be very effective in helping design systems for time-critical, high-performance tasks (e.g., Foyle & Hooey, 2008). Because many multitasking scenarios also require high efficiency, there should be a good synergy between cognitive modeling and multitasking studies.

The goal of this paper is to show that cognitive modeling is an important method for validating and advancing multitasking theories in both science and engineering. The next, second, section introduces the basic concepts of cognitive modeling. Two main cognitive architectures are discussed in detail as they are those most frequently used in the HCI research. The third section reviews theories that address two main research questions in multitasking: Where does the interference come from and how does strategic control affect performance? Cognitive modeling is extensively used in answering these questions. The fourth section discusses studies on multitasking while driving. Several modeling studies of driving are introduced in this section as well. Driving is selected as the main topic for applied multitasking research because it is a relatively demanding task and because improving driving efficiency and safety will broadly impact people's everyday lives. The last section summarizes the advantages of cognitive modeling over traditional experimental approach to multitasking research, and proposes possible future research directions for modeling.

Cognitive Modeling

Cognitive modeling as a methodology to improve user interface design was proposed by Card, Moran, and Newell (1983), who have introduced three modeling approaches: Keystroke Level Modeling (KLM), GOMS (Goals, Operators, Methods, and Selection rules), and Model Human Processor (MHP). The goal of these modeling approaches is to provide predictions, primarily about time to complete a task, such that alternative user interfaces can be evaluated even before conducting extensive user studies. All three methods require the analysts to decompose the process of doing a task into a set of primitive operators, but MHP's operators are defined at a finer level than those of KLM and GOMS. For example, to represent mental processes, KLM has defined a single operator which always takes 1.2 s to complete, whereas MHP has defined several memory and goal manipulating operators. Although the simplification of KLM and GOMS is perhaps a good way to provide preliminary evaluations for a design, they are not sufficient for acquiring accurate estimates of task execution time.

While KLM and GOMS were designed for practical applications, MHP was more of a theoretical development for exploring the limits of cognitive science in simulating human behaviors. Following a computer metaphor, the MHP architecture characterizes human information processing with a set of memory storages and processors. The perceptual processors, the eye and ear processors, handle the inputs from the environment. They transform perceptual information into symbols and store them in visual and auditory working memory. The cognition processor then makes decisions based on the symbols presented in the working memory and in the long-term memory, and send commands to the motor processor. All the processors are parameterized by their cycle time, and the two working memory systems are parameterized by their storage capacity, memory decay time, and symbol code type (such as visual or semantic). MHP is the first of its kind to integrate decades of psychological theories on memory, perception, and motor into one unified system.

Although MHP was not computationally realized, it influenced the development of computational cognitive architectures, with the two most pertinent to HCI being EPIC (Kieras & Meyer, 1997) and ACT-R/PM (Byrne & Anderson, 2001). The following two sections introduce the features of EPIC and ACT-R, and discuss how these features could be useful for modeling multitasking scenarios.

EPIC

Figure 1 summarizes the overall structure of EPIC (Executive Process-Interactive Control). Similar to MHP, EPIC consists of several perceptual and motor processors as well as a cognitive processor. However, EPIC's implementation of the perceptual and motor processors encompasses more recent and better developed theories. In addition, EPIC uses a formal programming language, production rules, to express the processes executed by the cognitive processor. To model a task, the analyst needs to build a simulated task environment in a program and codify the presumed users' task strategies in production rules. The cognitive architecture uses the production rules to run through the simulation and produce various predictions. Although EPIC in its current form is still too complex to be used in routine user interface design, it is well positioned for studying the general effects that occur in a class of tasks, particularly multimodal multitasking tasks.

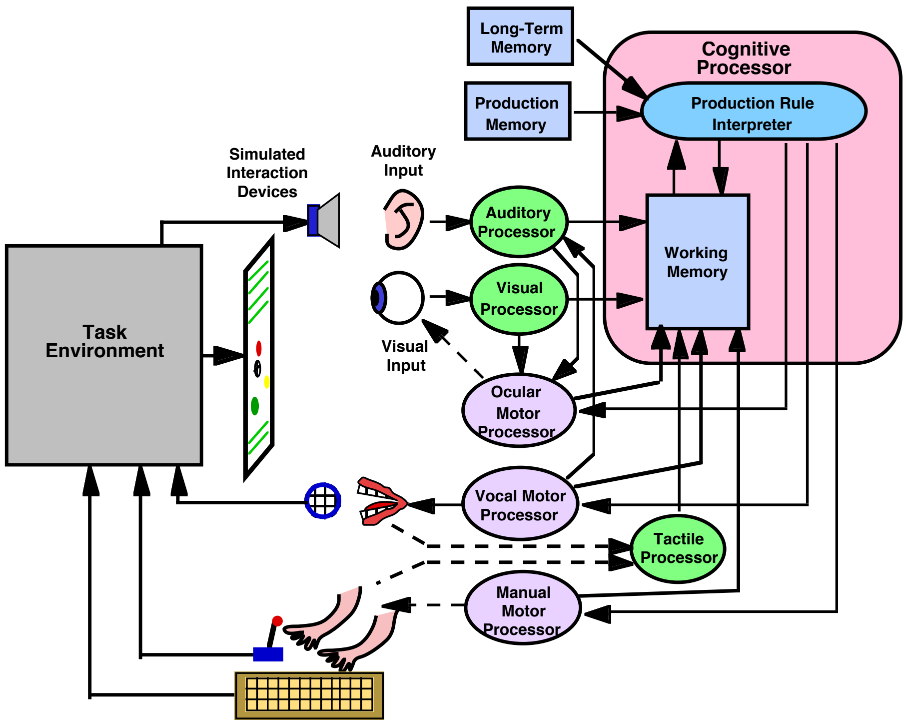


Figure 1. The overall structure of EPIC. The right side of the diagram shows the components of the architecture, and the left side shows the simulated task environment and the input and output devices. Image from Kieras (2004).

EPIC's visual processor and oculomotor processor capture some basic properties of the human eyes, making it possible to simulate the complex visual interactions involved in multimodal multitasking scenarios. EPIC's visual processor is perhaps the only computational implementation that takes into account the physical property of the eyes, i.e. the varying density of the receptors on the retina. EPIC's simulated retina consists of three zones, each with a standard radius: the fovea (1º), the parafovea (7.5º), and the periphery (60º). Consistent with past research (Sanders & McCormick, 1987), in EPIC some properties of the stimuli such as color are made available to the whole retina, whereas other properties such as textual information are only made available to smaller zones. This visual perceptual feature allows the architecture to correctly implement parallel processing of the visual information. For example, it allows a driving model to attend to the front roadway while also noticing the red octagonal stop sign in the periphery. EPIC's oculomotor processor further enhances the visual system by simulating eye movements to orient the high-resolution vision to appropriate stimuli. The implementation of the oculomotor enables the architecture to predict eye movement patterns, which can be a very helpful guidance for user interface design.

EPIC's motor system permits a certain degree of concurrency among different movements, which can be used to model highly practiced, expert performance in some multitasking scenarios. Except for aimed movements, motor movements in EPIC such as keypresses generally involve three stages: movement preparation, movement initiation, and movement execution (Kieras & Meyer, 1997). Although each stage only allows processing of one movement, different stages can process different movements in parallel. Due to this pipelined processing mechanism, a chain of keypresses would take less time than the same number of separated keypresses, allowing for modeling of expert-level performance.

EPIC's support for multitasking not only exists at the perceptual and motor level, but also at the cognitive level. While most cognitive architectures permit only one production rule to fire during a cycle in order to prevent conflicts among multiple rules, EPIC chooses to drop this limitation and allows any number of production rules to fire during each cycle. Although this architectural decision makes it more difficult to build models because conflicts may arise from incompatible rules, it is perhaps a more suitable assumption for modeling multitasking performance as has been suggested by some studies (Meyer & Kieras, 1997a, 1997b).

None of the above mechanisms were specifically designed for modeling multitasking, but by accurately representing the current knowledge about human performance at the cognitive, perceptual and motor level, EPIC was able to tackle some very complex multitasking scenarios (e.g., Kieras & Meyer, 1997). ACT-R takes a similar approach, and it is widely used to model learning, psychological laboratory tasks, and some large scale HCI tasks.

ACT-R

Figure 2 shows the overall structure of ACT-R (Adaptive Control of Thought–Rational). Similar to EPIC, ACT-R also consists of a set of perceptual and motor modules (the analogs of processors) as well as a production system to simulate cognition. Each module communicates with the production system through a buffer. ACT-R's buffers are similar to EPIC's working memory in that the contents in the buffers determine which rule should be fired. The difference is that a buffer can hold only one memory chunk, whereas the working memory in EPIC has no capacity limit.

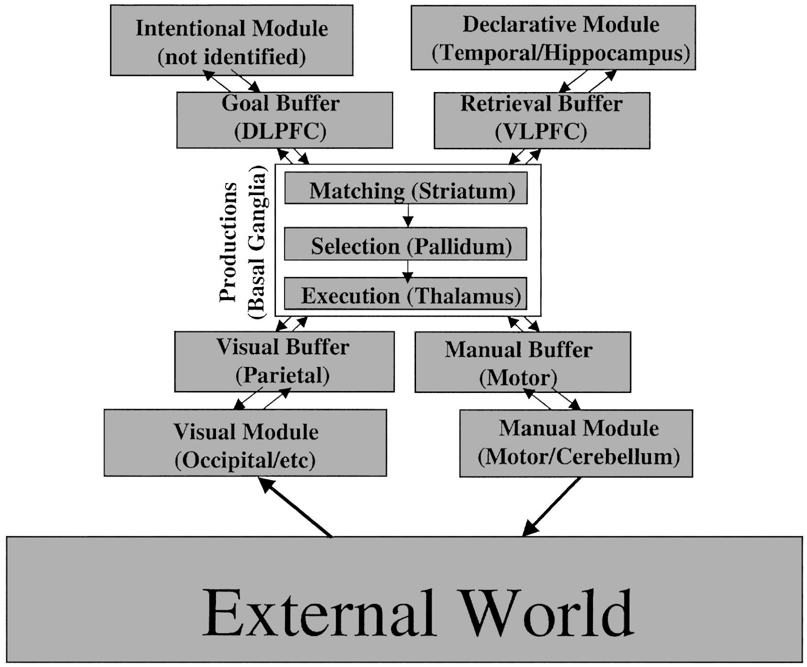


Figure 2. The overall structure of ACT-R, noted with presumed corresponding brain regions for each component. Image from Anderson et al. (2004).

