Queueing Network Modeling of Human Performance and Mental Workload in Perceptual-Motor Tasks

by

Changxu Wu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Industrial and Operations Engneering)
in The Universty of Michigan
2007

Doctoral Committee:
Associate Professor Yili Liu, Chair
Associate Professor Nadine B. Sarter
Associate Professor Jun Zhang
Adjunct Assistant Professor Omer Tsimhoni

To my parents

Acknowledgements

I would like to gratefully acknowledge the enthusiastic supervision of Dr. Yili Liu for his inspiration and tremendous help and support in this thesis work at the University of Michigan.

Many thanks also to Dr. Omer Tsimhoni at the University of Michigan Transportation Research Institute for his great help in this thesis work. I would also like to express gratitude to Dr. Nadine Sarter at the Department of Industrial and Operations Engineering and Dr. Jun Zhang at the Department of Psychology who gave me valuable suggestions in this thesis and a lot of help in my study and research at the University of Michigan.

Much credit is also awarded all of professors, staff and students at the Department of Industrial and Operations Engineering at the University of Michigan and the University of Michigan Transportation Research Institute for their great assististance in this thesis work.

Last but not least, I am forever indebted to my family and friends for their love, understanding, endless encouragement and support.

Table of Contents

List of Figures

List of Tables

Abstract

Integrated with mathematical modeling approaches, this thesis uses Queueing Network-Model Human Processors (QN-MHP) as a simulation platform to quantify human performance and mental workload in four representative perceptual-motor tasks with both theoretical and practical importance: discrete perceptual-motor tasks (transcription typing and psychological refractory period) and continuous perceptual-motor tasks (visual-manual tracking and vehicle steering with secondary tasks). The properties of Queueing networks (Queueing/waiting in processing information, serial and parallel information processing capability, overall mathematical structure, and entity-based network arrangement) allow QN-MHP to quantify several important aspects of the perceptual-motor tasks and unify them into one cognitive architecture. In modeling the discrete perceptual-motor task in a single task situation (transcription typing), QN-MHP quantifies and unifies 32 transcription typing phenomena involving many aspects of human performance—interkey time, typing units and spans, typing errors, concurrent task performance, eye movements, and skill effects—providing an alternative way to model these basic and common activities in human-machine interaction. In quantifying the discrete perceptual-motor task in a dual-task situation (psychological refractory period), the Queueing network model is able to account for various experimental findings in PRP including all of these major counterexamples of existing models with less or equal number of free parameters and no need to use task-specific lock/unlock assumptions, demonstrating its unique advantages in modeling discrete dual-task performance. In modeling the human performance and mental workload in the continuous perceptual-motor tasks (visual-manual tracking and vehicle steering), QN-MHP is used as a simulation platform and a set of equations is developed to establish the quantitative relationships between Queueing networks (e.g., subnetwork's utilization and arrival rate) and P300 amplitude measured by ERP techniques and subjective mental workload measured by NASA-TLX, predicting and visualizing mental workload in real-time.

Moreover, this thesis also applies QN-MHP into the design of an adaptive workload management system in vehicles and integrates QN-MHP with scheduling methods to devise multimodal in-vehicle systems. Further development of the cognitive architecture in theory and practice is also discussed.

Chapter 1.
Introduction

Chapter Summary

This chapter introduces the main theme and structure of the thesis—modeling the perceptual-motor tasks with Queueing network approaches. After the practical and theoretical importance of modeling perceptual-motor tasks are introduced, a taxonomy of perceptual-motor tasks with its representative tasks in each category is proposed and following chapters in the thesis model each of these representative tasks. The Queueing Network-Model Human Processors (QN-MHP) is described in detail in this chapter, as a modeling platform of the following chapters.

1. Introduction

Different from high-level cognitive tasks (e.g., problem solving and language comprehension), perceptual-motor tasks are those tasks to use sensory information to guide physical movements, i.e., linking perceptual input to motor output including identification of stimuli, selecting responses and executing motor response (Chua, Weeks and Goodman, 1996). In practice, many tasks in human-machine interaction belong to perceptual-motor tasks by nature, for instance, tracking a moving target on a display using a trackball, listening to someone's vocal messages and typing the messages into a computer, pressing a button on a user interface with a mouse, grasping a virtual object in virtual reality wearing a data glove, driving a car on a road using a steering wheel, accelerator and break, manipulating a joystick to increase the altitude of an airplane, and using communication and control system to perform a remote and endoscopic surgery.

The most critical characteristic of perceptual-motor tasks is the coordination among different components (perceptual, cognitive and motor) in the cognitive system (Glaser, 1965; Redding & Wallace, 1996; Holden, Flach, & Donchin, 1999; Ohta, Luo, & Ito, 1997). Unlike the high-level cognitive tasks which might mainly focus on the cognitive stage in processing information, perceptual-motor behavior is an interaction among perceptual, cognitive and motor stages in information processing or at least an interaction between perceptual and motor stages: the perceptual stage provides information for the response selection process at the cognitive stage and response execution process at the motor stage, while the information processing rate that the sensory/perceptual stage depends on the processing rate that the cognitive and motor stages. This characteristic of perceptual-motor tasks is consistent with Allen Newell's main comments on psychology research—"You Can't Play 20 Questions with Nature and Win" (Newell, 1973), in which he advocates the development of unified theories of cognition (UTC) and made theoretical unification of micromodels and theoretical constructs an immediate and principal goal. Decomposing the cognitive system into many elements may allow researchers to study one of them thoroughly and accurately, however, it may also miss the important links and interactions among different elements while some and even the major properties of the cognitive system can be studied only if a model includes the major parts of the cognitive system together. Starting from 1980's, several important models and architectures have been developed along the line of UTC proposed by Newell. These models include the Model Human Processor (MHP) and the GOMS family of models (Card, Moran, Newell, 1983; John and Kieras, 1996a, 1996b; Olson and Olson, 1990), ACT-R and ACT-R/PM (Anderson and Lebiere, 1998), SOAR (Newell, 1990; Laird, Newell, Rosenbloom, 1987), CAPS (Just and Carpenter, 1992), and EPIC (Meyer and Kieras, 1997a, 1997b). Perceptual-motor tasks are one of the major groups of target tasks modeled by these architectures.

2. A Taxonomy of Perceptual-Motor Tasks with its Representative Tasks

In the study of perceptual-motor tasks, perceptual-motor behavior are typically classified into two categories: discrete and continuous motor activity (Schmidt, 1988; Health, Roy & Weir, 1999; Elliott, Carson, Goodman, & Chua, R. 1991): discrete

activities are those having a recognizable beginning and end and require that a single action be performed on each trial (e.g., pressing a key on a keyboard, throwing or shifting gears in a car); continuous activities have no recognizable beginning or end and may last for an arbitrary or a predefined period of time (e.g., steering a car, and flying a simulated aircraft).

In addition, the number of perceptual-motor tasks to be processed concurrently at the cognitive system is another important aspect to study perceptual-motor tasks. Single perceptual-motor tasks allow researchers to study the coordination among perceptual, cognitive and motor components for the same task without interference from the other tasks (Salthouse, 1986); while dual or multiple perceptual-motor tasks provide researchers a scenario to study the interference between the two tasks and the constraints of the cognitive system in processing the dual-task information (Meyer & Kieras, 1996a, 1996b).

Based on these two dimensions of perceptual-motor tasks (discrete vs. continuous motor activities and single vs. dual perceptual-motor tasks), a taxonomy of perceptual-motor tasks including the representative tasks is descried in

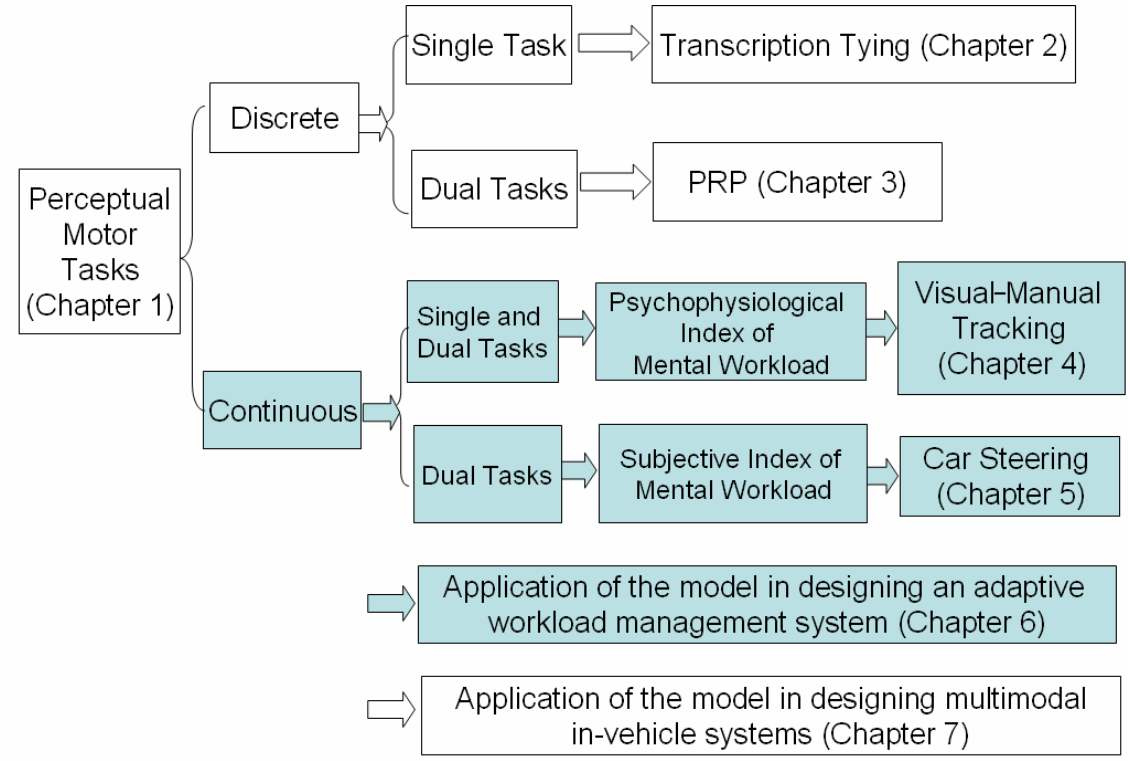
Table 1-1. Newell (1990) regards transcription typing as a representative discrete perceptual-motor task since transcription typing is composed of a series of discrete/single motor execution in pressing each key on the keyboard. Psychological refractory period (PRP) (subjects respond to stimuli of two choice reaction tasks which are presented within a very short period) is a basic dual perceptual-motor task (Meyer & Kieras, 1996a, 1997b). When motor responses are executed in a continuous manner with a joystick or a steering wheel, car steering and visual-manual tracking become representative tasks of continuous perceptual-motor tasks (Horrey & Wickens, 2004; Yucel, Petty, McCarthy, & Belger, 2005). In addition, these continuous perceptual-motor tasks often involve higher mental workload compared with discrete perceptual-motor behavior (Fenter, 2002). Therefore, both human performance and mental workload need to be considered in these continuous perceptual-motor behaviors.

Table 1-1 Representative Tasks in the Taxonomy of Perceptual-motor Tasks

Number of Tasks	Motor Responses	
	Discrete	**Continuous**
Single	Transcription Typing (Newell, 1990)	Visual-Manual Tracking (Yucel, et al., 2005) Car Steering (Horrey & Wickens, 2004)
Dual	Psychological Refractory Period (PRP) (Meyer & Kieras, 1996a, 1996b)	Visual-Manual Tracking with a secondary Task (Yucel, et al., 2005) Car Steering with a secondary Task (Horrey & Wickens, 2004)

3. Thesis Structure

To model the representative tasks in the taxonomy of perceptual-motor tasks, in the following section in this Chapter 1, a new computational architecture—Queueing Network Model-Human Processors (QN-MHP) is described in detail in this chapter. In the following chapters the thesis, integrating with mathematical modeling approaches, QN-MHP is applied to model the representative tasks in the taxonomy of perceptual-motor tasks and design adaptive multimodal in-vehicle systems (see Figure 1-1).



Shadowed Area: Model both human performance and mental workload

Figure 1-1 Structure of the Thesis

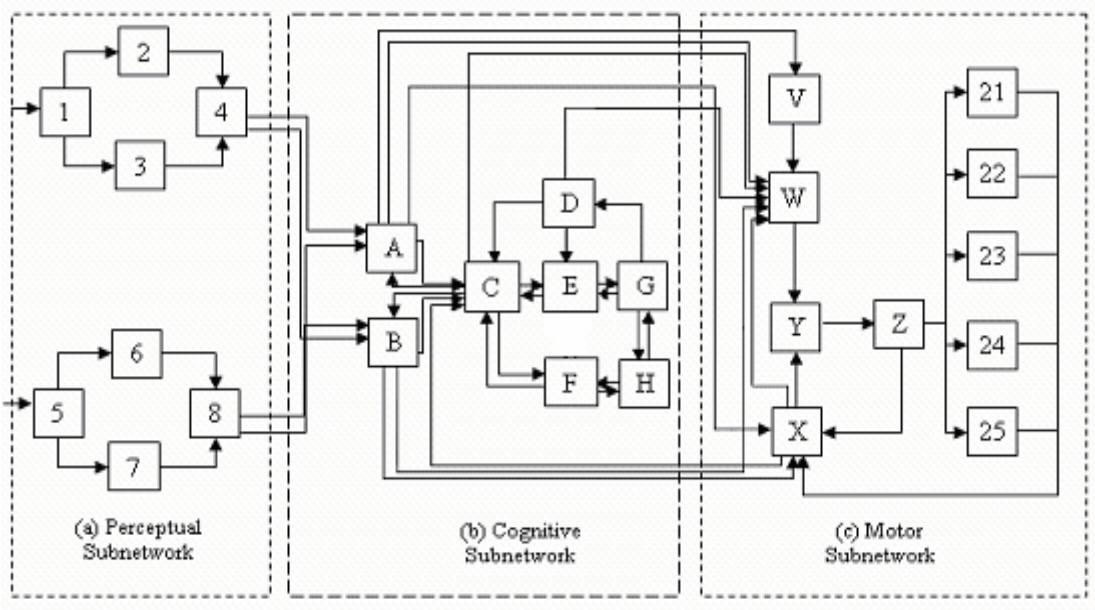
For the discrete perceptual-motor tasks, Chapter 2 and 3 focus on computational modeling of human performance in the transcription typing and psychological refractory period (PRP), respectively. Chapter 4 and 5 model both human performance and mental workload in the continuous perceptual-motor tasks: Chapter 4 describes how to model a psychophysiological index of mental workload and human performance in both single and dual-task situations in visual-manual tracking task; Chapter 5 quantifies subjective index of mental workload and human performance in driving (car steering while performing a secondary task). Chapter 6 applies the model developed in previous chapters to design an adaptive workload management system to reduce driver workload. Chapter 7 focuses on how to integrate the model with scheduling methods to design multimodal in-vehicle systems. Chapter 8 summarizes the major finding in this thesis and proposes topics for future research.

4. QN-MHP—A Computational Architecture to Model Perceptual-Motor Tasks

In modeling human performance, computational models based on Queueing networks have successfully integrated a large number of mathematical models in response time (Liu, 1996) and in multitask performance (Liu, 1997) as special cases of Queueing networks. Moreover, it unifies the two isolated major groups in reaction time models (e.g., Cascade model, program-evaluation-and-review-technique/PERT networks) and response accuracy models (e.g., accumulator, diffusion models)(Liu, 2005). A simulation model of a Queueing network mental architecture, called the Queueing Network-Model Human Processor (QN-MHP), has been developed to represent information processing in the mental system as a Queueing network on the basis of neuroscience and psychological findings. Ample research evidence has shown that major brain areas with certain information processing functions are localized and connected with each other via neural pathways (Bear, Connors, & Paradiso, 2001; Faw, 2003; Roland, 1993; Smith & Jonides, 1998), which is highly similar to a Queueing network of servers that can process entities traveling through the routes serially or/and in parallel depending on specific network arrangements. Therefore, brain regions with similar functions can be regarded as servers and neural pathways connecting them are treated as routes in the Queueing network (see Figure 1-2 and Figure 1-3). Further, it has been discovered that information processed in

the brain is coded in population spike trains (Rieke, Warland, R.S., & Bialek, 1997); depending on different tasks and learning stages, the to-be-processed information represented by these spike trains sometimes are processed by the brain regions (servers) immediately; sometimes they have to be maintained in certain regions to wait for the previous spike trains finishing their processing (Smith & Jonides, 1998; Taylor et al., 2000). Hence, these spike trains can be regarded as one type of entities in the Queueing network. Population spike trains transmitting through different brain regions require various neurotransmitters so that the electrical responses in the presynaptic neuron populations can be sent to the postsynaptic neuron populations (Bear et al., 2001; Haines, 2002); these neurotransmitters are regarded as the second type of entities in the Queueing network.

QN-MHP is a task-independent cognitive architecture and it has been successfully used to generate human behavior in real time, including simple and choice reaction time (Feyen, 2002), transcription typing (C. Wu & Y. Liu, 2004), psychological refractory period (C. Wu & Y. Liu, 2004b), visual search (Lim & Liu, 2004), and driver performance (Liu, Feyen, & Tsimhoni, in press). Moreover, QN-MHP is able to account for the brain imaging data in the transcription typing task (C. Wu & Y. Liu, 2004a).



Perceptual Subnetwork	Cognitive Subnetwork	Motor Subnetwork
1. Common visual processing 2. Visual recognition 3. Visual location 4. Visual recognition and location integration 5. Common auditory processing 6. Auditory recognition 7. Auditory location 8. Auditory recognition and location integration	A. Visuospatial sketchpad B. Phonological loop C. Central executive D. Long-term procedural memory E. Performance monitor F. Complex cognitive function G. Goal initiation H. Long-term declarative & spatial memory	V. Sensorimotor integration W. Motor program retrieval X. Feedback information collection Y. Motor program assembling and error detecting Z. Sending information to body parts 21-25: Body parts: eye, mouth, left hand, right hand, foot

Figure 1-2 The general structure of the Queueing network model

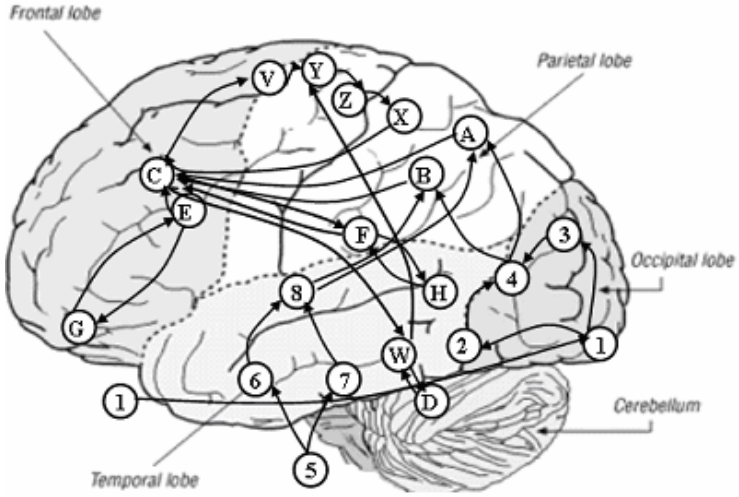


Figure 1-3 Approximate mapping of servers in the Queueing network model onto human brain

The QN-MHP consists of three subnetworks: perceptual, cognitive, and motor subnetworks as described in the following sections.

4.1 Perceptual Subnetwork

The perceptual subnetwork includes a visual and an auditory perceptual subnetwork, each of which is composed of four servers. In the visual perceptual subnetwork, light waves (represented by numerical codes) are transmitted to neuron signal (represented by information entities) at the eye, the lateral geniculate nucleus, the superior colliculus, the primary visual cortex, and the secondary visual cortex (represented by Server 1) (Bear et al., 2001). Then, these entities are transmitted in parallel visual pathways—the parvocellular stream (represented by Server 2) and the magnocellular stream (Server 3) where the object content features (e.g., color, shape, labeling etc.) and location features (e.g., spatial coordinates, speed etc.) are processed (Bear et al., 2001; Feyen, 2002; Simon et al., 2002; Smith & Jonides, 1998). The distributed parallel area (represented by Server 4)—including the neuron connections between V3 and V4 as well as V4 and V5, the superior frontal sulcus, and the inferior frontal gyrus—integrates the information of these features from the two visual pathways and generates integrated perception of the objects (Bear et al., 2001; Feyen, 2002).

The auditory perceptual subnetwork also contains four servers: the middle and the inner ear (represented by Server 5[1]) transmits sound to parallel auditory pathways, including the neuron pathway from the ventral cochlear nucleus to the superior olivary complex (represented by Server 7) and the neuron pathway from the dorsal and ventral cochlear nuclei to the inferior colliculus (Server 6) where location, pattern and other aspects of the sound are processed (Bear et al., 2001). The auditory information in the auditory pathways is integrated at the primary auditory cortex and the planum temporale (represented by Server 8) (Mustovic et al., 2003).

[1] Since the middle ear is located behind the eardrum and the inner ear is located in the temporal bone, the location of Server 5 is marked outside of the picture of the brain in Figure 3.

4.2 Cognitive Subnetwork

The cognitive subnetwork includes a working memory system, a goal execution system, a long-term memory system and a complex cognitive processing system.

Following Baddeley's working memory model, there are four components in the working memory system: a visuospatial sketchpad (Server A), representing the right-hemisphere posterior parietal cortex; a phonological loop (Server B), standing for the left-hemisphere posterior parietal cortex; a central executor (Server C), representing the dorsolateral prefrontal cortex (DLPFC), the anterior-dorsal prefrontal cortex (ADPFC), and the middle frontal gyrus (GFm); and a performance monitor (Server E), standing for the anterior cingulate cortex (ACC). The visuospatial sketchpad and the phonological loop store and maintain visuospatial and phonological information in working memory (Smith & Jonides, 1998).

The long-term memory system represents two types of long-term memory in the human brain: 1) declarative (facts and events) and spatial memory (Server H), standing for the medial temporal lobe including the hippocampus and the diencephalons which store various kinds of production rules in choice reaction, long-term spatial information, perceptual judgment, decision making, and problem solving; 2) nondeclarative memory (procedural memory and motor program) (Server D), representing the striatal and the cerebellar systems which store all of the steps in a task procedure and the motor programs related to motor execution (Bear et al., 2001).

The goal execution system (Server G) represents the orbitofrontal region, brain stem including the locus coeruleus-norepinephrine (LC-NE) system, and the amygdala complex which are typically involved in goal initiation and motivation (Rolls, 2000). And it sends the neurotransmitter entities to other servers following the NE output function in Nieuwenhuis et al.'s model (2005) (Nieuwenhuis, 2005).

The complex cognitive processing system (Server F) stands for the brain areas performing complex cognitive functions—multiple-choice decisions, phonological judgments, spatial working memory operations, visuomotor choices, and mental calculations. These brain areas include the intraparietal sulcus (IPS), the superior frontal gyrus (SFS), the inferior frontal gyrus (GFi), the inferior parietal cortex and the

ventrolateral frontal cortex, the intraparietal sulcus and the superior parietal gyrus (Fletcher & Henson, 2001; Manoach et al., 1997; Smith & Jonides, 1998).

4.3 Motor Subnetwork

The motor subnetwork includes 5 servers corresponding to the major brain areas in retrieval, assembling, and execution of motor commands as well as sensory information feedback. First, Server V represents the premotor cortex in Brodmann Area 6 which plays an important role in sensorimotor and sensory cue detection (Kansaku, Hanakawa, Wu, & Hallett, 2004; Mitz, Godschalk, & Wise, 1991; Roland, 1993). Second, the basal ganglia (Server W) retrieves motor programs and long term procedural information from long term procedural memory (Server D) (Bear et al., 2001; Cook & Woollacott, 1995; Gilbert, 2001). Third, the supplementary motor area and the pre-SMA (Server Y) have the major function of assemblying motor programs and ensuring movement accuracy (Gordon & Soechting, 1995). Fourth, the function of the primary motor cortex (Server Z) is to address the spinal and bulbar motor neurons and transmit the neural signals to different body parts as motor actuators (mouth, left and right hand, left and right foot server etc., (Roland, 1993)). Fifth, the S1 (the somosensory cortex, Server X) collects motor information of efference copies from the primary motor cortex (Server Z) and sensory information from body parts and then relay them to the prefrontal cortex (Server C) as well as the SMA (Server Y) (Roland, 1993).

In the following chapters, integrating with mathematical modeling approaches, QN-MHP is applied to model the representative tasks in the taxonomy of perceptual-motor tasks and design adaptive multimodal in-vehicle systems.

Reference

Anderson, J. R., & Lebiere, C. (1998). *The Atomic Components of Thought*: Lawrence Erlbaum Associates.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience, 28*, 403-450.

Bear, M. F., Connors, B. W., & Paradiso, M. A. (2001). *Neuroscience: exploring the brain* (8th ed.). Baltimore, MD: Lippincott Williams & Wilkins,.

Card, S., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hinsdale, NJ: Lawrence Erlbaum.

Chua, R., Weeks, D. J., & Goodman, D. (2002). Perceptual-Motor Interaction: Some Implications for Human-Computer Interaction. In J. A. Jacko & A. Sears (Eds.), *Human-Computer Interaction Handbook* (pp. 23-34): Lawrence Erlbaum Associates, INC.

Chubb, G. P., Laughery, K. R., & Pritsker, A. B. B. (1987). Simulating manned systems. In G. Salvendy (Ed.), *Handbook of Human Factors & Ergonomics*. New York: Wiley.

David, O., & Friston, K. J. (2003). A neural mass model for MEG/EEG: coupling and neuronal dynamics. *Neuroimage, 20*, 1743-1755.

David, O., Harrison, L., & Friston, K. J. (2005). Modeling event-related respones in the brain. *Neuroimage, 25*, 756-770.

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

Elliott, D., Carson, R. G., Goodman, D., & Chua, R. (1991). Discrete vs. continuous visual control of manual aiming. *Human Movement Science, 10*, 393-418.

Faw, B. (2003). Pre-frontal executive committee for perception, working memory, attention, long-term memory, motor control, and thinking: A tutorial review. *Consciousness and Cognition, 12*(1), 83-139.

Fenter, P. (2002). Understanding the Role of Practice in Learning for Geriatric Individuals. *Topics in Geriatric Rehabilitation, 17*(4), 11-32.

Glaser, R. (1965). *Training Research and Education*: University of Pittsburgh Press.

Gratton, G., Coles, M. G. H., Sirevaag, E. J., Eriksen, C. W., & Donchin, E. (1988). Pre- and post-stimulus activation of response channels: A psychophysiological analysis. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 331-344.

Gray, C. M., Freeman, W. J., & Skinner, J. E. (1986). Chemical dependencies of learning in the rabbit olfactory bulb: acquisition of the transient spatial pattern change depends on norepinephrine. *Behavioral Neuroscience, 100*(4), 585-596.

Haines, D. E. (2002). *Fundamental Neuroscience*. New York: Churchill Livingstone.

Hamilton, D. B., & Bierbaum, C. R. (1990). *Task analysis/workload (TAWL): AMethodology for predicting operator workload.* Paper presented at the Proceedings of the Human Factors Society 34th Annual Meeting, Santa Monica.

Harris, R. M., Glenn, F., Iavecchia, H. P., & Zaklad, A. (1986). Human operator simulator. In W. Karwoski (Ed.), *Trends in Ergonomic: Human Factors III (Part A)*. Amsterdam: North-Holland.

Heath, M., Eric, A., Patricia, R., & Weir, L. (1999). Visual-Motor Integration of Unexpected Sensory Events in Young and Older Participants: A Kinematic Analysis. *Developmental Neuropsychology, 16*(2), 197-211.

Holden, J. G., Flach, J. M., & Donchin, Y. (1999). Perceptual-motor coordination in an endoscopic surgery simulation. *Surgical Endoscopy, 13*(2), 127-132.

Horrey, W. J., & Wickens, C. D. (2004). Cell Phones and Driving Performance: A Meta-Analysis, *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*.

Jansen, B. H., & Rit, V. G. (1995). Electroencephalogram and Visual-Evoked Potential Generation in a Mathematical-Model of Coupled Cortical Columns. *Biological Cybernetics, 73*(4), 357-366.

John, B. E., & Kieras, D. E. (1996a). Using GOMS for user interface design and evaluation: Which technique? *ACM Transanctions on Human-Computer Interaction, 3*(4), 287-319.

John, B. E., & Kieras, D. E. (1996b). The GOMS family of user interface analysis techniques: Comparison and contrast. *ACM Transanctions on Human-Computer Interaction, 3*(4), 320-351.

Just, M. A., & Carpenter, P. N. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review, 99*, 122-149.

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence, 33*, 1-64.

Levison, W. H. (1979). A model for mental workload in tasks requiring continuous information processing. In N. Moray (Ed.), *Mental Workload: Its Theory and Measurement* (pp. 189-218). New York: Plenum.

Lim, J., & Liu, Y. (2004). *A queueing network model of menu selection and visual search.* Paper presented at the Proceedings of the 48 Annual Conference of the Human Factors and Ergonomics Society, New Orleans, Louisiana, USA.

Liu, Y. (1996). Queueing network modeling of elementary mental processes. *Psychological Review, 103*(1), 116-136.

Liu, Y. (1997). Queueing network modeling of human performance of concurrent spatial and verbal tasks. *IEEE Transactions on Systems Man and Cybernetics Part a-Systems and Humans, 27*(2), 195-207.

Liu, Y. (2005). *Queueing Network Modeling of Mental Architecture, Response Time, and Response Accuracy: Reflected Multi-dimensional Diffusions.* Paper presented at the Annual Meeting of Mathematical Psychology Society, Memphis, TN.

Liu, Y., Feyen, R., & Tsimhoni, O. (in press). Queueing Network-Model Human Processor (QN-MHP): A Computational Architecture for Multitask Performance. *ACM Transaction on Human Computer Interaction*.

Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance .1. Basic mechanisms. *Psychological Review, 104*(1), 3-65.

Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance .2. Accounts of psychological refractory-period phenomena. *Psychological Review, 104*(4), 749-791.

Moray, N. (1988). Mental workload since 1979. *International Review of Ergonomics, 2*, 123-150.

Moray, N., Dessouky, M. I., Kijowski, B. A., & Adapathya, R. (1991). Strategic Behavior, Workload, and Performance in Task-Scheduling. *Human Factors, 33*(6), 607-629.

Neff, N. H., Spano, P. F., Groppetti, A., Wang, C. T., & Costa, E. (1971). A simple procedure for calculating the synthesis rate of norepinephrine, dopamine and serotonin in rat brain. *The Journal of Pharmacology and Experimental Therapeutics, 176*, 701-709.

Nellgard, B. M. G., Miura, Y., Mackensen, G. B., Pearlstein, R. D., & Warner, D. S. (1999). Effect of intracerebral norepinephrine depletion on outcome from severe forebrain ischemia in the rat. *Brain Research, 847*(2), 262-269.

Newell, A. (1990). *Unified Theories of Cognition*: Harvard University Press.

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (in press). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin*.

Nieuwenhuis, S., Gilzenrat, M. S., Holmes, B. D., & Cohen, J. D. (2005). The role of the locus coeruleus in mediating the attentional blink: A neurocomputational theory. *Journal of Experimental Psychology-General, 134*(3), 291-307.

North, R. A., & Riley, V. A. (1989). W/INDEX: A predictive model of operator workload. In G. R. McMillan, D. Beevis, E. Salas, M. H. Strub, R. Sutton & L. V. Breda (Eds.), *Applications of Human Performance Models to System Design* (pp. 203-218). New York: Plenum.

Nunez, P. L. (1981). *Electric Fields of the Brain*. New York: Oxford University Press.

Ohta, K., Luo, Z. W., & Ito, M. (1997). Human perceptual-motor coordination in unknown dynamicalenvironments, *IEEE International Conference on Systems, Man, and Cybernetics 'Computational Cybernetics and Simulation'*.

Olsen, G., & Olsen, J. R. (1990). The Growth of Cognitive Modeling in Human-Computer Interaction Since GOMS. *Human-Computer Interaction, 5*(2&3), 221-265.

Olson, J. R., & Olson, G. M. (1990). The growth of cognitive modeling in human-computer interaction since GOMS. *Hum.-Comput. Inter, 5*, 221-265.

Parasuraman, R. (1990). Event-related brain potentials and human factors research. In J. W. Rohrbaugh, R. Parasuraman & R. Johnston, Jr. (Eds.), *Event-related brain potentials*. London, UK: Oxford University Press.

Parasuraman, R. (2003). Neuroergonomics: research and practice. *Theoretical Issues in Ergonomics, 4*(1-2), 5-20.

Parks, D. L., & Boucek, G. P., Jr. (1989). Applications of human performance models to system design. In G. R. McMillan, D. Beevis, E. Salas, M. H. Strub, R. Sutton & L. V. Breda (Eds.), *Applications of Human Performance Models to System Design* (pp. 203-218). New York: Plenum.

Paton, J. F. R., Foster, W. R., & Schwaber, J. S. (1993). Characteristic Firing Behavior of Cell-Types in the Cardiorespiratory Region of the Nucleus-Tractus-Solitarii of the Rat. *Brain Research, 604*(1-2), 112-125.

Piechulla, W., Mayser, C., Gehrke, H., & Konig, W. (2003). Reducing drivers' mental workload by means of an adaptive man-machine interface. *Transportation Research, Part F, 6*, 233-248.

Redding, G. M., & Wallace, B. (1996). Adaptive Spatial Alignment and Strategic Perceptual-Motor Control. *Journal of Experimental Psychology: Human Perception and Performance 22*(2), 379-394.

Rieke, F., Warland, D., R.S., R., & Bialek, W. (1997). *Spikes: Exploring the Neural Code (Computational Neuroscience)*: MIT Press.

Roland, P. E. (1993). *Brain activation*. New York, NY: Wiley-Liss.

Rolls, E. T. (2000). Memory systems in the brain,. *Annual Review of Psychology, 51*, 599-630.

Rouse, W. B. (1980). *Systems Engineering Models of Human-Machine Interaction*. New York: North Holland.

Rouse, W. B., Edwards, S. L., & Hammer, J. M. (1993). Modeling the Dynamics of Mental Workload and Human-Performance in Complex-Systems. *IEEE Transactions on Systems Man and Cybernetics, 23*(6), 1662-1671.

Rugg, M. D., & Coles, M. G. H. (1995). *Electrophysiology of Mind*. New York: Oxford University Press.

Schmidt, R. A. (1988). *Motor control and learning: A behavioral emphasis*. Champaign, IL: Human Kinetics.

Smith, E. E., & Jonides, J. (1998). Neuroimaging analyses of human working memory. *Proc. Natl. Acad. Sci. USA, 95*, 12061-12068.

Taylor, J., Horwitzc, B., Shaha, N. J., Fellenzb, W. A., Mueller-Gaertnera, H.-W., & Krausee, J. B. (2000). Decomposing memory: functional assignments and brain traffic in paired word associate learning. *Neural Networks, 13*, 923-940.

Tsang, P. S., & Vidulich, M. A. (2003). *Principles and practice of aviation psychology*. Mahwah, NJ.: Lawrence Erlbaum Associates.

Wickens, C. (1990). Applications of event-related potentials to problems in human factors. In J. W. Rohrbaugh, R. Parasuraman & R. Johnston, Jr. (Eds.), *Event-related brain potentials*. London, UK: Oxford University Press.

Wickens, C., Kramer, A., Vanasse, L., & Donchin, E. (1983). Performance of Concurrent Tasks - a Psychophysiological Analysis of the Reciprocity of Information-Processing Resources. *Science, 221*(4615), 1080-1082.

Wu, C., & Liu, Y. (2004a). *Modeling Behavioral and Brain Imaging Phenomena in Transcription Typing with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA.

Wu, C., & Liu, Y. (2004). *Modeling human transcription typing with Queueing network-model human processor.* Paper presented at the Proceedings of the 48th Annual Meeting of Human Factors and Ergonomics Society, New Orleans, Louisiana, USA.

Wu, C., & Liu, Y. (2004b). *Modeling Psychological Refractory Period (PRP) and Practice Effect on PRP with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA.

Wu, C., & Liu, Y. (2006a). *Modeling fMRI BOLD Signal and Reaction Time of a Dual Task with a Queueing Network Modeling Approach.* Paper presented at the 28th Annual Conference of the Cognitive Science Society, Vancouver, BC, Canada.

Wu, C., & Liu, Y. (2006). *Queueing Network Modeling of a Real-time Psychophysiological Index of Mental Workload—P300 Amplitude in Event-Related Potential (ERP).* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.

Wu, C., & Liu, Y. (2006b). *Queueing Network Modeling of Age Differences in Driver Mental Workload and Performance.* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.

Wu, C., & Liu, Y. (2006c). *Queueing Network Modeling of Driver Workload and Performance.* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.

Wu, C., & Liu, Y. (2006d). *Queueing Network Modeling of Reaction time, Response Accuracy, and Stimulus-Lateralized Readiness Potential Onset Time in a Dual Task.* Paper presented at the 28th Annual Conference of the Cognitive Science Society, Vancouver, BC, Canada.

Yucel, G., Petty, C., McCarthy, G., & Belger, A. (2005). Graded Visual Attention Modulates Brain Responses Evoked by Task-irrelevant Auditory Pitch Changes. *Journal of Cognitive Neuroscience, 17*(12), 1819-1828.

Chapter 2
Queueing Network Modeling of Transcription Typing

Chapter Summary

Transcription typing is one of the basic and common activities in human-machine interaction and 34 transcription typing phenomena have been discovered involving many aspects of human performance—interkey time, typing units and spans, typing errors, concurrent task performance, eye movements, and skill effects. Newell (1990) regarded transcription typing as one of the major tasks to be modeled by cognitive architectures. Based on the Queueing network theory of human performance (Liu, 1996, 1997) and current discoveries in cognitive and neural science, this paper extends and applies the Queueing Network-Model Human Processor (QN-MHP, Liu, Feyen and Tsimhoni, 2006) to model 32 transcription typing phenomena. The Queueing network model of transcription typing offers new insights into the mechanisms of cognition and human-computer interaction. Its value in proactive ergonomics design of user interfaces is illustrated and discussed.

1. Introduction

Despite the increasing popularity of speech recognition and handwriting systems (Wu, et al., 2003), typing is still one of the common activities in human-computer interaction (John and Newell, 1989; Lyons,et al. 2004). For example, people sometimes need to transcribe a manually written document into a computer using a standard keyboard. Drivers sometimes need to manually input the address of their target destination into a GPS system that is presented to them via a computer screen or voice messages from other people. Pilots need to manually input some textual flight control information into the

aircraft system based on voice messages from the air traffic controller. Police officers often need to input the plate number of a skeptical vehicle via a regular keyboard while they are driving. Cashiers have to type a UPC bar code of a product into a system using a keypad if the bar code can not be read automatically by a scanner.

Transcription typing involves intricate and complex interactions of concurrent perceptual, cognitive, and motoric processes (Salthouse, 1986a). Numerous studies in psychology (Salthouse, 1983; Salthouse, 1984a, 1984b, 1985, 1986a, 1986b; Salthouse and Saults, 1987), human-computer interaction (Card, et al., 1983; Duric, et al., 2002; Fish, et al., 1997; John and Newell, 1989; Pearson and van Schaik, 2003), and neural science (Gordon, et al., 1998) have been conducted to quantify transcription typing behavior and explore its underlying mechanisms. Several decades of research have identified numerous robust transcription typing phenomena including concurrent tasks, typing errors, visuomotor coordination, and skill acquisition. Salthouse (1986a) reviewed a majority of the experimental studies and summarized their findings as a list of 29 transcription typing phenomena (referred to as the Salthouse phenomena in this article). The availability of a wide range of experimental data and an extensive list of phenomena makes transcription typing one of the best candidate tasks to test theories and models of human performance. Modeling this rich and coherent set of behavior data and phenomena with the same set of assumptions and mechanisms is an important challenge to any theory or model of human performance. In practice, many human-computer interaction tasks involve the interaction of the perceptual, cognitive and motoric processes. Once a model can generate the interaction of these three processes and account for a wide range of phenomena in transcription typing, it can serve as a step towards modeling other tasks in human-computer interaction.

Inspired by Allen Newell's dream of unified theories of cognition (UTC) (Newell, 1973), researchers have developed several important UTCs or harbingers to UTCs, including the Model Human Processor (MHP) and the GOMS family of models (Card, Moran, Newell, 1983; John and Kieras, 1996a, 1996b; Olson and Olson, 1990), ACT-R (Anderson and Lebiere, 1998), SOAR (Newell, 1990; Laird, Newell, Rosenbloom, 1987), CAPS (Just and Carpenter, 1992), and EPIC (Meyer and Kieras, 1997a, 1997b). Newell (1990) regarded transcription typing as one of the major tasks to be modeled by cognitive

architectures. Although these architectures have been successfully applied to modeling a variety of tasks, it seems that this extensive set of 34 phenomena in transcription typing has not been modeled by any of these major existing cognitive architectures. John et al., (2004) mentioned that "skilled typing is not approximated well in the current implementation of ACT-R…, ACT-R currently moves its finger back to the home row after each keypress, resulting in much longer typing time than skilled typists normally achieve".

In this article we describe the application of a Queueing network based theory of cognition (Liu, 1996, 1997; Liu, Feyen, and Tsimhoni, 2006) in modeling transcription typing. Our model not only successfully accounted for a wide range of transcription typing phenomena, but can be used to simulate and analyze typing behavior and interfaces.

This article is organized as follows. In the remaining part of this introduction section, we first summarize the rich list of phenomena in transcription typing, followed by a summary of existing models. In the second section, we describe the Queueing network model in general and its application in typing modeling in particular. In the third section, we describe the mechanisms and results of simulating transcription typing with the Queueing network model. In the fourth section, we illustrate some of the potential applications of the model in HCI interface design, and the implications of the research are further discussed in the final section.

2. Phenomena in Transcription Typing and Existing Models

2.1 Phenomena in Transcription Typing

After Salthouse's (1986a) review of the 29 behavioral phenomena in transcription typing, additional phenomena have been identified and summarized. John (1988) summarized 2 behavioral phenomena discovered by other researchers (Gentner, 1983; John, 1988; Salthouse and Saults, 1987). In addition, three eye movements phenomena and one neural imaging pattern in transcription typing have been discovered (Inhoff, et al., 1992; Rayner, 1998). These 34 phenomena are introduced in Table 2-1 as six categories

(Salthouse, 1986a), including basic phenomena, units of typing, typing error, skill effects, and eye movements phenomena (see Table 2-1).

Table 2-1 Phenomena in Transcription Typing

Category	Phenomena	Category	Phenomena
1) Basic	1. Typing is faster than choice reaction time 2. Typing is slower than reading 3. Typing skill and comprehension are independent * 4. Typing rate is independent of word order 5. Typing speed is slower with random character order 6. Rate of typing is severely impaired by restricted preview window 7. Alternate-hand keystroke are faster than the same-hand keystroke * 8. More frequent character pairs are typed more quickly 9. Interkey time is independent of word length 10. The first keystroke in a word is slower than subsequent keystrokes 12. A concurrent task does not affect typing	3) Errors	18. 40%-70% of typing errors are detected without reference to the typed text 19. Many substitution errors involve adjacent keys 20. Intrusion error mostly short interkey time 21. Omission error mostly long interkey time 22. Transposition error mostly occur cross-hand
		4) Skill Effects	23. Two-finger digram improves faster than one-finger digram * 24. Repetitive tapping rate increases with skill 25. Variability decrease with skill 26. Eye-hand span increases with skill 27. Replacement span increases with skill 28. Copy span is depend on skill 29. Stopping span increases with skill §* 31. Learning curve follows power law of practice
2) Units of Typing	13. Copy span is 7-40 characters 14. Stopping span is one or two characters 15. Eye-hand span is 3-8 characters 16. Eye-hand is smaller for meaningless material than for then normal text 17. Replacement span is about 3 characters §30. Detection span is about 8 characters	5) Eye Move-ments	§32. Gaze duration per character decreased with enlarging of preview window size §33. Mean saccade size is about 4 characters §34. Fixation duration is around 400 ms

§ Phenomenon beyond Salthouse's review (1986a).

* Qualitative phenomena: existing experimental studies only reported the significance levels of comparisons between different conditions rather than detailed values of dependent variables.

2.1.1 Basic Phenomena

The following 12 behavioral phenomena are categorized as the basic phenomena in transcription typing by Salthouse (1986a) and they are related to the major factors affecting the interkey time, comparison of transcription typing with other tasks, and concurrent tasks in typing. Interkey time refers to the interval between two adjacent keystrokes, and is regarded as the basic measurement of human performance in transcription typing.

- Phenomenon 1: Typing is faster than choice reaction time.

Salthouse (1984a) reported that the median interkey time for skilled typists was 177 ms, whereas the typical reaction time for a two choice reaction time task is about 300-400 ms. Based on Hick's law on choice reaction time, a typical binary choice reaction time is $150+170\times\log_2(2)=320$ ms (Schmidt, 1988).

- Phenomenon 2: Typing is slower than reading.

Salthouse (1984a) found that the reading speed of the typists in his experiments was 253 words per minute (wpm), but their typing speed was only 58 words per minute.

- Phenomenon 3: Typing skill and comprehension are independent.

Involvement of comprehension is optional while typing (Salthouse, 1986a). Nonsignificant correlations were reported between net typing speed and comprehension scores obtained in typing (Salthouse, 1984a).

- Phenomenon 4: The rate of typing is nearly the same for random words as it is for meaningful text.
- Phenomenon 5: The rate of typing is slowed as the material approaches random.

The difference between phenomena 4 and 5 is that the former refers the order of the words being randomized while the latter refers to the order of characters within each word being randomized. Hershman (1965) found that the average interkey time in typing increased to 454 ms when subjects are typing materials composed of words with random characters (Hershman and Hillix, 1965).

- Phenomenon 6: The rate of typing is severely impaired by restricted preview of the to-be-typed material.

Decreasing the number of characters to-be-typed in the restricted preview increased the interkey time and severely impaired the typing rate.

- Phenomenon 7: Alternate-hand keystrokes are faster than the same-hand keystrokes (called the alternate-hand advantage).

Successive keystrokes from fingers on alternate hands are 30-60 ms faster than successive keystrokes from fingers on the same hand.

- Phenomenon 8: Digram (letter pairs) that occur more frequently in normal language are typed faster than less frequent digram (called the digram frequency effect).

The significant difference in typing speed between the low-frequency digrams and the high-frequency digrams has been reported in numerous studies (Salthouse, 1984a, 1984b).

- Phenomenon 9: Interkey time is independent of word length.

Salthouse (1986a) summarized several experiments in transcription typing and found no significant difference between the interkey time in typing long words and short words.

- Phenomenon 10: The first keystroke in a word is slower than the subsequent keystrokes (called the word initiation effect).

Salthouse (1986a) reviewed 5 researchers' experiments and found that the interval before the first keystroke in a word is approximately 20% (45 ms, Salthouse, 1984a) longer than that between the later keystrokes in the word.

- Phenomenon 11: The time for a keystroke is dependent on the specific context in which the character appears, especially for the topography of the keyboard (called the context phenomenon).

The "specific context" here refers to the character ahead of and behind the target character. The context phenomenon is a combination of the alternate-hand advantage (phenomenon 7), the digram-frequency effect (phenomenon 8), the word-initiation effect (phenomenon 10), and more specifically, the effect of topography of the keyboard in interacting with prior and subsequent keystrokes. For example, in typing the key sequence "r-e", the close proximity of the two keys "r" and "e" in the same row on a standard QWERTY keyboard allows the middle finger on the left hand to move toward the target "e" while the index finger on the left hand is typing character "r", which may save half of the movement distance of the middle finger from the home position "d" to the target position "e".

- Phenomenon 12: A concurrent task does not affect typing performance.

For highly skilled typists, a concurrent activity can be performed with little or no effect on the speed or accuracy of typing. Salthouse and Saults (1987) added a secondary task in parallel with the primary task of transcription typing: typists were asked to press a foot pedal as soon as they heard a tone signal (Salthouse and Saults, 1987). They found that the interkey time in this concurrent task situation was 185 ms, which was not significantly longer than that in a single task situation (transcription typing only, interkey time 181 ms).

2.1.2 Units of Typing

This group contains six phenomena related to the various "spans and units" of typing (defined below), five of which appeared on the original list of Salthouse (1986a) and the last one was identified after the list was published (Salthouse, 1987). It was regarded as one of the post 29 phenomena (phenomenon 30) by John (1996).

Phenomenon 13: Copying span is 2-8 words or 7-40 characters for all typists

Copying span is the amount of characters that can be typed accurately after a single inspection of the copy (Salthouse, 1986a). Without requiring the typists to commit the to-be-typed material to memory before typing or by randomizing the order of the words, Salthouse (1985) measured the copying span as the number of characters typed correctly after an unexpected disappearance of the copy and found that the copying span in normal transcription typing situation was 14.6 characters on average for the skilled typists.

- Phenomenon 14: Stopping span is one or two characters

Stopping span is the number of keystrokes typed after the subjects were requested to terminate their typing immediately after perceiving a stop signal. Using an auditory stop typing signal, Logan (1982) found that the stopping span was 2.16 characters when the typing materials were sentences.

- Phenomenon 15: Eye-hand span is 3-8 characters

Eye-hand span is defined as the number of characters intervening between the character whose key is currently being pressed and the character receiving the attention of the eyes (Salthouse, 1986a). Butsch (1932) found that the eye-hand span was 5 characters. The result is consistent with the other studies reviewed by Salthouse (1986a) who found that the range of eye-hand span is between 3 to 8 characters.

- Phenomenon 16: Eye-hand span is smaller for unfamiliar or meaningless material than for normal texts

When typists were typing a text and each word in it was composed of randomly ordered letters, Salthouse (1984a) found that their eye-hand span was only 1.75 characters.

- Phenomenon 17: Replacement span is about 3 characters

The subjects in Salthouse and Saults (1987) were asked to type exactly what appeared on the screen where one of the to-be-typed characters could be suddenly replaced by another character. Replacement span is defined as the keystroke-replacement interval corresponding to a 0.5 probability of typing the second (replaced, i.e., newly appeared) character. The replacement span was 2.9 characters on average (Salthouse, 1986a).

- Phenomenon 30: Detection span is about 8 characters

In the experiment of Salthouse (1987), subjects were asked to press the "/" key when they noticed a capital character on the line. The detection span is defined as the number of characters intervening between the capital character and the character currently being typed. The observed mean detection span was 8 characters approximately.

2.1.3 Errors in Transcription Typing

Salthouse (1986a) classified the vast majority of typing errors into four categories: substitution (e.g., work for word), intrusion (e.g., worrd for word), omission (e.g., wrd for word) and transposition (e.g., wrod for word). He summarized five major typing error phenomena related to these four categories of errors.

- Phenomenon 18: 40%-70% of typing errors are detected without reference to the typed text.

After reviewing three studies in transcription typing, Salthouse (1986a) summarized that about 40%-70% of typing errors are detected without reference to the typed copy. In his review (1986a), Salthouse suggested that typing errors include: a) undetected errors which can be postulated to originate at earlier levels of processing (errors mainly caused by failure to preserve sequences in the sensory and working memory) and b) detected errors without reference to the typed copy which probably stem from later stages of processing (hand and finger movement) that are handled by the efferent response feedback.

- Phenomenon 19: Many substitution errors involve adjacent keys.

Experimental results from highly skilled typists indicated that 30.1% of substitution errors involved horizontally or vertically adjacent keys (Salthouse, 1986a).

- Phenomenon 20: Many intrusion errors involve extremely short interkey time in the immediate vicinity of the error

Nearly 38% of the intrusion error keystrokes had ratios (interkey time of an error keystroke divided by that of the regular interkey time) less than 0.1 of the average interkey time (Salthouse, 1986a) and over 54% of intrusion errors involved an adjacent key in the same row or the same column.

- Phenomenon 21: Many omission errors are followed by a keystroke interval approximately twice the overall median.

Salthouse (1986a) summarized this phenomenon based on Shaffer's study (1975) which found that the interkey time of the keystroke right after the omission error was 1.54 times longer than that of the average interkey time (Shaffer, 1975).

- Phenomenon 22: Transposition errors mostly occur cross-hand.

Salthouse (1986a) reported that 80% of the transposition errors were typed by the opposite hands.

2.1.4 Skill Effects in Transcription Typing

Salthouse (1986a) summarized seven phenomena related to the improvement of typing performance via practice. In addition, Gentner (1983) found another related phenomenon—the interkey time of transcription typing decreases with practice following the power law (Gentner, 1983), which is listed below as one of the post 29 phenomena (phenomenon 31).

- Phenomenon 23: Digrams typed with two hands (two-hand digrams) or with two different fingers of the same hand (two-finger digrams) exhibit greater changes with the skill level of typists than do digrams typed with one finger.

Salthouse (1984a) found that the slope of the regression equations relating the digram interval to typing speed of two-hand digrams (-2.08) and two-finger digrams (-2.38) were greater than that of one-finger digrams (-1.38 on average).

- Phenomenon 24: Repetitive tapping rate increases with the skill level of typists.

Salthouse (1984a) found a significant positive correlation between the tapping rate and the net typing speed ($p<.01$).

- Phenomenon 25: The variability of interkey time decreases with the skill level of typists.

Salthouse (1984a) found that two types of variability of the interkey time (75% quartile-25% quartile) decreased with an increase in typists' skill level: a) Inter-keystroke variability, which refers to the distribution of interkey time across different keystrokes and different contexts, correlated -.69 with the net typing speed; b) Intra-keystroke variability, which represents the distribution of interkey time for the same keystroke in the same context but across multiple repetitions, correlated -.71 with the net typing speed.

- Phenomenon 26: Eye-hand span is larger with increased skill level of typists.

In the Salthouse (1984b) studies, the correlation between the eye-hand span and net words per minute across 74 typists was significant with $p<.01$. There was an increase of between 0.5 and 1.2 characters with every 20 net words per minute increase in typing skill (Salthouse 1985, Salthouse and Saults 1987).

- Phenomenon 27: Replacement span is larger among more skilled typists

Salthouse and Saults' studies (1985) found that the correlation between net words per minute and the replacement span was 0.80 ($p<.01$).

- Phenomenon 28: Copying span is moderately related to the skill level of typists

The correlation coefficient between copying span and net words per minute ranges from 0.35 to 0.57 (however, the correlation is not significant, $p>.05$, Salthouse, 1985a; Salthouse and Saults, 1985).

- Phenomenon 29: Fast typists have larger stopping spans than slow typists.

The experimental results of phenomenon 29 are not conclusive. Salthouse and Saults (1985) reported a correlation of 0.57 between the typing speed and the stopping span. However, another study of Salthouse (1985) did not find any significant correlation between these two variables ($p>.05$).

- Phenomenon 31: Interkey time of transcription typing decreases with practice following the power law of practice (Gentner, 1983).

Typing speed of an unskilled typist can be improved to that of a skilled typist. According to the learning curve of the single typist in the study of Gentner (1983), the improvement of interkey time follows the power law of practice.

2.1.5 Eye movements Phenomena

Although not included in the Salthouse list of phenomena, eye movements are one of the most important aspects of human behavior in eye-hand coordination tasks including transcription typing. Among the various variables in eye movements data, fixation duration (the length of time for one fixation of the eye movements), saccade size (the number of characters or the degrees of visual angle between two fixation points) and gaze duration per character (equals fixation duration divided by saccade size) are the major parameters in determining eye movements in transcription typing (Inhoff and Wang, 1992). Three recently discovered eye movements phenomena related to transcription typing are listed below.

- Phenomenon 32: Gaze duration per character decreases with increased preview window size.

Inhoff and Wang (1992) found that the gaze duration per character decreased from 280 ms to 182 ms when the preview window size increased from 1 to 11 characters.

- Phenomenon 33: The mean saccade size is about 4 characters (Rayner, 1998).
- Phenomenon 34: The mean fixation duration in transcription typing is 400 ms (Rayner, 1998).

2.2 Existing Models of Transcription Typing

Several quantitative and qualitative models have been proposed to analyze transcription typing behavior. The quantitative models includes a central control model (Terzuolo and Vivianai, 1979, 1980), a composite model (Gentner, 1987), an activation-trigger-schema model (Rumelhart and Norman, 1982), and a PERT-network based model (John, 1988, 1996). The model proposed by Salthouse (1984a, 1986a) is a qualitative model.

Terzuolo and Vivianai (1979, 1980) proposed a central control model of timing in transcription typing and they suggest that interkey time is generated in parallel from centrally stored, word-specific timing patterns. Gentner provided experimental evidence against this central model (Gentner, 1982) and proposed a composite model composed of both central and peripheral mechanisms (Gentner, 1987).

Rumelhart and Norman (1982) proposed a model based upon an activation-trigger-schema system in which a hierarchical structure of schemata directs the selections of the

characters to be typed and controls the hand and finger movements by a cooperative algorithm. The model reproduces several major phenomena of typing including the interkey time and the patterns of transposition errors found in skilled typists.

John (1988, 1996) proposed a model called "TYPIST" (an acronym for TheorY of Performance In Skilled Typing) which uses the Project Evaluation and Research Technique (PERT) method of scheduling to quantify the parallel activities of typing performed by the three perceptual, cognitive, and motor processors in the Model Human Processor (MHP) (Card, et al. 1983). TYPIST is thus far the most extensive quantitative model of transcription typing and it covers 19 of the 34 phenomena in transcription typing, including 17 phenomena reviewed by Salthouse (1986) and 2 additional phenomena found by Gentner (1983) and Salthouse and Saults (1987) (phenomena 31 and 30 reviewed above).

Salthouse (1984a, 1986a) proposed a qualitative model of transcription typing which consists of 4 components: input (convert text into chunks), parsing (decompose chunks into ordinal strings of characters), translation (convert characters into movement specifications) and execution (implement movement in ballistic fashion). It is a synthesis of many previous works and provides a basic conceptual framework in transcription typing. However, because it is a qualitative model, it does not simulate or generate typing behavior, or make quantitative predictions.

In the following section, we describe a Queueing network model of human performance and its application in modeling transcription typing. The model captures the nature of transcription typing as a parallel process—the typist looks ahead at the words on a display while executing the motor responses for the current characters (John, 1989). The model analyzes time and error simultaneously with the same underlying cognitive structure, and generates typing behavior as observable behavioral manifestations of the underlying cognitive Queueing network at work.

3. Learning Mechanisms in the Queueing Network

Based on the description of QN-MHP in Chapter 1, to model the skill effect and related phenomena in transcription typing, three learning mechanisms in the QN-MHP were proposed in this section. These learning mechanisms were another important improvement of the Queueing network model. At the level of networks, the probability that entities take different routes may change in the learning process, representing the change of connection strengths and rewiring of neural pathways in the brain network (Van Mier, et al., 1998; Petersen, et al., 1998). At the level of individual servers, server processing time decreases and information processing in servers can also be optimized via trial-and-error, reflecting the improvement of information processing efficiency of individual brain regions via a learning process (Braus, 2004; Boettiger and D'Esposito, 2005).

3.1 Change of Routing Probability

It is well recognized that the human brain is not only a network of brain regions, but also a system that is able to change itself dynamically in the process of development and learning (Chklovskii, et al., 2004; Habib, 2003). On the one hand, the "brain traffic" concept in neuroscience suggests that information flow represented by spike trains in the brain exhibit features of traffic flow in the network—spike trains (represented by entities in the model) form the information flow among brain regions. Depending on different tasks and learning stages, these information flows sometimes can be processed by the brain regions (servers) immediately, but sometimes they have to be maintained in certain regions to wait for the previous flow being processed (Bullock, 1968; Eagleman, Jacobson, & Sejnowski, 2004; E. E. Smith & Jonides, 1998; Taylor et al., 2000; Braus, 2004; Chklovskii, et al., 2004; Habib, 2003). On the other hand, different brain areas are activated during the visual-motor learning process (Van Mier, et al., 1998; Petersen, et al., 1998; Aizawa et al. 1991). This plasticity aspect of the human brain concerns the change of synaptic connection strength between neurons and rewiring among neural pathways—spike trains change from one neural pathway to anther one with stronger synaptic connection strength and higher efficiency in information processing. This rapid regulation is related to a brain derived neurotrophic factor (BDNF) regarded as a signal of

synaptic plasticity in adults (Black, 1999; Braus, 2004), and Black (1999) proposed a model explaining the role of BDNF in its regulation of the synaptic plasticity.

Equation 1 is developed based on Black (1999)'s model and the "brain traffic" concept above (see Appendix 1 for its derivation), where routing probability (P_i) stands for the probability that spike trains (represented by entities) pass through a certain neural pathway (route i) in a total of U multiple routes; and sojourn time (S_i) is defined as the sum of waiting time (W_i) and processing time (T_i) of these spike trains (entities) along that neural pathway.

$$P_i = \frac{1/S_i}{\sum_{j=1}^{U} \frac{1}{S_j}} \tag{1}$$

3.2 Reduction of Server Processing Time

Besides the change of connection strengths and rewiring of pathways at the network level, individual brain regions also exhibit improvements in information processing speed in the learning process (Braus, 2004). Moreover, ample research has demonstrated that exponential functions characterize the learning processes in memory search, motor learning, visual search, and mathematic operation tasks better than the power law (Heathcote, et al., 2000). Accordingly, exponential functions are employed in the Queueing network model to characterize the learning process in the individual servers (see Equation 2), with the exception of the six perceptual servers (servers 1-3 and 5-7) that are only related to neural signal transmissions which are relatively stable in the learning process.

$$T_i = A_i + B_i \, Exp(-\alpha_i N_i) \qquad \text{(Heathcote, et al., 2000)} \tag{2}$$

In Equation 2, T_i stands for the processing time in each server; A_i represents the expected minimal processing time (Ti) at server i after intensive practice (Feyen, 2002). B_i is the change in the expected processing time from the beginning to the end of practice; α_i represents the learning rate of server i (e.g., $\alpha_i = .001$, Heathcote et al., 2000); and N_i is the number of entities processed by server i; for example, N_i in servers A, B, C, and F refers to the number of chunks the server processed, while N_i in server W refers to the

number of retrievals of a certain motor program in general (e.g., in transcription typing, N_i in server W refers to the number of retrievals for a certain digram).

3.3 Optimization of Information Processing via Trial-and-Error

Numerous studies have found that mammals including human beings optimize their movement and behavior via the learning process (Alexander, 1993; Borghese and Calvi, 2003; Laureys, et al., 2001). For example, mammals optimize the movement of their legs to run quickly with the smallest amount of energy. Among these optimization processes, trial-and-error is one of the major formats of learning (Boettiger and D'Esposito, 2005; Bustillos and de Oliveira, 2004; Ghilardi, et al., 2000; Sakai, et al., 1998)—mammals may try many actions until one of them satisfies their goal. For human beings, trial-and-error is also an important aspect of motor learning (Ghilardi, et al., 2000) and optimization in information processing in working memory (Asari, et al., 2005; Baltes, et al., 1999; Bor, et al., 2004; Genovesio, et al., 2005; Krampe, et al., 2003; Schmuck and WobkenBlachnik, 1996), and it involves the activation of the frontal cortex (represented by server A, B, C) and the pre-supplementary motor area (pre-SMA, represented by server Y) (Boettiger and D'Esposito, 2005; Nakamura, et al., 1998). Typically, this trial-and-error learning is simulated via Monte Carlo simulation (Bustillos and de Oliveira, 2004) whose nature is a trial-and-error process of using random numbers to reach a solution. In general, this Monte Carlo learning mechanism can be implemented in any of the QN-MHP servers, but for transcription typing modeling, it is only implemented in server B and server Y since they are most relevant to learning of motor skill and keyboard characteristics.

4. Simulating Transcription Typing With QN-MHP: Mechanisms and Results

Simulation of any human-machine interaction task requires the specification of three components: a human model, the machine or the environment with which the human model interacts, and the task input to the human model. These three components correspond to the QN-MHP, a typewriter, and a display presenting the to-be-typed text, respectively, in the context of the transcription typing task.

The general human model of QN-MHP is described in the previous section. To possess the basic knowledge of typing requires the QN-MHP to have the corresponding procedure knowledge rules stored in its long-term procedure memory server. Thus, following the general method of QN-MHP simulation (Liu, Feyen, and Tsimhoni, 2006), a 5-step NGOMSL-style task description of transcription typing is developed (see Table 2-2) and stored in server D as the long-term procedure knowledge of typing in the model (also called as operator or command entity). Step 1 (watch for < > on < >) defines how the model samples visual information (e.g., the characters) on a certain user interface (e.g., the display) via the visual perceptual subnetwork following a Queueing process—the number of entities leaving server A or B at one time, forming a chunk (a meaningful information unit, chunk size=*x*), determines the number of entities sampled by the servers in the visual perceptual subnetwork at one time. After the stimuli are retained in the working memory (step 2), step 3 defines how the model presses a certain control device on a user interface (e.g., keys on a QWERTY keyboard) with defined body parts (e.g., hands). Finally, when the model reaches the end of the text (step 4), it stops typing (step 5). All of these steps or operators have two properties. First, they are defined in a task-independent manner; task-specific information is treated as their parameters. Second, even though these steps are listed in a serial manner in the NGOMSL-style task description, they can run in the model in parallel because of the parallel processing property of the Queueing network. For example, the perceptual subnetwork is able to "watch for" new stimuli (step 1) while the motor subnetwork is still executing the simulated actions (step 3).

Table 2-2 NGOMSL-Style Task Description of Transcription Typing Task

GOAL: Do transcription typing task

Method for GOAL: Do transcription typing task
Step 1. Watch for <the characters> on <the display>
Step 2. Retain <the characters >
Step 3. Press <keys> on <a QWERTY keyboard> with <hands>
Step 4. Decide: If <the characters> is <the end of text>, then move to step 5
Else move to step 1
Step 5. Cease //task completed

Method for GOAL: Press <keys> on <a QWERTY keyboard> with <hands>
Step 1. Decide: If location of <keys> in memory, then move to step 3
Else move to step 2
Step 2. Visual search for <locations> of <keys> on <a QWERTY keyboard>
Step 3. Reach <keys> on <a QWERTY keyboard> with <hands>
Step 4. Return with goal accomplished

Method for GOAL: Visual search for <locations> of <keys> on <a QWERTY keyboard>
Step 1. Recall <characters> from <working memory> as <the target characters>
Step 2. Watch for <key labels> on <a QWERTY keyboard>
Step 3. Compare <key labels> with <the target characters>
Step 4. Decide: If match, then move to step 5
Else move to step 2
Step 5. Retain <the location> of <key labels>
Step 6. Return with goal accomplished

To define a typewriter with which the QN-MHP interacts, a software module called m-hQWERTY was implemented to represent a QWERTY keyboard, the most commonly used keyboard in the English-speaking world. This module defines the size and location of each key and the distance between each pair of the keys on the keyboard. We selected the same text source employed in Salthouse's study (1984a, 1984b, 1987): the Nelson-Denny Reading Test. A module in the simulation software (Promodel®) is designed to represent the display containing the position and content of the text characters. In each run, the model types 1,000 letters from the Nelson-Denny Reading Test; and the model performed 10 simulation runs with different standard random number series in the Promodel software (Promodel, 2004).

In the following, the simulation mechanisms and results are described in detail for each of the six groups of phenomena reviewed above. In each group, we describe how the corresponding phenomena are generated based on the mechanisms in the Queueing network. Simulation results were validated with the same error estimation calculation

method employed in John (1988, 1996), including the percentage of relative error $=|Y-X|/X \cdot 100\%$, Y: simulation result; X: experimental result, which was summarized at the end of this section.

4.1 Basic Phenomena

4.1.1 Simulation Mechanisms

The ten phenomena in this group are modeled with three fundamental mechanisms of the QN-MHP: parallel processing (phenomena 1, 7, 12), motor processing (phenomena 4, 5, 6, 8, 9, 11), and visual processing (phenomenon 10).

1) Parallel Processing (Phenomena 1, 7, 12)

Phenomena 1, 7, and 12 emerge naturally as the result of parallel processing in the Queueing network. Typing is faster than choice reaction time (phenomenon 1) because the servers in the visual perceptual subnetwork of the QN-MHP can process visual entities (watch for the remaining letters to be typed) at the same time while the motor servers execute tying actions. This is in contrast to a choice reaction time task which requires a single response execution to follow stimuli perception in a serial manner. In the QN-MHP, the two hand servers can process information in parallel, while each hand can only process information serially, producing phenomenon 7—alternate-hand keystrokes are faster than same hand keystrokes. Similarly, a concurrent task does not affect typing (phenomenon 12), when it involves the servers and routes that can be performed concurrently with the typing task, as in the case of the tone-pedal pressing task (see Table 2-3 for its NGOMSL task description).

Table 2-3 NGOMSL-style Task Description of Tone-pedal Press Task

GOAL: Do tone-pedal pressing task Method for GOAL: Do tone-pedal pressing task Step 1. Listen to <the tone> from <the speaker> Step 2. Retain <the tone> Step 3. Compare: <the tone> with <the target tone> in memory Step 4. Decide: If match, then go to step 5 Else move to step 1 Step 5. Press <the pedal> on <the floor> with <one foot> Step 6. Return with goal accomplished

2) Motor Processing (Phenomena 4, 5, 6, 8, 9, 11)

The motor subnetwork in the Queueing network model is able to generate these 6 phenomena in a natural and consistent manner. In the motor subnetwork, motor programs of high-frequency digrams are retrieved more often by server W from server D, requiring less processing time than low-frequency digrams and producing the digram frequency effect according to Equation 2 (phenomenon 8). Correspondingly, if all of the letters to be typed are composed of random ordered letter pairs, this digram frequency effect disappears and the interkey time increases (phenomenon 5). Similarly, if the model can only sample one or two characters at one time via the preview window, it increases the chance that motor programs of high-frequency digrams are decomposed and therefore attenuates this digram frequency effect, producing phenomenon 6—typing rate is impaired by the restricted preview window. In contrast, if only the order of the word is randomized but the order of the letters in each word remains unchanged (phenomenon 4) or the number of letters in each word increases (phenomenon 9), this digram frequency effect is not affected since the digrams in each word are still preserved, generating phenomena 4 and 9—interkey time is independent of word order and its length. In addition, step 3 in the NGOMSL-style task description (press < > on < > with < >)—a task-independent operator treating task-specific information such as keyboard layout as its parameters, specifies how the two hand servers interact with a QWERTY keyboard (implemented in the m-hQWERTY module) and generates the movement distance of fingers according to the topography of the keyboard; then, the hand servers in the model are able to produce the movement time of fingers (see Appendix 3), producing phenomenon 11—the keystroke time depends on the specific context. It is important to note that there is no free parameter in the formula to simulate the experimental results.

3) Visual Processing (Phenomenon 10)

Phenomenon 10 is produced by the model naturally via its visual sampling process defined in the “watch for” operator. The hunt-feature production which is employed by ACT-R and implemented in QN-MHP, facilitates the servers in the visual perceptual subnetwork to locate the fixation point at the feature of a meaningful unit—the middle

point of the first half of a word (Rayner, 1998) in the text viewing condition. This process indicates that the first character in each word is the expected first character in each chunk (see calculations in Appendix 2), which increases the processing time of the first character of each word by the time needed in encoding visual stimulus into chunks, producing phenomenon 10—the keystroke of the first character is longer than that of other keystrokes in a word.

4.1.2 Simulation Results

QN-MHP showed an average interkey time of 176 ms, which was shorter than choice reaction time (the typical two choice reaction time is 320 ms, Feyen 2002) (phenomenon 1, estimation error= 0.56%). In these keystrokes, the simulated alternate-hand strokes were 40 ms shorter than the same-hand strokes on average (phenomenon 7, estimation error= 11%). The simulated average interkey time in concurrent task situation was 174 ms which was not affected by the pedal pressing task and no significant difference in the number of typing errors (Kolmogorov-Smirnov Test, Z=0, df=18 (10 runs for the single and concurrent task conditions), p=1>.05) was found between the simulated single and dual task situations (phenomenon 12, estimation error= 5.95%).

When the order of the words was randomized but the order of letters in each word remained unchanged, the simulated interkey time did not show significant change compared to that in the "normal text" typing condition (Independent Sample T-Test, df=19998, p=.11>.05) (phenomenon 4). However, when the order of letters within each word was randomized, the simulated average interkey time increased to 354 ms (phenomenon 5, estimation error=22%); as the size of the preview window decreased, the simulated interkey time also increased (R square of simulated interkey time is .97, see Figure 2-1) (phenomenon 6, estimation error= 10.98%). In addition, the simulated interkey time of high-frequency digrams was significantly shorter than that of low-frequency digrams (Independent Sample T-Test, df=398, p=.024<.05) (phenomenon 8) but no significant difference of simulated interkey time was found between the long and short words (Independent Sample T-Test, df=1998, p=.148>.05) (phenomenon 9). The simulated interkey time of the first keystroke in a word was 14% longer than that of the subsequent keystrokes (phenomenon 10, estimation error=30%).

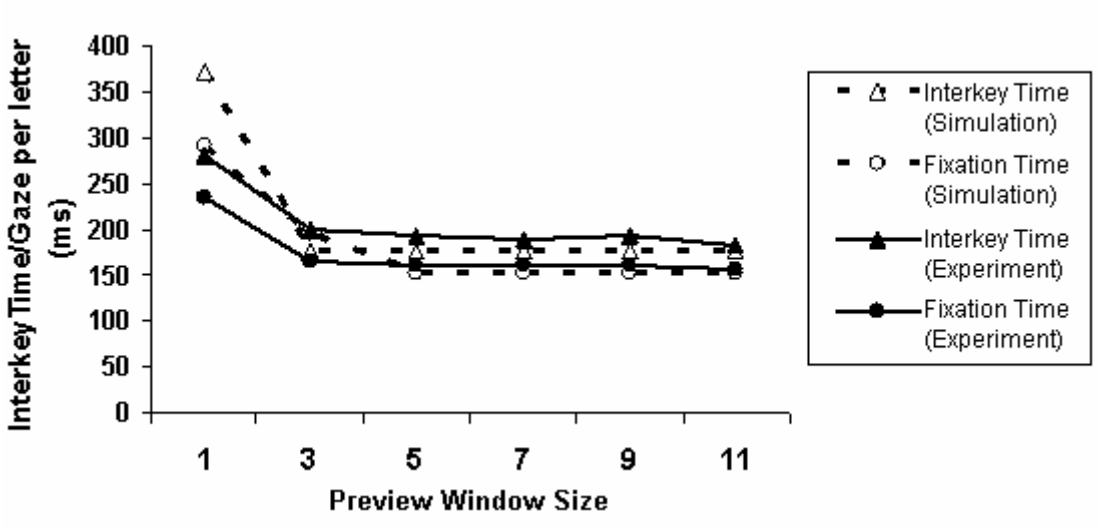


Figure 2-1 Comparison of simulated interkey time and gaze duration per character with those of experimental results (Inhoff & Wang, 1992) in different preview window sizes (unit of size: character)

The simulated movement time and interkey time of the same letter pairs modeled by TYPIST were summarized in Table 2-4, which showed that interkey time depended on the specific context (phenomenon 11).

Table 2-4 Simulation Results of Interkey Time of the Letter Pairs

Keys	Observed (ms)	Simulation Results		Absolute of relative error %
		Distance (cm)	Interkey time(ms)	
e-e	165	0	165.0	0.00
d-e	201	2	240.2	19.49
c-e	215	4	249.4	16.00
r-e	145	1	151.2	4.26
t-e	159	1.5	154.7	2.70
f-e	168	2	157.7	6.14
g-e	178	3	162.7	8.61
v-e	178	3	162.7	8.61
b-e	195	4	166.9	14.42
Average of relative percentage of error				8.91

4.2 Units of Typing in Transcription Typing

4.2.1 Simulation Mechanisms

The six phenomena in this group are modeled with two fundamental mechanisms in the model: entity-based information processing (phenomena 13, 15, 16, 17, 30) and parallel processing (phenomenon 14).

1) Entity-based Information Processing (Phenomena 13, 15, 16, 17, 30)

An entity is a basic piece of information processed in the Queueing network model, which allows us to observe the activity of entities in the network during the simulation and count the number of these entities in various parts of the network with simple calculations based on the definitions of tying units. According to the definition of copying span—the number of characters typed correctly after an unexpected disappearance of the copy, once the input to the model is suddenly stopped, the total number of entities (characters) held and processed in the model equals copying span and its expected value is 10 characters (phenomenon 13) (see Appendix 5 for its estimation). Moreover, since the visual sampling process defined in the "watch for" operator allows x characters to enter the model at one time, when the input to the model is suddenly stopped, these x sampled characters are already in the model and thus counted as part of the copying span (see Figure 2-2). As shown in Figure 2-2, eye-hand span (the number of characters between the fixation point and the character currently being typed) equals the expected copying span minus the $x/2$ characters on the right side of the fixation point excluding the character being pressed. Given the optimal x value via the optimization process (x_{opt}=4, see Appendix 4), the expected eye-hand span=expected copying span-$x/2$-1=10-4/2-1=7 characters (phenomenon 15). When the to-be-typed text is composed of random letters, similar to the simulation mechanism of phenomenon 5, the digram frequency effect disappears and each pair of entities takes a longer processing time in the model. Since entities in each subnetwork decay in the model, this reduces the amount of entities held and processed in the model, producing smaller copying spans and eye-hand spans (phenomenon 16). Moreover, when the to-be-typed text is composed of random letters, the chunk size of each pseudo word decreases, thus increasing the amount of time in perceiving each pseudo word. In addition, as shown in Figure 2-2, the detection span (characters between the capital character and the character currently being typed) is the sum of the eye-hand span plus the radius of foveal vision excluding the capital character (the central 2 degree vision, 1 degree as radius= 4 characters, Rayner, 1998). Thus, expected detection span=expected eye-hand span+4-1=7+4-1=10 characters (phenomenon 30). Finally, once one of the characters in the to-be-typed text is suddenly replaced by another character, the model is able to detect this change as long as the

entities have not left server Y because server Y is the server for detecting errors and reassembling the motor program in the motor subnetwork. Thus, the total number of entities in the servers after server Y (server Z and the two hand servers) is the replacement span and its expected value equals 3.6 characters (see Appendix 5 for its estimation) (phenomenon 17). In addition, due to the stochastic property of the model (e.g., exponentially distributed processing time of the servers), there are possible differences between these predicted values and simulation results.

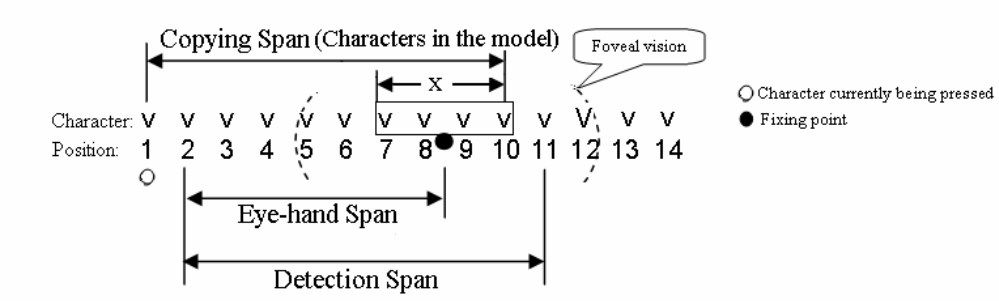


Figure 2-2 Graphical illustration of the expected copying span, eye-hand span and detection span

2) Parallel Processing (Phenomenon 14)

Similar to phenomenon 12, the Queueing network model is able to process the entities representing the stopping span task as well as those of the transcription typing task at the same time. Table 2-5 listed the NGOMSL task procedure of the stopping span task as a secondary task. Consistent with the definition of the stopping span, the number of entities typed by the model during the processing period of a tone is regarded as the simulated stopping span.

Table 2-5 NGOMSL-Style Task Description of Stopping Span Task

GOAL: Do stopping span task
Method for GOAL: Do stopping span task
Step 1. Listen to <the tone> from <the speaker>
Step 2. Retain <the tone>
Step 3. Compare: <the tone> with <the target tone> in memory
Step 4. Decide: If match, then go to step 5
Else move to step 1
Step 5. Cease //task complete

4.2.2 Simulation Results

The simulated average copying span, eye-hand span, and detection span were 9.4, 6.4, and 9.4 characters, respectively (phenomenon 13, estimation error=35.6%; phenomenon 15, estimation error=28%; phenmonon 30, estimation error=17.5%). When the to-be-typed text was composed of random letters, the simulated eye-hand span decreased to 1.4 characters on average (phenomenon 16, estimation error=20%). The simulated average stopping span and replacement span were 2.5 and 3.5 characters respectively (phenomenon 14, estimation error=15.7%; and phenomenon 17, estimation error=20.7%).

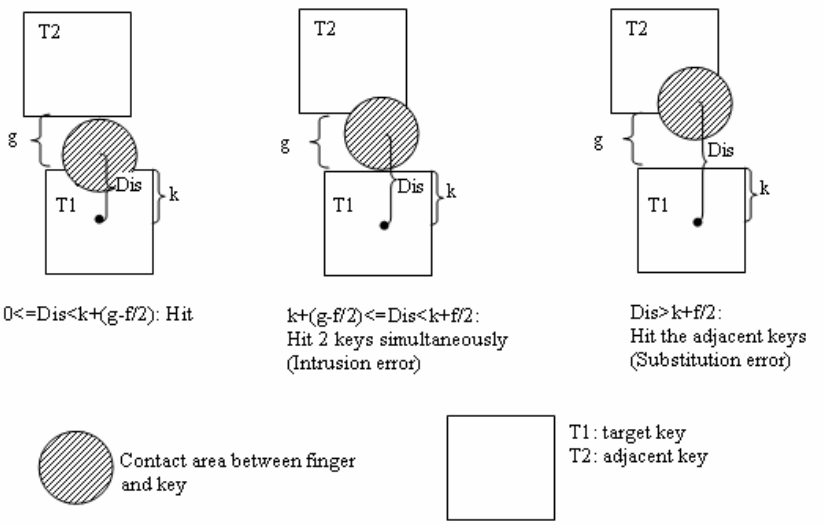
4.3 Errors in Transcription Typing

4.3.1 Simulation Mechanisms

The five phenomena in this group are simulated with two mechanisms of the Queueing network model: distribution of movement distance and force (phenomena 19, 20, 21) and optimized motor processing (phenomenon 22). Phenomenon 18 can be modeled by the further calculation of simulation results of phenomena 19-22.