One of the strengths of ACT-R is its highly developed declarative module, which simulates a human's long-term declarative memory system. Most psychologists believe that there are two types of long-term memory: procedural memory, which stores skills, and declarative memory, which stores facts and events. Procedural memory is represented by production rules in both ACT-R and EPIC. Declarative memory, however, is handled differently in the two architectures. In ACT-R, declarative memory is organized as chunks. Each chunk has several attributes that store properties of a symbol and its relations to other symbols. A chunk has an activation level, which determines its likelihood to be retrieved and its retrieval latency. Using the activation level, ACT-R can model several phenomena of the long-term declarative memory system. For example, memory decay is characterized by the logarithmic decrease of the activation level across time. The implementation of the declarative module enables ACT-R to model learning processes. By contrast, EPIC does not have such comprehensive facilities for handling declarative memory. In EPIC, declarative memory is simply initialized and put into working memory at the beginning of a simulation. For modeling highly practiced, expert performance, however, this simplified declarative memory system is less problematic because, in these cases, long-term memory retrieval tends to be fast and is close to working memory retrieval time (25 ms in EPIC).

The central difference between ACT-R and EPIC, which also leads to distinct approaches to modeling multitasking, is ACT-R's commitment to a central bottleneck theory. ACT-R's production system can only fire one rule within a cycle. If multiple rules' conditions are matched, the rule with the highest utility is selected to fire. Although an ACT-R model can still have temporal-overlapping processes across different modules, it cannot make multiple cognitive decisions at the same time. By contrast, EPIC has no restriction on how many rules can fire within a cycle. This distinction underlies a long-lasting debate regarding how the observed interference among multiple tasks should be interpreted. Issues related to this topic will be discussed in the next section.

Both EPIC and ACT-R are well developed cognitive architectures. Although currently they are still too complex to be used in routine user interface designs, they are, as shown next, very useful for exposing the implications of various multitasking theories.

Theories of Multitasking

To design user interfaces that better support multitasking, two questions need to be answered: (a) What are the invariable factors that cause interferences among multiple tasks, and (b) how can strategic control influence performance? The first question relates to humans' fundamental information processing capacities, which is always a central concern of human performance research. Knowing the answer to this question will help establish design guidelines that can effectively reduce multitasking interference. The second question is unique to multitasking, as people tend to have a great deal of flexibility in choosing how to allocate efforts and when to switch tasks in a multitasking scenario. Knowing the answer to this question can help design interfaces that motivate users to more efficiently and safely (essential for some tasks such as driving) complete a task.

Cognitive modeling has played a key role in the research endeavor to answer both questions. As shown next, even a very simple laboratory dual-task experiment may involve complex interactions that entail a cognitive architecture integrating theories at all levels. Moreover, the production system adopted by most cognitive architectures provides a formal approach to studying task strategies involved in multitasking scenarios.

Invariable Factors in Multitasking

Although many studies (cf. Navon & Gopher, 1979) have shown that multitasking performance can vary considerably even for an individual, there are some invariable, structural factors that can affect even the simplest multitasking situation. The psychological refractory period (PRP) experimental paradigm has been frequently used to study such invariable factors. In PRP experiments, participants are required to do two choice-reaction tasks concurrently, with Task 2's stimulus appearing slightly after Task 1's. The reaction time (RT) of Task 2 or of both tasks are often found longer than that of the single task performance (Welford, 1952). These results suggest that at some point during a PRP trial, which usually lasts less than two seconds, the two tasks must have interfered with each other in some way.

Early theories such as the global single channel hypothesis (Craik, 1948) and the unitary-resource theory (Kahneman, 1973) attempted to define the invariable factors in multitasking, but they failed to account for many interference effects observed in empirical studies. The global single channel hypothesis states that all stages involved in information processing constitute a single channel, which is occupied by a task in an all-or-none fashion. This hypothesis was rejected by some PRP studies (e.g., Karlin & Kestenbaum, 1968), which showed that when the two tasks are initiated at exactly the same time, Task 2's RT increases by an amount that is less than Task 1's RT, suggesting that the two tasks must have overlapped for a period. This overlapping, however, can be explained by the unitary-resource theory, which postulates a single resource that has a fixed amount of capacity and can be shared by parallel tasks. The unitary-resource theory predicts that all processes, if they draw similar amount of resources, should introduce similar interferences to another task. However, studies have shown that just by changing the response modality of one task in the PRP experiments while maintaining the same level of task difficulty, the interference between the two tasks can be substantially changed

(McLeod, 1977). This result indicates the possibilities of structural factors that affect different stages or aspects of the task execution process.

The current prevalent multitasking theory is the multiple resource theory (Navon & Gopher, 1979; Wickens, 2002), which incorporates several factors at the perceptual-motor and the cognitive level to account for multitasking interferences. As an extension of the unitary-resource theory, the multiple resource theory postulates a structure of several resource pools, with each pool having its own divisible capacity. Based on the past empirical findings, Wickens (1980) proposed a three-dimensional taxonomy of the resources, with the three dimensions being modalities, codes, and stages. Tasks that utilize the same resources have the greatest interferences. For example, this theory predicts that two auditory tasks presented at the same time (e.g., dichotic listening) produce a greater interference than an auditory task combined with a visual task.

Figure 3 illustrates Wickens' three-dimensional taxonomy. In this diagram, the horizontal axis represents the stage dimension, the vertical axis represents the modality dimension, and the axis pointing out represents the code dimension. The stage dimension consists of three phases: perception, cognition, and responding. But in this taxonomy, the perception and cognition stages share the same resources, as can be seen in the diagram that there is no division between them. The modality dimension, which consists of visual and auditory modalities, only extends to the perception stage, i.e. at the cognition and responding stage, visual and auditory information is assumed to be processed by the same set of resources. Because of this, the modality dimension

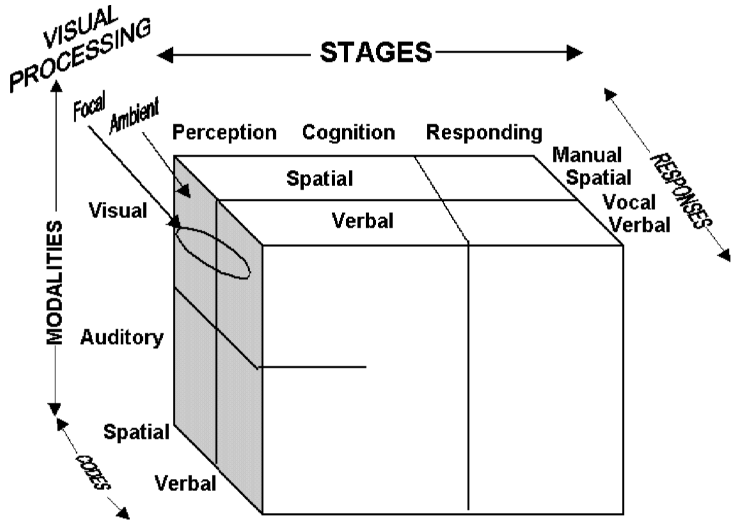


Figure 3. A taxonomy of multiple resources. Image from Wickens (2002).

is also referred to as the perceptual modality dimension. The code dimension, which consists of spatial and verbal codes, exists at all three stages. Wickens assumed that at the responding stage, manual responses only use spatial codes and vocal responses only use verbal codes. Thus, the two responding resources are called manual spatial and vocal verbal resources. Besides the three major dimensions, Wickens also proposed a fourth dimension that only exists at the visual perception stage, the dimension of focal vs. ambient visual processing. The discussion in this section will be structured around the three major dimensions, and the fourth dimension is regarded as a component of the perceptual processing stage.

Although there are still many debates around the multiple resource theory, it servers as a good framework to navigate through a plethora of possible structural factors revealed by the past research. The rest of this section discusses those factors along each dimension and shows how cognitive architectures implement them and use them to study complex multitasking scenarios that cannot be easily tackled by the multiple resource framework.

The perceptual modality dimension

The perceptual modality dimension is divided into visual and auditory modalities. Many studies have shown that concurrent cross-modal (e.g., auditory-visual) perceptual processing can be done more efficiently than concurrent unimodal processing (e.g., visual-visual). For example, Rollins and Hendricks (1980) showed that messages presented only through auditory stimuli are more difficult to process than messages presented partly through visual stimuli and partly through auditory stimuli. The fact that humans have dedicated organs (eyes and ears) and brain regions (Kandel, Schwartz, & Jessell, 2000) for processing the two types of sensory stimuli also suggests that there should be less interference across the two perceptual modalities. The question, though, is what interferences exist for information delivered within the same modality.

Parts of the intra-modal interferences arise from the physical constraints of the sensory organs. For vision, the main constraint is the limited size of the fovea region. While certain information (e.g., color and motion) can be perceived in the periphery, some other information (e.g., text) requires an eye movement to center the foveal vision to that object. Because an eye movement typically takes about 200 ms to initiate and 10-100 ms to complete (Duchowski, 2007), this lag alone can cause considerable interference between visual tasks (Meyer & Kieras, 1997b). For hearing, the physical constraint has to do with how sound is perceived by the mechanoreceptors in the ears (Kandel et al, 2000). Particularly, sounds with similar frequencies have strong interferences as they stimulate the same set of mechanoreceptors (Moore, 1986).

Selective attention is another source of the intra-modal interferences. Selective visual attention can bias processing towards one part of the visual periphery (Moore & Armstrong, 2003; Sperling & Melchner, 1978), and selective auditory attention can bias processing towards one side of the ears (Cherry, 1953). It is still not clear whether attention filters raw sensory information (also known as early selection model, Broadbent, 1958) or just affects higher-level analysis such as semantic analysis (also known as late-selection model, Deutsch & Deutsch, 1963), but the common implication of these two possible mechanisms is that information from the stimuli that have not been attended may be lost.

The higher efficiency of cross-modal perceptual processing and the interferences of the unimodal processing are implemented in ACT-R and EPIC. To implement parallel cross-modal processing, both architectures have separate processors for handling visual and auditory stimuli, and these processors can work concurrently with no interference. To implement the intra-modal interferences, however, the two architectures have taken different paths. ACT-R relies on selective attention to access information from the outside world. Since this attention mechanism can only be directed to one location at a time, an ACT-R model cannot simultaneously perceive multiple stimuli. By contrast, EPIC's implementation of intra-modal interferences focuses on the physical properties of the perceptual modalities. For example, visual information from the stimuli is filtered by the graded resolution of the retina, and is "selected" by eye movements rather than internal attention. EPIC omits the effect of selective visual attention because in most real world tasks, it is tightly coupled with eye movements (Findlay & Gilchrist, 2003).