1) Distribution of Movement Distance and Force (Phenomena 19, 20, 21)

Based on Tanaka (1994)'s equations in quantifying the root-mean-square error (RMSE) of movement directions generated by population vectors in the primary motor cortex, the distribution of movement distance of fingers follows a normal distribution (Dis $\sim N(\overline{Dis}, 0.317)$) (unit: cm) (see its derivation in Appendix 3), which allows the Queueing network model to generate intrusion and substitution errors depending on the range of finger movement distance (Dis) in three possible conditions (see Figure 2-3): i) 0<=Dis<k+(g-f/2): when the contact area between a finger and a target key does not contact with the other keys surrounding the target key; ii) k+(g-f/2)<=Dis<k+f/2: when this area contacts both the target key and an adjacent key, i.e., the finger hits 2 keys simultaneously (intrusion error, phenomenon 20); iii) Dis>k+f/2: when this area falls in to the area of an adjacent key but not the target key (substitution error, phenomenon 19).



g: the average width of gaps between adjacent keys (0.55 cm, measured on a standard QWETY keyboard)

k: half of the average key width (0.65 cm)

f: the average diameter of the finger-key contact area, which equals to the half width of the key (0.65 cm)

Figure 2-3 Three possible conditions for the range of finger movements in pressing the target key

According to the distribution of the pressing force of fingers (see Appendix 3) and the typical key activation force (0.28 N, Gerard et al., 1999), the model is also able to generate omission errors in phenomenon 21. The simulated omission errors are categorized into two types: type A—an omission error occurs and no simulated finger movement is recorded (the omission error is caused by the failure to preserve sequences in the sensory and working memory); type B—an omission error occurs and the movement of a finger is recorded but the simulated finger pressing force on the target key is less than 0.28 N (the omission error is caused by an insufficient depression of a keystroke).

2) Optimized Motor Processing (Phenomenon 22)

The coordination of bimanual movements in motor processing is optimized via the optimization of EPD (spell out) (cross-hand error prevention duration, i.e., waiting duration of between two entities belonging to different hands, see Appendix 4) in Monte Carlo simulation—if EPD is too long, the interkey time becomes very long which

deteriorates the typing performance; if EPD is too short, the model has to spend extra time in correcting the typing errors (see Appendix 4).

Since server Y in the Queueing network model is able to detect errors via the tactile feedback from the two hand servers and server X, typing errors caused by the hand movements including the deviated movement direction and finger force as well as the insufficient waiting time between the two hands, can be detected without reference to the typed copy. The ratio of typing errors detected without reference to the typed copy over the total number of errors is calculated based on the simulation results of phenomena 19-22 (phenomenon 18).

4.3.2 Simulation Results

It was found that in typing 10,000 characters, 1) 41.3% of the substitution errors involved horizontally or vertically adjacent keys (phenomenon 19, estimation error=37.7%); 2) 35.4% of the intrusion errors involved keystrokes with less than 10% of the average interkey time and 57.1% of them involved an adjacent key in the same row or the same column (phenomenon 20, estimation error=6.4%); 3) the average interkey time of the keystrokes right after an omission error occurred was 253 ms, which was 1.44 times of the simulated average interkey time (176 ms) (phenomenon 21, estimation error=6.6%); 4) 68% of the transposition errors were made by the alternate hands (phenomenon 22, estimation error=6.6%). In typing 10,000 characters, 74.5% of the errors were caused by the hand and finger movements and detected without reference of the typed copy (phenomenon 18, estimation error=6.4%).

4.4 Skill Effect in Transcription Typing

4.4.1 Simulation Mechanisms

The eight phenomena in this group are modeled by the three general learning mechanisms of the Queueing network model described earlier in this paper: change of routing probability), optimized motor processing, and general effects of learning.

1) Change of Routing Probability (Phenomena 24, 25)

During the learning process of transcription typing, after entities arrive at server B from the perceptual subnetwork, entities can either take route 1 (go to server C and F for visual guidance and then go to server W without long-term motor program information retrieved from server D) or route 2 (go to server W directly with the long-term motor program information retrieved from server D). At the beginning stage of the learning process, server D has not stored sufficient motor program information and server W is not able to retrieve these motor programs effectively from server D, prolonging the sojourn time of route 2 and decreasing the routing probability of taking route 2, based on Equation 1. With the number of practice increases, more and more motor programs of digrams as well as the location information of keys are stored at server D. Once the sojourn time of route 2 decreases with a higher efficiency in retrieving motor program in server D and its value is lower than that of route 1, the majority of entities start to travel via route 2. In other words, at this stage, the model does not have to perform visual search for each digram and the route of visual search (Server C→F→C) is skipped by the majority of entities, forming a new route starting from the servers in the visual perceptual subnetwork to Server B→ W→ Y→ Z→ Two Hand Servers. This simulation mechanism is consistent with fMRI studies in transcription typing and other motor control tasks. At the beginning stage of learning, a visuomotor control task including transcription typing mainly activates the DLPFC (dorsal lateral prefrontal cortex) (Server C) and the basal ganglia (Server W) (Jueptner and Weiller, 1998; Sakai, et al., 1998). In the well-learned stage (skilled typist in Gordon et al.'s study, 1995), in typing normal texts (multi-digit sentence), activation of the DLPFC disappeared and stronger activations were observed in the SMA (supplementary motor area) (Server Y), the basal ganglia (Server W) and the primary motor cortex (M1) (Server Z).

Since the server processing times follow the exponential distribution in QN-MHP (Liu, Feyen and Tsimhoni, 2006), if Y_1.. Y_k are k independent exponential random variables, their sum X follows an Erlang distribution (see Equation 3). Via rewiring of routes in the learning process, servers C and F are skipped by the majority of entities, i.e., parameter k in Equations 4 and 5 decreases. If k' after practice is smaller than k before practice, then the expected overall processing time and its variance decrease, producing phenomena 24 and 25 (see Equation 6).

$$X = \sum_{i=1}^{k} Yi \tag{3}$$

$$E[X] = E[\sum_{i=1}^{k} Yi] = \sum_{i=1}^{k} E[Yi] = k\frac{1}{\lambda} \tag{4}$$

$$Var[X] = Var[\sum_{i=1}^{k} Yi] = \sum_{i=1}^{k} Var[Yi] = k\frac{1}{\lambda^2} \tag{5}$$

$$If\ k' < k,\ then\ E[X'] < E[X];\ Var[X'] < Var[X] \tag{6}$$

2) Optimized Motor Processing (Phenomenon 23)

Based on the optimization process of the hand and finger movements in the learning process (see Appendix 4), the interkey time of the two-hand (2H) digrams and two-finger (2F) digrams decrease via the optimization of both EPD (cross-hand error prevention duration) and 2FC (two-finger coordination time), while the interkey time of the digrams of the one-finger (1F) digrams is reduced only by the optimization of 1FW (one-finger waiting time). Since the sum of the magnitude of EPD and 2FC's reduction is greater than that of 1FW, the model produces phenomenon 23—the reduction of the 2F or 2H digrams' interkey time is greater than that of 1F digrams.

3) General Effect of the Learning Process (Phenomena 26, 27, 28, 29, 31)

The increase in the size of the typing units (copying span, eye-hand span, stopping span, replacement span in phenomena 26-29) is due to several factors in the learning process: 1) the processing speed increases in each server (see Equation 2); 2) the route of the majority of entities rewires and servers C and F are skipped from the route (see simulation mechanism of phenomena 24 and 25), this rewiring process reduces the amount time for each entity spent in the model (see Equation 6 in the simulation mechanism of phenomenon 24); 3) the optimization of the motor process reduces waiting time in movement (see Appendix 4). Since every subnetwork has certain decay functions

in the model (Card, et al. 1983; Ito, 1991), the shorter each entity spends in the network, the greater is the amount of entities held and processed in the model, increasing the value of these typing units (see the simulation mechanism of phenomena in the "units of typing" group) and decreasing the interkey time (phenomenon 31). However, the random effect in the Monte Carlo simulation in the optimization process (see Appendix 4) as well as the stochastic property of the whole model may attenuate the increase of these typing units via the learning process.

4.4.2 Simulation Results

The model's simulation of its learning process[2] showed that the simulated tapping rate and the typing speed during the learning process was significantly correlated (Pearson correlation coefficient=0.784, N=8, p=.021<.01) (phenomenon 24). The change of the quartile range (75% quartile-25% quartile) of the interkey time, i.e., the inter-keystroke variability, was correlated with the change of the simulated interkey time with the Pearson correlation coefficient -0.911; the intra-keystroke variability simulated by the model correlated -0.795 with the simulated typing speed (phenomenon 25, estimation error = 22%). The average slope of regression equations relating the simulated digram interval to the simulated typing speed were -2.03 and -1.71 for 2H and 2F digrams respectively, while the average slope of 1F digrams was -1.65 (phenomenon 23, estimation error = 17.9%).

For the eye-hand span, significant correlation between the eye-hand span and the net words per min was found in the simulation results (Pearson correlation coefficient=.721, N=8, p=.044<.05). The eye-hand span of the model increased by 0.87 characters on average with every 20 net words per minute increase in skill level (phenomenon 26, estimation error = 2.6%). For the replacement span, the Pearson correlation coefficient between net words per minute and the replacement span was .867 (N=8, p=.005<.01) (phenomenon 27, estimation error = 8.4%). For the copying span, the Pearson correlation coefficient between the simulated copying span and net words per minute was 0.704 (N=8, p=.05) (phenomenon 28, estimation error=23.5%). For the stopping span, the

[2] The number of keystrokes typed by QN-MHP and the number of training stages in the simulation of all of the 8 phenomena in skill effects were set according to those in Gentner's experimental study (1983): a total of about 15,000,000 letters were typed in eight training weeks.

correlation efficient between the simulated stopping span and net words per minute was 0.868 (Pearson correlation, N=8, p=.004<.05) (phenomenon 29, estimation error=44.6%). After the model finished its learning process, the simulated interkey time reduced from 385 ms to 176 ms, which followed the power law of practice (R square=0.84 with significant correlation, N=8, p=.005 <.001) (phenomenon 31).

4.5 Phenomena in Eye Movements

4.5.1 Simulation Mechanism

All of the three phenomena in this group emerged as the natural outcomes of the Queueing mechanism in the Queueing network model. First, similar to the simulation mechanism for phenomenon 6, when the preview window size is very small (1 or 2 characters), motor programs of the high-frequency digrams are decomposed, which increases their retrieval time at server W from server D. Since information entities flow in the model in a Queueing process, slower information processing in the motor subnetwork in turn slows down information processing in the perceptual subnetwork. Therefore, gaze duration per character increases because servers in the visual perceptual subnetwork have to wait for the motor subnetwork to catch up, producing phenomenon 32—gaze duration per character decreases with an enlarged preview window size.

Second, following the Queueing process in visual sampling—the number of entities (the number of chunks c multiplied by the chunk size x) that leave server B at one time determines the number of entities sampled by servers in the visual perceptual subnetwork at one time, the expected saccade size (s) (the number of entities entering the visual perceptual subnetwork at one time) equals the product of c and x. Through the optimization process (see Appendix 4), the expected optimal value of c and x is 1 and 4, respectively, indicating that the expected saccade size is 4 characters (phenomenon 33).

Third, the average fixation duration in phenomenon 34 is the average gaze duration per character without the preview window (phenomenon 32) multiplied by the average saccade size (phenomenon 33).

4.5.2 Simulation Results

Figure 2-1 shows the simulated gaze duration per character (R square of the simulated fixation time is .94) (phenomenon 32). The simulated gaze duration per character without

the preview window was 136 ms on average. The average saccade size generated by the model was 3.18 characters (phenomenon 33, estimation error= 20.5%) and the average fixation duration was 483 ms (phenomenon 34, estimation error= 20.8%).

5. Potential Applications of the Model in User Interface Design

In practice, the Queueing network model can be applied in designing user interface when users of a system need to input multimodal textual information, transcribe information using a mobile device, or enter address information in an in-vehicle system. Besides the simulation of human performance, QN-MHP is also able to simulate mental workload in inputting the textual information.

5.1 Simulation of Human Performance in Multimodal Textual Information Input

By modifying the arrival of stimuli and using appropriate interface modules, the Queueing network model can simulate human performance in inputting textual information via multimodal human-computer interfaces. Table 2-6 summarizes the 9 possible combinations of input modalities and output devices (3*3=9) which can be simulated by QN-MHP. One the one hand, text (entities) can be set to arrive at Server 1 (visual modality), Server 5 (auditory modality), or central executive (Server C, from long-term memory Server D or H) for simulating human performance in inputting textual information from these different sources (looking while typing, listening while typing and thinking while typing). On the other hand, if the interface/device module is replaced by the modules of different keyboards (e.g., changing the distance between different keys) (see Figure 2-4) or modules of a handwriting recognizer (Wu et al., 2003), or when the route of entities are changed from the hand servers to the Mouth server, then the model can simulate human performance in typing on different keyboards, handwriting (handwriting recognition), and reading aloud (voice recognition).

Table 2-6 Extension of the model in simulating human performance in inputting textual information via multimodal human-computer interaction

Model Input (source of text)	Model Output devices
• Watching (display) • Listening (speaker) • Thinking (LTDSM)	• Typing (different keyboards) • Handwriting (hand recognition) • Reading aloud (voice recognition)

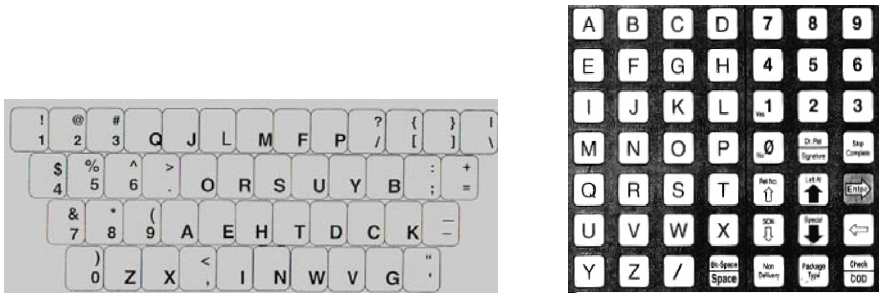


Figure 2-4 Different keyboards can be simulated in different device modules in QN-MHP

For example, in model a reading aloud task, if the motor sub-network may process the information entities faster than typing (assume *w,* duration of processing each chunk at motor sub-network, equals 0.3 sec on average) and R (repairing time equals 0.3 sec on average), the optimized chunk size (*x*) (x_{opt}=12 letters, see Figure 2-5) and saccade size (s=min(S_{max},x_{opt})=8 letters), fixation duration, and overall speed of reading aloud speed can be estimated by the model. These simulation results can be used in the design of voice recognition interfaces in the trade-off in recognition accuracy and error repairing, moving-letters display, and the reduction of the user's workload (by reducing the number of entities (*cx*) at Server B per unit of time) esp. in concurrent tasks situations.

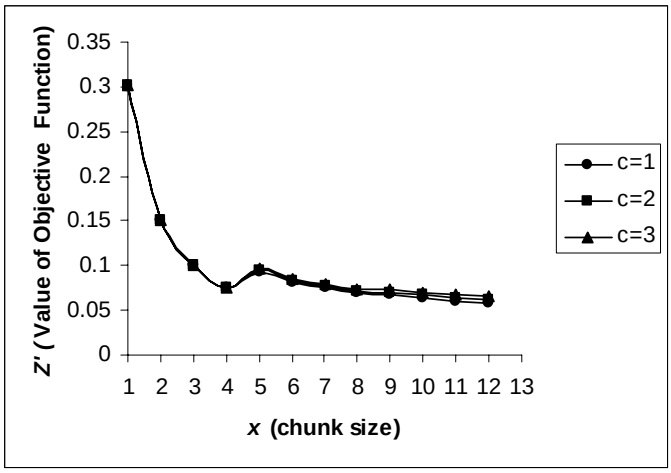


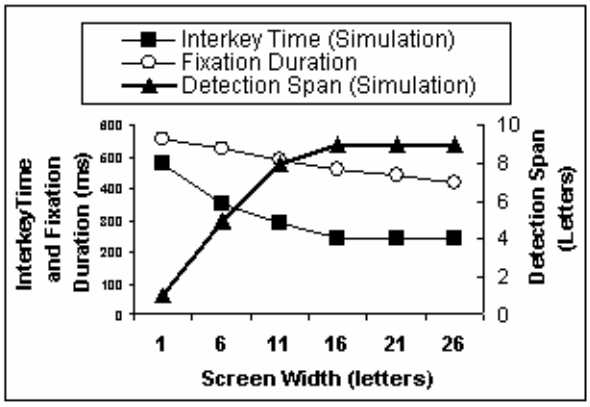
Figure 2-5 The change of objective function value (Z') with chunk size (x) and number of chunks at Pho server (R=.3 sec, *w*=.3 sec)

5.2 Design of Mobile Device based on Simulation Results (Single Task)

The current model can also be used to design a mobile device where textual information is presented on a small screen and users of the device need to type the

information via the soft keyboard on the interface (see Figure 2-6). The optimal screen width (distance *A* (in letter unit) in Figure 2-6a) can be determined by the maximal simulated detection span and minimal interkeytime (See Figure 2-6b), so that users of the mobile device can maximize their utilization of foveal vision and minimize the interkey time. If the presentation speed of the textual information is controlled by the device itself, the minimal pacing time (duration between two different lines of a text shown on the display separately) might be longer than the simulated fixation time generated by the model (See Figure 2-6b).

a. An example interface of a mobile device



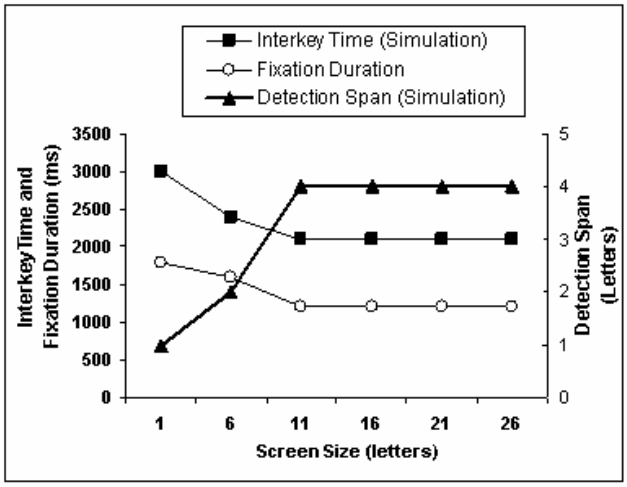
b. An example simulated interkey time and detection span in using a mobile device

Figure 2-6 An example interface of a mobile device and the simulated human performance with QN-MHP

5.3 Design of Data Entry Device in Driving based on Simulation Results (Dual Task)

The same logic in single task situations can be applied into modeling of a dual task—steering a car while inputting an address to a GPS system at the same time (See Figure 2-7). Since QN-MHP have successfully modeled the steering task (Liu, et al., 2006), by adding the arrival of another source of entities to the model without writing the third program to coordinate the two types of entities (steering and typing), QN-MHP is able to simulate human performance in dual task including interkey time and detection span.

a. An example data entry device in the driving context (Tsimhoni et al., 2004)



b. An example simulated interkey time and detection span in using a data entry device while driving

Figure 2-7 An example interface of a data entry device in the driving context and the simulated human performance with QN-MHP

The sample simulation results in Figure 2-7b show that users' detection span becomes smaller in a dual task condition than in a single task condition, which informs designers of the data entry interface that they might need to enlarge the font size on the screen to maximize the utilization of the screen presenting the address information.

5.4 Simulation of Mental Workload in Inputting Textual Information

QN-MHP is also able to model and generate both mental and motor workload by using the subnetworks' utilization levels as workload indexes (Liu, Feyen, and Tsimhoni, 2006; Wu and Liu, 2006; Rouse, 1980). In the simulation results in modeling the learning phenomena, it was found that the utilization of the cognitive subnetwork is lower than that of the perceptual and motor subnetworks in the well-learned situation of the model. This indicated that mental workload of skill typists is mainly allocated at the perceptual and motor subnetworks, which is consistent with the experimental results in phenomenon 3—skilled typist can perform reading comprehension (a high level of mental workload at the cognitive subnetwork) and transcription typing at the same time with very little interference. Moreover, server utilizations in the simulation results suggested that the physical workload (utilization) on the left hand server is significantly higher than that of the right hand server, and it is also consistent with the experimental results of QWERTY

keyboard studies which found that the left hand is used more often than the right hand in typing tasks (Goldstein, et al., 1999). The authors' recent work in modeling driver's workload and performance using QN-MHP provides a detailed description on quantifying mental workload with the Queueing network model (Wu & Liu, 2006c).

6. Conclusions

Working as a single cognitive architecture with the same set of assumptions and mechanisms, the Queueing network model is able to simulate diverse aspects of human performance in this typical human computer interaction task—interkey time, typing units and spans, errors, skill acquisition, and eye movements (32 of the 34 transcription typing phenomena). Furthermore, the Queueing network model offers an alternative way of understanding the mechanisms of cognition and human-computer interaction.

The Queueing network model is able to simulate and analyze design concepts related to information processing capacity (e.g., various typing units and spans). Using an intrinsic feature of Queueing networks—entity-based information processing, the model is able to not only quantify but also visualize the various spans in typing, which has potential value for HCI interface comparison and analysis. In addition, Queueing or waiting is part of our intuitive daily experience, both in general and in HCI tasks, and the Queueing network model emphasizes the importance of this aspect and explicitly incorporates the Queueing process as one of the major mechanisms in human-machine interaction (e.g., in simulating phenomenon 32, the eyes are waiting for the hands to catch up).

In summary, our model offers an alternative method in modeling and quantifying a diverse range of phenomena in typing. QN-MHP offers not only theoretical insights into typing performance, but is a step toward developing proactive multi-purpose analysis tools for textual data-entry tasks in human-computer interaction.

Acknowledgments This article is based upon work supported by the National Science Foundation under Grant No. NSF 0308000. However, any opinions, findings and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation (NSF).

Reference

Aizawa, H., Inase, M., Mushiake, H., Shima, K., & Tanji, J. (1991). Reorganization of Activity in the Supplementary Motor Area Associated with Motor Learning and Functional Recovery. *Experimental Brain Research, 84*(3), 668-671.

Alexander, R. M. (1993). Optimization of Structure and Movement of the Legs of Animals. *Journal of Biomechanics, 26*, 1-6.

Anderson, F. C., & Pandy, M. G. (2001). Static and dynamic optimization solutions for gait are practically equivalent. *Journal of Biomechanics, 34*(2), 153-161.

Anderson, J. R., & Lebiere, C. (1998). *The Atomic Components of Thought*: Lawrence Erlbaum Associates.

Anderson, J. R., Qin, Y. L., Sohn, M. H., Stenger, V. A., & Carter, C. S. (2003). An information-processing model of the BOLD response in symbol manipulation tasks. *Psychonomic Bulletin & Review, 10*(2), 241-261.

Anderson, J. R., Qin, Y. L., Stenger, V. A., & Carter, C. S. (2004). The relationship of three cortical regions to an information-processing model. *Journal of Cognitive Neuroscience, 16*(4), 637-653.

Armstrong, T. J. (2004). *Applied Anthropometry*: http://ioe.engin.umich.edu/ioe491Anthrodata.htm accessed on May 21, 2004.

Asari, T., Konishi, S., Jimura, K., & Miyashita, Y. (2005). Multiple components of lateral posterior parietal activation associated with cognitive set shifting. *Neuroimage, 26*(3), 694-702.

Baddeley, A. D. (1992). Working memory. *Science, 255*(5044), 556-559.

Baltes, P. B., Staudinger, U. M., & Lindenberger, U. (1999). Lifespan psychology: Theory and application to intellectual functioning. *Annual Review of Psychology, 50*, 471-507.

Bear, M. F., Connors, B. W., & Paradiso, M. A. (2001). *Neuroscience: exploring the brain* (8th ed.). Baltimore, MD: Lippincott Williams & Wilkins,.

Beehler, P. J. H., & Wright, J. P. (1999). Effects of age and task complexity when performing manual aiming movements: A test of Hick's law using physically active adults. *Research Quarterly for Exercise and Sport, 70*(1), A66-A66.

Black, I. B. (1999). Trophic regulation of synaptic plasticity. *Journal of Neurobiology, 41*(1), 108-118.

Boettiger, C. A., & D'Esposito, M. (2005). Frontal networks for learning and executing arbitrary stimulus - Response associations. *Journal of Neuroscience, 25*(10), 2723-2732.

Bor, D., Cumming, N., Scott, C. E. L., & Owen, A. M. (2004). Prefrontal cortical involvement in verbal encoding strategies. *European Journal of Neuroscience, 19*(12), 3365-3370.

Borghese, N. A., & Calvi, A. (2003). Learning to maintain upright posture: What can be learned using adaptive neural network models? *Adaptive Behavior, 11*(1), 19-35.

Braus, D. F. (2004). Neurobiology of learning - The basis of an alteration process. *Psychiatrische Praxis, 31*, S215-S223.

Bullock, T. (1968). Representation of information in neurons and sites for molecular participation. *Proc Natl Acad Sci U S A, 60*(4), 1058-1068.

Bustillos, A. T., & de Oliveira, P. M. C. (2004). Evolutionary model with genetics, aging, and knowledge. *Physical Review, 69*(2).

Butsch, R. L. C. (1932). Eye movements and the eye-hand span in typewriting. *Journal of Educational Psychology, 23*, 104-121.

Card, S., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hinsdale, NJ: Lawrence Erlbaum.

Chklovskii, D. B., Mel, B. W., & Svoboda, K. (2004). Cortical rewiring and information storage. *Nature, 431*(7010), 782-788.

Cohen, M. S. (1997). Parametric analysis of fMRI data using linear systems methods. *Neuroimage, 6*(2), 93-103.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204-256.

Cook, A. S., & Woollacott, M. H. (1995). *Motor Control: Theory and Practical Applications*. Philadelphia, PA: Williams& Wilkins.

Drury, C. G. (1975). Application of Fitt's Law to foot-pedal design. *Human Factors, 17*, 368-373.

Duric, Z., Gray, W. D., Heishman, R., Li, F. Y., Rosenfeld, A., Schoelles, M. J., et al. (2002). Integrating perceptual and cognitive modeling for adaptive and intelligent human-computer interaction. *Proceedings of the IEEE, 90*(7), 1272-1289.

Eagleman, D., Jacobson, J., & Sejnowski, T. (2004). Perceived luminance depends on temporal context. *Nature, 428*(6985), 854-856.

Eberts, R. (1997). Cognitive Modeling. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (pp. 1329-1373). New York: John Wiley.

Faw, B. (2003). Pre-frontal executive committee for perception, working memory, attention, long-term memory, motor control, and thinking: A tutorial review. *Consciousness and Cognition, 12*(1), 83-139.

Feyen, R. (2002). *Modeling Human Performance using the Queueing Network ---Model Human Processor (QN-MHP).* University of Michigan, Ann Arbor, MI.

Fish, L. A., Drury, C. G., & Helander, M. G. (1997). Operator-specific model: An assembly time prediction model. *Human Factors and Ergonomics in Manufacturing, 7*(3), 211-235.

Fletcher, P. C., & Henson, R. N. A. (2001). Frontal lobes and human memory - Insights from functional neuroimaging. *Brain, 124*, 849-881.

Gan, K., & Hoffman, E. R. (1988). Geometrical conditions for ballistic and visually controlled movment. *Ergonomics, 31*, 829-839.

Genovesio, A., Brasted, P. J., Mitz, A. R., & Wise, S. P. (2005). Prefrontal cortex activity related to abstract response strategies. *Neuron, 47*(2), 307-320.

Gentner, D. R. (1983). The acquisition of typewriting skill. *Acta Psychologica, 54*, 233-248.

Georgopoulos, A. P., Taira, M., & Lukashin, A. (1993). Cognitive Neurophysiology of the Motor Cortex. *Science, 260*(5104), 47-52.

Gerard, M., Armstrong, T. J., Franzblau, A., Martin, B. J., & Rempel, D. M. (1999). The effect of keyswitch stiffness on typing force, finger electromyography, and subjective discomfort. *American Industrial Hygiene Association Journal, 60*, 762-769.

Ghilardi, M. F., Ghez, C., Dhawan, V., Moeller, J., Mentis, M., Nakamura, T., et al. (2000). Patterns of regional brain activation associated with different forms of motor learning. *Brain Research, 871*(1), 127-145.

Gilbert, P. F. C. (2001). An outline of brain function. *Cognitive Brain Research, 12*(1), 61-74.

Goldstein, M., Book, R., Alsio, G., & Tessa, S. (1999). *Non-Keyboard QWERTY Touch Typing: A Portable Input Interface For the Mobile User.* Paper presented at the CHI.

Gordon, A. M., Lee, J. H., Flament, D., Ugurbil, K., & Ebner, T. J. (1998). Functional magnetic resonance imaging of motor, sensory, and posterior parietal cortical areas during performance of sequential typing movements. *Experimental Brain Research, 121*(2), 153-166.

Gordon, A. M., & Soechting, J. F. (1995). Use of tactile afferent information in sequential finger movements. *Experimental Brain Research, 107*, 281-292.

Habib, M. (2003). Rewiring the dyslexic brain. *Trends in Cognitive Sciences, 7*(8), 330-333.

Hansen, M. H., Hurwitz, W. N., & Madow, W. G. (1953). *Sample Survey Methods and Theory* (Vol. I). New York: John Wiley & Sons, Inc.

Hasegawa, M., Carpenter, P. A., & Just, M. A. (2002). An fMRI study of bilingual sentence comprehension and workload. *Neuroimage, 15*(3), 647-660.

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review, 7*(2), 185-207.

Heiss, D. G., & Pagnacco, G. (2002). Effect of center of pressure and trunk center of mass optimization methods on the analysis of whole body lifting mechanics. *Clinical Biomechanics, 17*(2), 106-115.

Hershman, R. L., & Hillix, W. A. (1965). Data-Processing in Typing - Typing Rate as a Function of Kind of Material and Amount Exposed. *Human Factors, 7*(5), 483-492.

Hobbs, J. B. (1993). *Homophones and homographs: An American Dictionary*. Jefferson, NC: McFarland & Company, Inc.

Inhoff, A. W., Briihl, D., Bohemier, G., & Wang, J. (1992). Eye-Hand Span and Coding of Text During Copytyping. *Journal of Experimental Psychology-Learning Memory and Cognition, 18*(2), 298-306.

Inhoff, A. W., Topolski, R., & Wang, J. (1992). Saccade Programming During Short Duration Fixations - an Examination of Copytyping, Letter Detection, and Reading. *Acta Psychologica, 81*(1), 1-21.

Inhoff, A. W., & Wang, J. (1992). Encoding of text, manual movement planning, and eye-hand coordination during copytyping. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 437 -448.

Ito, M. (1991). Short-term retention of a constructed motor program. *Perceptual and motor skill, 72*(1), 339-347.
Izzetoglu, K., Bunce, S., Onaral, B., Pourrezaei, K., & Chance, B. (2004). Functional optical brain imaging using near-infrared during cognitive tasks. *International Journal of Human-Computer Interaction, 17*(2), 211-227.
Janelle, C. M., Champenoy, J. D., Coombes, S. A., & Mousseau, M. B. (2003). Mechanisms of attentional cueing during observational learning to facilitate motor skill acquisition. *Journal of Sports Sciences, 21*(10), 825-838.
John, B. E. (1988). *Contributions to Engineering models of Human-Computer Interaction.* Carnegie-Mellon University, Pittsburgh, Pennsylvania.
John, B. E. (1996). TYPIST: A theory of performance in skilled typing. *Human-Computer Interaction, 11*(4), 321-355.
John, B. E., & Kieras, D. E. (1996). The GOMS family of user interface analysis techniques: comparison and contrast. *ACM Transactions on Human-Computer Interaction, 3*(4), 320-351.
John, B. E., & Kieras, D. E. (1996). Using GOMS for user interface design and evaluation: Which technique? *ACM Transactions on Human-Computer Interaction, 3*(4), 287-319.
John, B. E., & Newell, A. (1989). *Cumulating the science of HCI: From S-R Compatibility to transcription typing.* Paper presented at the CHI'89 Proceedings.
Jueptner, M., & Weiller, C. (1998). A review of differences between basal ganglia and cerebellar control of movements as revealed by functional imaging studies. *Brain, 121*, 1437-1449.
Jurgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews, 26*(2), 235-258.
Just, M. A., & Carpenter, P. N. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review, 99*, 122-149.
Kansaku, K., Hanakawa, T., Wu, T., & Hallett, M. (2004). A shared neural network for simple reaction time. *Neuroimage, 22*(2), 904-911.
Karlin, L., & Kestenbaum, R. (1968). Effects of Number of Alternatives on Psychological Refractory Period. *Quarterly Journal of Experimental Psychology, 20*, 167-178.
Kaufer, D. I., & Lewis, D. A. (1999). *Frontal Lobe Anatomy and Cortical Connectivity. In The Human Frontal Lobes: Functions and Disorders.* New York, NY: The Gullford Press.
Kieras, D. E. (1994). *GOMS Modeling of User interfaces using NGOMSL.* Paper presented at the Human Factors in Computing Systems, CHI'94, Boston, Massachusetts, USA.
Krampe, R. T., Rapp, M. A., Bondar, A., & Baltes, R. B. (2003). Allocation of cognitive resources during the simultaneous performance of cognitive and sensorimotor tasks. *Nervenarzt, 74*(3), 211-218.
Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence, 33*, 1-64.
Langolf, G., D. , Chaffin, D., & Foulke, J. (1976). An Investigation of Fitts' Law Over a Wide Range of Movement Amplitudes. *Journal of Motor Behavior, 8*(2), 113-128.
Laureys, S., Peigneux, P., Phillips, C., Fuchs, S., Degueldre, C., Aerts, J., et al. (2001). Experience-dependent changes in cerebral functional connectivity during human rapid eye movement sleep. *Neuroscience, 105*(3), 521-525.
Leonard, J. A., & Carpenter, A. (1964). On the correlation between a serial choice reaction task and subsequent achievement at typewriting. *Ergonomics, 7*(2), 197-204.
Li, Z. M., Zatsiorsky, V. M., Li, S., Danion, F., & Latash, M. L. (2001). Bilateral multifinger deficits in symmetric key-pressing tasks. *Experimental Brain Research, 140*, 86-94.
Lim, J., & Liu, Y. (2004a). *A queueing network model for eye movement.* Paper presented at the Proceedings of the 2004 International Conference on Cognitive Modeling, Pittsburg, PA.
Lim, J., & Liu, Y. (2004b). *A queueing network model of menu selection and visual search.* Paper presented at the Proceedings of the 48 Annual Conference of the Human Factors and Ergonomics Society, New Orleans, Louisiana, USA.
Liu, Y., Feyen, R., & Tsimhoni, O. (in press). Queueing Network-Model Human Processor (QN-MHP): A Computational Architecture for Multitask Performance. *ACM Transaction on Human Computer Interaction*.
Liu, Y. L. (1996). Queueing network modeling of elementary mental processes. *Psychological Review, 103*(1), 116-136.

Liu, Y. L. (1997). Queueing network modeling of human performance of concurrent spatial and verbal tasks. *IEEE Transactions on Systems Man and Cybernetics Part a-Systems and Humans, 27*(2), 195-207.
Logan, G. D. (1982). On the Ability to Inhibit Complex Movements - a Stop-Signal Study of Typewriting. *Journal of Experimental Psychology-Human Perception and Performance, 8*(6), 778-792.
Manoach, D. S., Schlaug, G., Siewert, B., Darby, D. G., Bly, B. M., Benfield, A., et al. (1997). Prefrontal cortex fMRI signal changes are correlated with working memory load. *Neuroreport, 8*(2), 545-549.
McIntosh, A. R. (1999). Mapping cognition to the brain through neural interactions. *Memory, 7*(5-6), 523-548.
McIntosh, A. R. (2000). Towards a network theory of cognition. *Neural Networks, 13*(8-9), 861-870.
Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance .1. Basic mechanisms. *Psychological Review, 104*(1), 3-65.
Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance .2. Accounts of psychological refractory-period phenomena. *Psychological Review, 104*(4), 749-791.
Michael, E. B., Keller, T. A., Carpenter, P. A., & Just, M. A. (2001). fMRI investigation of sentence comprehension by eye and by ear: Modality fingerprints on cognitive processes. *Human Brain Mapping, 13*(4), 239-252.
Mitz, A. R., Godschalk, M., & Wise, S. P. (1991). Learning-Dependent Neuronal-Activity in the Premotor Cortex - Activity During the Acquisition of Conditional Motor Associations. *Journal of Neuroscience, 11*(6), 1855-1872.
Murata, A. (1996). Empirical Evaluation of Performance Models of Pointing Accuracy and Speed with a PC Mouse. *International Journal of Human Computer Interaction, 8*(4), 457-469.
Murdock, B. B. (1961). The retention of individual items. *Journal of Experimental Psychology, 62*, 618-625.
Mustovic, H., Scheffler, K., Di Salle, F., Esposito, F., Neuhoff, J. G., Hennig, J., et al. (2003). Temporal integration of sequential auditory events: silent period in sound pattern activates human planum temporale. *Neuroimage, 20*(1), 429-434.
Nakamura, K., Sakai, K., & Hikosaka, O. (1998). Neuronal activity in medial frontal cortex during learning of sequential procedures. *Journal of Neurophysiology, 80*(5), 2671-2687.
Nazir, T. A., Ben-Boutayab, N., Decoppet, N., Deutsch, A., & Frost, R. (2004). Reading habits, perceptual learning, and recognition of printed words. *Brain and Language, 88*(3), 294-311.
Nazir, T. A., Decoppet, N., & Aghababian, V. (2003). On the origins of age-of-acquisition effects in the perception of printed words. *Developmental Science, 6*(2), 143-150.
Neptune, R. R. (1999). Optimization algorithm performance in determining optimal controls in human movement analyses. *Journal of Biomechanical Engineering-Transactions of the Asme, 121*(2), 249-252.
Newell, A. (1973). You cann't play 20 questions with nature and win. Projective comments on the papers of the symposium. In W. G. Chase (Ed.), *Visual Information processing*. Washington, D.C.: Academic Press.
Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
Ohbayashi, M., Ohki, K., & Miyashita, Y. (2003). Conversion of working memory to motor sequence in the monkey premotor cortex. *Science, 301*(5630), 233-236.
Olson, J. R., & Olson, G. M. (1990). The growth of cognitive modeling in human-computer interaction since GOMS. *Human-Computer Interaction, 5*, 221-265.
Pandy, M. G., Anderson, F. C., & Hull, D. G. (1992). A Parameter Optimization Approach for the Optimal-Control of Large-Scale Musculoskeletal Systems. *Journal of Biomechanical Engineering-Transactions of the Asme, 114*(4), 450-460.
Pearson, R., & van Schaik, P. (2003). The effect of spatial layout of and link colour in web pages on performance in a visual search task and an interactive search task. *International Journal of Human-Computer Studies, 59*(3), 327-353.
Penfield, W., & Rasmussen, T. (1950). *The cerebral cortex of man: A clinical study of localization of function*. New York, NY: MacMillen.
Petersen, S. E., van Mier, H., Fiez, J. A., & Raichle, M. E. (1998). The effects of practice on the functional anatomy of task performance. *Proceedings of the National Academy of Sciences of the United States of America, 95*(3), 853-860.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422.
Reinkensmeyer, D. J., Iobbi, M. G., Kahn, L. E., Kamper, D. G., & Takahashi, C. D. (2003). Modeling reaching impairment after stroke using a population vector model of movement control that incorporates neural firing-rate variability. *Neural Computation, 15*(11), 2619-2642.
Roland, P. E. (1993). *Brain activation*. New York, NY: Wiley-Liss.
Rolls, E. T. (1994). Neurophysiology and Cognitive Functions of the Striatum. *Revue Neurologique, 150*(8-9), 648-660.
Rolls, E. T. (2000). Memory systems in the brain. *Annual Review of Psychology, 51*, 599-630.
Rothkopf, E. Z. (1980). Copying Span as a Measure of the Information Burden in Written Language. *Journal of Verbal Learning and Verbal Behavior, 19*(5), 562-572.
Rouse. (1980). *Systems Engineering Models of Human-Machine Interaction* (Vol. 6). New York: North Holland.
Rudell, A. P., & Hu, B. (2001). Does a warning signal accelerate the processing of sensory information? Evidence from recognition potential responses to high and low frequency words. *International Journal of Psychophysiology, 41*(1), 31-42.
Rumelhart, D. E., & Norman, D. A. (1982). Simulating a Skilled Typist - a Study of Skilled Cognitive-Motor Performance. *Cognitive Science, 6*(1), 1-36.
Sadato, N., Ibanez, V., Campbell, G., Deiber, M. P., LeBihan, D., & Hallett, M. (1997). Frequency-dependent changes of regional cerebral blood flow during finger movements: Functional MRI compared to PET. *Journal of Cerebral Blood Flow and Metabolism, 17*(6), 670-679.
Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Sasaki, Y., & Putz, B. (1998). Transition of brain activation from frontal to parietal areas in visuomotor sequence learning. *Journal of Neuroscience, 18*(5), 1827-1840.
Salthouse, T. (1983). Why Is Typing Rate Unaffected by Age. *Gerontologist, 23*, 68-68.
Salthouse, T. A. (1984a). Effects of Age and Skill in Typing. *Journal of Experimental Psychology-General, 113*(3), 345-371.
Salthouse, T. A. (1984b). The Skill of Typing. *Scientific American, 250*(2), 128-135.
Salthouse, T. A. (1985). Anticipatory Processing in Transcription Typing. *Journal of Applied Psychology, 70*(2), 264-271.
Salthouse, T. A. (1986a). Effects of Practice on a Typing-Like Keying Task. *Acta Psychologica, 62*(2), 189-198.
Salthouse, T. A. (1986b). Perceptual, Cognitive, and Motoric Aspects of Transcription Typing. *Psychological Bulletin, 99*(3), 303-319.
Salthouse, T. A., & Saults, J. S. (1987). Multiple Spans in Transcription Typing. *Journal of Applied Psychology, 72*(2), 187-196.
Salvucci, D. D. (In press). Modeling Driver Behavior in a Cognitive Architecture. *Human Factors*.
Schmidt, R. A. (1988). *Motor control and learning*. Champaign, IL: Human Kinetics Publishers.
Schmuck, P., & WobkenBlachnik, H. (1996). Behavioral flexibility and working memory. *Diagnostica, 42*(1), 47-66.
Schubotz, R. I., Friederici, A. D., & von Cramon, D. Y. (2000). Time perception and motor timing: A common cortical and subcortical basis revealed by fMRI. *Neuroimage, 11*(1), 1-12.
Schubotz, R. I., & von Cramon, D. Y. (2001). Functional organization of the lateral premotor cortex: fMRI reveals different regions activated by anticipation of object properties, location and speed. *Cognitive Brain Research, 11*(1), 97-112.
Shaffer, L. H. (1975). Control processes in typing. *Quarterly Journal of Experimental Psychology, 27*, 419-432.
Simon, S. R., Meunier, M., Piettre, L., Berardi, A. M., Segebarth, C. M., & Boussaoud, D. (2002). Spatial attention and memory versus motor preparation: Premotor cortex involvement as revealed by fMRI. *Journal of Neurophysiology, 88*(4), 2047-2057.
Siok, W. T., Jin, Z., Fletcher, P., & Tan, L. H. (2003). Distinct brain regions associated with syllable and phoneme. *Human Brain Mapping, 18*(3), 201-207.
Smith, E. E., & Jonides, J. (1998). Neuroimaging analyses of human working memory. *Proc. Natl. Acad. Sci. USA, 95*, 12061-12068.

Steyvers, M., Etoh, S., Sauner, D., Levin, O., Siebner, H. R., Swinnen, S. P., et al. (2003). High-frequency transcranial magnetic stimulation of the supplementary motor area reduces bimanual coupling during anti-phase but not in-phase movements. *Experimental Brain Research, 151*(3), 309-317.

Tanaka, S. (1994). Numerical Study of Coding of the Movement Direction by a Population in the Motor Cortex. *Biological Cybernetics, 71*(6), 503-510.

Tanaka, S., Honda, M., & Sadato, N. (2005). Modality-specific cognitive function of medial and lateral human Brodmann area 6. *Journal of Neuroscience, 25*(2), 496-501.

Taylor, J. (2003). Paying attention to consciousness. *Progress in Neurobiology, 71*(4), 305-335.

Taylor, J., Horwitzc, B., Shaha, N. J., Fellenzb, W. A., Mueller-Gaertnera, H.-W., & Krausee, J. B. (2000). Decomposing memory: functional assignments and brain traffic in paired word associate learning. *Neural Networks, 13*, 923-940.

Trainor, P. G. S., McLachlan, K. R., & McCall, W. D. (1995). Modeling of Forces in the Human Masticatory System with Optimization of the Angulations of the Joint Loads. *Journal of Biomechanics, 28*(7), 829-843.

Tsimhoni, O., & Liu, Y. (2003). *Modeling steering using the queueing network model human processor (QN-MHP).* Paper presented at the Proceedings of the Human Factors and Ergonomics Society 47th annual meeting, Santa Monica, CA.

Tsimhoni, O., Smith, D., & Green, P. (2004). Address entry while driving: speech recognition versus a touch-screen keyboard. *Human Factors, 46*(4), 600-610.

Van Mier, H., Tempel, L. W., Perlmutter, J. S., Raichle, M. E., & Petersen, S. E. (1998). Changes in brain activity during motor learning measured with PET: Effects of hand of performance and practice. *Journal of Neurophysiology, 80*(4), 2177-2199.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature, 383*(6597), 254-256.

Welford, A. T. (1968). *Fundamentals of skill*. London: Methuen.

Wu, C., & Liu, Y. (2004a). *Modeling Behavioral and Brain Imaging Phenomena in Transcription Typing with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA.

Wu, C., & Liu, Y. (2004). *Modeling human transcription typing with Queueing network-model human processor.* Paper presented at the Proceedings of the 48th Annual Meeting of Human Factors and Ergonomics Society, New Orleans, Louisiana, USA.

Wu, C., & Liu, Y. (2004b). *Modeling Psychological Refractory Period (PRP) and Practice Effect on PRP with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA.

Wu, C., Zhang, K., & Hu, Y. (2003). Human performance modeling in temporary segmentation Chinese character handwriting recognizers. *International Journal of Human Computer Studies, 58*, 483-508.

Zysset, S., Muller, K., Lehmann, C., Thone-Otto, A. I. T., & von Cramon, D. Y. (2001). Retrieval of long and short lists from long term memory: a functional magnetic resonance imaging study with human subjects. *Neuroscience Letters, 314*(1-2), 1-4.

Appendix

Appendix 1. Derivation of Equation 1

Equation 1 is derived based on Black (1999)'s model as well as other neuroscience findings. Black (1999) proposed a model to explain the role of BDNF (brain derived neurotrophic factor) in its regulation of synaptic plasticity in adults—BDNF increases the activity of NMDA (N-methyl-D-aspartate) receptors, increases neuron channel open probability by increasing opening frequency, and then increases the velocity of spikes trains travel (V) through these neuron channels (Black, 1999). Hence, the stronger synaptic connection strength (the amount of presynaptic transmitter released and the degree of postsynaptic responsiveness) of an individual route, the greater the probability (P_i) that spikes trains (represented by entities) travel through that route (Black, 1999; Braus, 2004; Chklovskii, et al., 2004; Habib, 2003) (see Equation 7 and Figure 2-8).

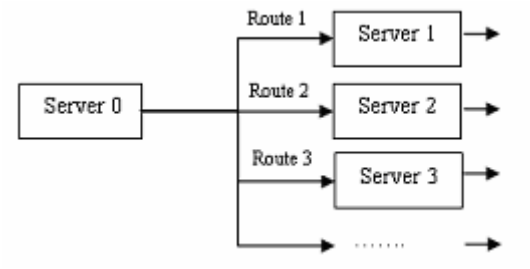


Figure 2-8 Multiple routes for one location in Queueing network (server 0 has *U* multiple routes as output)

$$P_i = \frac{ST_i}{\sum_{j=1}^{U} ST_j} \tag{7}$$

In Equation 7, the numerator (ST_i) stands for the standardized synaptic connection strength of route i ($ST_i \in [0, 1]$). The denominator represents the sum of the standardized synaptic connection strength of all the multiple routes starting from the original brain region (server 0 in Figure 2-8). Moreover, the standardized synaptic connection strength of route i (ST_i) is in direct ratio with the standardized velocity (Vi) that the spikes trains travel through that route (Black, 1999; Bullock, 1968; Chklovskii, et al., 2004) (see Equation 8).

$$ST_i = r_0 V_i \tag{8}$$

In Equation 8, r_0 is a parameter stands for the ratio between ST_i and V_i.

Since the Queueing network is able to capture several properties of information processing in the human brain—spikes trains carrying information (represented by entities) travel through different brain regions and form a "brain traffic" including possible waiting of the previous information flow to be processed (see the first learning mechanism in the Queueing network), the travel time of the spikes trains (represented by entities) in route i is composed of both waiting and processing time and therefore this travel time can be regarded as the sum of waiting time (W_i) and processing time (T_i) of entities, i.e. sojourn time (S_i) in that route. Furthermore, this sojourn time or travel time (sum of waiting and processing time) is in inverse ratio with the standardized velocity (V_i) of the travel process (see Equation 9).

$$V_i = \gamma\left(\frac{1}{W_i + T_i}\right) = \frac{\gamma}{S_i} \tag{9}$$

In Equation 9, γ is a parameter represents the inverse ratio between W_i+T_i and Vi.

Combining Equation 7-9, Equation 10 and 11 quantify the probability (P_i) that the spikes trains (entities) pass through route i in totally U multiple routes.

$$P_i = \frac{ST_i}{\sum_{j=1}^{U} ST_j} = \frac{r_0 V_i}{\sum_{j=1}^{U} r_0 V_j} = \frac{r_0\left(\frac{\gamma}{S_i}\right)}{\sum_{j=1}^{U}\left[r_0\left(\frac{\gamma}{S_j}\right)\right]} = \frac{1/S_i}{\sum_{j=1}^{U}\frac{1}{S_j}} \tag{10}$$

Thus, we have:

$$P_i = \frac{1/S_i}{\sum_{j=1}^{U}\frac{1}{S_j}} \tag{11}$$

In short, learning process increases the synaptic connection strength, which improves the effectiveness of the information processing of brain regions in the neuron pathway (route) and then changes the probability that the majority of spikes trains (entities) enter one of multiple neuron pathways (routes). If the majority of entities change their route from one to another, rewiring of neuron pathways (routes) occurs.

Appendix 2. Calculation of the Expected Position of the First Character in Each Chunk in a Word

The expected position of the first character in each chunk can be estimated by using the following logic. Suppose the position of characters of a word is starting from 1, based on the definition of different units in typing (see Figure 2-2), the expected position of the first character in each chunk ($E(FC)$) can be quantified into Equation 12.

$$E(FC) = E(FP) - \left\lceil \frac{x_{opt} - 1}{2} \right\rceil \tag{12}$$

In Equation 12 above, $E(FP)$ stands for the expected position of the fixation point and $\left\lceil \frac{x_{opt} - 1}{2} \right\rceil$ refers to the half-range of each chunk under extensive practice condition. Since the average word length is 5 characters (John, 1988, 1996), the expected fixation point is located at the middle point of the first half of a word (see the simulation mechanism of phenomenon 10), i.e. the second character ($E(FP)$=2). In addition, the optimal chunk size is 4 characters (x_{opt}=4, see Appendix 4). Therefore, $E(FC)$ equals 1 (see Equation 13), i.e. the expected position of the first character in each chunk equals the first character in each word in transcription typing.

$$E(FC) = E(FP) - \left\lceil \frac{x_{opt} - 1}{2} \right\rceil = 2 - \left\lceil \frac{4 - 1}{2} \right\rceil = 1 \tag{13}$$

Appendix 3. Processing Logic of Hand and Foot Servers

This section describes the context-free processing logic of the hand and foot servers in detail.

1. Hand Servers

The processing logic of hand servers includes three aspects: simulated movement time, distribution of movement distance and the pressing force of the fingers in the hand servers.

1) Movement Time

The simulated movement time of hands including their fingers is estimated depending on whether the movement is executed with visual guidance or not. If the movement is executed with the visual guidance, a variant of Fitt's Law (Welford, 1968) is used to estimate the horizontal movement time (*MT*) of the hands including their fingers (see Equation 14).

$$MT = I_m \log_2(Dis/S+0.5) \quad \text{(Welford, 1968)} \tag{14}$$

In Equation 14, *Dis* is the movement distance; *S* refers to the size of a key or button (*S*=1.3 cm for a standard QWERTY keyboard); and I_m is a parameter corresponding to different parts of the hands, e.g., for fingers, I_m=1000/38=26.3 (Langolf et al., 1976).

If the movement can be executed without visual guidance (e.g. ballistic movements), e.g., movements in typing after extensive practice, the Queueing network model uses the formula proposed by Gan and Hoffman (1988) to estimate the movement time:

$$MT = a + b\sqrt{Dis} \quad \text{(Gan and Hoffman, 1988)} \tag{15}$$

a and *b* are constants depending on number of components in the movement (e.g., a=52.95, b=15.72 for the movement composed of single component, Gan and Hoffman, 1988).

2) Distribution of Movement Distance

The distribution of movement distance is estimated based on the findings in neurological studies which discovered that the movement direction of body parts can be predicted by the action of motor cortical neurons in the primary motor cortex (Georgopoulos, et al., 1993). When individual cells in the primary motor cortex are represented as vectors, they make weighted contributions along the axis of their preferred direction and the resulting vector (population vector) is the sum of all of these cell vectors. Tanaka (1994) quantified the RMSE (room-mean-square error) of the movement direction ($RMSE_\theta$) of certain body part as a function of the population size (*M*) of corresponding brain area in the primary motor cortex (see Equation 16).

$$RMSE_\theta = 97.3M^{-1/2} - 0.1 \quad \text{(Tanaka, 1994)} \tag{16}$$

Since RMSE in general can be quantified into Equation 17 (Hansen, et al., 1953), where $\bar{\theta}-\tilde{\theta}$ refers the difference between the expected value of the sample mean ($\bar{\theta}$) and the true value of θ ($\tilde{\theta}$) (unit of θ is degree).

$$RMSE_\theta = \sqrt{SD_\theta^2 + (\bar{\theta}-\tilde{\theta})^2} \qquad \text{(Hansen, et al., 1953)} \tag{17}$$

According to the law of large numbers in statistics, when the value of the sample size increases to a great value (e.g., sample size>1000), $\bar{\theta}$ is closing to $\tilde{\theta}$, i.e. $(\bar{\theta}-\tilde{\theta})^2 \to 0$. Thus,

$$RMSE_\theta = \sqrt{SD_\theta{}^2 + (\bar{\theta}-\tilde{\theta})^2} = \sqrt{SD_\theta{}^2} = SD_\theta$$

Moreover, since Tanaka (1994) found the distribution of θ follows normal distribution, combining Equation 16 and 17, the distribution of θ can be quantified into Equation 18 where SD_θ stands for the standard deviation of the distribution.

$$\theta \sim N(\bar{\theta}, SD_\theta)$$

$$\text{i.e. } \theta \sim N(\bar{\theta},\ 97.3M^{-1/2}-0.1) \tag{18}$$

Based on Equation 18, given that the movement distance (Dis) is the product of the $2\pi\times$movement radius (RD) and θ/360 (i.e. $Dis=(\theta/360)\times 2\pi RD$), the distribution of movement distance can be estimated via Equation 19.

$$Dis \sim N\{\overline{Dis},\ [(97.3M^{-1/2}-0.1)/360]\times 2\pi RD\} \tag{19}$$

Based on the value of M measured in neuroscience studies and the value of RD measured in anthropometry studies, Equation 19 can be used to estimate the distribution of movement distance of different body parts including hand and fingers. For example, given that the population size (M) of the brain area corresponding to each finger (M =7300 on average, Reinkensmeyer et al., 2003; Penfield and Rasmussen, 1950) and movement radius (RD) in typing (17.5 cm on average, since the hands of the typist are moved to reach different keys with the wrist as an axis and the average distance from the wrist to the tip of fingers is 17.5 cm (Armstrong, 2002)), the distribution of movement distance of each finger on average follows Equation 20.

$$\begin{aligned} &Dis \sim N\{\overline{Dis},\ [(97.3\times 7300^{-1/2}-0.1)/360]\times 2\pi\times 17.5\} \\ &i.e.\ Dis \sim N(\overline{Dis},\ 0.317) \qquad \text{(unit: cm)} \end{aligned} \tag{20}$$

3) Distribution of Finger Pressing Force

Table 2-7 is directly quoted from Li et al. (2001)'s study which summarized the mean and standard deviation in the distribution of fingers' pressing force (F~N(M, SD)) in a key pressing task under bilateral multi-finger condition.

Table 2-7 Finger Force and its Variability in a Key Pressing Task (Li, et al., 2001)

	Right hand				Left hand			
Mean & SD	Little	Ring	Middle	Index	Little	Ring	Middle	Index
M (Newton)	6.2	9.8	18.5	17.4	7.8	9.9	15.1	19.4
SD	2.2	2.5	3.0	2.7	1.8	1.2	3.2	2.8

The forces of the 8 fingers are implemented in the model's two hand servers as 8 variables which follow the normal distribution with mean and standard deviation in Table 2-7.

2. Foot Server

The foot server executes the simulated movement to press a pedal and its movement time (MT_{foot}) can be estimated by the formula proposed by Drury (1975) (Equation 21), where S refers to the shoe width (10 cm, Armstrong, 2002); W is the pedal width (10 cm, same with the shoe width) and A stands for the movement distance (3 cm, typical movement distance for a foot pedal).

$$MT_{foot}=(1/1.64)[0.1874+0.0854\times log_2(A/(W+S)+0.5)] \quad \text{(Drury, 1975)} \quad (21)$$

Appendix 4. Optimization of the Parameters of the Queueing Network

To simulate the trial-and-error learning in the motor learning process, Monte Carlo simulation[3] is performed in server B and Y to find the optimal value of five parameters in transcription typing task: chunk size (x), number of chunks (c), EPD (cross-hand error prevention duration), 2FC (two-finger coordination time), and 1FW (one-finger waiting time).

1. Chunk size (x) and Number of Chunks (c)

[3] The length of the Monte Carlo simulation (number of letters typed by the model) is the same with the approximate number of letters typed during the learning process of typing (10,000,000 letters, Genter, 1983) (There are totally 50 runs for the Monte Carlo simulation). The random numbers used in each run as the stochastic input to the model are the standard random number series in Promodel software (Promodel, 2004).

In processing the normal text in human-machine interaction (e.g., typing and handwriting), the chunk size (x) and the number of chunks (c) at server B are determined by the optimization process which trades off the time required to correct the wrongly processed entities caused by failure in preserve the characters at server B with the time saved by increasing the size of each chunk and number of chunks.

Definitions:

x: chunk size	R: average duration to correct an error caused by a wrongly processed entity or character
x_{opt}: expected optimal chunk size	
N: total number of entities processed	
w: overall duration of processing each chunk at servers after server B	N/x: total chunks of a normal text which composed of N entities or characters
c: current number of chunks at server B	cx: current number of entities at server B
e_{pho}: rate of retrieval failure at server B	w/x: duration of processing each entity or character
	w(N/x): overall duration of processing N entities or characters:

Objective function:

$$Z= Min\ [w(N/x)+ e_{pho}NR]=Min\ \{N\ [(w/x)+ e_{pho}R]\} \quad (22)$$

i.e. $$Z'= Min\ \ [(w/x)+ e_{pho}R] \quad (23)$$

On aspect of typing out the chunks and fixing the errors in retrieval of these chunks, the objective function (Equation 22) of the task completion time in this aspect is composed of two parts: a) typing time ($w(N/x)$, i.e., overall duration of processing each chunk at servers (w) multiply by the total chunks of a normal text which composed of N (N/x)); b) fixing time (rate of retrial failure of entities (e_{pho}) multiplied by total number of entities processed and average duration to correct a wrongly processed entity).

Constraints:

a) The average preservation duration of each character at server B (B_p) is quantified in Equation 24:

$$B_p = \frac{1}{cx}\sum_{n=1}^{n=cx} n\ (w/x) = 0.5(1+cx)(w/x) \quad (24)$$

For example, suppose 3 characters (L1, L2, L3) enter server B with order L1 to L3, and the duration of L3 being preserved at server B equals (w/x) waiting for the current character to exit the model so that L3 can enter server W. Similarly, duration of L2 preserved at server B is 2 (w/x) and L1 is 3(w/x). Thus, the average preservation duration of each character is $[(1+2+3)/3]\times(w/x)$.

b) Based on the decay rate of characters at server B (Card, et al., 1983):

i) If $1 \le cx \le 4$ (one word condition; average word length is 4 for the most frequent used words in Murdock's experiment, 1961):

$$e_{pho} = .0065 \times 0.5(1+cx)(w/x) \quad (25)$$

ii) If $5 \le cx \le 8$ (2 words condition, deducted from 1 and 3 words condition):

$$e_{pho} = .0403 \times 0.5(1+cx)(w/x)+0.1 \quad (26)$$

iii) If $9 \le cx \le 13$ (3 words condition):

$$e_{pho} = .074 \times 0.5(1+cx)(w/x)+0.1 \quad (27)$$

Therefore, the objective functions in three different conditions are:

i) If $1 \le cx \le 4$: $Z' = (w/x) + 0.0065 \times 0.5(1+cx)(w/x)R$ (28)

ii) If $5 \le cx \le 8$: $Z' = (w/x) + 0.0403 \times 0.5(1+cx)(w/x)R + 0.1R$ (29)

iii) If $9 \le cx \le 13$: $Z' = (w/x) + 0.074 \times 0.5(1+cx)(w/x)R + 0.1R$ (30)

In the learning process of the model, the optimal value of c and x are selected via Monte Carlo simulation based on the objective functions in the three different conditions. For example, given the range of w ($.5 \le w \le 5$ sec) in typing normal text and R=2726 ms (determined by simulation results of the model in correcting a typing error), by simulating the objective functions based on the constraints, we obtained the optimal value of x and c in typing condition: c_{opt}=1 chunk, x_{opt}=4 characters (see Figure 2-9 as a graphical illustration). Based on Equation 25, e_{pho}=0.3 %. In general, Equation 22-30 are not task-specific and they can be applied into modeling other text processing tasks including reading, handwriting, and typing with other keyboards.

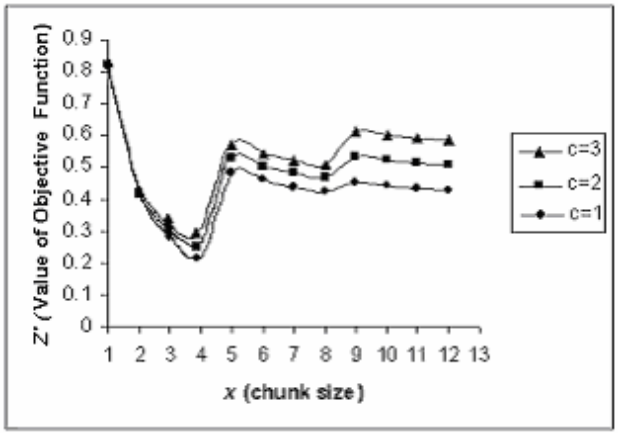


Figure 2-9 The change of objective function value (Z') with chunk size (*x*) and number of chunks (*c*) (*w*=0.8 sec based on the simulation results in typing normal text at well-learned situation; the curves of *c*>3 conditions are located above the curve c=3 condition, following the same pattern)

2. EPD (Cross-hand Error Prevention Duration)

According to the Queueing structure of the two hands, the entities or characters belong to different hands have to wait EPD to prevent the frequent occurrence of the transposition error, otherwise the transposition error always occurs when the interkey time of previous keystroke is longer than that of the current keystroke in this 2H situation. The improvement of overlapping movement of the two hands is quantified as the reduction of EPD via its optimization process.

The optimization process of EPD is a trade-off between the time in typing and the time in error correcting—reducing the value of EPD causes: 1) more efficient overlapping of the movements of the two hands, reducing the interkey time; 2) higher probability in making transposition error, increasing the time in error correcting. This trade-off can be quantified in the following equations.

The time (Y) saved by optimization of EPD is:

$$Y=N(EPD_0-EPD)- eNR_t \quad (31)$$

In Equation 31, N is the number of characters typed; e refers to the error rate of the transposition error made by reducing of EPD; R_t specifies how long to correct one transposition error; and EPD_0 is the original value of EPD at beginning of learning.

Hence, the optimization of EPD can be quantified with the following equations:

Objective function:

$$Max\ (Y) = Max\ [N(EPD_0\text{-}EPD)\text{-}\ eNR_t] \quad (32)$$

Constraints: e=f (EPD) (f is the function which represents the relationship between EPD and e) (0<=e<=1); 0<=EPD<= 176 ms (the maxim value of EPD is lower than one interkey time on average).

Specification of parameters and constraints:

a) N=1000 characters as the total number of characters in the sample text.

b) EPD_0 = 354 ms as the 2 times of an average interkey time (a sensitivity analysis indicates that this initial value of EPD does not affect simulation results)

c) R_t = 3112 ms on average which is determined by simulation results of the model in correcting a transposition error.

d) e=f (EPD), the relationship between e and EPD is set via the curve estimation of the simulation results (R square=.996) (see Equation 33).

$$e=0.16\text{-}0.034ln(EPD) \quad (33)$$

Consequently, the objective function can be simplified into the following Equation 34:

$$\begin{aligned} &Max\ (Y) = Max\ [N(EPD_0\text{-}EPD)\text{-}\ eNR_t] \\ &= Max\ \{1000\times(354\text{-}EPD)\text{-}\ [0.16\text{-}0.034ln(EPD)] \\ &\times 1000\times 3122\} \end{aligned} \quad (34)$$

Monte Carlo simulation was performed during the learning process of the model. During the learning process, the value of EPD was updated after typing every 50 characters. A better value of EPD which generated a greater value of Y replaced the original value of EPD and therefore we obtained the optimal value of EPD and its range (EPD_{opt}=108±10 ms) to maximize Y value of the objective function (see Figure 2-10 for the curve of objective function) via the learning process.

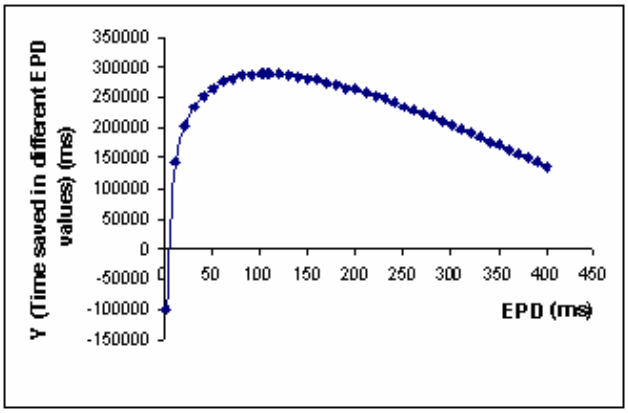


Figure 2-10 The relationship of EPD and Y value of the objection function in Monte Carlo simulation results

3. Two-finger Coordination Time and One-finger Waiting Time

The value of 2FC (two-finger coordination time) and 1FW (one-finger waiting time) are set based on the similar Monte Carlo simulation logic during the learning process—the two parameters are updated during the learning process to minimize the interkey time. The obtained optimal values of the two parameters were 0 ms.

Appendix 5. Calculation of the Expected Copying span and Replacement Span

The copying span and replacement span can be estimated based on following mechanisms. For the copying span, i) in the motor subnetwork, since the half-life of entities in the motor subnetwork is 1000 ms, the last entity in the motor subnetwork decays at the end of the 1000 ms with .5 of chance when the input to the model is stopped; therefore, including this last entity, the total expected number of entities exited from the motor subnetwork is 1000/interval of leaving = 1000/simulated interkey time = 1000/176 ≈ 6 entities, i.e. the expected number of entities in the motor subnetwork is 6; ii) in the cognitive subnetwork, only server B is in the route of entities (see simulation mechanism of phenomenon 24 and 25), and it holds 1 chunk (x_{opt}=4 characters, see Appendix 4); iii) in the perceptual subnetwork, when 4 entities in the motor subnetwork leave the model (it takes 4×176=704 ms on average), which allows a chunk to leave server B and entities from the perceptual subnetwork enter server B, all of the entities in the perceptual

subnetwork have already decayed since the half-life of information in the perceptual subnetwork is only 200 ms. In sum, the expected copying span is 6 characters in the motor subnetwork plus 1 chunk (4 characters) in the cognitive subnetwork, i.e. 10 characters.

For the replacement span, the 6 characters in the motor subnetwork are distributed in the 5 servers in the motor subnetwork (server W, Y, Z, and two hand servers) and each server holds or processes 6/5=1.2 characters on average. Accordingly, the expected number of entities in server Z and 2 hand servers is 1.2×3=3.6 characters, i.e. the expected replacement span is 3.6 characters.

Appendix 6. Sources of Equations and their Parameters (Table 2-8)

Table 2-8 Equations and Sources of Equations and Parameters

Equations and their Parameters	Sources of Equations and Parameters
Equation 1	Black, 1999; Bullock, 1968; Chklovskii, et al., 2004
Si (Sojourn time of route *i*)	Value obtained during the simulation of the model (sum of waiting time (W_i) and processing time (T_i) of entities, Feyen 2002; Liu, et al., 2006)
Equation 2	Heathcote, et al., 2000
$A_{i,}$ (the expected minimal processing time (*Ti*) at server i after intensive practice) B_i (change of expected processing time from the beginning to the end of practice)	Feyen, 2002; Liu, et al., 2006; Rektor, et al., 2003
α_i (0.001, learning rate of server i)	Heathcote, et al., 2000
N_i (10,0000, number of entities processed by server i)	Nelson-Denny Reading Test used (Salthouse's study (1984a, 1984b, 1987)
Equations 3-6	Gross, 1988; Fundamentals of queueing theory
Equations 3-6 are served to prove the change of expected interkey time and its variation.	
Equations 7-11	Black, 1999; Bullock, 1968; Chklovskii, et al., 2004
Equations 3-11 are served to derive Equation 1.	
Equations 12-13	Definition of units of typing (Salthouse, 1986; 1984a, 1984b, 1987)
FP (2, expected position of the fixation point in a word)	Rayner, 1998
X_{opt} (4, optimal chunk size)	Derived based on Equations 22-30 in Appendix
Equation 14	Welford, 1968
Im (26.3, a parameter corresponding to different parts of the hands)	Langolf et al., 1976
Dis (movement distance)	Standard QWERTY keyboard and averaged anthropometric data of hands (Armstrong, 2002)
S (1.3 cm, size of each key)	Standard QWERTY keyboard

Equation 15	Gan and Hoffman, 1988
a=52.95, b=15.72 constants for the movement composed of single component of fingers	Gan and Hoffman, 1988
Equations 16-18	Tanaka, 1994; Hansen, et al., 1953;
Equations 16-18 are used to develop Equations 19-20	
Equations 19-20	Equations 16-18
M (7300 clusters, population size (*M*) of the brain area corresponding to each finger)	Reinkensmeyer et al., 2003; Penfield and Rasmussen, 1950
RD (17.5 cm, movement radius)	Armstrong, 2002
Equation 21	Drury, 1975
S (10 cm, shoe width), *W* (10 cm, the pedal width)	Anthropometric data of foot (Armstrong, 2002)
A (3 cm, movement distance of foot)	Typical movement distance for a foot pedal. Based a sensitivity analysis, when A varies from 3-10 cm (maximum of foot movement on a pedal), it did not affect the simulation results of current task.
Equations 22-24 See Definition of the parameters in Appendix 4 and their values are set during the optimization process	Developed based on the nature of the composition of task completion time and preservation duration of each character at a server (see paragraph right below those two equations)
Equation 25-34	Developed based on Equations 22-24

Chapter 3
Queueing Network Modeling of Psychological Refractory Period (PRP)

Chapter Summary

PRP (psychological refractory period) is a basic but important form of human information processing in dual tasks. Existing models of PRP including response selection bottleneck (RSB), EPIC-SRD, and ACT-R/PM, regarding cognition either as a serial or parallel process, each encounters at least one experimental counterexample(s) to their predictions or modeling mechanisms. Based on corresponding neuroscience evidence, Queueing Network-Model Human Processor (QN-MHP)—a computational architecture that quantifies the cognitive process with both serial and parallel properties—is able to account for various experimental findings in PRP including all of these major counterexamples of existing models with less or equal number of free parameters and no need to use task-specific lock/unlock assumptions required by both EPIC and ACT-R/PM, thus demonstrating its unique advantages in modeling dual-task performance. Theoretical implications of the model as well as its extension in the future research are discussed.

1. Introduction

Performing multiple tasks at the same time is common in daily life, e.g., drivers can steer a car and at same time talk with friends in the car; telephone operators can answer customer phone calls and type textual information into a computer. Among these multiple tasks, psychological refractory period (PRP) is one of the most basic and simplest forms of a dual-task situation. PRP has been studied in laboratories over 100 years from the behavioral level (Solomons & Stein, 1896; Welch, 1898; Creamer, 1963; Kantowitz, 1974; Oberauer & Kliegl, 2004; Pashler, 1984a, 1984b, 1994b; Schumacher et al., 1999;

Welford, 1952) to the neurological level (Jiang, Saxe, & Kanwisher, 2004; Sommer et al., 2001). It is also the subject of extensive theoretical work and the focal point of an important theoretical controversy between several computational models of cognition—there are several important cognitive models of PRP, including response selection bottleneck (RSB) or central bottleneck model proposed by Pashler (1984, 1990, 1994), executive-process interactive control model-strategic response deferment (EPIC-SRD) proposed by Meyer & Kieras (1997) and ACT-R/perceptual-motor system (ACT-R/PM) proposed by Byrne & Anderson (2001). Each of these models is able to account for some of the important aspects of PRP; however, each appears to encounter at least one experimental counterexample to its predictions from either behavioral experiments or electrophysiological or brain imaging studies (Jiang, Saxe, & Kanwisher, 2004; Meyer & Kieras, 1997a, 1997b; Oberauer & Kliegl, 2004; Ruthruff, Pashler, & Klaassen, 2001). Therefore, the question remains on how to model these experimental results, unify the discoveries both in behavioral and neurosciences studies, and provide a deeper understanding the mechanism of dual-task performance.

This article takes further steps toward addressing this important question with a queueing network based computational cognitive architecture. First, we introduce the major experimental results in PRP studies and the major PRP effects. Second, the major existing models of PRP are described, including their advantages and their counterexamples. Third, we introduce the major assumptions and components of the queueing network model. In the fourth section, we describe how the queueing network model models the basic PRP paradigm and the counterexamples via mathematic modeling. Finally, we discuss the theoretical implications of the model as well as its extension in the future research.

2. Experimental Studies in PRP

In the following section, we introduce the major findings in experimental studies of PRP including the basic PRP experiment paradigm and the major effects related to the theoretical controversy—subadditivity difficulty effect, response grouping effect, practice effect and brain imaging pattern in PRP.

Basic PRP Experiment Paradigm

The basic PRP experiment paradigm requires subjects to perform two tasks called task 1 (T1) and task 2 (T2) concurrently. The delay between the presentation of the stimulus of T1 and T2 is called stimulus onset asynchrony (SOA). Two stimuli (S1 and S2) are presented to subjects in rapid succession and each requires a quick response (R1 and R2). Reaction time of each task (RT1 and RT2) is measured from the time when the stimulus is presented to the time when the corresponding response is made. In the basic PRP paradigm in which the tasks are choice reaction time tasks (Kantowitz, 1974) and subjects do not receive extensive practice on the dual tasks, typically, responses to the first stimulus (S1) are unimpaired, but responses to the second stimulus (S2) are slowed by 300 ms or more at short SOA conditions (Ruthruff, Pashler, & Klaassen, 2001) (Figure 3-1 as the experimental results of Schumacher et al., 1999).

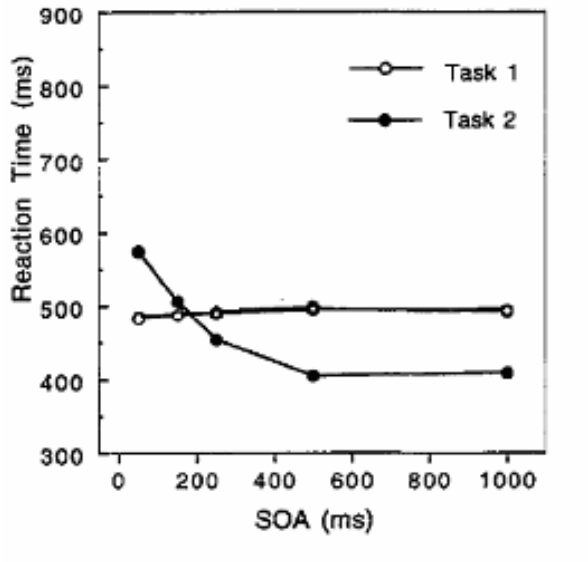


Figure 3-1 Typical experimental results in the basic PRP experiment paradigm (Schumacher et al., 1999)

Subadditivity Difficulty Effect

Several experiments in PRP (Karlin & Kestenbaum, 1968; Hawkins, et al., 1979; Schumacher et al., 1999; Sommer et al., 2001) found if the difficulty level of T2 at its central processing stage (response selection stage occurring after perceptual process and before motor process) is manipulated, the difference of RT2 between easy and hard T2

under short SOA conditions is smaller than that under long SOA conditions. This pattern or effect in these experimental results is called the subadditive difficulty effect.

Karlin & Kestenbaum's Experimental Study (1968)

Karlin & Kestenbaum (1968) found the subadditivity difficulty effect by manipulating the difficulty level of T2 via the number of S-R pairs—one was a simple reaction task and the other is a two choice reaction task.

In their experiment, T1 was a visual-manual task: subjects were asked to respond to the digits (1 to 5) on a visual display by pressing the fingers on the left hand corresponding to the digits beginning with the number one for the little finger. T2 was an auditory-manual task where the index and middle fingers of the right hand of subjects are used to respond to high and low tones respectively. Their experimental results clearly demonstrated the pattern of subadditivity difficulty effect (see Figure 3-2).

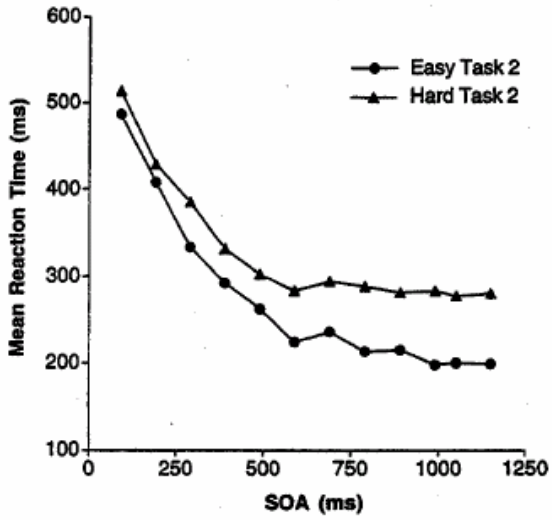


Figure 3-2 Experimental results of subadditive difficulty effect (Karlin & Kestenbaum, 1968)

Hawkins, Rodriguez, & Reicher's Experimental Study (1979)

The subadditive difficulty effect can also be found in the experimental results of Hawkins et al.'s study (1979). The difficulty level of T2 in Hawkin et al. (1979) was manipulated by the number of stimuli in a category in making the same response: in the

easy T2 condition, the stimuli were the digits 2 and 3, and the responses were keypresses with the right-hand index and middle fingers, respectively (one stimulus in the category corresponding one response); in the hard T2 condition, the stimuli were the digits 2-9—four of them (2, 5, 6, 9) belong to the first category and the left four digits belong to the second category, and subjects were asked to press the right-hand index or middle finger key once they saw one of these four digits in the first or the second category.

Schumacher et al.'s Experimental Study (1999)

Schumacher et al. (1999) also found the subadditive difficulty effect in their Experiments 3 and 4. The level of difficulty of T2 was manipulated via the degree of compatibility of task 2. With different perceptual and motor modalities, they found an underadditive interaction between SOA and the response-selection difficulty effect on mean Task 2 RTs (i.e., the subadditive difficulty effect). In their Experiment 3, task 1 (T1) was an auditory-manual task: participants heard either an 1120-Hz or a 1450-Hz tone and responded by pressing the left middle-finger or left index-finger on a keypad, respectively. T2 was a visual-manual task. In each trial of Task 2, an O replaced one of four dashes in a horizontal row centered on the display monitor. In the compatible situation of T2, participants pressed the right index, middle, ring, or little finger keys when the O appeared in the far left, middle left, middle right, or far right spatial positions, respectively. In the incompatible situation of T2, participants pressed the right index, middle, ring, or little finger keys when the O appeared in the middle left, far right, far left, or middle right positions, respectively. The stimuli for the two tasks were separated by one of five SOAs: 50, 150, 250, 500, or 1,000 ms. The only difference between their Experiments 3 and 4 was that the subject made vocal rather than manual responses to the stimuli in T1.