Although neither ACT-R nor EPIC implements the frequency interference that occurs in the auditory modality, their auditory modules are capable of explaining another interference which may occur more frequently in HCI tasks. This interference is that when there is too much auditory information presented at the same time (e.g., listening to the GPS instructions while having a conversation with a passenger), only a certain amount can be processed and retained. In ACT-R and EPIC, the processed auditory information are stored in working memory and decays within a short span of time, usually around 2 seconds. To delay a memory item's decay in the auditory working memory, the model needs to employ a rehearsal strategy, which was demonstrated by Kieras, Meyer, Mueller, and Seymour (1999) using EPIC's auditory and vocal modules. The rehearsal strategy periodically subvocalizes items in the auditory working memory and then stores them again in the memory through a shortcut channel that exists in the architecture. Even with the rehearsal strategy, however, the amount of auditory information that can be retained is still limited because if there are too many chunks in the auditory working memory, some chunks may decay before they get subvocalized. This mechanism may be useful for modeling tasks that involve extensive dialogs, such as the telephone operator task (Kieras, Wood, & Meyer, 1997).

With the above perceptual processing characteristics implemented in the EPIC architecture, Kieras and Meyer (2000) successfully modeled a dual task experiment that employed extensive auditory and visual stimuli. The dual task consisted of a tracking task and a choice-reaction task presented on separated displays. The tracking task required the participants to center the tracking cursor on a moving target with a joystick. The choice-reaction task presented series of moving stimuli accompanied by spatialized auditory alerts to signal their change of status. Participants could only respond to a stimulus when it became active (denoted by a color different from white or black). EPIC's implementation of the visual and auditory processing is key to correctly modeling participants' performance of this task. For example, EPIC's implementation of the retina zones allows the model to simulate how participants efficiently identified the location of the active stimuli with peripheral vision. EPIC's implementation of the auditory processing allowed the model to simulate how participants switched tasks using auditory alerts as cues. This dual task was difficult to analyze using traditional statistical methods, but with cognitive modeling Kieras and Meyer were able to extract and study the effects of the visual and auditory interference between the tasks.

Overall, the perceptual modules of EPIC and ACT-R are capable of modeling very complex multitasking scenarios. Though some aspects still need improvement, the current implementation can automatically take into account a variety of human multitasking characteristics such as the independent processing of visual and auditory information, the capacity of visual processing over the entire visual field, and the capacity of auditory working memory. These implementations were shown to be vital for studying complex tasks that cannot be easily examined through traditional statistical methods. The following section shows that in the response modalities, cognitive modeling have achieved a similar success.

The code dimension

In Figure 3, the code dimension of Wickens' taxonomy consists of spatial and verbal codes. These two codes represent two hypothesized types of memory storage employed in cognitive and perceptual-motor processing. Under the assumption of separated spatial and verbal codes, the multiple resource theory predicts that two tasks utilizing the same codes would incur strong interference, whereas a combination of a spatial task and a verbal task would incur small interference. This prediction has been verified by a few studies. For example, Brooks (1968) found that when recalling information about a line diagram, a task that requires spatial thinking, it takes much longer for the participant to point at the response in a spatial array than to speak the response. By contrast, when recalling information about a sentence, it takes longer to speak the response than to point at it. Research on driving (Recarte & Nunes, 2000) also shows that spatial-imagery tasks would greatly reduce a driver's visual scan area, whereas verbal tasks would cause a much smaller impairment. These results suggest that spatial thinking and verbal thinking may indeed operate on different resources.

Although there is some evidence to support the separation between spatial and verbal codes at the perception and cognition stages, it is questionable whether this separation should be extended to the response stage. Wickens (2002) argued that the strong interference occurred between responses made by the same modality because they shared the same code. However, this interference could also be explained in that the responses share the same modality resources such as hands or voice. The latter explanation is arguably more parsimonious because it represents the physical and biological separation that exists instead of a hypothetical construct, that of processing codes.

Preferring the more parsimonious theory, EPIC and ACT-R adopted the shared-modality view in explaining interference that occurs at the response stage. Their implementation of the response modalities largely follows Rosenbaum's (1980) motor programming framework. In this framework, motor processing—including manual, ocular, and vocal processing—generally goes through three stages: preparation, initiation, and execution. Each stage only allows processing of one movement at a time. This movement-production bottleneck is the key to explaining the interference between concurrent responses. For example, it predicts strong interference between concurrent manual responses because both hands share the same manual motor processor (Kieras & Meyer, 1997), and the execution or preparation of one movement has to wait until that of the other movement is completed. A manual response and a vocal response, however, can be processed more efficiently because they are handled by separate processors and can be executed in parallel without interferences.

EPIC and ACT-R's implementation of motor programming can be used to explain many effects that were attributed to spatial-verbal codes, and can even predict when and how the interference may occur. For example, Martin-Emerson & Wickens (1997) conducted an experiment that consisted of a tracking task and an arrow-discrimination task. The arrow-discrimination task requires the participant to press a key in response to a left or right arrow that periodically appeared above the tracking task. The study found a considerable interference between the two tasks, and both tracking error and reaction time (RT) increase as the separation between the two tasks' display increases. Wickens (2002) attributed the interference to the sharing of the spatial processing codes. However, Kieras, Meyer, Ballas and Lauber (2000) found that a model with just the movement production bottleneck can also explain the effect. Their model shows that the two tasks, though both fall into Wickens' category of spatial tasks, do not always interfere with each other on all three processing stages. In fact, the response selection stage of the arrow-discrimination task was done while tracking was in progress. The multiple resource theory cannot make such detailed inference, nor can it predict how exactly the interaction unfolds in such multitasking scenarios.

The stage dimension

In the stage dimension, Wickens' taxonomy separates resources between the response stage and the perceptual-cognitive stages, but it does not separate resources between the perceptual stage and cognitive stage. Isreal, Chesney, Wickens, and Donchin (1980) provided some support for this way of resource separation. They showed that the manipulation of the response difficulty does not change the event-related potential (ERP) response that is assumed to be correlated with the perceptual-cognitive activities.

EPIC and ACT-R went one step further and assume that perceptual and cognitive processing stages also use different resources. In these two cognitive architectures, the cognitive processor and the perceptual processors operate independently without interferences. This separation between cognitive and perceptual processing resources is supported by many brain imaging and brain lesion studies (see Kandel et al., 2000), which show that decision-making mainly relies on the prefrontal cortex, whereas perception relies on other brain structures.

The prior discussion on perceptual modality dimension and code dimension has suggested that concurrent perceptual processes or concurrent responding processes would interfere with each other if they share the same modalities, but it remains to be answered whether there is interference among concurrent cognitive processes. ACT-R and EPIC take different positions on the issue of cognitive interferences: ACT-R assumes that the cognitive processor can only execute one rule within a cycle (a cognitive bottleneck), whereas EPIC assumes no limitations. This difference reflects a controversy on this issue that exists in the past few decades of psychological refractory period (PRP) studies.

The next section discusses modeling of PRP experiments. The PRP modeling studies serve as a great example of exploiting the various sensorimotor and cognitive functions implemented in cognitive architectures. These studies also reveal the potential of computational modeling in advancing cognitive science.

Putting together all the invariable factors

Figure 4 shows the structure of a typical PRP experiment and illustrates how the interaction of the two tasks might unfold under the view of two competing theories. In this graph, Task 1 and Task 2 start in close succession, with the stimulus of Task 2 appearing slightly after that of Task 1. The interval between the presentation of the two tasks' stimuli is called the stimulus onset asynchrony (SOA), and it is a primary factor manipulated in PRP experiments. In PRP experiments, the reaction time of Task 2 is often found longer than its single task performance, and two major theories have been proposed to explain this phenomenon. The response-selection bottleneck (RSB) theory (Pashler, 1989), depicted as Task 2 RSB in Figure 4, assumes that the response-selection stage is a single channel and only allows response selection for one task at a time. By contrast, the movement-production bottleneck (MPB) theory (Keele, 1973), depicted as Task 2 MPB, allows concurrent perception and response selection, but permits only one movement production at a time.

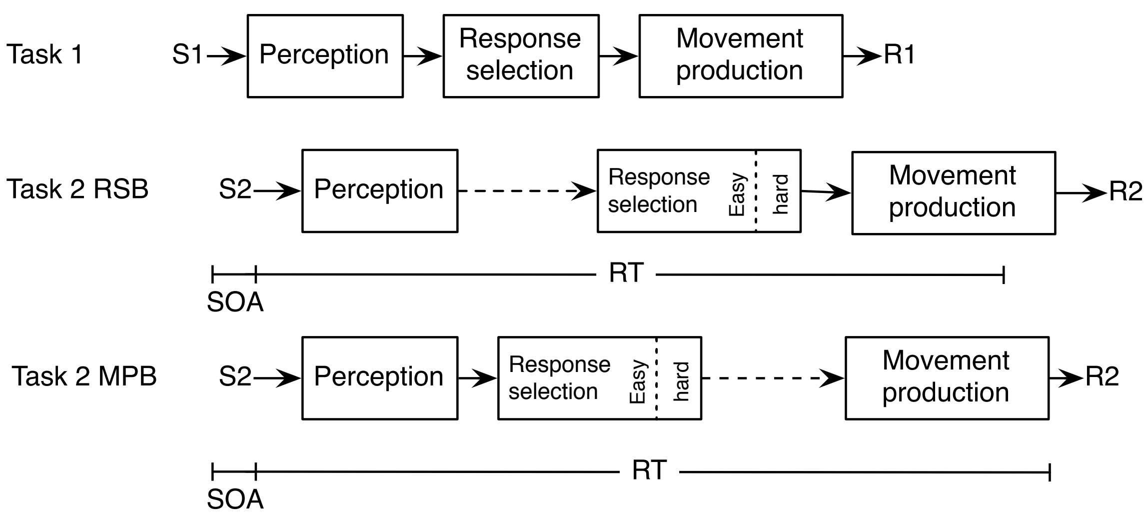


Figure 4. A stage model of the processing involved in PRP experiments. Task 2 RSB shows how the processing of Task 2 might proceed under the response-selection bottleneck hypothesis, and Task 2 MPB shows the view of the movement-production bottleneck theory.