Sommer et al.'s Electrophysiological Study (2001)

Sommer et al. (2001) replicated Karlin & Kestenbaum's experiment (1968) and used the ERP (event-related potential) techniques to measure stimulus-lateralized readiness potential (LRP) that is a measure of response preparation or activation at the cerebral motor cortex including the premotor cortex (Sommer et al., 2001; Wildgruber,

Ackermann, & Grodd, 2001). They found that when T2 was a simple reaction task, there was an early onset of stimulus-LRP before S2 was presented (its onset time was 15 ms before the S2 is presented), which meant that the motor component of the cognitive system started to prepare the processing of S2 before it was presented. Moreover, their results suggested that the percentage of negative RT2 (RT2 <0 ms) increased with an increase of SOA, which was consistent with the results of another behavioral study (Van Selst & Jolicoeur, 1997) that also found the percentage of negative RT2 increased with an increase of SOA.

Brain Imaging Patterns (Jiang et al., 2004)

Jiang et al. (2004) conducted the first fMRI study closely following the paradigm of PRP and did not find the increase of activation in the brain regions corresponding to "executive control" in short SOA conditions compared with long SOA conditions. Jiang et al. (2004) tested a large number of subjects in an effort to find neural correlates of the basic PRP with fMRI techniques. In their experiment, both task 1 and 2 were two choice reaction tasks. Task 1 was a visual-manual task: square or circles were presented on a display, and subjects pressed ''1'' for a square and ''2'' for a circle with the left hand. Task 2 was also a visual-manual task and there were two groups of subjects to perform task 2: the first group's task 2 was a letter discrimination task—subjects responded to a letter "A" or "B" by pressing the number "3" or "4" on a keypad, respectively; the second group's task 2 was a color discrimination task—subjects responded red or green crosses by pressing "3" or "4" on the keypad. Jiang et al. (2004) measured all of the activation of the brain areas related to the possible executive control, including the DLPFC (dorsal lateral prefrontal cortex), the ACC (anterior cingulate cortex), the GFi (inferior frontal gyrus), the ADPFC (anterior-dorsal prefrontal cortex), the SPL (superior parietal lobule), and the GFm (middle frontal gyrus). However, they found there was virtually no increase of activation in these brain regions in the short SOA conditions compared with the long SOA conditions.

Response Grouping Effect (Ruthruff et al., 2001)

Ruthruff et al. (2001) designed a new PRP experiment paradigm in which subjects were asked to emit both responses at the same time (grouping the responses, hence, called "response grouping effect" in this article) and stimuli of T1 and T2 were presented at same time (SOA=0) with the same emphasis. They found that the RT1 and RT2 in dual tasks were significantly longer than those in single task situations. In their experiment, one of the two tasks was a tone counting task—subjects were asked to count the number of tones (one or two). In the easy version, the subjects said this number aloud; in the hard version, they said the opposite number aloud (i.e., "two" if they heard one tone, "one" if they heard two tones). The other task was a spatial working memory task—subjects responded whether upside-down letter stimuli were normal or mirror images by pressing the "j" and "k" keys, respectively. In single task blocks, either a tone or a letter appeared (chosen at random). In dual-task blocks, the tone and the letter always appeared simultaneously. The subjects were instructed to emit both responses at about the same time (response grouping). In addition, to test the cost of response grouping, Ruthruff et al. (2001) also conducted several control experiments which found that the cost of response grouping was only 21 ms (Ruthruff et al., 2001). In addition, subjects in Ruthruff et al.'s experiment received relatively extensive practice in both single and dual-task situations (2 blocks of single task practice and 2 blocks of dual-task practice, each block include 20 warm-up trials and 80 practice trails).

Practice Effect on PRP (Oberauer et al., 2004)

Oberauer et al.'s experimental study (2004) found that people can perform two complex cognitive tasks at the same time after hundreds of trials in dual-task practice (Oberauer & Kliegl, 2004). In their experiment, one experimental task was a spatial working memory task: spatial operations were indicated by red arrows with a length of 2.5 cm, displayed in the central cell of a 5×5 grid and pointing in one of eight possible directions horizontally, vertically, or diagonally; subjects were asked to mentally shift a dot from its current location (cell 1) to another location (cell 2) in the indicated direction. Another task was a numerical operations task in which a given digit from 1-9 is presented at the center of the

grid. The task was indicated by 50 ms tones—a high tone (800 Hz) requiring an addition of 2 and a low tone (200 Hz) requiring a subtraction of 1 from the current value of the digit. There were 7-9 operations in each trial. After the subjects completed the current mental and/or numerical operation(s), they were asked to hit the space bar and the stimuli of the next operation (or pair of operations) was displayed immediately. At the end of the trials, the subjects were asked the final digit value and the final spatial position in a random order.

There were two practice groups of subjects in the experiment—a single-task and a dual-task practice group. In the single-task practice group, subjects only practiced one task at one time and there was no trial in which two stimuli of the two tasks were presented at the same time. In the dual-task practice group, two stimuli of the two tasks were always presented at the same time. Each group received a pretest and a posttest for their performance before and after extensive practice (12 sessions). There were two conditions in each test—a sequential and a simultaneous condition. In the sequential condition, the test started with a complete sequence of operations of one kind (either spatial or numerical, selected at random from trial to trial), followed by the same number of successive operations of the other kind. The maximum reaction time of numerical and spatial task was regarded as the reaction time of the two tasks in this sequential condition. In the simultaneous condition, one numerical and one spatial operation were displayed at the same time. The reaction time of the two tasks was the duration between when the stimuli of the two tasks appeared on the display and when the space-bar was hit in each operation.

How can these major effects in PRP be modeled by existing cognitive models? After reviewing all of these major effects in PRP, it is important to review existing models including their assumptions.

3. Existing Models of PRP

Over the past decades, several computational models were developed to model the experimental results in PRP, including response selection bottleneck model (RSB) or central bottleneck model (Pashler, 1984, 1990, 1994), executive-process interactive control model-strategic response deferment (EPIC-SRD) (Meyer & Kieras, 1997) and ACT-R/perceptual-motor system (ACT-R/PM) (Byrne & Anderson, 2001). Among these representative models, RSB and ACT-R/PM assume the serial processing at the cognitive process, while EPIC-SRD assumes that the cognitive process is parallel. Each of these models successfully accounted for one or several experimental studies reviewed above; however, it appears that each of them also encounters at least one experimental study as counterexample(s) to their predictions or modeling mechanisms (see Table 3-1).

Table 3-1 Coverage of experimental studies and effects by the existing models

Experimental Results	Coverage of Existing Cognitive Models		
	Serial in Cognitive Process		Parallel in Cognitive Process
	RSB	ACT-R/PM	EPIC-SRD
1. Basic PRP	Yes	Yes	Yes[1]
2. Subadditive Difficulty Effect			
Schumacher et al. (1999)	CR	Yes	-
Hawkins et al. (1979)	CR	-	Yes
Karlin & Kestenbaum (1968)	CR	-	Yes[2]
Sommer et al. (2001)	-	-	CR
3. Brain Imaging Pattern			
Jiang, et al. (2004)	-	-	CR
4. Response Grouping Effect			
Ruthruff et al. (2001)	-	-	CR
5. Practice Effect on PRP			
Oberauer et al. (2004)	CR	CR	-

Yes: modeled the experimental results
CR: experimental results might be contradictory to the prediction or basic assumptions of the model
-: experimental results are not modeled but the they may not contradictory to the basic assumptions or the prediction of the model
1. But based on strategic scheduling mechanism which appears to be contradictory to the fMRI study (Jiang, et al., 2004)
2. But based on post-response selection bottleneck mechanism which seems to be contradictory to the ERP study (Sommer et al., 2001).

Response Selection Bottleneck Model (RSB)

The RSB model was developed by Pashler (1984, 1990, 1994) based on the work of other researchers in dual-task performance (Smith, 1967; Welford, 1952). The basic assumption of the RSB model is that multiple stimuli may be identified simultaneously and stored in short-term memory, but the process of response selection (i.e., converting symbolic stimulus codes to symbolic response codes, also called the "central process", "cognitive process", or "central stage") is able to accommodate only one task at a time.

Overall, the RSB model is a very parsimonious model and its serial processing assumption at the central stage of cognition is able to explain the experimental result of the basic PRP paradigm, the effect of increasing the difficulty level of task 1 (increased

the difficulty level of T1 prolongs RT2), and the effect of precentral task 2 manipulation (when the difficulty level of T2 at the pre-central processing increases, the difference of RT2 between easy and hard T2 under short SOA conditions is smaller than that in the long SOA conditions) (Pashler, 1984a, 1984b, 1989, 1990, 1994a, 1994b, 1994c; Pashler, Johnston, & Ruthruff, 2001; Pashler et al., 1994).

However, based on the review of experimental studies in the previous section, it appears that there are two major effects in PRP study that might be contradictory to the prediction or assumptions of the RSB model.

First, the subadditive difficulty effect indicates that if we manipulate the difficulty level of T2 at its central process stage, the difference of RT2 between easy and hard T2 under short SOA conditions is smaller than that under long SOA conditions. However, according to the prediction of the RSB model, if the duration of cognitive process of T2 increases, the difference between the easy and hard T2 under short SOA conditions should be the same as that under long SOA conditions (see the detailed description in Byrne & Aderson, 1998).

Second, the practice effect of PRP in Oberauer et al. (2004) demonstrated that people can perform two complex cognitive tasks at the same time after several sessions of dual-task practice, which might be contradictory to the serial cognitive processing as the core assumption of the RSB model. Moreover, the PRP effect did not disappear with the same amount of single-task practice and the central bottleneck can be eliminated only through dual-task practice. Even though Pashler, Johnston, and Ruthruff (2000) have recently argued that the bottleneck need not be structural, this effect might remain a challenge for the RSB model to explain why the disappearance of PRP can be achieved only by dual-task practice rather than by single-task practice.

ACT-R/PM

ACT-R/PM was developed by Byrne & Anderson (2001) based on the original structure of ACT-R (Anderson & Lebiere, 1998) and the perceptual and motor part of EPIC. At the central processing stage, ACT-R/PM assumes the cognition (production firing) is serial and this is the most crucial assumption in ACT-R/PM. To support this assumption, Byrne & Anderson (2001) conducted a series of experiments that provided

evidence that people can not perform two complex cognitive tasks at the same time. At the motor processing stage, since the motor part of ACT-R/PM is adapted from EPIC which assumes serial processing at each motor module, the second serial processing or bottleneck in ACT-R/PM is located at its motor module (movement feature preparation) (Byrne & Anderson, 1998).

ACT-P/PM offers accurate quantification of the basic PRP paradigm and the subadditive difficulty effect in PRP studies. First, ACT-R/PM simulated the result of the basic PRP paradigm in Karlin & Kestenbaum's experiment (1968) with the bottleneck at the motor module (Byrne & Anderson, 1998). In this simulation, ACT-R/PM acts as the RSB model in which a structural bottleneck in the motor module (motor feature preparation) produces the experimental results in the basic PRP paradigm. Furthermore, the multiple-bottleneck assumption enables ACT-R/PM to model the subadditive difficulty effect in Experiment 3 of Schumacher et al.'s study (1999) with the same multiple-bottleneck mechanism originally proposed by Dejong (Dejong, 1993). The subadditive difficultly effect found in Experiment 4 of Schumacher et al.'s study (1999) is simulated by ACT-R/PM with similar scheduling strategy in EPIC-SRD.

ACT-R/PM used an objective and systematic method to set the values of free parameters: free parameters are adjusted in long SOA conditions to fit the simulation results with experimental results, and then without changing their value, these parameters are used in the model at short SOA conditions to generate the simulation results. Therefore, in this parameter setting method, there is no free parameter to fit the experiment result at short SOA conditions.

Similar to the counterexamples to the RSB model, since ACT-R/PM regards the seriality of production execution as a constraint of the cognitive architecture and all of the response selection in ACT-R is allocated at the production processor (Anderson et al., 2004), without adding some "jump cables" to connect stimulus and response directly (Anderson et al., 2004), it might be difficult for the current version of ACT-R/PM to explain the experimental result of Oberauer et al. (2004) which demonstrated that people can perform two complex cognitive tasks at same time after 10 sessions of dual-task practice. Moreover, similar to the RSB model, it appears to be a challenge for the current

version of ACT-R/PM to explain naturally why the disappearance of PRP can be achieved only by dual-task practice rather than by single-task practice.

EPIC-SRD

EPIC is one of the most comprehensive cognitive architectures in quantifying and simulating the dual-task performance (Meyer & Kieras, 1997a, 1997b, 1999). In contrast to the RSB model, the basic assumptions of EPIC in modeling PRP are that—there is no "hardware" bottleneck in the cognitive process and the observed bottleneck in PRP is the result of strategic or voluntary control in the cognitive process as well as the serial processing in peripheral motor output process (i.e., each motor processor processes information in a serial manner). Within this architecture, Meyer and Kieras developed the strategic response deferment (SRD) (called as "executive control") to model PRP phenomena—the executive control in EPIC needs to monitor the progress of both T1 and T2, using complex scheduling strategies to lock and unlock T1 and T2 in certain processing stage to produce the effects in PRP.

The most significant advantage of EPIC-SRD model is its strategic scheduling methods which provide great flexibility to simulate several major effects in PRP, e.g., the basic PRP, and the subadditive difficulty effect as introduced in the previous section. Moreover, EPIC itself is also a comprehensive architecture that unifies many findings in psychological studies and it has been successfully employed to model dual-task performance in both theoretic research (Meyer & Kieras, 1997a, 1997b) and applied research (Kieras & Meyer, 1997).

In comparing the EPIC-SRD's basic assumption with the major effects reviewed in the introduction section, it appears that EPIC-SRD also has several counterexamples raised in the existing experimental studies.

First, the experimental results in response grouping effect (Ruthruff, et al., 2001) might contradict EPIC's prediction: Ruthruff et al.'s experiment eliminates two possible bottlenecks: i) strategic postponement or bottleneck by emphasizing the two tasks equally, SOA=0, and urging subjects to emit both responses at the same time; and ii) perceptual bottleneck by using different perceptual modalities. If there is no strategic bottleneck at the central processing stage, according to the assumption of EPIC, the central processor

can process the information of the two tasks at the same time; plus there is no other non-central interference (interferences in perceptual or motor process), dual-task performance should be similar to single-task performance. However, they still found that the RT1 and RT2 in the dual tasks were significantly longer than in single task situations. This experimental result appears to be contradictory to the prediction of EPIC in central processing.

Second, according to the assumptions of EPIC, dual-task interference arises when SOA becomes shorter, since subjects need to monitor the progress of T1, halt T2, and resume T1 etc. Therefore, the brain areas corresponding to central executive should increase their activation level in short SOA conditions compared to long SOA conditions. However, the fMRI study of Jiang (2004) found that there was virtually no increase in these brain regions in the short SOA conditions compared with the long SOA conditions. "These data suggest that passive Queueing, rather than active monitoring, occurs during the PRP" (Jiang et al., 2004, p390).

Third, Sommer et al. (2001) replicated Karlin & Kestenbaum's experiment (1968) with ERP techniques and they measured the reaction time, percentage of negative RT2 and lateralized related potential (LRP). Ample research in ERP studies found that onset time of LRP reflects the starting stage of motor preparation in the cognitive system (Leuthold & Jentzsch, 2001; Leuthold & Jentzsch, 2002; Ulrich, Leuthold, & Sommer, 1998). When T2 was a simple reaction time task, Sommer et al. (2001) found that: first, the average onset time of LRP occurred before S2 was presented, which indicated that the cognitive system started to prepare the processing of S2 before it was presented; second, the percentage of negative RT2 (RT2<0 ms) increased as SOA increased. However, according to the current EPIC-SRD model, S2 is processed by the model after it is presented (Mayer & Kieras, 1997b, p756: "When Task 2 involves just one S-R pair [i.e., it is a simple-reaction task, detection of the auditory stimulus triggers the SRD model's Task 2 production rules…]"), the onset time of LRP in this simple reaction condition should not be observed before the arrival of S2 and RT2 at the behavioral level should be positive. Based on these two contradictions, Sommer et al. believed that their experimental results may reject the modeling mechanism of EPIC-SRD (a post-response

selection bottleneck) in generating the reaction time of Karlin & Kestenbaum's experiment with some confidence (Sommer et al., 2001, p87).

In summary, each existing model encounters at least one experimental study that can be regarded as a counterexample(s) to their predictions from either behavioral experiments or electrophysiological and brain imaging studies. And a new computational architecture with different modeling mechanism is expected to quantify these experimental results.

4. Modeling Mechanisms and Results

Based on the major assumptions of QN-MHP, in the following section we describe how a Queueing network model is able to model all of the counterexamples of the existing models raised in experimental studies of PRP as well as the basic PRP effect. In each effect or pattern, we introduce the experimental results in detail, the route of entities in the network, the corresponding modeling mechanism and mathematical equations in quantifying the effect or pattern, the parameter setting method, and the comparison between the modeling and experimental results.

4.1 Basic PRP

Modeling the experimental results in the basic paradigm of PRP gives an introduction to the modeling approach of the Queueing network model in this article, including how the routes of entities in the network are selected, how the mathematical models are developed, how the parameters in the model are set, and how the modeling results are validated with the experimental results.

One of the typical experimental results in the basic paradigm of PRP can be found in one of the experimental results of Schumacher et al.'s study (1999) (compatible Task 2 condition in Experiment 4): the reaction time of Task 1 is not affected by Task 2 but the reaction time of T2 is slower at short SOA conditions than at long SOA conditions. Similar patterns of these experimental results can be found in many other PRP studies (e.g., Karlin & Kestenbaum, 1968). In the compatible T2 condition of Experiment 4 in Schumacher et al.'s study (1999), Task 1 (T1) was an auditory-vocal task: participants heard either a 1120-Hz or a 1450-Hz tone and responded by saying "low" or "high", respectively. T2 was a visual-manual task: on each trial, a letter "O" replaced one of four dashes in a horizontal row centered on the display monitor, and participants pressed the right index, middle, ring, or little finger keys when the "O" appeared in the far left, middle left, middle right, or far right spatial positions, respectively.

Routes of Entities

The route of entities in the network is determined based on the previous Queueing network modeling work in modeling the connectivity of brain regions (Wu and Liu,

2004b): in general, depending on the task to be performed, servers whose function is related to the target task are included in the route of entities, which is consistent with the concept of "functional connectivity" discovered in neuroscience—a functional network for a particular cognitive task is defined by specifying the brain regions comprising the network as well as the anatomical links between these regions (Horwitz et al., 1999; McIntosh, 1999, 2000; Sporns, Tononi, & Edelman, 2000; Taylor, 2003; Taylor et al., 2000).

In Task 1, entities representing the auditory stimulus enter the auditory perceptual subnetwork first (Server 5->6/7->8); then, they are transformed into the cognitive subnetwork including Servers B, C, and F to make the phonological judgment. After that, since T1 involves vocalization (produces vocal response to the tones), Servers Y, Z and Mouth are in the route of entities according to the neuron pathways among the brain regions represented by these servers (see the introduction section in this article). Therefore, according to the connection of these brain regions the route of Task 1 is:

T1: 5->6/7->8->B->C->F->C->Y->Z->Mouth

For Task 2, entities representing the visual stimulus enter the visual perceptual subnetwork first (Servers 1->2/3->4). Via Server 4, the entities are transformed into the cognitive subnetwork including Servers A, C, and F in which the judgment is made. After that, they travel to the motor subnetwork (Servers W, Y, Z and hand server) to retrieve motor programs, assemble the motor programs, and initiate the motor response. As a result, according to the connection of these brain regions, the route of Task 2 is:

T2: 1->2/3->4-> A->C->F->C->W->Y->Z->Hand

Modeling Mechanisms

Entities of both tasks went through Server F which works as a structural bottleneck based on the assumption of QN-MHP in the cognitive subnetwork. Accordingly, in simulating Experiment 4 of Schumacher et al. (1999), the entities of Task 1 are not delayed but the entities of Task 2 have to wait for the entities of T1 to leave Server F before they can be processed at Server F (see Figure 3-3).

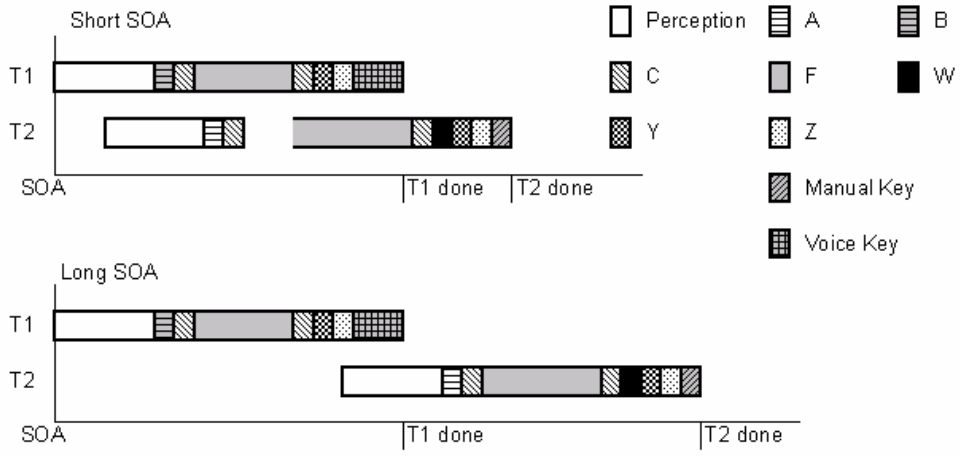


Figure 3-3 Modeling mechanisms of the basic PRP with QN-MHP

Mathematical Modeling of the Expected Reaction Time

The expected reaction time of T1 (*E(RT1)*) can be predicted by the sum of servers' processing time in the route of entities of T1 since no previous entities occupy any of the servers in the route (see T1 in Figure 3-3 and Equation 1).

$$E(RT1)=T_{1,AP}+T_{1,B}+T_{1,C}+T_{1,F}+T_{1,C}+T_{1,Y}+T_{1,Z}+T_{1,V} \quad (1)$$

where $T_{1,AP}$ is the processing time of the auditory perceptual subnetwork; $T_{1,B}$, $T_{1,C}$, $T_{1,F}$, $T_{1,Y}$, $T_{1,Z}$, and $T_{1,V}$ represent the processing time of Servers B, C, F, Y, Z and Mouth, respectively.

The expected reaction time of T2 (*E(RT2)*) depends on the comparison between a) the difference between SOA and the time point when entities of T1 exit Server F ($T_{1,AP}+T_{1,B}+T_{1,C}+T_{1,F}$ -*SOA*) and b) the duration of the processing time before entities of T2 enter Server F (the sum of processing time at the visual perceptual subnetwork, Servers A and C, $T_{2,VP}+T_{2,A}+T_{2,C}$) (see Equation 2): if part a is longer than part b (short SOA conditions), entities of T2 have to wait for the entities of T1 to finish their processing at Server F (short SOA conditions in Figure 3-3); hence, the waiting time of T2's entities is the difference between SOA and the time point when entities of T1 exit Server F ($T_{1,AP}+T_{1,B}+T_{1,C}+T_{1,F}$ -*SOA*). The duration of time that entities of T1 spend in the perceptual subnetwork as well as in Servers A and C is absorbed during this waiting process since other servers in the network can process the entities of T1 and T2 at the same time. As a result, the total processing time of T2's entities in this condition is the

waiting time ($T_{1,AP}+T_{1,B}+T_{1,C}+T_{1,F}$-$SOA$) plus the sum of the processing time of Server F and the following servers. Similarly, if part a is shorter or equal to part b (long SOA conditions), entities of T2 enter the vacant Server F immediately and the reaction time of T2 is the sum of the processing time of T2's entities at the perceptual subnetwork and at Servers A and C ($T_{2,VP}+ T_{2,A} +T_{2,C}$) as well as the processing time in the servers that follow Server C.

$$E(RT2) = max(T_{1,AP}+T_{1,B}+T_{1,C}+T_{1,F}-SOA,\ T_{2,VP}+T_{2,A}+T_{2,C})+T_{2,F}+T_{2,C}+T_{2,W}+T_{2,Y}+T_{2,Z}+T_{2,K} \quad (2)$$

Equation 2 above can be rewritten into:

$$E(RT2)=\begin{cases} T_{1,AP}+T_{1,B}+T_{1,C}+T_{1,F}-SOA+T_{2,F}+T_{2,C}+T_{2,W}+T_{2,Y}+T_{2,Z}+T_{2,K} \\ \qquad SOA < T_{1,AP}+T_{1,B}+T_{1,C}+T_{1,F} - (T_{2,VP}+T_{2,A}+T_{2,C}) \\ T_{2,VP}+T_{2,A}+T_{2,C}+T_{2,F}+T_{2,C}+T_{2,W}+T_{2,Y}+T_{2,Z}+T_{2,K} \\ \qquad SOA \geq T_{1,AP}+T_{1,B}+T_{1,C}+T_{1,F} - (T_{2,VP}+T_{2,A}+T_{2,C}) \end{cases} \quad (3)$$

Figure 3-4 shows the expected pattern of RT1 and RT2 with an increase of SOA. Based on Equation 1 of RT1, the expected RT1 keeps constant and its value is independent of SOA; based on Equation 3 of RT2, when $SOA<T_{1,AP}+T_{1,B}+T_{1,C}+T_{1,F}-(T_{2,VP}+T_{2,A}+T_{2,C})$, a negative linear relationship between SOA and RT2 is expected (slope= -1, see the first part of Equation 3); when SOA $\geq T_{1,AP}+ T_{1,B}+T_{1,C}+T_{1,F}-(T_{2,VP}+ T_{2,A} +T_{2,C})$, the expected RT2 keeps constant and its value is independent of SOA.

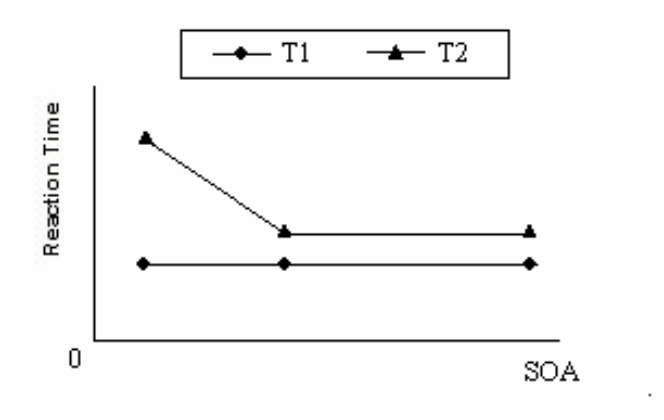


Figure 3-4 The expected pattern of reaction time in the basic PRP based on QN-MHP's prediction

Parameter Setting

The parameter setting method of QN-MHP in modeling all of the experimental studies in this article follows the same parameter setting method as in ACT-R/PM (Byrne & Anderson, 2001) (see introduction of ACT-R/PM in this article). For example, in modeling the basic PRP, only two free parameters in QN-MHP (processing time of Server F in T1 and T2) are set to fit the experimental data at long SOA conditions. The same values of these parameters are used at short SOA conditions to predict the RT1 and RT2. Therefore, at short SOA conditions, there are no free parameters to fit the experimental results in QN-MHP. Moreover, the value of these free parameters set at long SOA conditions are also constrained by the task properties: the processing times of T1 and T2 at Server F are close to each other (range of difference≤120 ms) since the difficulty levels of T1 and T2 are similar (see Table 3-2).

The other processing times are set following the parameter setting in the original setting of QN-MHP which models a wide range of human performance in various tasks (Feyen, 2002; Liu, Feyen, & Tsimhoni, 2005). In addition, the setting of the key closure time and voice key closure time are directly based on the study of Byrne and Anderson (2001) (see Table 3-2).

Table 3-2 Parameter setting in modeling of the basic PRP (Experiment 4, Schumacher et al, 1999)

Parameter	Value	Description	Source
$T_{1,AP}$	126 ms	Time for auditory perception time (42 ms at Servers 1,2/3, and 4, i.e., 42*3=126)	Liu et al. (in press)
$T_{2,VP}$	126 ms	Time for visual perception time (42 ms at Servers 1,2/3, and 4)	Liu et al. (in press)
$T_{2,A,}$ $T_{2,B}$	18 ms	Processing time at Server A and B	Liu et al. (in press)
$T_{1,C,}$ $T_{2,C}$	18 ms	Processing time at Server C	Liu et al. (in press)
$T_{1,F}$	179 ms	Processing time at Server F (T1)	Value estimated at long SOA conditions
$T_{2,F}$	165 ms	Processing time at Server F (T2)	Value estimated at long SOA conditions
$T_{2,W}$	24 ms	Processing time at Server W	Liu et al. (in press)
$T_{1,Y,}$ $T_{2,Y}$	24 ms	Processing time at Server Y	Liu et al. (in press)
$T_{1,Z,}$ $T_{2,Z}$	24 ms	Processing time at Server Z	Liu et al. (in press)
$T_{2,K}$	10 ms	Key closure time	Byrne & Anderson (2001)
$T_{1,V}$	100 ms	Voice key closure time	Byrne & Anderson (2001)

Modeling Results and its Validation

Figure 3-5 shows the modeling results compared with experimental results. The R square of the model is .95 and the RMS=22 ms.

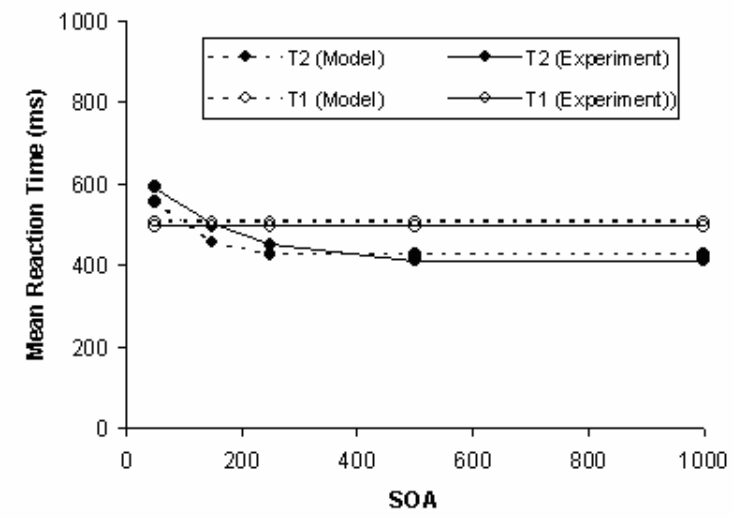


Figure 3-5 Mean reaction time in the basic PRP effect in Experiment 4 of Schumacher et al. (1999) (compatible T2 condition) (solid line) compared with modeling results (dashed lines)

4.2 Subadditive Difficulty Effect

The modeling of the subadditive difficulty effect includes modeling of 4 experimental studies: Schumacher et al. (1999), Hawkins et al. (1979), Karlin & Kestenbaum (1968)

and Sommer et al. (2001), and the difficulty level of T2 is manipulated by the degree of compatibility (compatible vs. incompatible, Schumacher et al., 1999), the number of stimuli in a category in making the same response (one stimulus vs. 4 stimuli in pressing either the left or the right index finger key, Hawkins et al., 1979), or the number of stimulus-response pairs (simple vs. choice reaction, Karlin & Kestenbaum, 1968; Sommer et al., 2001).

4.2.1 Schumacher et al.'s Experiments (1999)

Schumacher et al. (1999) conducted a series of experiments to provide evidence for adaptive executive control of task scheduling in PRP. The level of difficulty of T2 was manipulated via the degree of compatibility of Task 2. With different perceptual and motor modalities, they found a subadditive or underadditive interaction between SOA and response-selection difficulty effects on Task 2's mean reaction times. In their Experiment 3, T1 was an auditory-manual task and T2 was a visual-manual task with 2 levels of stimulus-response compatibility. The only difference between their Experiments 4 and 3 was that the subject made vocal response to the stimuli in T1 rather than using hands (see the detailed experiment description in the introduction part of this article).

Routes of Entities

Routes of Entities in Experiment 3

Similar to the routes in the modeling mechanism of the basic PRP, based on the task and corresponding function of the brain areas, the routes of Task 1 and Task 2 (both compatible and incompatible condition) are:

T1: 5->6/7->8-> B->C->F->C->W->Y->Z->Hand

T2: 1->2/3->4-> A-> C-> F->C->W->Y->Z->Hand

Routes of Entities in Experiment 4

Following the same route setting logic, the routes of entities in Experiment 4 are:

T1: 5->6/7->8-> B->C->F->C->Y->Z-> Mouth

T2: 1->2/3->4-> A-> C-> F ->C->W->Y->Z->Hand

Modeling Mechanisms

The Queueing network model is able to generate the subadditive difficulty effect difficulty effect naturally without using scheduling or lock/unlock strategies. In short SOA conditions, entities in both compatible and incompatible conditions of Task 2 have to wait in front of Server F until entities of T1 complete the processing at Server F. Based on the function of Server C in inhibiting incompatible responses, it takes Server C one additional cycle to inhibit the incompatible responses. Server C takes less time to process the entities of compatible T2 than those of incompatible T2; as a result, entities of compatible T2 wait a longer time than entities of incompatible T2. However, since this extra waiting time of compatible T2 in front of Server F is absorbed in Server F's processing time in short SOA conditions, the difference between compatible and incompatible reaction time is smaller than that in long SOA conditions in which this extra waiting of compatible T2 is not absorbed (see Figure 3-6 and Figure 3-7). Therefore, it is predicted that the difference between compatible and incompatible reaction time in short SOA conditions is smaller than that in long SOA conditions, which is consistent with the subadditive difficulty effect. The same mechanism can also be applied to explain the experimental results of Experiment 4 in which different motor output modalities were used. Therefore, the subadditive difficulty effect can be modeled as a natural outcome of the interactions of the servers providing service to entities in the Queueing network, without using any lock/unlock strategies or executive control processes.

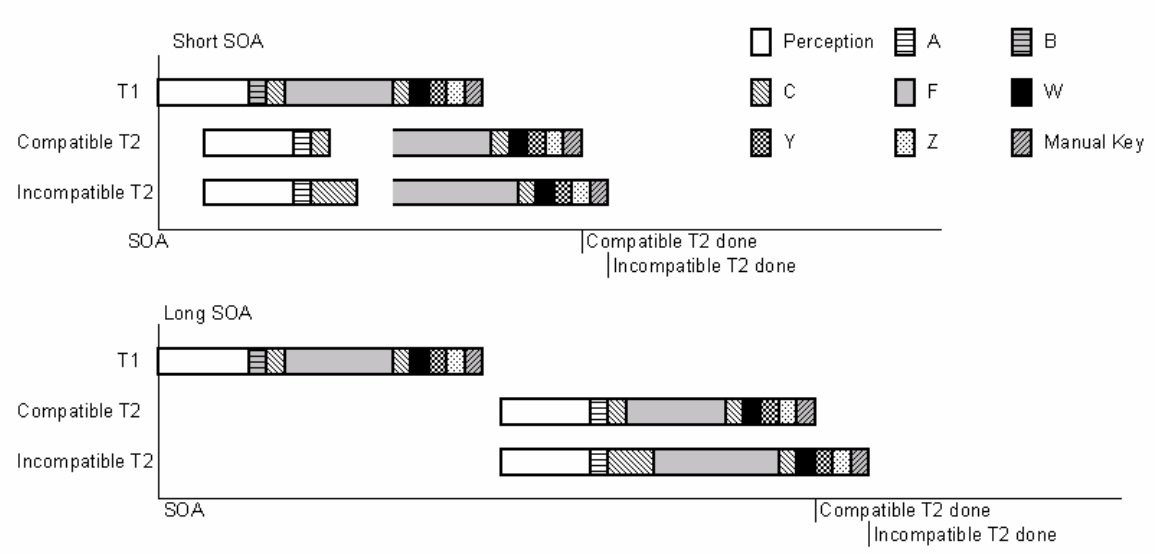


Figure 3-6 Modeling Mechanisms of Experiment 3 of Schumacher et al (1999)

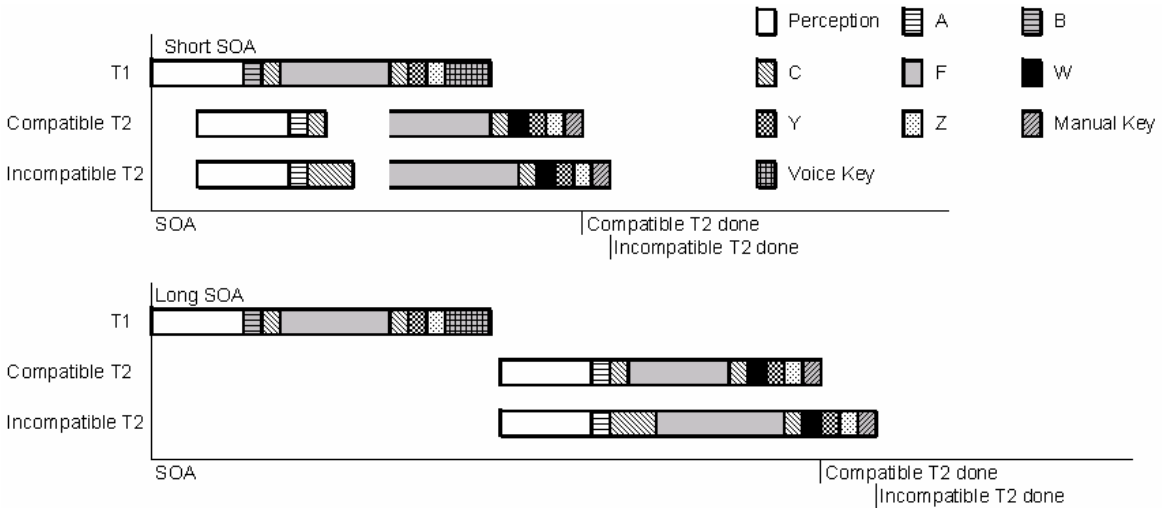


Figure 3-7 Modeling Mechanisms of Experiment 4 of Schumacher et al (1999)

Mathematical Modeling of the Expected Reaction Time

Similar to the equations in quantification of the reaction time in the basic PRP, the mathematical quantification of the expected RT1 and RT2 in Experiment 3 is:

$$E(RT1) = T_{1,AP} + T_{1,B} + T_{1,C} + T_{1,F} + T_{1,C} + T_{1,W} + T_{1,Y} + T_{1,Z} + T_{1,K} \quad (4)$$

$$E(RT2_{comp}) = max(T_{1,AP} + T_{1,B} + T_{1,C} + T_{1,F} - SOA,\ T_{2,VP} + T_{2,A} + T_{2,C\text{-}comp}) + T_{2,F\text{-}comp} + T_{2,C\text{-}comp} + T_{2,W} + T_{1,Y} + T_{2,Z} + T_{2,K} \quad (5)$$

$$E(RT2_{incomp}) = max(T_{1,AP} + T_{1,B} + T_{1,C} + T_{1,F} - SOA,\ T_{2,VP} + T_{2,A} + T_{2,C\text{-}incomp}) + T_{2,F\text{-}incomp} + T_{2,C\text{-}incomp} + T_{2,W} + T_{1,Y} + T_{2,Z} + T_{2,K} \quad (6)$$

The mathematical quantification of the expected RT1 and RT2 in Experiment 4 is:

$$E(RT1) = T_{1,AP} + T_{1,B} + T_{1,C} + T_{1,F} + T_{1,C} + T_{1,Y} + T_{1,Z} + T_{1,V} \quad (7)$$

$$E(RT2_{comp}) = max(T_{1,AP} + T_{1,B} + T_{1,C} + T_{1,F} - SOA,\ T_{2,VP} + T_{2,A} + T_{2,C\text{-}comp}) + T_{2,F\text{-}comp} + T_{2,C\text{-}comp} + T_{2,W} + T_{2,Y} + T_{2,Z} + T_{2,K} \quad (8)$$

$$E(RT2_{incomp}) = max(T_{1,AP} + T_{1,B} + T_{1,C} + T_{1,F} - SOA, T_{2,VP} + T_{2,A} + T_{2,C\text{-}incomp}) + T_{2,F\text{-}incomp} + T_{2,C\text{-}incomp} + T_{2,W} + T_{2,Y} + T_{2,Z} + T_{2,K} \quad (9)$$

where $T_{2,C\text{-}comp}$, $T_{2,C\text{-}incomp}$, $T_{2,F\text{-}comp}$, and $T_{2,F\text{-}incomp}$ are the processing time of Server C and F in the compatible and incompatible conditions of T2, respectively.

Based on the equations developed above, the expected pattern of reaction time in the compatible and incompatible conditions of T2 in Schumacher et al (1999) is shown in Figure 3-8. The difference of RT2 between the compatible and incompatible conditions is

larger in long SOA conditions ($2T_{2,C\text{-}incomp}$-$2T_{2,C\text{-}comp}$) than in short SOA conditions ($T_{2,C\text{-}incomp}$-$T_{2,C\text{-}comp}$) and the slope of T2 in short SOA conditions is -1.