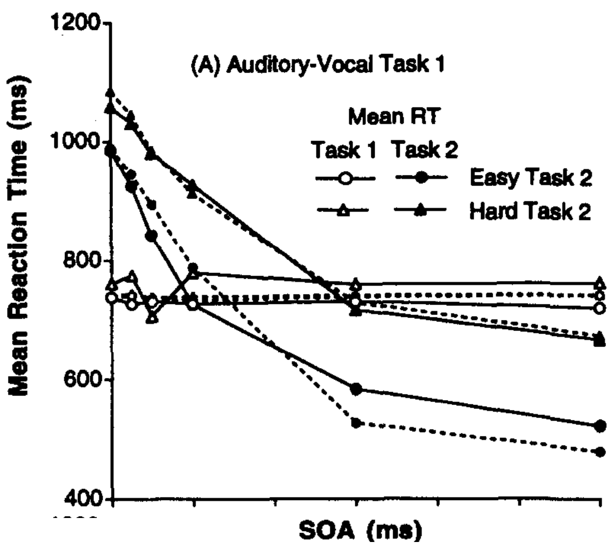
Though both theories can explain why Task 2 RTs are longer in dual task conditions than in single task conditions, they lead to different predictions when the difficulty of Task 2's response selection is varied. Specifically, the RSB theory predicts that the difference between the RT of an easy Task 2 and a hard Task 2 will be equal to the difference in the easy and hard response selection times, regardless of the duration of the SOA. In other words, if the RTs of the easy and hard Task 2 conditions were plotted against SOA, the RSB theory would predict that the curve of the easy condition should parallel the curve of the hard condition. By contrast, the MPB theory predicts diverging curves: At a short SOA, the easy and hard Task 2 would have the same RTs because the post-selection slack would absorb the difference in the two conditions' response selection times; at a long SOA, Task 2 Hard would have longer RT than Task 2 Easy because the post-selection slack would no longer exist.

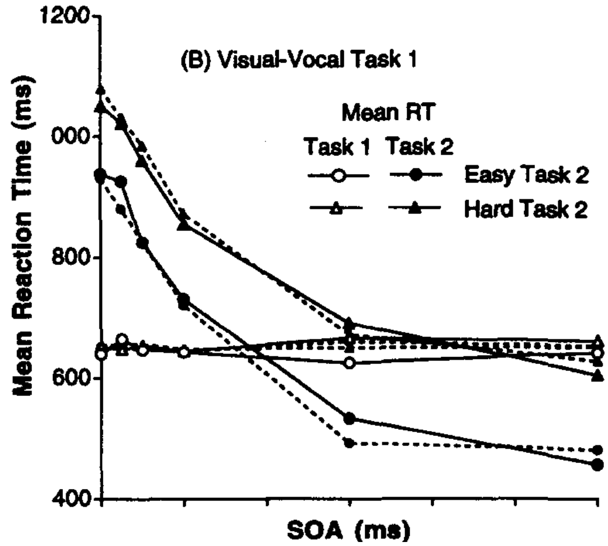
Given the predictions of the two theories, it would seem easy to reject one theory or another through experimentation, yet support for each theory has been observed (e.g., Becker, 1976 observed parallel curves, and De Jong, 1993 observed diverging curves) and no decisive conclusions could be drawn. The inconsistent results found in the PRP literature suggest that the simple flowcharts used by the RSB and MPB theories (such as those shown in Figure 4) could not adequately encompass the details of an experiment and could not provide accurate predictions to support either theory: a more comprehensive approach of modeling the PRP procedure was needed. Computational cognitive modeling seemed to just fit this task, because of its ability to produce quantitative predictions and to simulate the fine details of the environment, the task procedure, and human information processing.

Using EPIC, Meyer and Kieras (1997a, 1997b) constructed a series of computational models to account for various effects observed in several PRP experiments. The EPIC cognitive architecture does not assume a response-selection bottleneck or any other cognitive bottleneck, but it has implemented the movement-production bottleneck by allowing each modality to prepare or execute only one movement at a time. Meyer and Kieras's models adopted a strategy that intentionally defers executing Task 2's response when Task 2's response selection finishes earlier than Task 1's. This strategical response-deferment model guarantees that Task 2 finishes after Task 1, as is required by most PRP experiments.

With the movement-production bottleneck and the strategical response-deferment model, Meyer and Kieras found that the seemingly inconsistent PRP results can all be accounted for through carefully capturing the variations in the experimental setup. For example, they found that the reason that Becker's (1976) experiment observed parallel RT curves for the two Task 2 conditions is because Task 1's response selection was too easy (Task 1 in Becker's study was a binary choice task). When Task 1's response selection is very easy, there is essentially no slack time for Task 2 because Task 1 may have completed well before Task 2's movement production begins. Since there is no slack time, the difference between the RTs of Task 2 easy and hard conditions would be constant across SOAs. Similarly, parallel curves can also occur when perceiving Task 2's stimulus requires an eye movement, which delays the actual perception process and effectively increases SOA by about 125 ms (Meyer & Kieras, 1997b). Because of this delay, Task 1 may finish before Task 2's movement production begins even at short SOAs, which again results in zero slack time for Task 2 and produces parallel curves.

Figure 5 shows the results of Meyer and Kieras' models for Hawkins, Rodriguez and Reicher's (1979) three experiments. These experiments differed in the modalities used for delivering Task 1's stimuli and producing the responses. The first experiment presents Task 1's stimuli through audio, and requires participants' vocal responses. The second experiment uses a visual-vocal





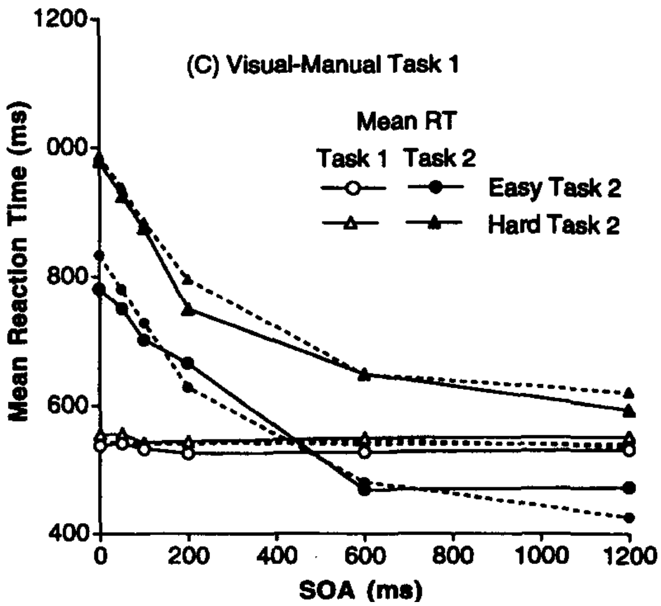


Figure 5. Results from Meyer and Kieras' (1997a) models for Hawkins et al.'s (1979) experiments. Three experiments were modeled, each with a different stimulus-response modality combination for Task 1. The solid lines are the empirical data, and the dashed lines are the model predictions. The open symbols represent the data for Task 1, and the solid symbols for Task 2.

stimulus-response combination, and the third experiment uses a visual-manual combination. As can be seen in the graph, the RTs of Task 2 show very different patterns across the experiments. The two curves (solid lines, solid symbols) of the Task 2 diverge in panel (A), but are almost parallel in panel (B) and (C). Such patterns have all been captured by the models. As can be seen in the graph, the predicted RTs closely followed the empirical data points in all three experiments for both Task 1 and Task 2. All models have yielded a very good fit, with most R^2 greater[1] than .95.

Several ACT-R models were also developed to explain the PRP effects (Anderson, Taatgen, & Byrne, 2005; Byrne & Anderson, 2001), and they have achieved a similar success. The ACT-R architecture has incorporated EPIC's motor programming framework, but ACT-R features a central cognitive bottleneck—only one production rule can fire during a cognitive cycle. However, the cognitive bottleneck had a very weak influence on the predictions of the PRP task performance, because the ACT-R models all assumed a short response selection process (50 to 100 ms, compared to EPIC models' 100 to 250 ms). Although the ACT-R models brought back some controversies to the issue of the cognitive bottleneck, they also confirmed that the assumption of the movement-production bottleneck is necessary (Howes, Lewis, & Vera, 2009).

Through computational modeling, researchers found strong support for the MPB hypothesis and have reconciled the seeming inconsistencies in the PRP empirical data. The computational models showed that even procedures as simple as PRP experiments may involve complex

[1] R^2 is a measure of goodness of fit, which indicates the proportion of variance in the response explained by the model. It ranges from 0 to 1, with 1 representing perfect fit between model predictions and the response data. It is calculated by subtracting from 1 the ratio between the residual sum of squares and the total sum of squares.

interactions that cannot be examined through simple flowcharts and qualitative estimations. Subtle changes in the experimental design can impact the results and only quantitative modeling can take into account these effects. Cognitive architectures have substantially eased the effort needed for quantitative modeling, because they take care of many basic human performance characteristics so that a modeler can consider the task on a higher level of abstraction such as task strategies.

In sum, the above discussion show EPIC and ACT-R have accounted for the major invariable factors that impact multitasking performance at various processing stages and modalities. Their implementation is largely consistent with the multiple resource theory. When the empirical findings are not conclusive, the two architectures have made reasonable assumptions that were validated by modeling laboratory tasks such as the PRP experiments. These models also demonstrate the advantages of computational modeling: It enables researchers to represent any detail of the experiment and to make quantitative predictions about the tasks, which in turn can be used to validate researchers' theoretical assumptions.

The invariable factors are just one aspect of the multitasking performance. The strategies with which the users carry out the tasks can also greatly shape the performance.

The Role of Strategic Control

Early theories regarding human multitasking believe that people adjust their performance through resource allocation. For example, the analysis of *performance resource function* (Norman & Bobrow, 1976) investigates the amount of resource invested in each task to determine the likelihood of interference. However, these early theories were developed based on the concept of a universal, homogeneous resource (unitary-resource theory) instead of the structural, heterogenous resources assumed by the multiple resource theories. Because the unitary-resource theories have already been replaced by the multiple resource theory, a new explanation of strategy control should be based on the control of concrete, structural factors that are embraced by the multiple resource theory.

Recently, more and more studies show that strategies that interleave tasks dictate performance (Brumby, Salvucci, & Howes 2009; Meyer & Kieras, 1997a, 1997b; Monsell, 2003). These strategies include decisions about when and how to switch tasks. In a PRP study, Schumacher et al. (2001) showed that just by instructing the participants to take a “daring” strategy to execute Task 2’s response as soon as possible, the SOA effect could be reduced and even eliminated for some participants. Different task-interleaving strategies can dramatically change the performance.

The production system used by most cognitive architectures are particularly suitable for studying strategies. This approach was first used as a formal method to describe humans’ problem solving processes by Newell and Simon (1972). A production rule is a condition-action statement. During each cognitive-processor cycle, which typically lasts a simulated period of 50 ms for most cognitive architectures, every production rule's conditions are tested to determine if they match the content in the working memory (or buffers in the case of ACT-R). If a rule's conditions are matched, the actions of the rule are then executed. Although building cognitive models in production rules is sometimes not as convenient as building models in general programming languages, a production system appropriately characterizes the human decision process as a collection of low-level stimulus-response (if-then) pairs. The cognitive processor's cycle time is set to 50 ms, appropriate for the scale of the processing that is simulated (Newell, 1994).