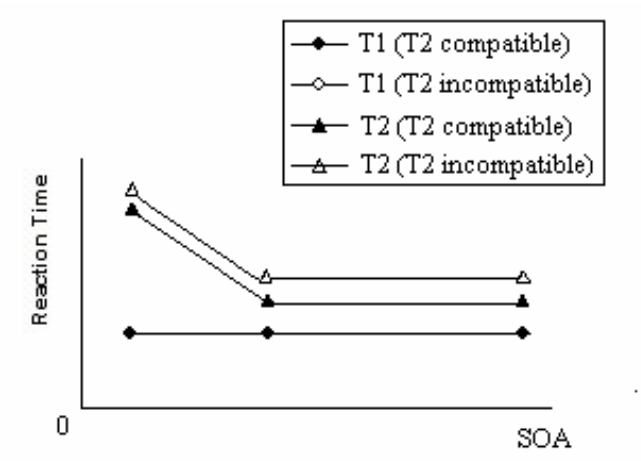


Figure 3-8 The expected pattern of reaction time in Schumacher et al.'s study (1999)

Parameter Setting

ACT-R/PM used 3 free parameters to model experimental results at long SOA conditions (tone recording time, activation of S-R mapping chunks in easy and hard T2 conditions in Experiments 3 and 4; see Table 1 in Bynre & Anderson, 2001). Using the same parameter setting method as in ACT-R/PM (see the parameter setting in modeling of the basic PRP), QN-MHP also used only 3 free parameters to model experimental results at long SOA conditions (see Table 3-3). Except for the 3 free parameters, the values of the other parameters are the same as those in Table 3-2. Moreover, the values of these free parameters at long SOA conditions are also constrained by the task properties: within each experiment, the processing time of entities at Server F in the compatible condition is shorter than that in the incompatible condition; across different experiments, in the two compatible conditions of T2 in Experiments 3 and 4, the processing times of entities at Server F are close to each other since the difficulty level of T1 in Experiment 3 is similar to that in Experiment 4 (range of difference≤120 ms); this constraint also applies to the processing time of T2 at Server F in the incompatible condition in Experiments 3 and 4.

Table 3-3 Parameter setting in modeling of Experiments 3 and 4 of Schumacher et al (1999)

Parameter	Mean Value	Description	Source
Parameters common in modeling Experiments 3 and 4 in Schumacher et al. (1999)			
$T_{1,AP}$	126 ms	Time for auditory perception time (42 ms at Servers 1,2/3, and 4)	Liu et al. (in press)
$T_{2,VP}$	126 ms	Time for visual perception time (42 ms at Servers 1,2/3, and 4)	Liu et al. (in press)
$T_{2,A,}$ $T_{2,B}$	18 ms	Processing time at Servers A and B	Liu et al. (in press)
$T_{1,C}$	18 ms	Processing time at Server C	Liu et al. (in press)
$T_{2,C\text{-}comp}$	18 ms	Processing time at Server C (compatible R-S)	Liu et al. (in press)
$T_{2,C\text{-}incomp}$	36 ms	Processing time at Server C (incompatible R-S)	Liu et al. (in press)
$T_{1,F}$	179 ms	Processing time at Server F (T1)	Value estimated at long SOA conditions
$T_{1,W,}$ $T_{2,W}$	24 ms	Processing time at Server W	Liu et al. (in press)
$T_{1,Y,}$ $T_{2,Y}$	24 ms	Processing time at Server Y	Liu et al. (in press)
$T_{1,Z,}$ $T_{2,Z}$	24 ms	Processing time at Server Z	Liu et al. (in press)
$T_{1,K}$	10 ms	Key closure time	Byrne & Anderson (2001)
$T_{1,V}$	100 ms	Voice key closure time	Byrne & Anderson (2001)
Parameters used in modeling Experiment 3 in Schumacher et al. (1999)			
$T_{2,F\text{-}comp}$	207 ms	Processing time at Server F (compatible R-S, T2)	Value estimated at long SOA conditions
$T_{2,F\text{-}incomp}$	295 ms	Processing time at Server F (incompatible R-S, T2)	Value estimated at long SOA conditions
Parameters used in modeling Experiment 4 in Schumacher et al. (1999)			
$T_{2,F\text{-}comp}$	165 ms	Processing time at Server F (compatible R-S, T2)	Value estimated at long SOA conditions
$T_{2,F\text{-}incomp}$	219 ms	Processing time at Server F (incompatible R-S, T2)	Value estimated at long SOA conditions

Modeling Results and their Validation

Figure 3-9 and Figure 3-10 show the modeling results compared with experimental results in Experiments 3 and 4: in Experiment 3, the R square of the model is .93 and the RMS=40.9 ms; in Experiment 4, the R square of the model is .95 and the RMS=18.7 ms.

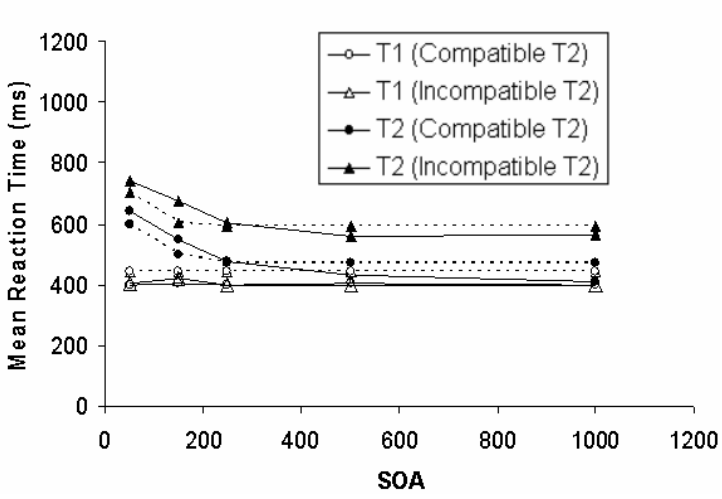


Figure 3-9 Mean reaction time of Experiment 3 in Schumacher et al. (1999) (solid lines) compared with modeling results (dashed lines)

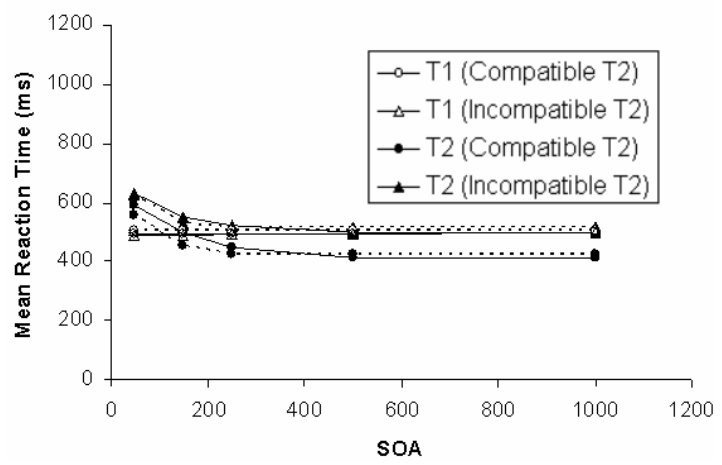


Figure 3-10 Mean reaction time of Experiment 4 in Schumacher et al. (1999) (solid lines) compared with modeling results (dashed lines)

4.2.2 Hawkins et al.'s Experiment (1979)

The difficulty level of T2 in Hawkins et al. (1979) was manipulated by the number of stimuli in a category making the same response: in the easy T2 condition, the stimuli were the digits 2 and 3 and the responses were keypresses with the right-hand index and middle fingers, respectively (one stimulus in the category corresponding to one responses); in the hard T2 condition, the stimuli were the digits 2-9: four of them (2, 5, 6, 9) belong to the first category, and the left four digits (3, 4, 7, 8) belong to the second

category. Subjects were asked to press the right-hand index or middle finger key once they saw one of these four digits in the first or the second category.

Routes of Entities

The 4 experiments in Hawkins et al.'s study (1979) included four different combinations of the perceptual and motor modalities of Task 1: auditory stimulus-manual response, auditory stimulus-vocal response, visual stimulus-vocal response, and visual stimulus-manual response. Following the same route selection method in the basic PRP, the routes of the entities in these four experiments are:

1) Auditory-manual response:

T1: 5->6/7->8-> B->C->F->C->W->Y->Z->Hand

2) Auditory-vocal response:

T1: 5->6/7->8-> B->C->F->C->Y->Z-> Mouth

3) Visual-vocal response:

T1: 1->2/3->4-> A ->C->F->C->Y->Z-> Mouth

4) Visual-manual response:

T1: 1->2/3->4-> A-> C-> F ->C->W->Y->Z->Hand

In all of the four experiments, T2 was a visual stimulus-manual response task. Hence, the route of entities of T2 in these 4 experiments is:

T2: 1->2/3->4-> A-> C-> F ->C->W->Y->Z->Hand

Modeling Mechanisms

Similar to the modeling mechanisms of Schumacher et al.'s experiments, based on the function of Server C in categorizing the stimuli, Server C takes additional cycles to locate the target stimulus into the right category (in the easy condition of T2 in Hawkins et al.'s experiment, 1 cycle time is needed since the category size is 1; in the difficult condition of T2, 4 cycle times are needed because the category size is 4); hence, Server C takes a longer time to process the entities of T2 that require this categorization process than entities of T1 without the categorization process. However, since this extra waiting time of compatible T2 in front of Server F is absorbed in Server F's processing time in short

SOA conditions, the difference between easy and hard reaction time is smaller than that in long SOA conditions.

Mathematical Modeling of the Expected Reaction Time

Similar to the mathematical modeling in the basic PRP, the equations quantifying the expected reaction time in the four experiments in Hawkins et al.'s experiment are developed as follows:

1) Auditory-manual response:

$$E(RT1) = T_{1,AP} + T_{1,B} + T_{1,C} + T_{1,F} + T_{1,C} + T_{1,W} + T_{1,Y} + T_{1,Z} + T_{1,K} \quad (10)$$

2) Auditory-vocal response:

$$E(RT1) = T_{1,AP} + T_{1,B} + T_{1,C} + T_{1,F} + T_{1,C} + T_{1,Y} + T_{1,Z} + T_{1,V} \quad (11)$$

3) Visual-vocal response:

$$E(RT1) = T_{1,VP} + T_{1,A} + T_{1,C} + T_{1,F} + T_{1,C} + T_{1,Y} + T_{1,Z} + T_{1,V} \quad (12)$$

4) Visual-manual response:

$$E(RT1) = T_{1,VP} + T_{1,A} + T_{1,C} + T_{1,F} + T_{1,C} + T_{1,W} + T_{1,Y} + T_{1,Z} + T_{1,K} \quad (13)$$

In all of the four experiments, T2 is a visual stimulus-manual response task; hence, the expected RT2 in these 4 experiments is:

$$E(RT2_{easy}) = max(T_{1,P} + T_{1,A|B} + T_{1,C} + T_{1,F} - SOA,\ T_{2,VP} + T_{2,A} + T_{2,C\text{-}easy}) + T_{2,F\text{-}easy} + T_{2,C\text{-}easy} + T_{2,W} + T_{2,Y} + T_{2,Z} + T_{2,K} \quad (14)$$

$$E(RT2_{diff}) = max(T_{1,P} + T_{1,A|B} + T_{1,C} + T_{1,F} - SOA,\ T_{2,VP} + T_{2,A} + T_{2,C\text{-}diff}) + T_{2,F\text{-}diff} + T_{2,C\text{-}diff} + T_{2,W} + T_{2,Y} + T_{2,Z} + T_{2,K} \quad (15)$$

where $T_{1,P}$ is the processing time of the perceptual subnetwork corresponding to the parameter in RT1's equations; $T_{1,A|B}$ is the processing time of Server A or B corresponding to the parameter in RT1's equations; $T_{2,C\text{-}easy}$, $T_{2,C\text{-}diff}$, $T_{2,F\text{-}easy}$, and $T_{2,F\text{-}diff}$ are the processing time of Server C and F in the easy and difficult T2 condition, respectively.

Since the modeling mechanisms of reaction time and the equations of RT1 and RT2 in Hawkins et al.'s experimental study are similar to that of Schumacher et al.'s experimental study, the expected pattern of reaction time in Hawkins et al.'s experiment

is similar to the Schumacher et al.'s experimental study except that the difficulty level of T2 includes easy and hard conditions rather than compatible and incompatible conditions.

Parameter Setting

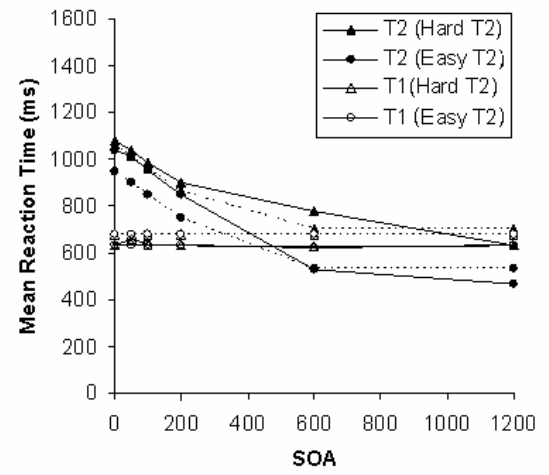
In modeling the 4 experiments of Hawkins et al. (1979), compared to EPIC which used 5 free parameters (unlocking onset latency, suspension waiting times under the easy and difficult conditions, preparation waiting time, and ocular orientation time, Meyer and Kieras, 1997a, p45), QN-MHP used only 3 free parameters ($T_{1,F}$, $T_{2,F\text{-}easy}$, and $T_{2,F\text{-}diff}$) with the same parameter setting method as in ACT-R/PM (see Table 3-4). Except for the 3 free parameters, the values of the other parameters are the same as those in Table 3-2 and Table 3-3. Moreover, the values of these free parameters set at long SOA conditions are also constrained by the task properties: within each experiment, processing time of entities at Server F in the easy condition is shorter than that in the hard condition; across different experiments, in the four easy conditions of T2 in Experiments 1-4, the processing times of T2 at Server F in Experiments 1-4 are close to each other (range of difference≤120 ms) since the difficulty levels of T2 are similar in these four experiments; this constraint also applies to the processing time of Server F in the four hard conditions of T2 and T1 in Experiments 1-4.

Table 3-4 Parameter setting in modeling of Experiments 1-4 of Hawkins et al. (1979)

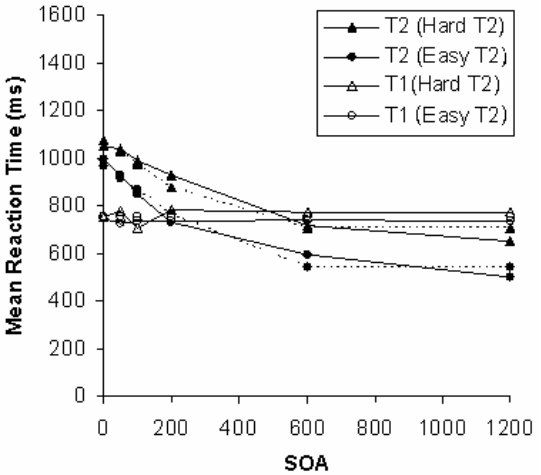
Parameter	Mean Value	Description	Source
Parameters common in modeling Experiments 1-4 of Hawkins et al's Study (1979)			
$T_{1,AP}$	126 ms	Time for auditory perception time (42 ms at Servers 1,2/3, and 4)	Liu et al. (in press)
$T_{2,VP}$	126 ms	Time for visual perception time (42 ms at Servers1,2/3, and 4)	Liu et al. (in press)
$T_{2,A,}$ $T_{2,B}$	18 ms	Processing time at Servers A and B	Liu et al. (in press)
$T_{1,C}$	18 ms	Processing time at Server C (T1)	Liu et al. (in press)
$T_{2,C\text{-}easy}$	18 ms	Processing time at Server C (T2, easy condition)	Liu et al. (in press)
$T_{2,C\text{-}diff}$	72 ms	Processing time at Server C (T2, difficult condition)	Liu et al. (in press)
$T_{1,W,}$ $T_{2,W}$	24 ms	Processing time at Server W	Liu et al. (in press)
$T_{1,Y,}$ $T_{2,Y}$	24 ms	Processing time at Server Y	Liu et al. (in press)
$T_{1,Z,}$ $T_{2,Z}$	24 ms	Processing time at Server Z	Liu et al. (in press)
$T_{1,K}$	10 ms	Key closure time	Byrne & Anderson (2001)
$T_{1,V}$	100 ms	Voice key closure time	Byrne & Anderson (2001)
Parameters used in modeling Experiment 1 in Hawkins et al. (1979) (auditory-manual condition)			
$T_{1,F}$	415 ms	Processing time at Server F (T1)	Value estimated at long SOA conditions
$T_{2,F\text{-}easy}$	270 ms	Processing time at Server F (T2, easy condition)	Value estimated at long SOA conditions
$T_{2,F\text{-}diff}$	330 ms	Processing time at Server F (T2, difficult condition)	Value estimated at long SOA conditions
Parameters used in modeling Experiment 2 in Hawkins et al. (1979) (auditory-vocal condition)			
$T_{1,F}$	423 ms	Processing time at Server F (T1)	Value estimated at long SOA conditions
$T_{2,F\text{-}easy}$	280 ms	Processing time at Server F (T2, easy condition)	Value estimated at long SOA conditions
$T_{2,F\text{-}diff}$	336 ms	Processing time at Server F (T2, difficult condition)	Value estimated at long SOA conditions
Parameters used in modeling Experiment 3 in Hawkins et al. (1979) (visual-vocal condition)			
$T_{1,F}$	346 ms	Processing time at Server F (T1)	Value estimated at long SOA conditions
$T_{2,F\text{-}easy}$	260 ms	Processing time at Server F (T2, easy condition)	Value estimated at long SOA conditions
$T_{2,F\text{-}diff}$	340 ms	Processing time at Server F (T2, difficult condition)	Value estimated at long SOA conditions
Parameters used in modeling Experiment 4 in Hawkins et al. (1979) (visual-manual condition)			
$T_{1,F}$	308 ms	Processing time at Server F (T1)	Value estimated at long SOA conditions
$T_{2,F\text{-}easy}$	224 ms	Processing time at Server F (T2, easy condition)	Value estimated at long SOA conditions
$T_{2,F\text{-}diff}$	280 ms	Processing time at Server F (T2, difficult condition)	Value estimated at long SOA conditions

Modeling Results and their Validation

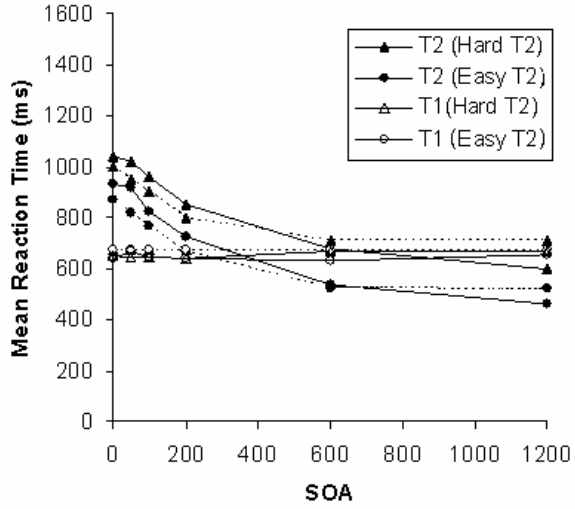
Figure 3-11 shows the modeling results compared with experimental results in Experiments 1 to 4: in Experiment 1, the R square of the model is .96 and the RMS=58.6 ms; in Experiment 2, the R square of the model is .96 and the RMS=49.1 ms; in Experiment 3, the R square of the model is .98 and the RMS=28.7 ms; and in Experiment 4, the R square of the model is .98 and the RMS=36.3 ms.



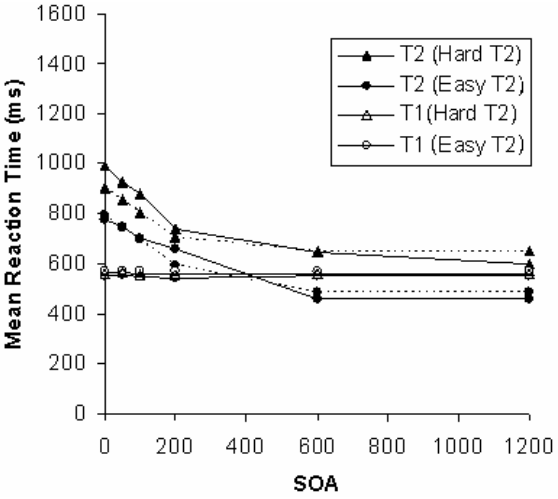
Experiment 1 (auditory-manual condition)



Experiment 2 (auditory-vocal condition)



Experiment 3 (visual-vocal condition)



Experiment 4 (visual-manual condition)

Figure 3-11 Mean reaction time in Experiments 1-4 of Hawkins et al. (1979) (solid lines) compared with modeling results (dashed lines)

4.2.3 Karlin & Kestenbaum's Experiment (1968)

Karlin & Kestenbaum (1968) manipulated the difficulty levels of T2 by the number of S-R pairs. T1 was a visual-manual choice reaction task and T2 was an auditory-manual reaction task including two difficulty levels: one was a simple reaction task and the other was a two choice reaction task (see the detailed experiment description in the introduction part of this article).

Routes of Entities

Similar to the routes in the modeling mechanisms of the basic PRP, based on the task and corresponding functions of the brain areas, the routes of T1 and T2 (choice reaction condition) are:

T1: 1->2/3->4->A->C->F->C->W->Y->Z-> Hand

T2 (choice reaction): 5->6/7->8-> B->C->F->C->W->Y->Z->Hand

Kansaku et al. (2004) found a specific neural network that characterizes simple reaction tasks irrespective of the input modalities and output effectors. This network includes the premotor cortex (Server V) and the right posterior superior temporal cortex (Server F). This is also consistent with the functions of the premotor cortex (included in Server V) and right posterior superior temporal cortex (included in Server F) since a simple reaction task involves sensorimotor processing/integration (Server V), sensory cue detection (Server V), working memory (phonological) information processing and judgment (Server F). Based on the connection between these brain regions, the route of entities in Task 2 under the simple reaction task condition is:

T2 (simple reaction): 5->6/7->8->B->C->F->C->V->Z->Hand

Modeling of the Expected Reaction Time

The expected reaction time in T1 and T2 (choice reaction condition) can be derived based on the same mechanism as in the basic PRP (see Equations 16 and 17).

$$E(RT1) = T_{1,VP} + T_{1,A} + T_{1,C} + T_{1,F} + T_{1,C} + T_{1,Y} + T_{1,W} + T_{1,Z} + T_{1,K} \quad (16)$$

$$E(RT2)\ (choice\ reaction) = max(T_{1,VP} + T_{1,A} + T_{1,C} + T_{1,F} - SOA,\ T_{2,AP} + T_{2,B} + T_{2,C}) + T_{2,F} + T_{2,C} + T_{2,Y} + T_{2,W} + T_{2,Z} + T_{2,K} \quad (17)$$

The expected RT2 in the simple reaction condition is modeled in Appendix 1 (see Equations 67 in Appendix 1).

$$E(RT2)_{(simple\ reaction)} = \begin{cases} pRT_{2,ANTI} + (1-p)RT_{2,NOAN} & t_a>0 \text{ (long SOA conditions)} \\ \max(T_{1,VP}+T_{1,A}+T_{1,C}+T_{1,F}-SOA,\ T_{2,AP}+T_{1,B}+T_{2,C}) & \\ +T_{2,F}+T_{2,C}+T_{2,V}+T_{2,Z}+T_{2,K} & t_a=0 \text{ (short SOA conditions)} \end{cases} \quad (18)$$

Based on the equations developed above, the expected pattern of the reaction time in Karlin & Kestenbaum's experiment is shown in Figure 3-12. Based on Equation 18, in short SOA conditions, RT2 (simple reaction) shows the same pattern as RT2 in the choice reaction condition; meanwhile, in long SOA conditions, according to Equation 68 in Appendix 1, there is a non-linear negative relationship between RT2 (simple reaction) and SOA.

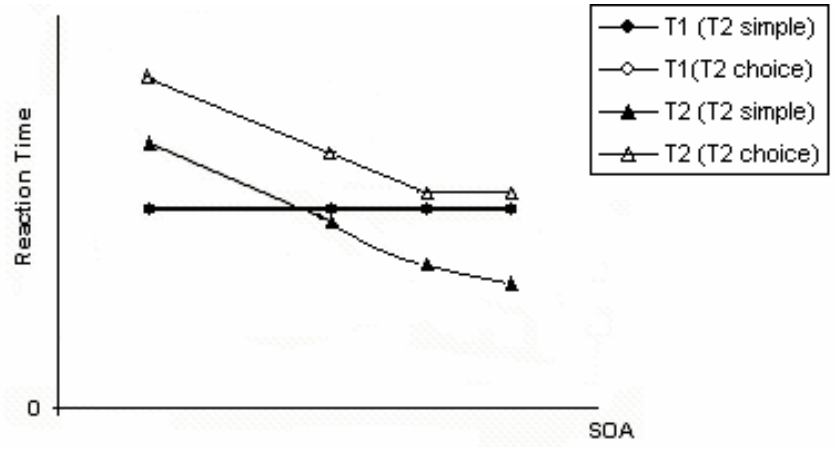


Figure 3-12 The expected pattern of reaction time in simple and choice reaction conditions in Karlin & Kestenbaum's experiment (1968)

This mechanism of the subadditive difficulty effect above is consistent with the behavioral and physiological studies of the same task. In the studies of Karlin and Kestenbaum (1968) and De Jong (1993), the stimulus of the second task was always presented after the first stimulus, increasing the conditional probability of the

presentation of the second stimulus with an increase of SOA (Luce, 1986; Naatanen, 1971; Niemi & Naatanen, 1981). Other behavioral studies (Nickerson, 1965, 1967) also found an increased probability of anticipatory reactions in a simple reaction task at long SOA conditions under such experimental conditions. The ERP study also confirmed the result of these behavioral studies (Sommer et al., 2001): lateralized readiness potential (LRP) in that study provided direct evidence that the subadditive difficulty effect is due to an increase of response anticipation in the simple response condition if the secondary task is a simple reaction task. In addition, fMRI studies (Brass & Cramon, 2002) also found the fronto-lateral frontal cortex at the junction of the precentral sulcus and the inferior frontal sulcus (represented by Server F) is the crucial frontal component in task preparation.

QN-MHP is able to model the subadditivity difficulty effect found in Karlin & Kestenbaum's experiment without using task-specific scheduling assumptions which are required by EPIC-SRD. When SOA is longer than certain duration, before S1 appears, Server F starts the anticipation process and prepares its response selection (see Appendix 1 for the modeling mechanism in detail). The response selection might occur earlier than the actual presentation of stimulus, which also allows QN-MHP to model the other dependent variables in Sommer et al.'s experiment (2001)—percentage of negative responses and onset time of the brain waves of LRP (see the modeling mechanism of Sommer et al.'s experiment and Appendix 2 in detail). Moreover, this modeling mechanism in stimulus anticipation has been successfully implemented in modeling a driving task in QN-MHP (Liu, et al., in press) and it is not an additional task-specific assumption added in the current study.

Parameter Setting

All of the parameters used in the model are listed in Table 3-5 below. Compared with EPIC which used 9 free parameters [4] in simulating the experiment of Karlin & Kestenbaum (1968) (see Table 1 in Meyer& Kieras, 1997b), QN-MHP used only 5 free

[4] In modeling Karlin & Kestenbaum 's experiment (compared with the parameters used by EPIC in modeling Hawkins et al.'s experiment) , the 9 free parameters used by EPIC were: auditory identification time, auditory detection time, visual identification time, number of selection cycles of T1 and T2, ocular orientation time, unlocking onset latency, suspension waiting time, and preparation waiting time.

parameters with the same parameter setting method as in ACT-R/PM. Except for the 5 free parameters, the values of the other parameters are the same as those in Table 3-2 to Table 3-4. The two parameters in Equation 56 in Appendix 1 are directly based on a psychophysical study (k=2, β=0.9, Wearden, Edwards, Fakhri, & Percival, 1998a), and they are not free parameters in this modeling process. Moreover, the values of these free parameters set at long SOA conditions are also constrained by the task properties: the processing time at Server F in the simple reaction condition is at least 18 ms (1 cycle time) less than that in the choice reaction condition.

Table 3-5 Parameter setting in modeling of Karlin & Kestenbaum's experiment (1968)

Parameter	Value	Description	Source
$T_{1,P}$	126 ms	Time for tone detection perception	Liu et al. (in press)
$T_{2,P}$	126 ms	Time for letter perception	Liu et al. (in press)
$T_{1,C,}$ $T_{2,C}$	18 ms	Processing time at Server C (compatible R-S)	Liu et al. (in press)
$T_{1,F}$	117 ms	Processing time at Server F (T1)	Value estimated at long SOA conditions
$T_{2,Fc}$	68 ms	Processing time at Server F (choice RT2 condition)	Value estimated at long SOA conditions
$T_{2,Fs}$	50 ms	Processing time at Server F (simple RT2 condition)	Value estimated at long SOA conditions
$T_{2,VS}$	140 ms	Processing time at Server V (T1)	Value estimated at long SOA conditions
$T_{1,W,}$ $T_{2,W}$	24 ms	Processing time at Server W	Liu et al. (in press)
$T_{1,Y,}$ $T_{2,Y}$	24 ms	Processing time at Server Y	Liu et al. (in press)
$T_{1,Z,}$ $T_{2,Z}$	24 ms	Processing time at Server Z	Liu et al. (in press)
$T_{1,X,}$ $T_{2,X}$	24 ms	Processing time at Server X	Liu et al. (in press)
T_k	10 ms	Key closure time	Byrne & Anderson (2001)
T_{perc}	350 ms	The duration between when the anticipation process starts and when the probability that subjects make the motor response equal to 1	Value estimated at long SOA conditions

Modeling Results and their Validation

Figure 3-13 shows the modeling results compared with experimental results. The R square of the model is .73 and the RMS=41.8 ms.

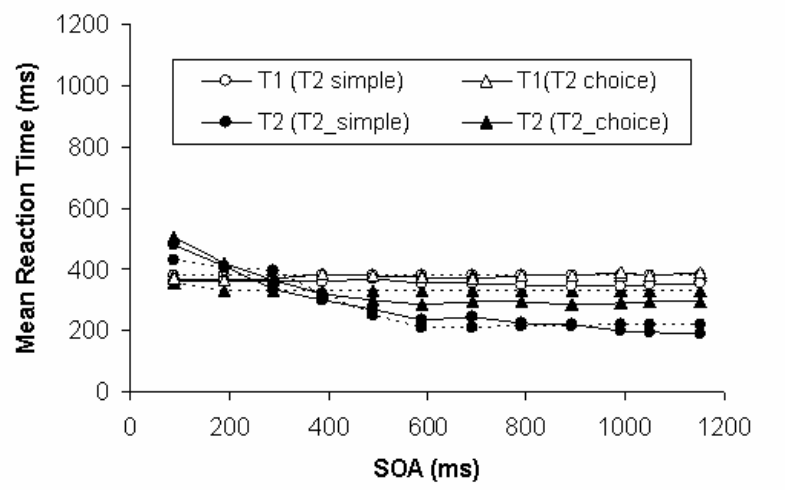


Figure 3-13 Mean reaction time in experimental results of Karlin & Kestenbaum (1968) (solid lines) compared with modeling results (dashed lines)

4.2.4 Sommer et al.'s Experiment (2001)

Sommer et al. (2001) replicated Karlin & Kestenbaum's experiment (1968) with ERP techniques, and they measured reaction time, percentage of negative RT2 and lateralized related potential (LRP). Their experiment setting (see detailed description of their experiment in the introduction section of this article) was the same as that of Karlin & Kestenbaum's experiment (1968) except that T1 in Sommer et al.'s experiment was an auditory-manual task and T2 was a visual-manual task. Therefore, the same modeling mechanisms are used in this article to model Sommer et al.'s experiment.

Routes of Entities

Similar to the routes of entities in Karlin & Kestenbaum' experiment, according to the functions and connections of these brain regions, the routes of tasks in modeling Sommer et al.'s experiment are:

T1: 5 ->6/7->8-> B-> C->F->C->W->Y->Z-> Hand

T2 (choice reaction condition): 1 ->2/3->4-> A-> C->F->C->W->Y->Z-> Hand

T2 (simple reaction condition): 1 ->2/3->4->A->C->F->C->V->Z-> Hand

Modeling of the Expected Reaction Time, Percentage of Negative Responses and S-LRP

1) Reaction Time

The expected reaction times in Sommer et al.'s experiment (2001) are modeled with the same formula as the modeling of Karlin & Kestenbaum's experiment (1968) except for the processing time of Server A, Server B and the perceptual subnetwork (see Equations 19-21).

$$E(RT1) = T_{1,AP} + T_{1,B} + T_{1,C} + T_{1,F} + T_{1,C} + T_{1,Y} + T_{1,W} + T_{1,Z} + T_{1,K} \quad (19)$$

$$E(RT2)\ (choice\ reaction) = max(T_{1,AP}+T_{1,B}+T_{1,C}+T_{1,F}-SOA,\ T_{2,VP}+T_{2,A}+T_{2,C}) + T_{2,F} + T_{2,C} + T_{2,Y} + T_{2,W} + T_{2,Z} + T_{2,K} \quad (20)$$

$$E(RT2)\ (simple\ reaction) = \begin{cases} p\,RT_{2,ANTI} + (1-p)\,RT_{2,NOAN} & t_a > 0 \text{ (long SOA conditions)} \\ max(T_{1,AP}+T_{1,B}+T_{1,C}+T_{1,F}-SOA,\ T_{2,VP}+T_{1,A}+T_{2,C}) & \\ \quad +T_{2,F}+T_{2,C}+T_{2,V}+T_{2,Z}+T_{2,K} & t_a = 0 \text{ (short SOA conditions)} \end{cases} \quad (21)$$

2) Percentage of Negative Responses

In the simple reaction condition of T2, the expected percentage of negative responses (P_n) is estimated according to the difference between SOA and the sum of $T_{2,C}$, $T_{2,V}$, $T_{2,Z}$, $T_{2,K}$, and T_{Fst} (see Equations 22 and 71 including their derivation in Appendix 2).

$$P_n = \begin{cases} 1 - \dfrac{T_{2,C}+T_{2,V}+T_{2,Z}+T_{2,K}}{SOA - T_{Fst}} & SOA \geq u + T_{Fst} \\ 0 & SOA < u + T_{Fst} \end{cases} \quad (22)$$

The expected pattern of the percentage of negative responses is shown in Figure 3-14. In the simple reaction condition, at short SOA conditions ($SOA < T_{2,C} + T_{2,V} + T_{2,Z} + T_{2,K} + T_{Fst}$), the expected percentage of negative responses is

0; at long SOA conditions, the expected percentage of negative responses increases as SOA increases following the inverse function in Equation 22. In T1 (choice RT) and the choice reaction condition of T2, the expected percentage of negative responses is 0.

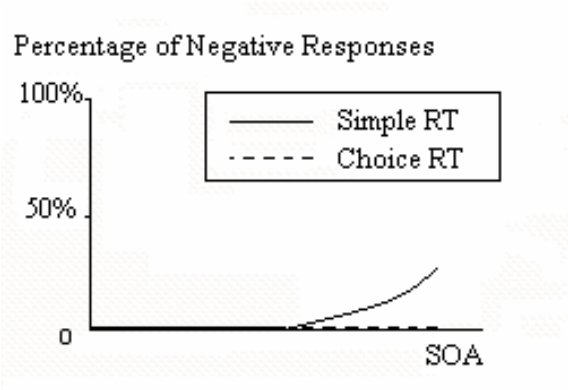


Figure 3-14 The expected pattern of the percentage of negative response in Sommer et al.'s experiment (2001)

3) Stimulus-LRP Onset Time

Mathematical Modeling of S-LRP Onset Time

1) Simple Reaction Time

Since LRP reflects motor preparation taking place within the premotor area (Server V) or the primary motor cortex (Server Z) (H. Leuthold & Jentzsch, 2001; H. Leuthold & Jentzsch, 2002; Ulrich et al., 1998), and Server V is located before Server Z in the route of the simple reaction task, the arrival time of entities into Server V is regarded as the LRP onset time in this simple reaction situation[5]. Based on Figure 3-30 in Appendix 1, the time that entities enter Server V (V_{st}) can be estimated in two conditions depending on the value of t_a (t_a= 0, short SOA conditions; t_a >0, long SOA conditions; see Equation 23):

$$V_{st}= \begin{cases} Max(T_{1,AP} + T_{1,B} + T_{1,C,} + T_{1,F},\ SOA + T_{2,VP} + T_{2,A} + T_{2,C}) + T_{2,F} + T_{2,C} & t_a=0 \\ T_{Fst} + t_a + T_{2,C} & t_a>0 \end{cases} \quad (23)$$

Since V_{st} starts from the arrival of S1 and S-LRP onset time (*S-LRP*) starts from the arrival of S2 (Sommer et al., 2001), *S-LRP* equals V_{st} - SOA:

[5] Because of the high temporary accuracy of ERP techniques (1 ms), the conduction time from the activation of brain regions to the time when the stimulus-LRP is observed can be ignored (Ilvonen et al., 2003).

$$S\text{-}LRP = \begin{cases} Max(T_{1,AP}+T_{1,B}+T_{1,C,}+T_{1,F},\ SOA+T_{2,VP}+T_{2,A}+T_{2,C}) & \\ +T_{2,F}+T_{2,C}\text{-}SOA & t_a=0 \\ & \\ T_{Fst}+t_a+T_{2,C}\text{-}SOA & t_a>0 \end{cases} \quad (24)$$

$$S\text{-}LRP = \begin{cases} T_{1,AP}+T_{1,B}+T_{1,C,}+T_{1,F}+T_{2,F}+T_{2,C}\text{-}SOA & \\ & SOA< T_{1,AP}+T_{1,B}+T_{1,C,}+T_{1,F}-(T_{2,VP}+T_{2,A}+T_{2,C}) \\ T_{2,VP}+T_{2,A}+2T_{2,C}+T_{2,F} & \\ & T_{1,AP}+T_{1,B}+T_{1,C,}+T_{1,F}-(T_{2,VP}+T_{2,A}+T_{2,C}) \le SOA \le T_{Fst} \\ T_{Fst}+(SOA^{\beta}-T_{Fst}{}^{\beta})^{1/\beta}+T_{2,C}\text{-}SOA & \\ & SOA>T_{Fst} \end{cases} \quad (25)$$

Since $t_a = \max[(SOA^{\beta}-T_{Fst}{}^{\beta})^{1/\beta}, 0]$, Equation 24 can be rewritten into Equation 25, which is shown in Figure 3-15 as the expected pattern of *S-LRP* with an increase of SOA.

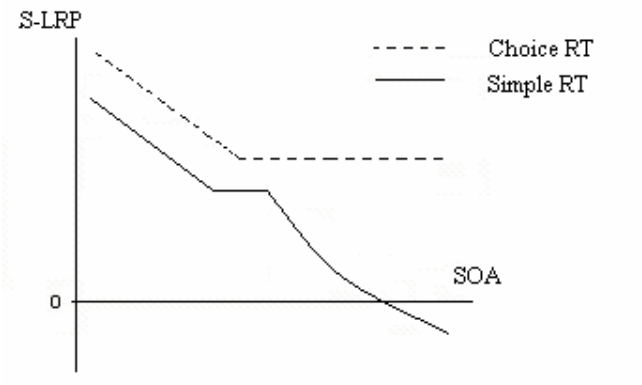


Figure 3-15 The expected pattern of S-LRP with an increase of SOA in T2

2) Choice Reaction Time

LRP reflects motor preparation taking place within the premotor area (Server V) or the primary motor cortex (Server Z). Hence, in the choice reaction condition of T2, the premotor area (Server V) is not in the route of entities; therefore, the arrival time of entities at Server Z is the expected S-LRP onset time (*S-LRP*) (see Equation 26 and Figure 3-15). The expected pattern of S-LRP is similar to the expected pattern of reaction time in this condition (see Figure 3-4).

$$S\text{-}LRP = max(T_{1,AP}+T_{1,B}+T_{1,C}+T_{1,F}-SOA,\ T_{2,VP}+T_{2,A}+T_{2,C}) + T_{2,F}+ T_{2,C}+ T_{2,Y}+ T_{2,W} \quad (26)$$

Parameter Setting

All of the parameters used in the model are listed in Table 3-6. Except for the 4 free parameters set at long SOA conditions following ACT-R/PM's parameter setting method, all of the values of the other parameters are the same as those in parameter setting tables in previous sections. The three dependent variables, including reaction time, ratio of correct response and stimulus-LRP onset time, are modeled based on the same set of parameters as in Table 3-2 to Table 3-5. Moreover, the values of these free parameters set at long SOA conditions are also constrained by the task properties: the processing time at Server F in the simple reaction condition is at least 18 ms (1 cycle time) less than that in the choice reaction condition.

Table 3-6 Parameter setting in modeling of the study of Sommer et al. (2001)

Parameter	Value	Description	Source
$T_{1,AP}$	126 ms	Time for auditory perception time (42 ms at Server 5,6/7, and 8)	Liu, et al. (in press)
$T_{2,VP}$	126 ms	Time for visual perception time (42 ms at Server 1,2/3, and 4)	Liu, et al. (in press)
$T_{2,A,}$ $T_{2,B}$	18 ms	Processing time at Server A	Liu, et al. (in press)
$T_{1,C,}$ $T_{2,C}$	18 ms	Processing time at Server C (compatible R-S)	Liu, et al. (in press)
$T_{1,F}$	338 ms	Processing time at Server F (T1)	Value estimated at long SOA conditions
$T_{2,F}$ *(choice RT)*	324 ms	Processing time at Server F (choice RT2 condition)	Value estimated at long SOA conditions
$T_{2,F}$ *(simple RT)*	293 ms	Processing time at Server F (simple RT2 condition)	Value estimated at long SOA conditions
T_{perc}	570 ms	The duration between when the anticipation process starts and when the probability that subjects make the motor response equal to 1	Value estimated at long SOA conditions
$T_{2,V}$	24 ms	Processing time at Server V (T1)	Liu, et al. (in press)
$T_{1,W,}$ $T_{2,W}$	24 ms	Processing time at Server W	Liu, et al. (in press)
$T_{1,Y,}$ $T_{2,Y}$	24 ms	Processing time at Server Y	Liu, et al. (in press)
$T_{1,Z,}$ $T_{2,Z}$	24 ms	Processing time at Server Z	Liu, et al. (in press)
$T_{1,X,}$ $T_{2,X}$	24 ms	Processing time at Server X	Liu, et al. (in press)

Modeling Results and their Validation

Figure 3-16 shows the modeling results in comparison with experimental results. The R square of the model is .84 and the RMS=53.9 ms.

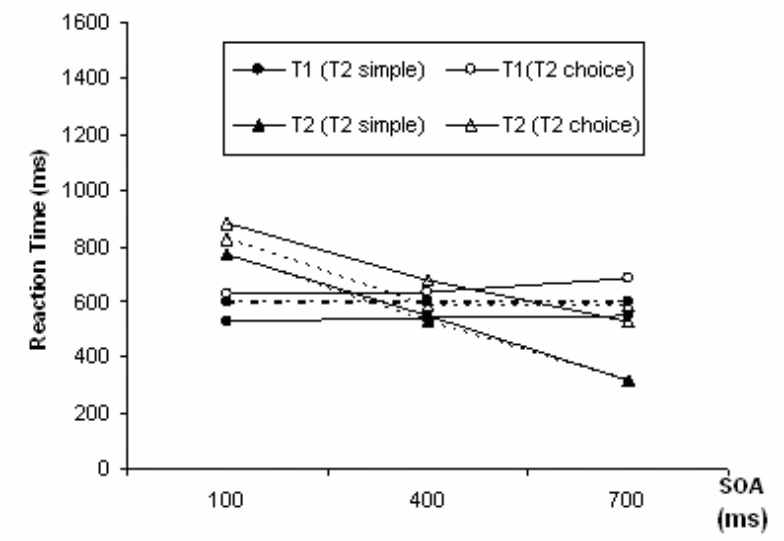


Figure 3-16 The reaction time in the study of Sommer et al. (2001) (solid lines) in comparison with the Queueing network modeling results (dashed lines)

The modeling results of the percentage of negative responses in comparison with the experiment result are shown in Figure 3-17. The R square of the model is 0.99 with RMS=.037. Moreover, it is found that at SOA=700 ms, the percentage of negative RT2 is 15%, which is consistent with the Sommer et al.'s experimental results (16%).

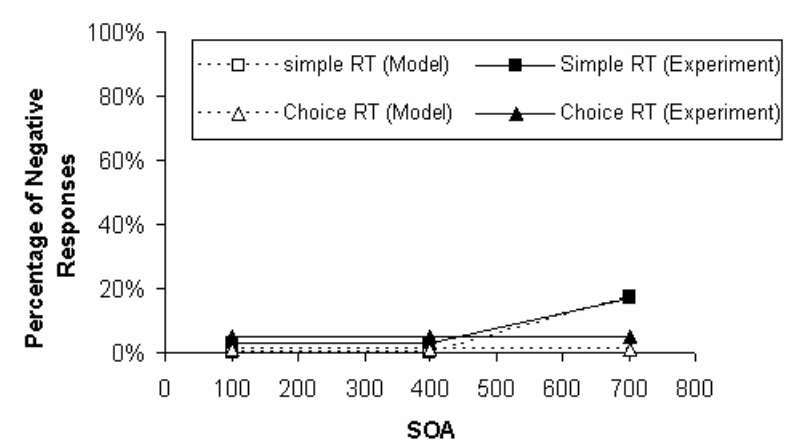


Figure 3-17 The percentage of negative responses in the study of Sommer et al. (2001) (solid lines) compared with the modeling results (dashed lines)

In addition, the modeling result of S-LRP exhibits a pattern similar to the experimental results (see Figure 3-18 for the comparison of the S-LRP onset time between the prediction of the model and the experimental results, R square=.96; RMS=127.5 ms).

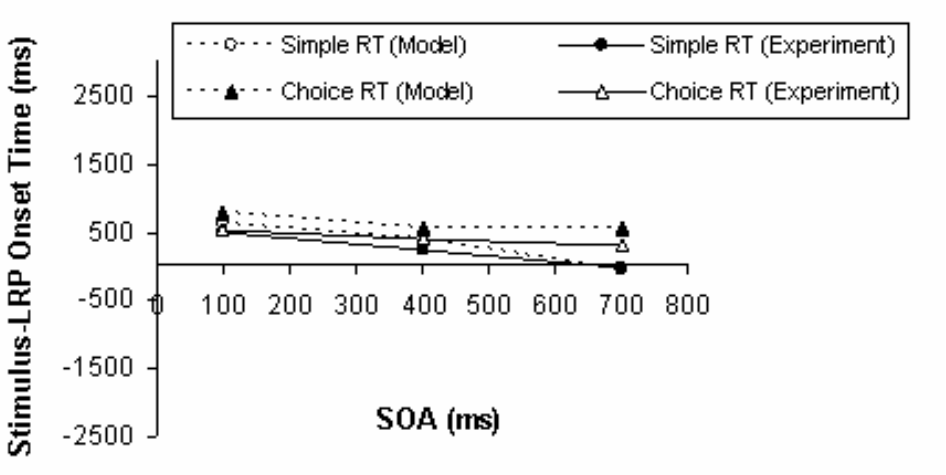


Figure 3-18 The S-LRP onset time in the study of Sommer et al. (2001) (solid lines) compared with the Queueing network modeling results (dashed lines)

4.3 The fMRI Study of PRP (Jiang, et al., 2004)

Jiang et al.'s experiment (2004) provides another challenging counterexample of the strategic scheduling mechanism in EPIC. In their experiment, both Task 1 and 2 are visual-manual choice reaction task (see detailed experiment description in the introduction of this article).

Routes of Entities

Similar to the routes in the modeling mechanisms of the basic PRP, based on the task and corresponding function of the brain areas, the routes of Task 1 and Task 2 are:

T1: 1->2/3->4-> A->C->F->C->W->Y->Z->Hand

T2: 1->2/3->4-> A-> C->F->C-> W->Y ->Z->Hand

Modeling of the Expected Reaction Time

Experiment 1 in Jiang et al.'s study (2004) can be modeled by the Queueing mechanism in the basic PRP directly: at short SOA conditions, entities of T2 will enter

Server F until entities of T1 leave Server F. In this process, Server F only processes the entities of individual tasks rather than scheduling the entities of two tasks actively. The mathematical equations of behavioral performance of this experiment are the same as those formulas used in modeling the basic PRP effect except for the change of the perceptual and motor subnetwork according to Jiang et al.'s experiment setting (see Equations 27 and 28). The expected pattern of the reaction time is the same as that in the basic PRP effect.

$$E(RT1)= T_{1,VP} + T_{1,A}+ T_{1,C} +T_{1,F}+ T_{1,C}+ T_{1,Y}+ T_{1,W}+ T_{1,Z}+ T_{1,K} \quad (27)$$

$$E(RT2)=max(T_{1,VP}+T_{1,A}+T_{1,C} +T_{1,F} –SOA,\ T_{2,VP}+T_{2,A}+T_{2,C}) + T_{2,F}+ T_{2,C}+ T_{2,Y}+ T_{2,W}+ T_{2,Z}+ T_{2,K} \quad (28)$$

Modeling of BOLD Signal and its Percentage of Change

The integrated BOLD signal (*CB(t)*) in the Queueing network model is modeled based on the prior fMRI signal modeling work of Cohen (1997) and Anderson et al. (2003) (see Equation 29 and Appendix 4 for its development).

$$CB(t)= \begin{cases} skM \int_{\frac{t-\eta}{s}}^{\frac{t}{s}} Y^a e^{-Y/b} dY & \frac{t-\eta}{s} \le Y \le \frac{t}{s} \\ 0 & \frac{t-\eta}{s} > Y \text{ or } Y > \frac{t}{s} \end{cases} \quad (29)$$

where *s, k, M, a,* and *b* came from the equations of Cohen (1997) and Anderson et al. (2003) determined by the properties of brain regions with certain fMRI measurement techniques; *t* is the duration of each trial; and η in Queueing networks can be quantified by Equation 30 (Gross and Harris, 1998):

$$\eta = \rho_i t = \frac{\lambda_i T_i}{Cap_i} t \quad (30)$$

where ρ_i is server *i*'s utilization (fraction of time a server is busy in total time of each trial); λ_i is the arrival rate (number of arrivals into sever *i* through *t*) and T_i and Cap_i are the processing time and capacity of server *i*, respectively.

For the same brain region, the percentage signal change (fMRI *PSC*) is the *CB(t)* of the experimental condition compared to the *CB(t)* of the baseline condition ($CB(t_0)$, e.g., fixation condition in Jiang et al. 2004) (see Equation 31) (Ben-Shachar, Hendler, Kahn, Ben-Bashat, & Grodzinsky, 2003).

$$PSC = \frac{CB(t) - CB(t_0)}{CB(t_0)} \tag{31}$$

Therefore, according to Equations 29 to 31, *PSC* at short and long SOA conditions (PSC_{long}, PSC_{short}) can be calculated if T_i, Cap_i, $\lambda_{i,}$, k, M, s, b, a, and t at these conditions are given.

For the same brain regions measured by the same fMRI techniques, s, k, M, a, T_i, Cap_i, and b are expected to be remained the same in short and long SOA conditions. Furthermore, since the length of each trial is fixed either at short or long SOA conditions, the value of t also remains the same in short and long SOA conditions. During each trial, the same amount of information through t arrived at the cognitive system; therefore, λ_i remains the same in short and long SOA conditions. Therefore, according to Equations 29-31 above, for the same brain region, the expected percentage of change of *CB(t)* keeps constant across different SOA conditions, i.e.:

$$\therefore CB(t)_{long} = CB(t)_{short}$$

$$\because PSC_{long} - PSC_{short} = \frac{CB(t)_{long} - CB(t_0)}{CB(t_0)} - [\frac{CB(t)_{short} - CB(t_0)}{CB(t_0)}] = \frac{CB(t)_{long} - CB(t)_{short}}{CB(t_0)}$$

$$\therefore PSC_{long} - PSC_{short} = 0$$

In other words, in this Queueing process, since the amount of information processed by each brain region remains the same in the short and long SOA conditions, the integrated BOLD signal remains the same in the short and long SOA conditions.

Parameter Setting

Except for the 2 free parameters set at the long SOA condition following the parameter setting method in ACT-R/PM, the values of the other parameters are the same as in those in Table 3-2 to Table 3-6 (see Table 3-7). Moreover, the values of the free parameters set at long SOA conditions are also constrained by the nature of the tasks: processing times

of entities at Server F in both T1 and T2 are similar (range of difference≤120 ms) since the difficulty levels of the two choice reaction tasks are close to each other.

Table 3-7 Parameters in modeling Jiang et al (2004)'s experiment

Parameter	Value	Description	Source
$T_{1,VP,}$ $T_{2,VP}$	126 ms	Time for visual perception time (42 ms at Servers 1,2/3, and 4)	Liu, et al. (in press)
$T_{2,A,}$ $T_{2,B}$	18 ms	Processing time at Server A	Liu, et al. (in press)
$T_{1,C,}$ $T_{2,C}$	18 ms	Processing time at Server C	Liu, et al. (in press)
$T_{1,F}$	408 ms	Processing time at Server F (T1)	Value estimated
$T_{2,F}$	376 ms	Processing time at Server F (T2)	Value estimated
$T_{1,W,}$ $T_{2,W}$	24 ms	Processing time at Server W	Liu, et al. (in press)
$T_{1,Y,}$ $T_{2,Y}$	24 ms	Processing time at Server Y	Liu, et al. (in press)
$T_{1,Z,}$ $T_{2,Z}$	24 ms	Processing time at Server Z	Liu, et al. (in press)
$T_{1,K}$	10 ms	Key closure time	Byrne & Anderson (2001)

Modeling Results and their Validation

Using the equations derived in the previous sections, the predicted results of both reaction time and the percentage of change of fMRI signal are presented and validated with the target experiment results.

1) Reaction Time

Figure 3-19 shows the modeling results in comparison with experimental results in reaction time: the R square of the model is .8 and the RMS=35.0 ms.

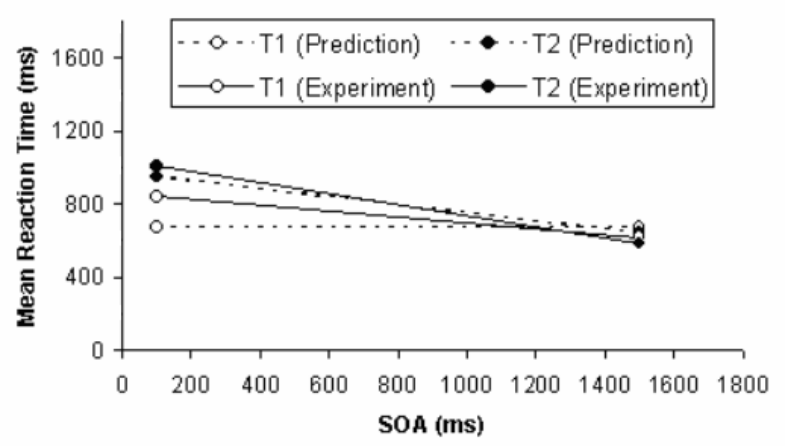


Figure 3-19 The reaction time in the study of Jiang et al. (2001) (solid lines) along with the Queueing network modeling results (dashed lines)

2) fMRI Signal

Figure 3-20 shows the modeling results in comparison with experimental results of the fMRI signal: the R square of the model is .70 and the RMS=0.03.

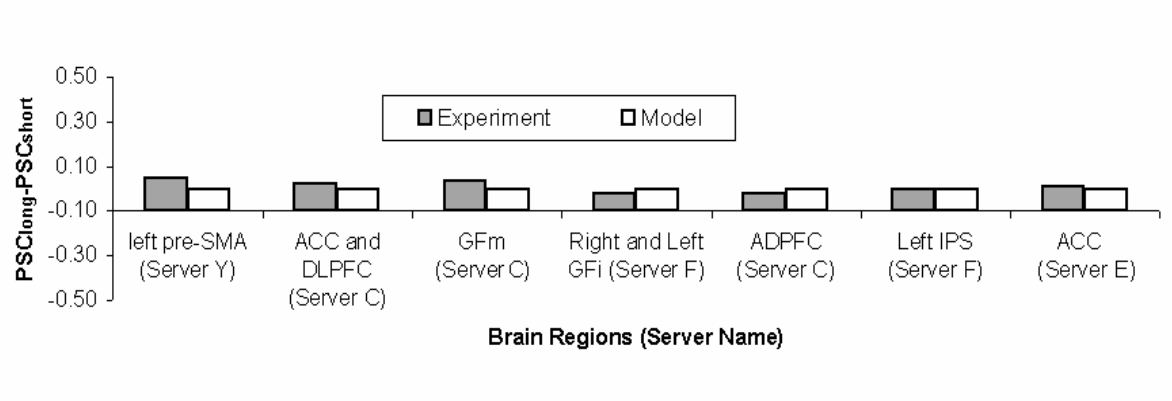


Figure 3-20 Difference of PSC between Long and Short SOA Conditions (PSC_{long}-PSC_{short}) in the study of Jiang et al. (2004)[6] along with the Queueing network modeling results

4.4 Response Grouping Effect

The experimental study of Ruthruff et al. (2001) found the response grouping effect by asking the subject to emit the responses of two tasks at the same time. The two tasks in their experiment are a tone counting task and a spatial working memory task (see the detailed experiment description in the introduction part of this article).

Routes of Entities

Spatial Working Memory Task

Based on the task and corresponding function of the brain areas, for the spatial working memory task, it is possible that the lateral BA 6 (represented by Server V), the superior frontal gyrus (SFS) (represented by Server F) or both of them can be involved in processing the spatial working memory task (see the introduction section of this article).

[6] Even though Jiang et al. (2004) found that there was a negative correlation between the fMRI PSC of the right GFi and the behavioral interference, there was no negative or positive correlation between a) the averaged fMRI PSC of the right and left GFi and b) the behavioral interference.

According to the neuron pathways and connections between these two brain regions and other regions, the two possible routes to process the spatial working memory task are as follows:

1) 1->2/3->4->A->C->F->C->W->Y->Z->Hand

2) 1->2/3->4->A->V->W->Y->Z->Hand

Tone Counting Task

According to the function of Server F (IPS in Server F is active for the mental calculation and numerical operation task), the route for the numerical operations task is:

5->6/7->8-> B-> C->F->C->W->Y->Z->Hand

Two modeling methods have been used to quantify the response grouping effect. In this section, we introduce the first method only considering the route without the effect of practice on performance (1->2/3->4->A->C->F->C->W->->Y->Z->Hand). In modeling the practice effect of PRP, we introduce the second method and it considers the practice effect on performance as well as the two possible routes in this spatial working memory task.

Modeling of the Expected Reaction Time

If the practice effect is not considered in Ruthruff et al.'s experiment (2001), entities of the spatial working memory task and the tone counting task both pass through the bottleneck server (Server F), which is similar to the basic PRP situation. In general, there are 5 possible conditions of the expected reaction time ($E(RT_{1\text{-}route})$) depending on the different values of *DIFF* (assuming $T_{1,P}+T_{1,A|B}+T_{1,C}-(T_{2,P}+T_{2,A|B}+T_{2,C})=DIFF$) (Because SOA=0 and the responses of the two tasks are emitted at the same time, the expected reaction times for the two tasks are equal) (see Figure 3-21).

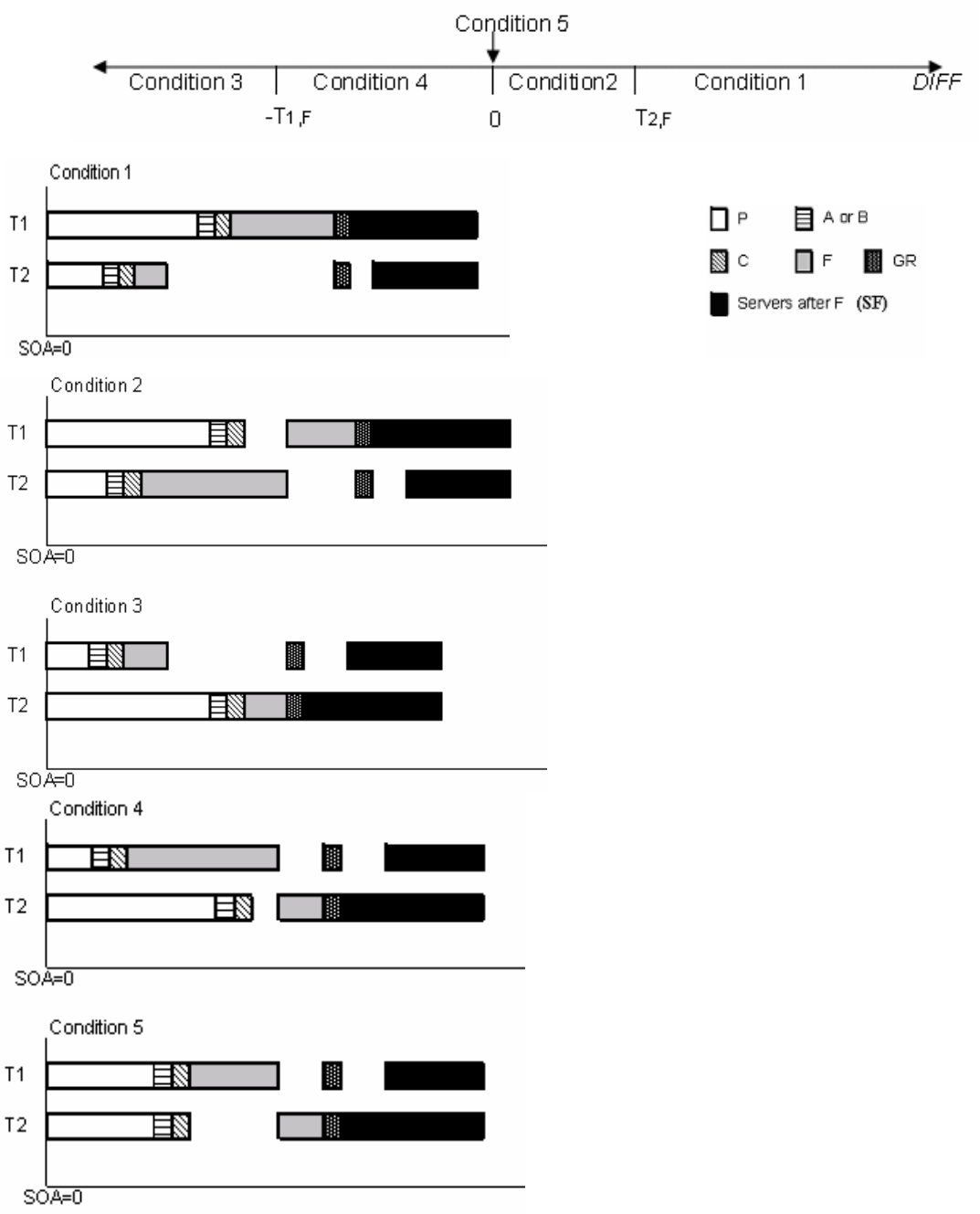


Figure 3-21Five possible conditions in estimating the $RT_{1\text{-route}}$ in modeling the response grouping effect

Condition 1 ($T_{1,P}+T_{1,A|B}+T_{1,C} > T_{2,P}+T_{2,A|B}+T_{2,C}$ and $T_{1,P}+T_{1,A|B}+T_{1,C}-(T_{2,P}+T_{2,A|B}+T_{2,C}+T_{2,F}) \geq 0$; i.e., $DIFF \geq T_{2,F}$):

$$E(RT_{1\text{-route}})=E(RT_1)=E(RT_2)= T_{1,P}+T_{1,A|B}+T_{1,C} +T_{1,F}+GR+ \max(SF_1,SF_2) \quad (32)$$

where $T_{1,A|B}$ is the processing time at Server A or B (depends on T1's route); *GR* refers to the time in grouping the two responses together; and SF_1 and SF_2 are the sum of processing time of servers after Server F (e.g., Servers Y, W, and Z) of T1 and T2, respectively.

Condition 2 ($T_{1,P}+T_{1,A|B}+T_{1,C} > T_{2,P}+T_{2,A|B}+T_{2,C}$ and $T_{1,P}+T_{1,A|B}+T_{1,C}-(T_{2,P}+T_{2,A|B}+T_{2,C}+T_{2,F})<0$; i.e., $0<DIFF< T_{2,F}$):

$$E(RT_{1\text{-route}})=E(RT_1)=E(RT_2)= T_{1,P}+T_{1,A|B}+T_{1,C} +T_{1,F}+T_{2,F} +GR+ max(SF_1,SF_2) \quad (33)$$

Condition 3 ($T_{2,P}+T_{2,A|B}+T_{2,C} >T_{1,P}+T_{1,A|B}+T_{1,C}$ and $T_{2,P}+T_{2,A|B}+T_{2,C} -(T_{1,P}+T_{1,A|B}+T_{1,C}+T_{1,F})>0$; i.e., $-T_{1,F} <DIFF<0$):

$$E(RT_{1\text{-route}}) =E(RT_1)=E(RT_2)= T_{2,P}+T_{2,A|B}+T_{2,C} +T_{2,F} +GR+ max(SF_1,SF_2) \quad (34)$$

Condition 4 ($T_{2,P}+T_{2,A|B}+T_{2,C} >T_{1,P}+T_{1,A|B}+T_{1,C}$ and $T_{2,P}+T_{2,A|B}+T_{2,C} -(T_{1,P}+T_{1,A|B}+T_{1,C}+T_{1,F})\leq 0$; i.e., $DIFF\leq -T_{1,F}$):

$$E(RT_{1\text{-route}})=E(RT_1)=E(RT_2)= T_{1,P}+T_{1,A|B}+T_{1,C} +T_{1,F}+T_{2,F} +GR+ max(SF_1,SF_2) \quad (35)$$

Condition 5 ($T_{2,P}+T_{2,A|B}+T_{2,C} =T_{1,P}+T_{1,A|B}+T_{1,C}$; i.e., $DIFF=0$):

$$E(RT_{1\text{-route}})=E(RT_1)=E(RT_2)= T_{1,P}+T_{1,A|B}+T_{1,C} +T_{1,F}+T_{2,F} +GR+ max(SF_1,SF_2) \quad (36)$$
$$= T_{2,P}+T_{2,A|B}+T_{2,C} +T_{1,F}+T_{2,F} +GR+ max(SF_1,SF_2)$$

Specifically, in Ruthruff et al.'s experimental study (2001), $T_{1,P}=T_{S,P}$, $T_{1,A|B}=T_{S,A}$, $T_{1,F}=T_{S,F}$, $SF_1= SF_S = T_{S,C}+ T_{S,W} +T_{S,Y}+T_{S,Z}+T_{S,k}$; $T_{2,P}=T_{T,P}$, $T_{2,A|B}=T_{T,B}$, $T_{2,F}=T_{T,F}$, $SF_2= SF_T = T_{T,C}+ T_{T,Y}+T_{T,Z}+T_{T,V}$. Based on Equation 32-36, $RT_{1\text{-route}}$ in Ruthruff et al.'s experimental study (2001) can be estimated:

Condition 1 ($DIFF\geq T_{T,F}$):

$$E(RT_{1\text{-route}})=E(RT_1)=E(RT_2) \quad (37)$$
$$=T_{S,P}+T_{S,A} +T_{S,C} +T_{S,F}+GR+ max(T_{S,C}+ T_{S,W} +T_{S,Y}+T_{S,Z}+T_{S,k}, T_{T,C}+ T_{T,Y}+T_{T,Z}+T_{T,V})$$

where *GR* refers to the time in grouping the two responses together.

Condition 2 ($0<DIFF< T_{T,F}$):

$$E(RT_{1\text{-route}})=E(RT_1)=E(RT_2)= T_{T,P}+T_{T,\ B}+T_{T,C} +T_{S,F}+T_{T,F} +GR+max(T_{S,C}+ T_{S,W} \quad (38)$$
$$+T_{S,Y}+T_{S,Z}+T_{S,k}, T_{T,C}+ T_{T,Y}+T_{T,Z}+T_{T,V})$$

Condition 3 ($-T_{S,F} < DIFF < 0$):

$$E(RT_{1\text{-route}}) = E(RT_1) = E(RT_2) \quad (39)$$
$$= T_{T,P} + T_{T,B} + T_{T,C} + T_{T,F} + GR + \max(T_{S,C} + T_{S,W} + T_{S,Y} + T_{S,Z} + T_{S,k},\ T_{T,C} + T_{T,Y} + T_{T,Z} + T_{T,V})$$

Condition 4 ($DIFF \leq -T_{S,F}$):

$$E(RT_{1\text{-route}}) = E(RT_1) = E(RT_2) = T_{S,P} + T_{S,A} + T_{S,C} + T_{S,F} + T_{T,F} + GR + \max(T_{S,C} + T_{S,W} + T_{S,Y} + T_{S,Z} + T_{S,k},\ T_{T,C} + T_{T,Y} + T_{T,Z} + T_{T,V}) \quad (40)$$

Condition 5 ($DIFF = 0$):

$$E(RT_{1\text{-route}}) = E(RT_1) = E(RT_2) = T_{S,P} + T_{S,A} + T_{S,C} + T_{S,F} + T_{T,F} + GR + \max(T_{S,C} + T_{S,W} + T_{S,Y} + T_{S,Z} + T_{S,k},\ T_{T,C} + T_{T,Y} + T_{T,Z} + T_{T,V}) \quad (41)$$

Based on the equations developed above and the modeling mechanism, the expected pattern of the reaction times in these 5 conditions is shown in Figure 3-22. The expected reaction time in the dual-task condition is longer than that in the single reaction time condition and the expected reaction time in Conditions 1 and 3 is shorter than that in Conditions 2, 4, and 5.

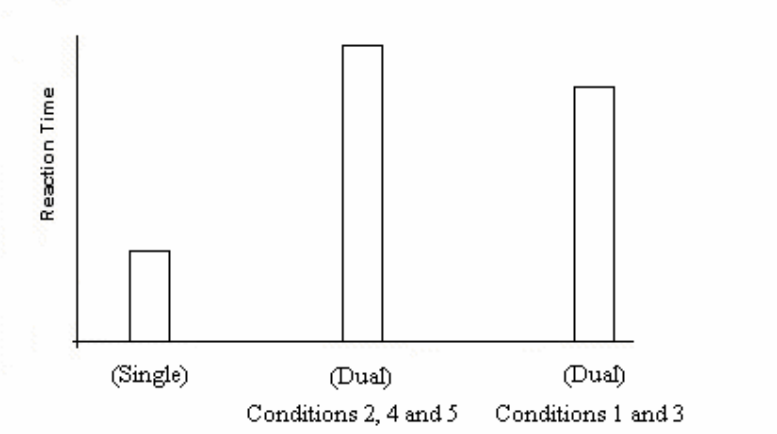


Figure 3-22 The expected patterns of the reaction time of single and dual task conditions in Ruthruff et al.'s experiment (2001) (without considering the practice effect)

Parameter Setting

Only two free parameters (F_{Tone} and F_{Letter}) are used and they are set based on the reaction time at the single task condition (see Table 3-8). Moreover, the value of these free parameters set at long SOA conditions are also constrained by the task properties: the

spatial working memory task involving mental rotations and the tone counting task involving mathematical addition are more complex than the choice reaction task in the previous experimental studies in PRP, increasing the number of cycles at Server F compared to the processing times at regular choice reaction conditions.

Table 3-8 Parameter setting in modeling of Ruthruff et al.'s experiment (2001) (practice effect not considered)

Parameter	Value	Description	Source
$T_{T,P}$	126 ms	Time for tone detection perception	Liu et al. (in press)
$T_{S,P}$	126 ms	Time for letter perception	Liu et al. (in press)
$T_{T,C,}$ $T_{S,C}$	18 ms	Processing time at Server C	Liu et al. (in press)
$T_{T,F}$	368 ms	Processing time of tone counting task at Server F	Value estimated at long SOA conditions
$T_{S,F}$	600 ms	Processing time of spatial working memory task at Server F	Value estimated at long SOA conditions
GR	21 ms	Time to group two responses	The control experiment in Ruthruff et al., (2001)
$T_{T,W,}$ $T_{S,W}$	24 ms	Processing time at Server W	Liu et al. (in press)
$T_{T,Z,}$ $T_{S,Z}$	24 ms	Processing time at Server Z	Liu et al. (in press)
$T_{S,K}$	10 ms	Key closure time	Byrne & Anderson (2001)
$T_{T,V}$	100 ms	Voice key closure time	Byrne & Anderson (2001)

Modeling Results and their Validation

Figure 3-23 shows the modeling results compared with experimental results in experiment: without considering the practice effect, the R square of the model is .99 and the RMS=25.5 ms.

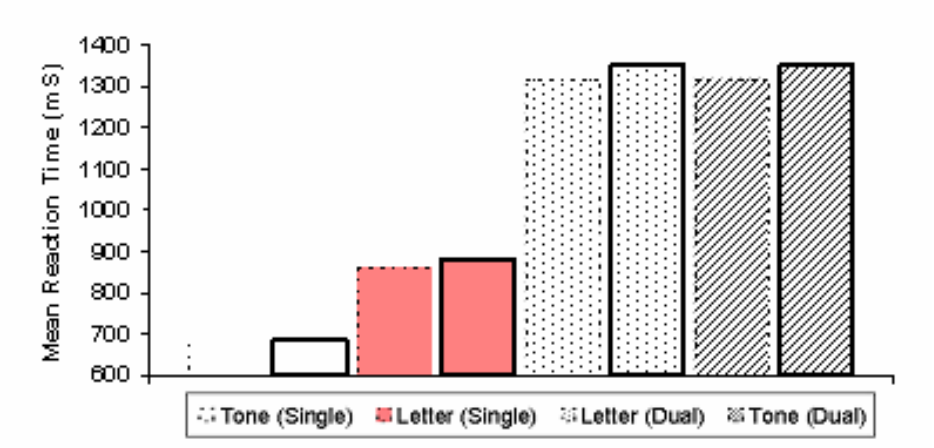


Figure 3-23 Mean reaction time in Ruthruff et al.'s experiment (2001) compared with modeling results (practice effect is not considered) (modeling result: columns with dashed line borders; experimental result: columns with solid line borders)

4.5 Practice Effect

In the experimental studies of Ruthruff et al. (2001) and Oberauer et al. (2004), subjects received relatively extensive practice, and there are two possible routes of the spatial working memory task in the cognitive system (see the two routes in modeling Ruthruff et al.'s study in the previous section). To model the practice effect of PRP in these two experiments, QN-MHP needs 2 additional task-independent assumptions: the first assumption quantifies the change of routing probability of entities in the network during the practice process; the second assumption quantifies the increase of servers' processing speed via practice.

Routes Rewiring via Practice

It is well recognized that the human brain is not only a network of brain regions, but also a system that is able to change itself dynamically in the process of development and learning (Chklovskii, et al., 2004; Habib, 2003). On the one hand, the "brain traffic" concept in neuroscience suggests that information flow represented by spike trains in the brain also exhibit features of traffic flow in the network. On the other hand, different brain areas are activated during the visual-motor practice process (Van Mier, et al., 1998; Petersen, et al., 1998; Aizawa et al. 1991). This plasticity aspect of the human brain concerns the change of synaptic connection strength between neurons and rewiring

among neural pathways—spike trains change from one neuron pathway to another with stronger synaptic connection strength and higher efficiency in information processing. This rapid regulation is related to a brain derived neurotrophic factor (BDNF) regarded as a signal of synaptic plasticity in adults (Black, 1999; Braus, 2004), and Black (1999) proposed a model explaining the role of BDNF in its regulation of the synaptic plasticity.

The routing probability equation (see Equation 42) in the Queueing network is developed based on Black's model (1999) and the "brain traffic" concept above (see Appendix 3 for its derivation), where routing probability (P_i) stands for the probability that spike trains (represented by entities) pass through a certain neuron pathway (route i) in total U multiple routes; additionally, sojourn time (S_i) is defined as the sum of waiting time (W_i) and processing time (T_i) of these spike trains (entities) at that neuron pathway.

$$P_i = \frac{1/S_i}{\sum_{j=1}^{U} \frac{1}{S_j}} \tag{42}$$

In other words, the synaptic connection strength, which depends on practice processes that improves the effectiveness of the information processing of brain regions in the neuron pathway (route), determines the probability that the spike trains (entities) enter one of multiple neuron pathways (routes). If the majority of entities change their route from one to another, rewiring of routes (neuron pathways) occurs.

Reduction of Processing Time at Individual Servers via Practice

Because exponential function fits the practice processes in various tasks, including memory search, motor learning, visual search, and mathematical operations, better than power law (Heathcote, Brown, & Mewhort, 2000), it is applied in order to model the practice process in each server (see Equation 43).

$$T_i = A_i + B_i \, Exp(-\alpha_i N_i) \tag{43}$$

where A_i is the minimal processing time of server i after intensive practice; B_i represents the change of expected value of processing time of server i from the beginning to the end of practice; α_i is the learning rate of server i; and N_i is the number of customers processed by server i.

4.5.1 Modeling Ruthruff et al.'s Experiment (2001) Considering the Practice Effect

In Ruthruff et al.'s experiment (2001), subjects received relatively extensive single and dual-task practice; hence, there are two multiple routes of entities in performing the same spatial working memory task (see two routes in modeling of Ruthruff et al.'s experiment). In the dual-task practice condition[7], sojourn time at Server C increased from the processing time to the sum of waiting time (entities maintained at Server C to wait the previous entities finished processing at Server F). Therefore, according to Equation 42 in the route rewiring, the probability of entities entering Server C decreases and the probability of entering Server V increases because of the increased sojourn time, which causes the processing time at Server V to decrease further. As a result, entities take 2 routes in the dual-task practice condition: A-> V or A->C->F. The higher the probability of entities taking the route from Server A to V, the higher the probability that Server F bottleneck is bypassed and the PRP effect disappears. Therefore, QN-MHP is able to model the disappearance of the PRP effect in dual-task practice naturally based on the queue network architecture, rather than relying on any executive control or strategic scheduling assumptions.

In the situation that entities of two tasks take two routes without conflicting with each other, the expected reaction time of both tasks ($E(RT_{2\text{-}route})$) can be estimated based on Figure 3-24 and Equation 44. Since the entities of T1 and T2 take two independent routes at the cognitive subnetwork and entities of T2 are never blocked by the entities of T1 in their route (see Figure 3-24), in general, the expected mean of reaction time of both tasks ($E(RT_{2\text{-}route})$) is quantified in Equation 45 (assuming Ω represents any server which can also have the function of Server F in the target task after practice).

[7] Modeling mechanism of Oberauer's experiment explains why only dual-task practice forms the 2 routes in processing the spatial working memory task.

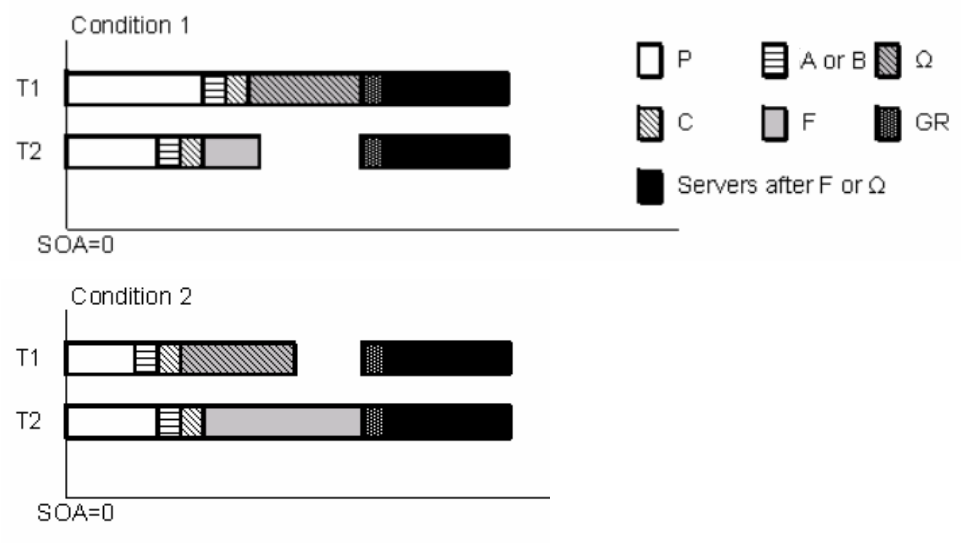


Figure 3-24 Two possible conditions in estimating the $RT_{2\text{-route}}$

$$E(RT_{2\text{-route}})=max(T_{1,P}+ T_{1,A|B}+ T_{1,\Omega},\ T_{2,P}+ T_{2,A|B}+ T_{2,C}+ T_{2,F})+GR+ max(SF_1,\ SF_2) \quad (44)$$

where $T_{1,\Omega}$ stands for the current processing time of Server Ω; $T_{1,A|B}$ is the current processing time at Server A or B (depends on T1's route); *GR* refers to the time in grouping the two responses together; and SF_1 and SF_2 are the sum of current processing time of servers after Server F (e.g., Servers Y, W, and Z) of T1 and T2, respectively.