The rest of the section focuses on strategies in the context of cognitive modeling. The ability to precisely specify strategies in a formal language has greatly advanced the research of strategic control.

Executive processes

In cognitive modeling, *executive processes* are the strategies responsible for interleaving tasks. In ACT-R, the executive processes are usually embedded in task processes. For example, a production rule that executes a response for one task might directly prompt the preparation for another task. ACT-R models use this schema because the architecture can only execute one rule within a cycle, and managing task transitions with multiple rules would likely inflate the predicted task-completion time. By contrast, the executive rules in EPIC are often decoupled from the task rules because the communication cost between executive and task processes can be absorbed by the parallel execution of the rules. This decoupling gives EPIC the opportunity to closely study the executive processes.

Kieras, Meyer, Ballas and Lauber (2000) identified three typical types of laboratory multitasking scenarios, each of which requires the executive processes to manage and coordinate resources in different ways to achieve good task performance.

The first multitasking scenario identified by Kieras et al. consists of two discrete tasks performed in succession. This procedure, also known as *task switching*, requires participants to perform basic choice-reaction tasks with the same sets of stimuli and responses, but with different Stimulus-Response (S-R) mappings across blocks. A common effect observed in these studies is that changing S-R mappings increases RTs for the initial trials of a block, presumably because of the need to activate a new set of mappings. With EPIC, Kieras et al. built a cognitive model to account for the task-switching costs observed in their experiment. The costs are primarily predicted by an executive process which cleans up working memory and sets up S-R mappings for the next block of trials. This account was further elaborated by Altmann and Gray (2008), who constructed an ACT-R model to show that the task-switching costs can be explained by the architecture's memory activation processes. These processes were required to switch from an already established set of S-R mappings to another set of mappings from block to block. This type of executive processing is potentially very useful to model the impact of inconsistent user interface designs such as using non-standard GUI elements.

The second multitasking scenario identified by Kieras et al. involves two discrete tasks performed in parallel. A good example of this type of task is the PRP paradigm. As discussed earlier, Meyer and Kieras (1997a, 1997b) modeled several PRP tasks with the assumption of a movement-production bottleneck and a strategy that ensures Task 1 finishes before Task 2. The strategy does so by caching the response for Task 2 and then retrieving and executing the response later when Task 2 is unlocked. Kieras et al. considered the ability to cache responses as the main function of the executive process for the discrete concurrent tasks scenario.

The third multitasking scenario discussed in Kieras et al. (ibid) involves one or more continuous tasks. The continuous tasks can cause many resource conflicts because of the tasks' constant demands on perceptual-motor resources. One commonly-used laboratory continuous task is the tracking task (e.g., Martin-Emerson & Wickens, 1997; Kieras & Meyer, 2000). To reach a very good performance level on this task, participants need to constantly monitor the tracking target and adjust the position of the tracking cursor. When performed concurrently with other tasks that also require foveal vision and manual motor control, the tracking task is inevitably interrupted. For example, Wickens (1976) showed that when a tracking task was performed concurrently with a force-application task that required constant output from the manual motor, the tracking error increased by 29%.

To model multitasking scenarios that involve continuous tasks, Kieras et al. (ibid) constructed two types of executive processes, each of which manages task-interleaving in a distinct way. The first type of executive imposes a strictly sequential order between the continuous task and another task, whereas the second type of executive allows as much overlap between the tasks as possible. Kieras et al. used these two executives to model Martin-Emerson and Wickens' (1997) dual task. As discussed in the code dimension section, this dual task consists of a tracking task and an arrow-discrimination task. In the sequential-ordering model, the tracking task is immediately suspended when the arrow appears on the screen, and is resumed only after a response has been made to the arrow. In the partial-overlap model, the tracking task is resumed at an earlier point in time, just when the visual information has been acquired from the arrow. Thus, the partial-overlap model can perform the tracking task while selecting and preparing the manual response for the arrow-discrimination task, which greatly reduces interruption to the tracking task. The results of these two models show that the partial-overlap model fits the empirical data better than the sequential-ordering model, demonstrating that participants likely have used the overlapping strategy to enhance performance on both tasks.

While the above analysis suggested that it is possible to resolve resource conflicts by selecting efficient task interleaving strategies, the resulting executive processes were often tailored to individual experiments and could not be readily adapted to other tasks. In addition, it is often difficult to build an executive to manage a continuous task because of the numerous ways in which a continuous task may interact with another task. Thus, a *general executive*, a set of production rules that can manage a variety of tasks in different contexts, is desirable.

Inspired by the way computer operating systems (OS) manage processes, Kieras et al. (ibid.) proposed two *general executives*, which were referred to as conservative general executive and liberal general executive. The conservative general executive manages motor resources using a first-come first-serve algorithm. It assumes that the task processes are “impolite” and grab resources without asking for the general executive’s permission. Once a task acquires a resource, the executive simply blocks access to that resource. If two tasks request the same resource simultaneously, the resource is granted to the one with a higher priority. By contrast, the liberal general executive works with polite task processes. A polite task process always requests the executive before using a resource and would not proceed until permission is granted. This schema gives the general executive greater control because it allows the executive to actively suspend a low-priority task’s use of a resource and gives it to a high-priority task, increasing the efficiency of the high-priority task. The general executives were tested by modeling Martin-Emerson and Wickens’ (1997) dual task. The results show that even without any task-specific knowledge, the models still fit the data reasonably well. Particularly, the liberal general executive model’s predicted RT for the arrow-discrimination task had only 7.8% mean absolute error. The executives could be improved further by adding some task-specific knowledge such as when to preallocate motor resources. It is reasonable to assume that with practice, participants would become familiarized with the task structure and prepare for the upcoming task in advance. Thus, the general executives not only can serve as the basis for building a specialized executive, but are also potential candidates for modeling non-expert performance.

Threaded cognition (Salvucci & Taatgen, 2008) is a general multitasking theory that has been used in some ACT-R models. Threaded cognition is very similar to the above general executive approach in that it also assumes an OS style process management. In this theory, task processes compete for resources in a greedy, polite manner. That is, a task process requests resources as soon as possible when needed and releases them immediately when the resources are no longer required. If a task process requests a resource that is already being used, this request is simply put on hold until the it is released. If multiple task processes wait for the same resource, threaded cognition ensures that the process that has least recently used the resource will acquire the resource first. This conflict resolution policy is certainly not the only way people interleave tasks, but it is a simple, preliminary solution for balancing processing among tasks.

To implement this conflict resolution policy, threaded cognition needs to be built into the ACT-R cognitive architecture. This is different from the way general executives are implemented in EPIC, which are constructed as modularized production rules just like other task strategies. Although embedding threaded cognition into the architecture reduces the programming effort for other modelers, the implementation is less malleable and makes it hard to explore other task-interleaving strategies.

The expressiveness of production rules has enabled a prolific exploration of executive processes, but this flexibility also exacerbates the problem of verifying the assumptions of strategies and cognitive architectures. If analysts were allowed to arbitrarily make up strategies, they might build models that overfit one set of data but lack real predictive power for a general class of tasks. Methods that provide constraints and guidance for strategy exploration are thus required.

Constraining strategy exploration

When building cognitive models, analysts often need to conduct an extensive exploration in order to find a strategy that fits the observed data reasonably well. Traditionally, a strategy exploration process involves three steps. First, the analyst conducts task analysis and decomposes the task into many small steps. Second, the analyst proposes strategies for accomplishing each step of the task, and implements the strategies in production rules. Third, the analyst runs the model, and compares the model's predictions with the observed data. If the predictions do not fit, the analyst needs to go back to the second step, and propose different strategies that might improve the goodness of fit.

The above strategy exploration process has two problems. First, this process is post-hoc, it needs empirical data to improve the model and the model's predictive power. As a result, the model might only make accurate predictions for a small range of tasks that require the same set of strategies as the original task. Second, revising production rules introduces many degrees of freedom to cognitive modeling (Newell, 1994), generating numerous models that cover a large range of predictions. Despite that the predictions may constitute a large space, researchers often choose to present one or two models that produce the best fit to the empirical data in order to show the success of the models or the validity of the architecture. Without proper justification as to why certain strategies are chosen and other strategies are ruled out, such a practice would reveal little useful information about the capacity of human information processing or about efficient user interface design.

Kieras and Meyer (2000) argued that instead of obsessively searching for the best-fitting model, useful information could be derived from the wide range of predictions that are outlined by just the fastest-possible and the slowest-reasonable strategies. These two strategies are built on a base strategy that performs the task in a very straightforward way. The slowest-reasonable strategy may include additional steps that represent some characteristics of novice performance such as always verifying the visual or auditory feedback given by the device. By contrast, the fastest-possible strategy would fully exploit the capacity of the architecture and use every possible way to speed up the performance. According to Kieras and Meyer's *bracketing heuristic*, the actual performance should fall between the two extremes. A designer can use this approach to determine what range of performance might be possible for a system, and then make design decisions early on. For example, if the performance of the fastest-possible strategy is still slower than required, then the system design may be seriously flawed and should be improved even before conducting any user studies.

If the bracketing heuristic is applied to analyze microstrategies on the scale of few hundred milliseconds, it can also help identify the potential flaws in a system design. Gray and Boehm-Davis (2000) found that when a participant is motivated and given enough practice, he or she will adopt microstrategies that produce the optimal performance. The authors showed that even for a task as simple as clicking and moving a mouse, the participants are able to change their strategies at the millisecond level in response to the subtle changes of the task. Gray and Boehm-Davis concluded that if users could not reach the optimal performance as predicted by the fastest-possible strategy, then perhaps some design flaws were preventing them from adopting the optimal strategy. For example, the authors found that in a target-acquisition task that consisted of several unit tasks, the participants did not reach the predicted optimal performance for steps that required moving the cursor to a position at which the target of the next unit task would appear. The authors concluded that the suboptimal performance resulted from a lack of clear indications in the UI about which unit task the participants were currently doing and which would be the next. As a result, the participants could not know exactly where the next target would appear even after repetitive practice, and they had to adopt a slower strategy for this step. When the bracketing heuristic is applied at the millisecond level, it is easier to infer which part of the task and UI is affecting the performance because there are many fewer factors to consider within a span of few hundred milliseconds.

One challenge of inferring the microstrategies used by participants is to collect behavioral data on the scale of tens to hundreds of milliseconds. Such data could help reveal the immediate states of cognition, which in turn narrow the strategies that need to be explored. Mouse movements, as shown above, provide one source of such data; another source is eye movement data because it has a high temporal resolution and it is closely related to visual attention.

In modeling a variation of the NRL dual task, Hornof and Zhang (2010) vividly demonstrated how eye movement data can be used to guide strategy exploration in a way that is much more reliable than using reaction time data. The authors' initial model for the dual task used a hierarchical-sequential strategy, i.e. the model performs only one of the two tasks at any point in time. Such a strategy is usually incorrect for predicting skilled multitasking performance because participants, if motivated, can easily find many opportunities to overlap two tasks (e.g. they could perceive the stimulus of one task while executing the response for another). It turned out, however, this initial model fit the reaction time data very well. But the authors did not stop there. They examined the model's goodness-of-fit to the eye movement data, and found that it was in fact far off from how participants performed the tasks: The predictions were off the time-course of the eye movement data by 91%. Further examination of the eye movement data revealed that participants adopted a strategy that extensively overlaps the two tasks in order to reach high performance on both tasks. The authors modified the strategy accordingly, and the resulting model's average absolute error was reduced to only 10% for the eye movement data, and 7% for the reaction time data. This synergy between cognitive modeling and eye tracking demonstrated by Hornof and Zhang provides a promising approach to analyzing and modeling multitasking performance.

Using behavioral data combined with the bracketing heuristic is just one way to constrain strategy exploration; other principled ways are still needed, especially for improving the effectiveness of predictive modeling. The key to constraining strategy exploration without relying on behavioral data is to realize that people do not change their strategies haphazardly; instead, under many circumstances, they behave rationally and adapt their strategies to achieve certain goals. In light of this, Howes et al. (2009) developed the *cognitive bounded rational*

(CBR) analysis to help constrain strategy exploration. The CBR analysis assumes that people always try to maximize subjective expected utility—which is an individual's judgement or instinct about the usefulness of an outcome—under the constraints imposed by the task environment and individuals' perceptual-motor and cognitive capacities. In the context of psychological experiments, maximizing expected utility of a task may involve maximizing monetary rewards or minimizing time cost. In the context of real world tasks such as driving, maximizing expected utility tends to mean achieving the optimal balance among many goals, e.g. maintaining driving safety, maintaining social contact, and reducing commute time.

Under the assumption of cognitive bounded rationality, the model that gains the most utility—given that the model and the participant have the same cognitive and perceptual-motor parameters—should adopt a similar strategy to what the participant would adopt. Thus, it should be possible for analysts to find the correct model by maximizing the model's utility rather than by fitting the performance data. Howes et al. demonstrated this analysis by modeling Schumacher et al. (1999)'s PRP experiment. Indeed, they found that strategies that maximize each individual's payoff, as determined by the same payoff evaluation function used in the original experiment, also fit the reaction time data very well. A similar finding was reported by Gray, Sims, Fu, and Schoelles (2006), who found that the strategies that minimize the time cost of a task accurately predict the observed performance. These studies demonstrated that the CBR analysis is potentially a very effective approach for predictive modeling. The main drawback of the CBR analysis, however, is that for some tasks, people might not be able to arrive at the best strategies on their own (Fu & Gray, 2006). Despite this limitation, the CBR analysis is a very valuable technique that greatly enhances the predictive power of cognitive modeling.

Before cognitive modeling, experimental psychology lacked an effective means to study the role of strategic control. Researchers may have had some ideas about how participants accomplish a task, but there was no way to formally express these ideas and test them. The production system adopted by most cognitive architectures changed this by providing a formal language to describe hypothesized task strategies. Because the production rules also directly serve as the script for a model to execute a task, task strategies expressed in production systems can be tested via the model predictions. This is a significant step towards understanding the role of strategic control, particularly for situations as complex as multitasking scenarios. Though the expressiveness of production rules introduces many degrees of freedom to modeling, several methods were developed to constrain and guide exploration of the task strategies. These methods also sparked new insights into issues of strategy selection such as skill acquisition and optimal strategic control.

Recently, cognitive modeling was applied to more practical areas such as driving. The next section summarizes some major findings from studies on multitasking while driving, and discusses the contribution of cognitive modeling to this research area.

Multitasking While Driving

Visual and Cognitive Interferences

Prior research has generally identified two types of tasks that impair driving performance: tasks that draw visual attention away from the road, and tasks that introduce heavy cognitive workload. Although it seems that the visual and cognitive interferences introduced by these tasks were already predicted by the multiple resource theory, understanding how and when such interferences may occur would still require a detailed examination of the interactions in a driving task.

The two oft-studied driving tasks—hazard detection and maintaining lane position—may both be affected by in-vehicle tasks that require visual attention. As predicted by the multiple resource theory, the visual interference between the in-vehicle task and the hazard detection task, such as responding to a sudden deceleration of a lead vehicle or detecting pedestrians on the roadway, can be mitigated by reducing the distance between the display of the in-vehicle device and the driver's forward line of sight. Thus, head-up displays (HUD) generally reduce hazard response time and increase hazard detection rates compared to head-down displays (HDD) (Horrey, Wickens, & Consalus, 2005). The effect of display separation may be modeled by EPIC through its graded retinal availability function. For example, the flashing brake lights may be set to be detectable only within the parafovea (7.5º), or within a fluctuating zone which gets larger as the brake lights get larger (cf. Kieras, 2010). This way, if the model fixates a HDD, the hazard events may fall outside of their availably zones and would likely be missed.

In-vehicle tasks may also impair the performance of maintaining lane position. Unlike hazard detection, the impact of in-vehicle tasks on lane keeping cannot be mitigated by reducing the degree of display separation. Horrey et al. (2005) found that when participants are allowed to freely scan between the display of the in-vehicle device and the roadway, both HUDs and HDDs increase the variability of the lane-keeping error by a similar amount.

Modeling the lane-keeping task and its interference with the in-vehicle task requires understanding the detailed mechanism of steering control. Wilkie and Wann (2003) proposed a locomotion control theory which suggests that drivers use the retinal image of the moving pattern of the texture elements, also known as optic flow (Gibson, 1958), to maintain lane position and negotiate curves. Locomotion control via optic flow was not a new concept. It has been shown that blowflies use optic flow to achieve incredibly fast flight control (Bialek, Rieke, de Ruyter van Steveninck, & Warland, 1991). As well, humans use optic flow to control walking (Warren, Kay, Zosh, Duchon, & Sahuc, 2001). Although optic flow is acquired from peripheral vision, it requires the eyes to fixate a certain location in order to achieve the optimal locomotion control. Wilkie and Wann found that participants generally choose to look ahead 1 to 2.5 s and to "fixate a point close to the desired future path" (p. 683). This strategy produces smaller steering error than looking at a fixed point or visually tracking the center of the roadway. Their finding is also supported by Land and Lee (1994), who found that when negotiating curves, drivers generally look at the tangent point on the inside of each curve 1 to 2 s before entering it.

Based on the above findings, Salvucci (2006) developed a steering control model under the ACT–R cognitive architecture. This model assumes that the steering angle is computed based on the visual angles of two points: a near point and a far point. The far point can be one of the three targets depending on the task condition: (a) the location at 2 s ahead, as suggested by Wilkie and Wann (2003); (b) the tangent point of the upcoming curve, as suggested by Land and Lee (1994); or (c) the lead vehicle. The near point is set at a distance of 10 m from the vehicle. The decision to take into account the near point is motivated by the results of Land and Horwood (1995), which showed that the near segment of the road needs to be visible for maintaining low error in lane position. Salvucci (2006) showed that this two-point steering control model produces results that largely match the participants' steering angle and lane position profiles for a curve negotiation task. Despite its success, this model may have overestimated the time needed to complete one steering angle adjustment. Due to the constraints of ACT-R, this steering control model needs to alternate visual attention between the far point and the near point. In EPIC, however, the fixation to the near point would be unnecessary because optic flow at the near segment of the road can be perceived through peripheral vision while maintaining the gaze on the far point, which is also consistent with the eye tracking data acquired from Wilkie and Wann (2003) and Land and Lee (1994).

Although it is easy to see how the visual demands of in-vehicle tasks affect driving performance, the influence of their cognitive demands is less clear and requires more thoughtful experiments to investigate. Many studies have shown that cognitively demanding tasks such as working memory tests, mental imagery rotation, and algorithmic tasks may increase lane departure and drivers' brake-response time (Alm & Nilsson, 1995; Horrey & Wickens, 2004; McKnight & McKnight, 1993; Recarte & Nunes, 2000). Some real-world in-vehicle tasks such as talking on a cell phone also have similar interferences (Strayer & Drews, 2007; Strayer & Johnston, 2001). A few studies have suggested some possible mechanisms regarding how cognitive tasks may affect driving. Recarte and Nunes's (2000) study indicates that the mental imagery rotation task may cause the eyes to freeze, which in turn reduces the visual scanning behaviors and the likelihood to detect hazard situations. Strayer and Drews (2007) conducted a series of experiments to show that cell phone conversation, while it may not alter the visual scanning behavior, can significantly reduce the recognition probability for the objects fixated during driving. Strayer argued that it is this inattentional blindness effect, rather than the manual control of a cell phone, that primarily impairs driving performance.

Besides the empirical experiments shown above, some cognitive modeling studies also suggested a few alternative explanations for how cognitive workload may interfere with driving. For instance, Salvucci and Beltowska (2008) built a model to account for the effects of a memory rehearsal task on lane-keeping and brake-response time. This model uses threaded cognition to interleave the driving task process and the memory rehearsal task process. As a result, much of the interference is accounted for by the central cognitive bottleneck. That is, because the memory rehearsal task occupies many cognitive-processor cycles, the rules for steering and responding to brake lights are often postponed, thus increasing the lane-keeping error and brake response time. However, this explanation is somewhat contradictory to the assumption of cognitive bounded rationality, because when facing resource conflicts, participants should have postponed the memory rehearsal task rather than the driving task since the latter is more important. The authors' model, however, does not account for this adaptation. Another weakness of this model is that it relies on a controversial assumption, the central cognitive bottleneck. It is relatively easy to attribute multitasking interference to a cognitive bottleneck because there is often no data to dispute such explanations. But as shown in PRP research, the assumption of a cognitive bottleneck may prove to be gratuitous, and there might be other explanations that rely on better-established mechanisms. For example, one alternative explanation is that memory rehearsal tasks may influence driving by deactivating driving related rules such as regularly checking potential hazards and lane deviations.

Strategic Adaptation in Driving

Despite the visual and cognitive interferences imposed by the in-vehicle tasks, drivers can still, to some degree, maintain their driving task performance by adopting good task-interleaving strategies. One strategy that people often use is to limit the duration of each glance on the display of the in-vehicle device. For example, Dingus, Hulse, Antin, & Wierwille (1989) found that in their experiments, when interacting with traditional in-vehicle devices such as speedometers and climate control gauges, drivers tended to limit their glance duration to 1.6 s. Tasks that took longer than this limit were completed by multiple glances with intermittent steering control to maintain a low steering error. Klauer, Dingus, Neale, Sudweeks, and Ramsey's (2006) report, in which the authors analyzed a large amount of data from 100 vehicles, also confirmed that eyes-off-road durations of greater than 2 seconds would significantly increase individual crash risk. Tsimhoni and Green (2001) found that the limit of glance duration decreases as visual demands of driving increase. It appears as if drivers can develop a sense of how long the gaze may be safely directed away from the road depending on the driving conditions. Nevertheless, Horrey et al. (2005) found that this self-monitoring strategy is not always reliable because the gaze duration sometimes exceeds the limit when the in-vehicle tasks become difficult.

The strategies adopted by drivers are often consistent with the basic assumptions of the cognitive bounded rational (CBR) analysis, i.e. people tend to select strategies that strike a good balance between optimizing performance of the driving task and the performance of other concurrent tasks. For example, Antin, Dingus, Hulse, and Wierwille (1990) found that a moving map draws a driver's gaze more often than does a paper map. In their task, the drivers need to balance between the performance of navigation, i.e. finding the route, and the performance of driving. The participants apparently achieved a good balance in both the moving map and the paper map conditions: When the cost of accessing the paper map is high, they chose to memorize the map in a few glances so that they can maintain relative good navigation and driving performance; when the cost of accessing the moving map is low, they chose to look at the map more frequently during driving to maintain low navigation error and short total task-completion time. This shifting from using knowledge in-the-head to using knowledge in-the-world as the cost of information-access decreases is also predicted by soft constraints hypothesis (Gray et al., 2006), a theory that is in principle very similar to the CBR analysis.

Another example of drivers adopting the optimal strategy in daily tasks can be found in phone number dialing while driving. By conducting a CBR analysis, Brumby, Salvucci, and Howes (2009) showed that the most common strategy for dialing a ten-digit phone number, entering digits in chunks of 3-3-4, is indeed more efficient than chunking the digits in any other ways such as 2-2-2-3. They found that if a driver steers the vehicle after dialing the first three digits and again after another three digits, it will strike a very good balance between lane position control and dialing speed. Other strategies will either result in a much longer dialing time or larger lane deviation.

Most of the results from the above driving studies are consistent with the multitasking theories discussed in the previous section, and it is apparent that cognitive models offer insights into the complex nature of applied multitasking research. Although there are so far only a few modeling studies on driving, they have already started to show some promising results. Because of its abilities to integrate a variety of psychological theories and to produce quantitative predictions, cognitive modeling will likely be applied increasingly often in driving research and other applied psychological studies.

Conclusions

This paper has shown how computational cognitive modeling integrates various theories of multitasking, enables extensive investigation on human strategic decisions, and contributes to advancing both theoretical cognitive science and applied human performance research. By implementing the invariable sensorimotor and cognitive factors, and by pioneering the research of task strategies with a formal language, cognitive modeling has earned its indispensable position in multitasking research.

Compared to the multiple resource theory, cognitive modeling offers at least three major advantages in studying multitasking performance. The first advantage is that cognitive modeling can fully account for the effects of subtle changes in the experimental design (e.g., the timing of stimulus appearance), which helps address some seeming inconsistencies among empirical observations. Meyer and Kieras (1997a) have demonstrated that depending on the timing, location, and type of the stimulus of a PRP experiment, there are at least four distinct ways that the interactions of the two tasks could play out. Many prior PRP experiments, however, only studied one of the four ways. This inevitably led to contradictory conclusions. With cognitive modeling, the dynamics of the tasks can be directly derived from the specified experimental condition and task strategies. It is thus much easier for researchers to correctly and comprehensively deduce the implications and predictions from a computational simulation than from the mere textual descriptions of the theories and the experimental design.

The second advantage of cognitive modeling over the multiple resource theory is that a comprehensively specified cognitive architecture enables researchers to study the cognitive system as a whole rather than each component individually. Traditionally, a psychological study usually addresses only one phenomenon and deals with only one aspect of cognition. As a result, cognitive science tends to postulate new cognitive constructs to account for effects that might actually be emergent properties of particular strategies (Gray, 2007). Again, taking the PRP research as an example, previous cognitive psychologists only focused on the response selection stage and the movement production stage because the two bottleneck theories, response-selection bottleneck (RSB) and movement-production bottleneck (MPB), only have to do with these processing stages. To reconcile the conflicting observations found in the PRP literature, a plethora of hypotheses were proposed such as Kahneman's (1973) unitary resource model and De Jong's (1993) hybrid-bottleneck model that consists of both the RSB and MPB. None of these hypotheses, however, was very successful. Meyer and Kieras' (1997a) solution did not rely on new assumptions, but merely integrated the effects of the perceptual stage. They showed that the stimulus perception time and even eye movements to the stimuli can impact the results. Thus, cognitive architectures make the properties of a whole cognitive system more tangible, and lead to more parsimonious theories.

The third advantage of cognitive modeling over the multiple resource theory is its ability to make quantitative predictions. Traditionally, the progress of cognitive science involves comparing empirical results with a theory's qualitative predictions. Such qualitative predictions can either be about the patterns of the data, or about comparisons between similar experimental conditions (e.g., which condition's RT is longer). Although this procedure of developing cognitive science is useful for examining simple phenomena, it has become increasingly insufficient for validating cognitive theories that do not have immediate behavioral consequences. For example, it is difficult to validate the RSB theory through its qualitative prediction—the RT curves of the easy and hard Task 2 conditions are parallel. Because this prediction not only depends on the cognitive state, but also depends on perceptual and motor processing as well as the interaction between the two tasks. In other words, qualitative predictions of such theories are unstable. With cognitive modeling, however, quantitative predictions can be derived and there is no need for these predictions to have a prominent shape, because the data itself can be directly compared with the empirical results. As psychological research advances deep into the cognition territory, perhaps more and more theories would require quantitative validations from computational cognitive models.

Currently, both EPIC and ACT-R have the necessary constructs to account for the majority of phenomena observed in multitasking studies, but there remain controversies and unanswered questions. Consistent with many psychological theories, EPIC and ACT-R have separate perceptual, cognitive, and motor processors, which can operate in parallel without interfering with each other. However, the two architectures sometimes hold different views regarding the specific mechanisms of the interferences within each processor. For example, EPIC explains visual interference as the result of the limited size of the focal vision, whereas ACT-R explains it as the result of limited selective attention. As well, ACT-R assumes that only one rule can be executed within a cognitive-processor cycle, whereas EPIC currently assumes no cognitive bottleneck. Each of these implementations has its own virtues and is perhaps more suitable for modeling some tasks than others. Although combining different solutions (e.g., combining the degraded visual acuity function and selective visual attention) might be able to explain more data, it may also reduce the parsimony of the architectures. Thus, more modeling studies are needed to resolve the differences between the two architectures and to reach a unified cognitive theory.

A good complementary approach to building better architectural modules is to constrain and validate strategies. As discussed in this paper, exploring strategies to fit empirical data at the scale of several seconds would likely prove ineffective, because there might be many alternative strategies to fit the same set of data. To address this problem, two methods were proposed: (a) fitting the model to empirical data of small timescales, and (b) developing strategies under certain principles such as the cognitive bounded rational analysis. Both methods are very effective and their usage has led to important insights into human decision-making. Therefore, instead of introducing more variability into theory validation, cognitive modeling in fact creates a unique opportunity for investigating the flexibility of human cognition.

Future research directions

As suggested by the CBR analysis, strategies can be shaped by altering the task utility, but people's subjective estimate of a task's utility may not match the task's actual utility, which may lead them to select a suboptimal strategy. In laboratory experiments, the utility or the importance of a task is often conveyed through a payoff schema designed by the researchers. But in reality, this objective measure is often unavailable. Gray et al. (2006) showed that when not paid, people often adapt their strategies to minimize the time cost of a task. This bias towards efficiency, if it also extends to real world tasks, may lead to people's neglect of other important utility such as personal safety. Future research can explore whether an objective task utility that appropriately incorporates safety, efficiency and many other factors can be conveyed through sensory stimuli such as visual indicators or buzzes. If this is possible, they might have a variety of important applications such as helping drivers to maintain attention to driving safety.

Regarding the practical applications of multitasking theories, driving will likely continue to be an important research subject. As discussed earlier, some modeling studies of driving have already revealed insights that had not been discovered by empirical studies, but these models only addressed a few aspects of the task. For instance, these models often do not consider the hazard detection task. To model the hazard detection task, EPIC is perhaps more suitable than ACT-R because EPIC allows parallel processing of the visual stimuli in the peripheral vision, whereas ACT-R has to take time to shift its visual attention to examine each stimulus. Modeling driving also puts various multitasking theories such as threaded cognition to the test because driving imposes great demands on the cognitive and visual resources. Thus, building computational models of driving will not only benefit user interface design of the in-vehicle devices, but will also greatly motivate theory development for cognitive modeling.

Allen Newell, a great scientist in computer science and cognitive psychology, once said: "You can't play 20 questions with nature and win." (Newell, 1973, p. 1). He argued that to advance our understanding of psychology, we should not stick to the old way of examining one problem at a time. Instead, we need a unified cognitive system (a cognitive architecture) and a complete process model (a production system), and we must analyze complex tasks. During the last twenty years, his vision has gradually come true. More researchers have begun to examine increasingly complex tasks with cognitive modeling. This, however, is still just a beginning. One day, cognitive modeling will perhaps become a routine approach for developing psychological theories, and a tutor for designing efficient user interfaces.

References

Alm, H. and Nilsson, L. (1995). The effects of a mobile telephone task on driver behaviour in a car following situation. *Accident Analysis & Prevention, 27*(5), 707–715.

Altmann, E. M. and Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review, 115*(3), 602.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., and Qin, Y. (2004). An integrated theory of the mind. *Psychological review, 111*(4), 1036–1060.

Anderson, J. R., Taatgen, N. A., and Byrne, M. D. (2005). Learning to achieve perfect time sharing: Architectural implications of Hazeltine, Teague, & Ivry (2002). *Journal of Experimental Psychology, 31*(4), 749–761.

Antin, J. F., Dingus, T. A., Hulse, M. C., and Wierwille, W. W. (1990). An evaluation of the effectiveness and efficiency of an automobile moving-map navigational display. *International Journal of Man-Machine Studies, 33*(5), 581–594.

Becker, C. A. (1976). Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 2*(4), 556–566.

Bialek, W., Rieke, F., De Ruyter Van Steveninck, R. R., and Warland, D. (1991). Reading a neural code. *Science, 252*(5014), 1854–1857.

Broadbent, D. E. (1958). *Perception and communication.* New York, NY: Oxford University Press.

Brooks, L. R. (1968). Spatial and verbal components of the act of recall. *Canadian Journal of Psychology, 22*(5), 349–368.

Brumby, D. P., Salvucci, D. D., and Howes, A. (2009). Focus on driving: How cognitive constraints shape the adaptation of strategy when dialing while driving. In *Proceedings of ACM CHI 2009: Conference on Human Factors in Computing Systems* 1629–1638.

Byrne, M. D. and Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time-sharing. *Psychological Review, 108*(4), 847–869.

Card, S. K., Moran, T. P., and Newell, A. (1983). *The psychology of human-computer interaction.* Psychology Press.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the acoustical society of America, 25*(5), 975–979.

Craik, K. J. W. (1948). Theory of the human operator in control systems. II. Man as an element in a control system. *British Journal of Psychology, 38*(142--148), .

Deutsch, J. A. and Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological review, 70*(1), 80–90.

Dingus, T. A., Hulse, M. C., Antin, J. F., and Wierwille, W. W. (1989). Attentional demand requirements of an automobile moving-map navigation system. *Transportation Research Part A: General, 23*(4), 301–315.

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice.* New York: Springer.

Findlay, J. M. and Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing.* New York, NY: Oxford University Press.

Foyle, D. C., & Hooey, B. L. (2008). *Human performance modeling in aviation*. CRC Press.

Fu, W.-T. and Gray, W. D. (2006). Suboptimal tradeoffs in information seeking. *Cognitive Psychology, 52*(3), 195–242.

Gibson, J. J. (1958). Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology, 49*(3), 182–194.

Gray, W. D. (2007). Composition and control of integrated cognitive systems. In W. D. Gray (Ed.), *Integrated models of cognitive systems* (pp. 3–12). Oxford University Press.

Gray, W. D. and Boehm-Davis, D. A. (2000). Milliseconds matter: An introduction to microstrategies and to their use in describing and predicting interactive behavior. *Journal of Experimental Psychology: Applied, 6*(4), 322–335.

Gray, W. D., Sims, C. R., Fu, W.-T., and Schoelles, M. J. (2006). The soft constraints hypothesis: A rational analysis approach to resource allocation for interactive behavior. *Psychological Review, 113*(3), 461–482.

Hawkins, H. L., Rodriguez, E., & Reicher, G. M. (1979). *Is time-sharing a general ability?* (ONR Tech. Rep. No. 3). Eugene, OR: University of Oregon.

Hornof, A. J. and Zhang, Y. (2010). Task-constrained interleaving of perceptual and motor processes in a time-critical dual task as revealed through eye tracking. Philadelphia, Pennsylvania: *Proceedings of the 10th International Conference on Cognitive Modeling*; 2010.

Horrey, W. J. and Wickens, C. D. (2004). Driving and side task performance: The effects of display clutter, separation, and modality. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 46*(4), 611–624.

Horrey, W. J., Wickens, C. D., and Consalus, K. P. (2005). The distracted driver: Modeling the impact of information bandwidth, in-vehicle task priority, and spatial-separation on driver performance and attention allocation. Technical report #AHFD-05-11/GM-05-02. University of Illinois at Urbana-Champaign.

Howes, A., Lewis, R. L., and Vera, A. (2009). Rational adaptation under task and processing constraints: Implications for testing theories of cognition and action. *Psychological Review, 116*(4), 717–751.

Isreal, J. B., Chesney, G. L., Wickens, C. D., and Donchin, E. (1980). P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. *Psychophysiology, 17*(3), 259–273.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice Hall.

Kandel, E., Schwartz, J., and Jessell, T. (2000). *Principles of neural science.* McGraw-Hill Medical.

Karlin, L. and Kestenbaum, R. (1968). Effects of number of alternatives on the psychological refractory period. *The Quarterly Journal of Experimental Psychology, 20*(2), 167–178.

Keele, S. W. (1973). *Attention and human performance.* Pacific Palisades, CA: Goodyear.

Kieras, D. E. (2004). EPIC architecture principles of operation. Available from http://ai.eecs.umich.edu/people/kieras/docs/EPIC/EPICPrinOp.pdf

Kieras, D. E. (2010). Modeling visual search of displays of many objects: The role of differential acuity and fixation memory. In *Proceedings of the 10th International Conference on Cognitive Modeling* 127–132.

Kieras, D. E. and Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human--Computer Interaction, 12*(4), 391–438.

Kieras, D. E., & Meyer, D. E. (2000). The role of cognitive task analysis in the application of predictive models of human performance. In J. M. Schraagen, S. F. Chipman, V. L. Shalin, & V. Shalin (Eds.), *Cognitive task analysis* (pp. 237–260). Mahwah, NJ: Erlbsum.

Kieras, D. E., Meyer, D. E., Ballas, J. A., and Lauber, E. J. (2000). Modern computational perspectives on executive mental processes and cognitive control: Where to from here. , 681–712.

Kieras, D. E., Meyer, D. E., Mueller, S., & Seymour, T. (1999). Insights into working memory from the perspective of the epic architecture for modeling skilled perceptual-motor and cognitive human performance. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 183–223). Cambridge University Press.

Kieras, D. E., Wood, S. D., & 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 (TOCHI), 4*(3), 230–275.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data.* Washington, DC: National Highway Traffic Safety Administration, USDOT.

Land, M. F. and Lee, D. (1994). Where we look when we steer. *Nature, 369*(30), 742–744.

Land, M. and Horwood, J. (1995). Which parts of the road guide steering? *Nature, 377*(28), 339–340.

Lee, J. D., Caven, B., Haake, S., and Brown, T. L. (2001). Speech-based interaction with in-vehicle computers: The effect of speech-based e-mail on drivers' attention to the roadway. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 43*(4), 631.

Martin-Emerson, R., & Wickens, C. D. (1997). Superimposition, symbology, visual attention, and the head-up display. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 39*(4), 581–601.

McKnight, A. J. and McKnight, A. S. (1993). The effect of cellular phone use upon driver attention. *Accident Analysis & Prevention, 25*(3), 259–265.

McLeod, P. (1977). A dual task response modality effect: Support for multiprocessor models of attention. *The Quarterly Journal of Experimental Psychology, 29*(4), 651–667.

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*(1), 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*(4), 749–791.

Monsell, S. (2003). Task switching. *Trends in cognitive sciences, 7*(3), 134–140.

Moore, B. C. J. (1986). *Frequency selectivity in hearing.* London: Academic Press.

Moore, T. and Armstrong, K. M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. *Nature, 421*(6921), 370–373.

Navon, D. and Gopher, D. (1979). On the economy of the human-processing system. *Psychological review, 86*(3), 214–255.

Newell, A. (1973). You can't play 20 questions with nature and win. In W. G. Chase (Ed.), *Visual information processing* (pp. 238–308). New York: Academic Press.

Newell, A. (1994). *Unified theories of cognition.* Harvard University Press.

Newell, A. and Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice Hall.

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on human factors in computing systems: Empowering people* (pp. 249–256).

Norman, D. A. and Bobrow, D. G. (1976). On the analysis of performance operating characteristics. *Psychological Review, 83*(6), 508–510.

Pashler, H. (1989). Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psychology, 21*(4), 469–514.

Recarte, M. A. and Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of Experimental Psychology: Applied, 6*(1), 31–43.

Richtel, M. (2010, March). Gadgets in Emergency Vehicles Seen as Peril. *The New York Times.* Retrieved from http://www.nytimes.com/2010/03/11/technology/11distracted.html

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? a comment on theory testing. Psychological Review, 107(2), 358–367.

Rollins, H. A. and Hendricks, R. (1980). Processing of words presented simultaneously to eye and ear. *Journal of Experimental Psychology: Human Perception and Performance, 6*(1), 99–109.

Rosenbaum, D. A. (1980). Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General, 109*(4), 444–474.

Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 48*(2), 362–380.

Salvucci, D. D. and Beltowska, J. (2008). Effects of memory rehearsal on driver performance: Experiment and theoretical account. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 50*(5), 834–844.

Salvucci, D. D. and Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review, 115*(1), 101–130.

Sanders, M. S. and McCormick, E. J. (1987). *Human factors in engineering and design.* New York: McGraw-Hill.

Schumacher, E. H., Lauber, E. J., Glass, J. M., Zurbriggen, E. L., Gmeindl, L., Kieras, D. E., and Meyer, D. E. (1999). Concurrent response-selection processes in dual-task performance: Evidence for adaptive executive control of task scheduling. *Journal of Experimental Psychology: Human Perception and Performance, 25*(3), 791–814.

Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., and Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological Science, 12*(2), 101–108.

Smith, M. C. (1967). Theories of the psychological refractory period. *Psychological Bulletin, 67*(3), 202–213.

Sperling, G. and Melchner, M. J. (1978). The attention operating characteristic: Examples from visual search. *Science, 202*(4365), 315–318.

Strayer, D., & Drews, F. (2007). Multitasking in the automobile. In A. F. Kramer, D. A. Wiegmann, & A. Kirlik (Eds.), *Attention: From theory to practice* (pp. 121–133). Oxford University Press.

Strayer, D. L. and Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological science, 12*(6), 462–466.

Tsimhoni, O. and Green, P. A. (2001). Visual demand of driving and the execution of display-intensive in-vehicle tasks. In *Proceedings of The Human Factors and Ergonomics Society Annual Meeting, 2*, 1586–1590.

Warren, W. H., Kay, B. A., Zosh, W. D., Duchon, A. P., and Sahuc, S. (2001). Optic flow is used to control human walking. *Nature Neuroscience, 4*(2), 213–216.

Welford, A. T. (1952). The psychological refractory period and the timing of high-speed performance. *British Journal of Psychology, 43*, 2–19.

Wickens, C. D. (1976). The effects of divided attention on information processing in manual tracking. *Journal of Experimental Psychology: Human Perception and Performance, 2*(1), 1–13.

Wickens, C. D. (1980). The structure of attentional resources. In R. S. Nickerson (Ed.), *Attention and performance viii (pp. 239–257)*. Psychology Press.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159–177.

Wierwille, W. W. (1993). Visual and manual demands of in-car controls and displays. In B. Peacock & W. Karwowski (Eds.), *Automotive ergonomics* (pp. 299–320). Philadelphia, PA: Taylor and Francis.

Wilkie, R. M. and Wann, J. P. (2003). Eye-movements aid the control of locomotion. *Journal of Vision, 3*(11), 677–684.