Specifically, $E(RT_{2\text{-route}})$ in Ruthruff et al.'s experiment can be estimated in Equation 45 in which V represents Ω.

$$E(RT_{2\text{-route}})=max(T_{S,P}+ T_{S,A}+ T_{S,VS},\ T_{T,P}+ T_{T,B}+ T_{T,C}+ T_{T,F})+GR+ max(T_{S,C}+ T_{S,W}+T_{S,Y}+T_{S,Z}+T_{S,k},\ T_{T,C}+ T_{T,Y}+T_{T,Z}+T_{T,V}) \quad (45)$$

Suppose $P_{2\text{-route}}$ is the possibility that the entities take the second route (A->V->Z) in the practice situation, i.e., the possibility that the entities of two tasks take two routes (see Equation 46); therefore, the expected reaction time ($E(RT_{practice})$) in the practice situation is:

$$E(RT_{practice})= P_{2\text{-route}}\ E(RT_{2\text{-route}}) + (1- P_{2\text{-route}})\ E(RT_{1\text{-route}}) \quad (46)$$

Based on the equations developed above and the modeling mechanism, Figure 3-25 shows the expected pattern of reaction time in single and dual task conditions (the expected reaction time of Conditions 1 and 2 in Figure 3-24 in the modeling mechanism overlap in Figure 3-25).

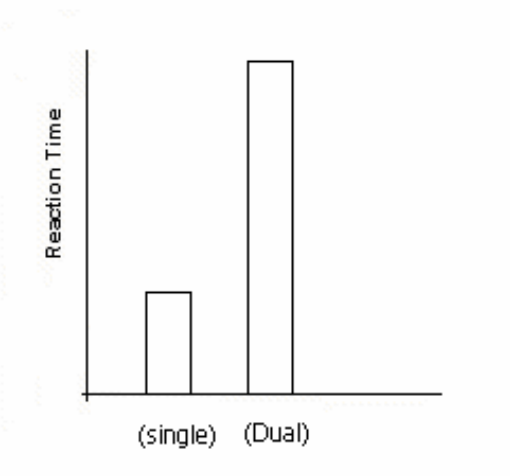


Figure 3-25 The expected pattern of reaction time of the single and dual task conditions in Ruthruff et al.'s experiment (2001) (considering the practice effect)

Parameter Setting

Only 3 free parameters are used in the parameter setting when the practice effect is taken into consideration (see

Table 3-9). The values of these free parameters set at long SOA conditions are also constrained by the task properties: the spatial working memory task involving mental rotations and the tone counting task involving mathematical addition are more complex than the choice reaction task in the previous experimental studies in PRP, increasing the number of cycles at Server F compared to the processing times at regular choice reaction conditions; since Server V is only able to process spatial working memory information after extensive practice (Inoue et al., 2000, see introduction section of this article), the processing time of Server V without extensive practice is longer than that after extensive practice.

Table 3-9 Parameter setting in modeling of Ruthuff et al.'s experiment (2001) (consider practice effect)

Parameter	Value before Extensive Practice (A_i+B_i) (Liu, et al, in press)	Minimal value (A_i) Feyen (2002)	Description and Practice Process Heathcote et al. (2000)
$T_{T,P}$	126 ms (42 ms mean perceptual cycle time at Servers 5, 6/7, 8, i.e., 42*3=126 ms)	75 ms (25 ms minimal perceptual cycle time at Servers 5, 6/7, 8, i.e., 25*3=75 ms)	Time for tone detection perception $A_i+B_i*Exp(-N\xi)$
$T_{S,P}$	126 ms (42 ms mean perceptual cycle time at Servers 1, 2/3, 4, i.e., 42*3=126 ms)	75 ms (25 ms minimal perceptual cycle time at Servers 1, 2/3, 4, i.e., 25*3=75 ms)	Time for letter perception $A_i+B_i*Exp(-N\xi)$
$T_{T,C,}$ $T_{S,C}$	18 ms	6 ms (minimal cognitive cycle time)	Processing time at Server C $A_i+B_i*Exp(-N\xi)$
*$T_{S,VS}$	1540 ms	10 ms (minimal motor cycle time)	Processing time at Server V (T1) $A_i+B_i*Exp(-N\xi)$
$T_{T,W,}$ $T_{S,W}$	24 ms	10 ms (minimal motor cycle time)	Processing time at Server W $A_i+B_i*Exp(-N\xi)$
$T_{T,Y,}$ $T_{S,Y}$	24 ms	10 ms (minimal motor cycle time)	Processing time at Server W $A_i+B_i*Exp(-N\xi)$
$T_{T,Z,}$ $T_{S,Z}$	24 ms	10 ms (minimal motor cycle time)	Processing time at Server Z $A_i+B_i*Exp(-N\xi)$
$T_{S,K}$	10 ms	10 ms	Key closure time
$T_{T,V}$	100 ms	100 ms	Voice key closure time
G	21 ms	6 ms	Time to group two responses $A_i+B_i*Exp(-N\xi)$
*$T_{T,F}$	439 ms	6 ms (minimal cognitive cycle time)	Processing time of the tone counting task at Server F $A_i+B_i*Exp(-N\xi)$
*$T_{S,F}$	746 ms	6 ms (minimal cognitive cycle time)	Processing time of the spatial memory task at Server F $Ai+Bi*Exp(-N\xi)$

* Value estimated at long SOA conditions; ξ=.001 (Heathcote et al., 2000)

Modeling Results and their Validation

Figure 3-26 shows the modeling results compared with the experimental results considering the practice effect: the R square of the model is .99 and RMS=0.1 ms. The modeling method considering practice effect generated a relatively better fit with the empirical data than that without considering the practice effect.

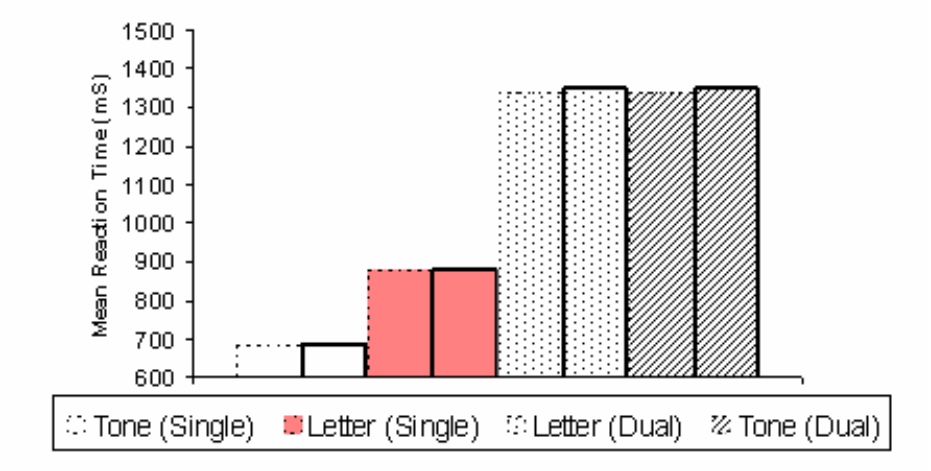


Figure 3-26 Mean reaction time in Ruthruff et al.'s experiment (2001) compared with modeling results (practice effect is considered) (modeling result: columns with dashed line borders; experimental result: columns with solid line borders)

4.5.2 Modeling Oberauer et al.'s Experiment (2004)

Oberauer et al.'s experimental finding (2004) is a counterexample of ACT-R/PM because they found that people can perform two complex cognitive tasks at the same time after hundreds of trials of dual-task practice (Oberauer & Kliegl, 2004). Two tasks in their experiment include a spatial working memory task and a numerical operations task. And there are two practice groups of subjects in the experiment: single and dual task practice group (see detailed description of their experiment in the introduction part).

Routes of Entities

Similar to the routes selection in the modeling of Ruthruff et al.'s experiment (2001), the two possible routes to process the spatial working memory task are:

1) 1->2/3->4->A->C->F->C->W->Y->Z->Hand

2) 1->2/3->4->A->V ->W->Y->Z->Hand

Numerical Operations Task

According to the function of Server F (IPS in Server F is active for the mental calculation task), the route for the numerical operations task is:

5->6/7->8->B->C->F->C->W->Y->Z->Hand

Modeling Mechanisms

Based on the possible two routes in processing the spatial memory task, starting at Server A, there are two multiple routes: route 1 (A->C->F->C->W->Y->Z->Hand) and route 2 (A->V->W->Y->Z->Hand). In the single-task practice condition, entities at Server A always enter Server C and then enter Server V because sojourn time at Server C is shorter than Server V (mean processing time at Server C=18 ms, sojourn time at Server C=18 ms; mean processing time at Server V=1540 ms; see the value of these parameters in Ruthruff et al.'s experiment (2001), sojourn time at Server V =1540 ms).

Since very few entities enter Server V and its processing time does not decrease much in the single-task practice task condition, the majority of entities still take the same route throughout the practice sessions. As a result, Server F bottleneck is not bypassed by the majority of entities and the PRP effect does not disappear under the single-task practice condition and the reduced the PRP effect is due to reduced processing time at Server F and the other servers.

In the dual-task practice condition, there is a 50% chance that sojourn time at Server C increases from the processing time to the sum of waiting time (entities maintained at Server C to wait the previous entities finished processing at Server F). Therefore, similar to the modeling mechanism in Ruthruff et al.'s experiment, according to Equation 42 above, the probability of entering Server C decreases and the probability of entering Server V increases because of the increased sojourn time, decreasing the processing time at Server V further. The higher the probability of entities taking the route from Server A to Server V, the higher the probability that Server F bottleneck is bypassed and the PRP effect disappears. Therefore, QN-MHP is able to account for the disappearance of the PRP effect in dual-task practice naturally based on the queue network architecture rather than relying on any strategic scheduling assumptions.

Mathematical Modeling of Expected Reaction Time

Sequential Test

In both the sequential pre-test and post-test, the expected reaction time ($E(RT_{pre\text{-}seq})$) in both the single-task practice and dual-task practice groups is the maximum of RT1 and RT2 (see introduction of their experiment in detail above) (see Equation 47)

$$E(RT_{pre\text{-}seq}) = max(T_{S,P}+ T_{S,A}+ T_{S,C} +T_{S,F} + T_{S,C} + T_{S,Y} +T_{S,W} +T_{S,Y}+ T_{S,Z} + T_{S,K,,} T_{T,P}+ T_{T,B}+ T_{T,C} +T_{T,F} + T_{T,C} + T_{T,Y} +T_{T,W} +T_{T,Y}+ T_{T,Z} + T_{T,K}) \quad (47)$$

where all of the processing times of servers refer to the current processing time determined by Equation 43.

Simultaneous Test

1) Pre-test RT

In the simultaneous test of the pre-test, the expected reaction time ($E(RT_{pre\text{-}simul})$) in the two practice groups is similar to that in the basic PRP paradigm (SOA=0); however, based on the experiment design in Oberauer et al. (2004), there is a 50% chance that either the spatial or numerical operation task can be treated as T1. Hence, the expected reaction time in this simultaneous test for both two groups is quantified into Equation 48.

$$E(RT_{pre\text{-}simul}) = 0.5[max(T_{S,P}+T_{S,A}+T_{S,C}+T_{S,F},\ T_{T,P}+T_{T,B}+T_{T,C})+ T_{T,F} +T_{T,C} +T_{T,Y}+ T_{T,W} +T_{T,Z} + T_{T,K}]+0.5[max(T_{T,P}+ T_{T,B}+ T_{T,C} +T_{T,F},\ T_{S,P}+T_{S,A}+T_{S,C})+ T_{S,F} +T_{S,C} + T_{S,Y} +T_{S,W} + T_{S,Z} + T_{S,K})] \quad (48)$$

2) Post-test RT

Dual-task practice Group

At the end of practice, there are two situations in estimating the expected reaction time in Oberauer's experiment: first, when the entities of the two tasks take two independent routes (A->V->Z and B->C->F->C->W->Y->Z), the expected reaction time in this condition ($E(RT_{2\text{-}route})$) is the maximum of reaction time of T1 and T2 (see Figure 3-27 and Equation 49).

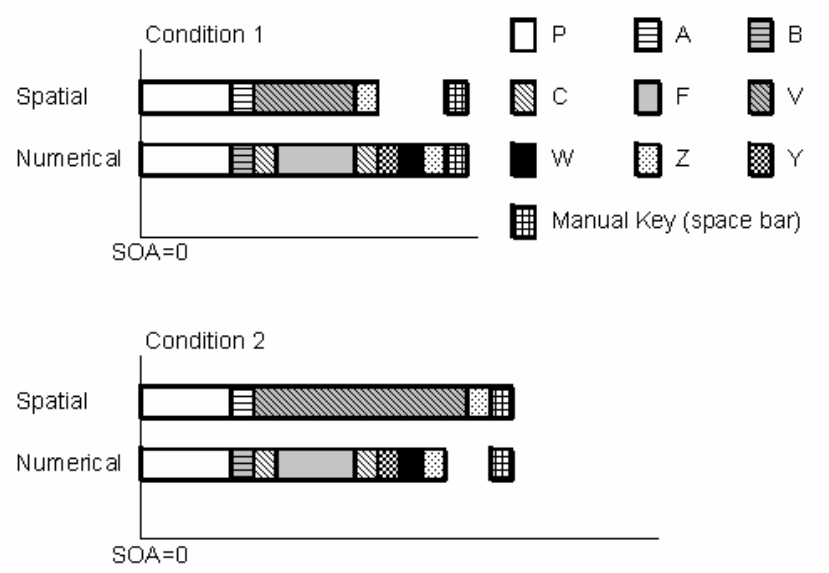


Figure 3-27 Two possible conditions of reaction time in the simultaneous post-test of the dual practice group (in both condition 1 and 2, the expected reaction time is the maximum of each individual task)

$$E(RT_{2\text{-}route})=max(T_{S,P}+T_{S,A}+T_{S,VS}+T_{S,Z}+T_{S,k},\ T_{T,P}+\ T_{T,B}+\ T_{T,C}+\ T_{T,F}+\ T_{T,C}+\ T_{T,Y}+\ T_{T,W}+\ T_{T,Z}+T_{T,K}) \quad (49)$$

Second, when the entities of the two tasks both enter Server F in the cognitive subnetwork, the expected reaction time in this condition ($E(RT_{1\text{-}route})$) can be quantified by Equation 50 except the processing time of servers is changed to the current processing time after practice.

$$E(RT_{1\text{-}route}) = 0.5[max(T_{S,P}+T_{S,A}+T_{S,C}+T_{S,F},\ T_{T,P}+T_{T,B}+T_{T,C})+\ T_{T,F}\ +T_{T,C}\ +T_{T,Y}+\ T_{T,W}\ +T_{T,Z}\ +\ T_{T,K}]+0.5[max(T_{T,P}+\ T_{T,B}+\ T_{T,C}\ +T_{T,F},\ T_{S,P}+T_{S,A}+T_{S,C})+\ T_{S,F}\ +T_{S,C}\ +\ T_{S,Y}\ +T_{S,W}\ +\ T_{S,Z}\ +\ T_{S,K}] \quad (50)$$

Therefore, similar to the quantification of the reaction time in Ruthruff et al.'s experiment, the expected reaction time in Oberauer's experiment ($E(RT_{practice\text{-}dual})$) can be quantified into Equation 51 where $P_{2\text{-}route\text{-}dual}$ (determined by Equation 42) refers to the probability that entities of spatial operation task take the second route (A->V->Z) in the dual-task practice group.

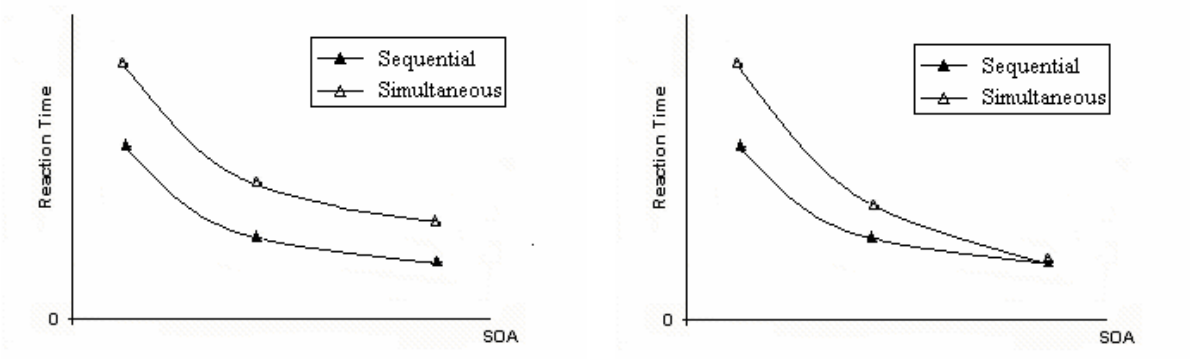
$$E(RT_{practice\text{-}dual}) = P_{2\text{-}route\text{-}dual}\ E(RT_{2\text{-}route}) + (1-\ P_{2\text{-}route\text{-}dual})\ E(RT_{1\text{-}route}) \quad (51)$$

Single-task practice Group

The expected reaction time of the single-task practice group ($E(RT_{practice\text{-}single})$) can also be quantified into Equation 52 where $P_{2\text{-}route\text{-}single}$ (determined by Equation 42) refers to the probability that entities of the spatial operation task take the second route (A->V->Z) in the single-task practice group.

$$E(RT_{practice\text{-}single}) = P_{2\text{-}route\text{-}single}\, E(RT_{2\text{-}route}) + (1 - P_{2\text{-}route\text{-}single})\, E(RT_{1\text{-}route}) \quad (52)$$

Based on the equations developed above and the modeling mechanism, Figure 3-28 shows the expected pattern of reaction time for the single-task and dual-task practice groups. The inverse relation between the probability of routing and sojourn time in Equation 42 produces a non-linear relation between the SOA and RT, and the slope of the curve of the dual-task practice group is greater than that of single-task practice group because of the higher probability of longer sojourn time at Server C in the dual-task practice condition than single-task practice condition (see the modeling mechanism in this section).



(Single-task practice group) (Dual-task practice group)

Figure 3-28 The expected pattern of reaction time of the single and dual-task practice groups in Oberauer et al.'s experiment (2004)

Parameter Setting

Only 2 free parameters were used in the parameter setting process (see Table 3-10): $T_{S,F}$ and $T_{T,F}$. The values of these free parameters set at long SOA conditions are also constrained by the task properties: the spatial working memory task involving mental updating of the current target and the numerical operation task involving mathematical addition or subtraction are more complex than the choice reaction task in the previous

experimental studies in PRP, increasing the number of cycles at Server F compared to the processing times at regular choice reaction conditions.

Table 3-10 Parameter setting in modeling of Oberauer et al.'s experiment (2004)

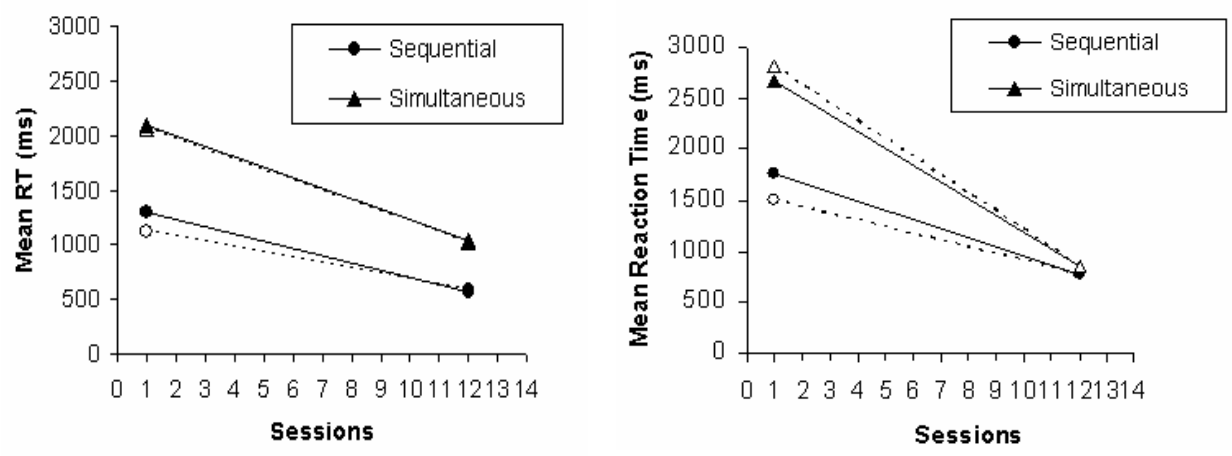
Parameter	Value before Extensive Practice (A_i+B_i) (Liu, et al, in press)	Minimal value (A_i) Feyen (2002)	Description and Practice Process Heathcote et al. (2000)
	Common parameters for both single and dual task groups		
$T_{T,P}$	126 ms (42 ms mean perceptual cycle time at Servers 5, 6/7, 8, i.e., 42*3=126 ms)	75 ms (25 ms minimal perceptual cycle time at Servers 5, 6/7, 8, i.e., 25*3=75 ms)	Time for tone detection perception $A_i+B_i*Exp(-N\xi)$
$T_{S,P}$	126 ms (42 ms mean perceptual cycle time at Servers 1, 2/3, 4, i.e., 42*3=126 ms)	75 ms (25 ms minimal perceptual cycle time at Servers 1, 2/3, 4, i.e., 25*3=75 ms)	Time for letter perception $A_i+B_i*Exp(-N\xi)$
$T_{T,C,}$ $T_{S,C}$	18 ms	6 ms (minimal cognitive cycle time)	Processing time at Server C $A_i+B_i*Exp(-N\xi)$
$T_{S,VS}$	1540 ms (See parameter setting in Ruthruff et al.'s experiment considering practice effect)	10 ms (minimal motor cycle time)	Processing time at Server V (T1) $A_i+B_i*Exp(-N\xi)$
$T_{T,W,}$ $T_{S,W}$	24 ms	10 ms (minimal motor cycle time)	Processing time at Server W $A_i+B_i*Exp(-N\xi)$
$T_{T,Y,}$ $T_{S,Y}$	24 ms	10 ms (minimal motor cycle time)	Processing time at Server W $A_i+B_i*Exp(-N\xi)$
$T_{T,Z,}$ $T_{S,Z}$	24 ms	10 ms (minimal motor cycle time)	Processing time at Server Z $A_i+B_i*Exp(-N\xi)$
$T_{S,K}$	10 ms	10 ms	Key closure time
	Parameters for single-task practice group[8]		
*$T_{S,F}$	1038 ms	6 ms (minimal cognitive cycle time)	Processing time of spatial memory task at Server F $A_i+B_i*Exp(-N\xi)$
*$T_{T,F}$	800 ms	6 ms (minimal cognitive cycle time)	Processing time of numerical operation task at Server F $A_i+B_i*Exp(-N\xi)$
	Parameters for dual-task practice group		
*$T_{S,F}$	925 ms	6 ms (minimal cognitive cycle time)	Processing time of spatial memory task at Server F $A_i+B_i*Exp(-N\xi)$
*$T_{T,F}$	1498 ms	6 ms (minimal cognitive cycle time)	Processing time of numerical operation task at Server F $A_i+B_i*Exp(-N\xi)$

[8] It is difficult to use a same set of $T_{S,F}$ and $T_{T,F}$ for both groups in session 1, since 500 ms difference in sequential and simultaneous conditions was found between those two groups before practice.

* Value estimated at long SOA conditions; ξ=.001 (Heathcote et al., 2000)

Modeling Results and their Validation

Figure 3-29 shows the modeling results compared with experimental results in the single and dual task practice groups. For the single-task practice group, the R square of the model is .99 and the RMS=46.4 ms[9]; in Experiment 4, the R square of the model is .99 and the RMS=43.7 ms.



(Single-task practice group) (Practice-task practice group)

Figure 3-29 Mean reaction time in Oberauer et al's experiment (2004) (solid lines) compared with modeling results (dashed lines)

5. Discussion

In this article, we described how the Queueing network model is able to account for various experimental findings in PRP, including those counterexamples to the existing models. In this modeling work, we demonstrated the unique features of QN-MHP in modeling multiple task performance.

The most important feature of the queueing network model compared with other models is its hybrid network structure with both serial and parallel information processing properties in its cognitive subnetwork. First, based on corresponding neuroscience

[9] Even though the absolute value of RMS is higher than the modeling result of previous studies, because the range of RT is around 0-3000 ms, the relative RMS compared to the whole range of RT is still small.

findings which found that different areas (e.g., DLPFC, IPS, and ACC) play different functions in cognitive information processing stage, QN-MHP decomposes the cognitive information processing component into a network and each server performs different roles in the cognitive information processing (e.g., Server C is responsible for automatic response suppression; Server F plays a role in phonological judgment). This network structure allows QN-MHP to quantify the experimental results of Schumacher et al. (1999) and Hawkins et al. (1979) in the subadditive difficulty effect. Second, the serial information processing at Server F allows QN-MHP to model the basic PRP including its brain imaging patterns, the response grouping effect, and also account for the reason why it is extremely difficult for people to perform two complex mental operations (e.g., mathematical calculations) simultaneously without extensive practice. Third, the parallel information processing in the other servers in the cognitive subnetwork enables QN-MHP to model the disappearance of the PRP effect after extensive practice.

The overall mathematical structure of the Queueing network model and entity-based information processing are also unique properties of the Queueing network model. These properties enable QN-MHP to quantify the interactions among the servers naturally without adding additional assumptions, including: a) sojourn time in Equation 42 quantifies how routes (representing the neuron pathways) of entities are rewired by extensive practice, accounting for the reason why the PRP effect disappears naturally only via dual-task practice; b) the two streams of entities representing T1 and T2 through the network without devising a complex task-procedure to either interleave production rules into a serial program (ACT-R) or for an executive process to interactively control task processes (EPIC) (Liu, et al., in press), indicating that the Queueing network model at least offers an alternative way to quantify various PRP phenomena without using complex executive control or lock/unlock strategies in dual-task performance. Furthermore, these features of the Queueing network model are also consistent with the findings of neuroscience in dual tasks: a) information are coded in spike trains (represented by entities in the Queueing network model) which are processed in different brain regions related to the perceptual, cognitive and motor processes; b) based on a comprehensive review of more than ten fMRI studies in dual tasks, Collette & Linden (2002) concluded that “no specific area could be associated with any cognitive processor

for dual-task performance and that dual-task performance coordination depends mainly on the interaction between cerebral areas already activated in the single task" (Collette & Linden, 2002); "there is no conductor in the brain.... [A] collection of neural ensembles that depict a memory trace, for example, may be activated through an external stimulus that evokes the memory trace or through internally initiated activity that similarly engages the trace. Such a property of the Queueing network model avoids the difficult situation of postulating a 'neural conductor' or an 'executive' that controls the unfolding of cognition. The temporal and spatial unfolding of cognition results from the dynamic interactions among several areas of the brain (e.g. Hebb, 1949)" (McIntosh, 2000).

The modeling mechanisms of QN-MHP are also consistent with the experimental results in several electrophysiological and brain imagining studies. It should be noticed that it is relatively easy to model overt behavioral data only by using various kinds of strategies or production rules in the model; however, it is challenging to keep the modeling mechanisms also consistent with the findings in electrophysiological and brain imagining studies reflecting the inner information processing in the brain. In this modeling work, the Queueing network model not only successfully quantifies the brain imaging results in the basic PRP effect reflecting the cognitive process in PRP along its spatial dimension in the brain areas, but also models the stimulus-lateralized readiness potential onset time in the subadditive difficulty effect reflecting the cognitive process in PRP along the temporal dimension. The modeling of these experimental results in electrophysiological and brain imagining studies are consistent with the modeling mechanisms in quantifying the behavioral performance of PRP.

QN-MHP modeled all of the major counterexamples to RSB, EPIC-SRD and ACT-R/PM with equal or less number of free parameters and there is no task-specific assumptions in the current modeling work. Compared with other factors, the number of free parameters in the cognitive modeling plays the most important factor in determining the power of a model in predicting the experimental result without the purposely tailoring the model itself. Researchers in mathematical psychology found that the number of free parameters in a model is in inverse proportion to a model's prediction power (Busemeyer, 2000). Furthermore, QN-MHP does not need any task-specific assumptions to quantify these PRP phenomena while both EPIC and ACT-R/PM need task-specific assumptions

to model the PRP phenomena: EPIC has to rely on several complex lock/unlock scheduling strategies and these scheduling strategies are task specific—researchers have to design different lock/unlock strategies to interleave the control process of the two individual tasks in different dual-task situations; ACT-R/PM also has to use 4 additional task-specific assumptions similar to the lock/unlock strategies in EPIC (Byrne & Anderson, 2001, p854) to generate the subadditive difficulty effect in Experiment 4 of Schumacher et al. (1999). Moreover, in developing mathematical models of various PRP phenomena with the Queueing network model, the similarity between the expected patterns of dependent variables (e.g., reaction time) and the patterns of these dependent variables in the experimental results suggests that that without using free parameters, the model is able to predict the major pattern of the empirical data. In addition, the values of free parameters are also constrained by the properties of tasks (see the parameter setting section in modeling each experimental study).

The settings of the entities' routes in modeling each experimental study in this modeling work are not task specific assumptions. The routes of entities are selected based on a task-independent rule—the function of brain regions (represented by the servers) determines which regions (servers) are selected in the route and the anatomical connections among these regions determine the directions of spike trains (represented by entities) in the route. This task independent rule is consistent with the method to determine the information flow in different brain regions in neuroscience (Taylor et al., 2000). In all of the choice reaction time conditions, the routes of entities in the cognitive subnetwork are the same except Server A or B based on the content of stimuli (phonological or visual); and their routes only at the perceptual and the motor subnetwork are changed according to the particular task setting. In the simple reaction time and spatial working memory task conditions, the routes are determined by the findings in brain imagining studies. Even though fMRI techniques are not good at measuring the brain activity along the time dimension, ample research shows that they can be used to determine the route of information processing in the brain with the anatomic connection of neuron pathways among these regions (Horwitz et al., 1999; McIntosh, 1999, 2000; Sporns et al., 2000; Taylor, 2003; Taylor et al., 2000).

While QN-MHP demonstrates its unique features in modeling the various PRP phenomena, a concern about this modeling work is that if entities of T2 arrive at Server F earlier than entities of T1 in short SOA conditions, Server F will process entities of T2 first and the order of the responses might be reversed. In the experimental studies in this article, except the experiment related to the response grouping effect and Hawkins et al.'s experiment (1979), SOAs in all of the other experiments are positive and greater than 50 ms. Moreover, according to the well-established cognitive model—MHP which uses the same parameter (τ_p) with the same value to represent the visual and auditory perception time. These indicate that the chance that entities of T2 arrive at Server F earlier than the entities of T1 is relatively small. Moreover, the chance that R2 is earlier than R1 becomes even smaller since it requires two conditions meet at the same time (entities of T2 arrive at Server F earlier than the entities of T1; and the sum of entities of T1's processing times at the cognitive and the motor subnetwork are longer or equal to those of the entities of T2). Furthermore, even if R2 is made earlier than R1, it still consistent with empirical results in PRP that subjects sometimes did reverse their responses' order (Pashler, 1990). In addition, QN-MHP can also model the performance in the extreme experimental situations that SOA is a big negative number (e.g., SOA= -2000 ms) and the instruction to the subjects still require they make the R1 earlier than R2. In these extreme situations, the entities of T2 arrive at the Server F much earlier than entities of T1 and it is expected that adult subjects can still wait/hold T2 until they perceive the arrival of T1. In this case, QN-MHP can assign the proprieties to the two different types of entities corresponding to T1 and T2 and simply block the low-priority of entities entering the Server F and even the perceptual subnetwork. Researchers of EPIC would argue this is similar to their lock/unlock strategies; however, the assignment of proprieties to task stimuli in QN-MHP is much simpler than the complex lock/unlock strategies in EPIC in terms of the degree of complexity in cognitive information processing.

This modeling work offers at least an alternative way in quantifying various PRP phenomena and it also demonstrates the unique advantages of QN-MHP in modeling the basic dual-task performance—not only models dual-task performance in PRP with equal or less number of free parameters and no task-specific assumptions, but also is consistent

with the findings in brain imaging and electrophysiological measurements reflecting the cognitive mechanisms in generating the basic dual-task performance in PRP.

Reference

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An Integrated Theory of the Mind. *Psychological review, 111*(4), 1036-1060.

Anderson, J. R., & Lebiere, C. (1998). *The Atomic Components of Thought*: Lawrence Erlbaum Associates.

Anderson, J. R., Qin, Y. L., Sohn, M. H., Stenger, V. A., & Carter, C. S. (2003). An information-processing model of the BOLD response in symbol manipulation tasks. *Psychonomic Bulletin & Review, 10*(2), 241-261.

Anderson, J. R., Qin, Y. L., Stenger, V. A., & Carter, C. S. (2004). The relationship of three cortical regions to an information-processing model. *Journal of Cognitive Neuroscience, 16*(4), 637-653.

Arnell, K. M., Helion, A. M., Hurdelbrink, J. A., & Pasieka, B. (2004). Dissociating sources of dual-task interference using human electrophysiology. *Psychonomic Bulletin & Review, 11*(1), 77-83.

Atwal, G. S. (2004). Dynamic plasticity in coupled avian midbrain maps. *Physical Review E, 70*(6).

Baddeley, A. D. (1992). Working memory. *Science, 255*(5044), 556-559.

Banks, J., Carson, J. S., Nelson, B. L., & Nicol, D. M. (2004). *Discrete-Event System Simulation*: Prentice Hall.

Ben-Shachar, M., Hendler, T., Kahn, I., Ben-Bashat, D., & Grodzinsky, Y. (2003). The neural reality of syntactic transformations: Evidence from functional magnetic resonance imaging. *Psychological Science, 14*(5), 433-440.

Black, I. B. (1999). Trophic regulation of synaptic plasticity. *Journal of Neurobiology, 41*(1), 108-118.

Brass, M., & Cramon, D. (2002). The role of the frontal cortex in the task preparation. *Cerebral Cortex, 12*, 908-914.

Braus, D. F. (2004). Neurobiology of learning - The basis of an alteration process. *Psychiatrische Praxis, 31*, S215-S223.

Bullock, T. (1968). Representation of information in neurons and sites for molecular participation. *Proc Natl Acad Sci U S A, 60*(4), 1058-1068.

Burle, B., Vidal, F., Tandonnet, C., & Hasbroucq, T. (2004). Physiological evidence for response inhibition in choice reaction time tasks. *Brain and Cognition, 56*(2), 153-164.

Busemeyer, J. R. (2000). Model Comparisions and Model Selections Based on Generalizaztion Methodolgy. *Journal of Mathematical Psychology, 44*, 171-189.

Byrne, M. D. (2001). ACT-R/PM and menu selection: applying a cognitive architecture to HCI. *International Journal of Human-Computer Studies, 55*(1), 41-84.

Byrne, M. D., & Anderson, J. R. (1998). Perception and Action. In J. R. Anderson & C. Lebiere (Eds.), *The Atomic Components of Thought* (pp. 167-200): Lawrence Erlbaum Associates.

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

Card, S., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hinsdale, NJ: Lawrence Erlbaum.

Chklovskii, D. B., Mel, B. W., & Svoboda, K. (2004). Cortical rewiring and information storage. *Nature, 431*(7010), 782-788.

Cohen, M. S. (1997). Parametric analysis of fMRI data using linear systems methods. *Neuroimage, 6*(2), 93-103.

Collette, F., & Linden, M. V. (2002). Brain Imaging of the Central Executive Component of Working Memory. *Neuroscience and Biobehavioral Review, 26*, 105-125.

Cook, A. S., & Woollacott, M. H. (1995). *Motor Control: Theory and Practical Applications*: Williams& Wilkins.

Creamer, L. R. (1963). Event Uncertainty, Psychological Refractory Period, and Human Data-Processing. *Journal of Experimental Psychology, 66*(2), 187-203.

Dejong, R. (1993). Multiple Bottlenecks in Overlapping Task-Performance. *Journal of Experimental Psychology-Human Perception and Performance, 19*(5), 965-980.

Dejong, R. (1995a). Perception-Action Coupling and S-R Compatibility. *Acta Psychologica, 90*(1-3), 287-299.

Dejong, R. (1995b). The Role of Preparation in Overlapping-Task Performance. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 48*(1), 2-25.

Dejong, R. (1995c). Strategical Determinants of Compatibility Effects with Task Uncertainty. *Acta Psychologica, 88*(3), 187-207.

Dejong, R., Coles, M. G. H., Logan, G. D., & Gratton, G. (1988). The Control of Response Processes in Speeded Choice Reaction Performance. *Psychophysiology, 25*(4), 442-442.
Dejong, R., Liang, C. C., & Lauber, E. (1994). Conditional and Unconditional Automaticity - a Dual-Process Model of Effects of Spatial Stimulus - Response Correspondence. *Journal of Experimental Psychology-Human Perception and Performance, 20*(4), 731-750.
Dejong, R., & Sweet, J. B. (1994). Preparatory Strategies in Overlapping-Task Performance. *Perception & Psychophysics, 55*(2), 142-151.
Eagleman, D., Jacobson, J., & Sejnowski, T. (2004). Perceived luminance depends on temporal context. *Nature, 428*(6985), 854-856.
Eisler, H. (1975). Subjective duration and psychophysics. *Psychological Review, 82*, 429-450.
Eisler, H. (1976). Experiments on subjective duration1968-1975 : a collection of power function exponents. *Psychological Bulletin, 83*, 1154-1171.
Faw, B. (2003). Pre-frontal executive committee for perception, working memory, attention, long-term memory, motor control, and thinking: A tutorial review. *Consciousness and Cognition, 12*(1), 83-139.
Feyen, R. (2002). *Modeling Human Performance using the Queueing Network ---Model Human Processor (QN-MHP)*. University of Michigan, Ann Arbor.
Fletcher, P. C., & Henson, R. N. A. (2001). Frontal lobes and human memory - Insights from functional neuroimaging. *Brain, 124*, 849-881.
Gilbert, P. F. C. (2001). An outline of brain function. *Cognitive Brain Research, 12*(1), 61-74.
Gordon, A. M., & Soechting, J. F. (1995). Use of tactile afferent information in sequential finger movements. *Experimental Brain Research, 107*, 281-292.
Grossman, M., Smith, E. E., Koenig, P., Glosser, G., DeVita, C., Moore, P., et al. (2002). The neural basis for categorization in semantic memory. *Neuroimage, 17*(3), 1549-1561.
Habib, M. (2003). Rewiring the dyslexic brain. *Trends in Cognitive Sciences, 7*(8), 330-333.
Harter, M. R. (1967). Excitability cycles and cortical scanning: a review of two hypotheses of central intermittency in perception. *Psychological Bulletin, 68*(1), 47-58.
Hawkins, H. L., Rodriguez, E., & Reicher, G. M. (1979). *Is time-sharing a general ability?* (No. Rep. No.3). Eugene: University of Oregon.
Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review, 7*(2), 185-207.
Herath, P., Klingberg, T., Young, J., Amunts, K., & Roland, P. (2001). Neural correlates of dual task interference can be dissociated from those of divided attention: an fMRI study. *Cerebral Cortex, 11*(9), 796-805.
Hornof, A. J. (1999). *Computational models of the perceptual, cognitive, and motor processes involved in the visual search of pull-down menus and computer screens.*, University of Michigan.
Horwitz, B., Tagamets, M. A., & McIntosh, A. R. (1999). Neural modeling, functional brain imaging, and cognition. *Trends in Cognitive Sciences, 3*(3), 91-98.
Ilvonen, T. M., Kujala, T., Kiesilainen, A., Salonen, O., Kozou, H., Pekkonen, E., et al. (2003). Auditory discrimination after left-hemisphere stroke - A mismatch negativity follow-up study. *Stroke, 34*(7), 1746-1751.
Jiang, Y. H. (2004). Resolving dual-task interference: an fMRI study. *Neuroimage, 22*(2), 748-754.
Jiang, Y. H., & Kanwisher, N. (2003). Common neural mechanisms for response selection and perceptual processing. *Journal of Cognitive Neuroscience, 15*(8), 1095-1110.
Jiang, Y. H., Saxe, R., & Kanwisher, N. (2004). Functional magnetic resonance imaging provides new constraints on theories of the psychological refractory period. *Psychological Science, 15*(6), 390-396.
John, B. E. (1996). TYPIST: A theory of performance in skilled typing. *Human-Computer Interaction, 11*(4), 321-355.
Jurgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews, 26*(2), 235-258.
Kansaku, K., Hanakawa, T., Wu, T., & Hallett, M. (2004). A shared neural network for simple reaction time. *Neuroimage, 22*(2), 904-911.
Kantowitz, B. (1974). Double Stimulation. In B. Kantowitz (Ed.), *Human Information Processing* (pp. 83-131).

Karlin, L., & Kestenbaum, R. (1968). Effects of Number of Alternatives on Psychological Refractory Period. *Quarterly Journal of Experimental Psychology, 20*, 167-178.
Kaufer, D. I., & Lewis, D. A. (1999). *Frontal Lobe Anatomy and Cortical Connectivity. In The Human Frontal Lobes: Functions and Disorders.* New York: The Gullford Press.
Kieras, D. E., & 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.
Lim, J., & Liu, Y. (2004a). *A queueing network model for eye movement.* Paper presented at the Proceedings of the 2004 International Conference on Cognitive Modeling, Pittsburg, PA.
Lim, J., & Liu, Y. (2004b). *A queueing network model of menu selection and visual search.* Paper presented at the Proceedings of the 47 Annual Conference of the Human Factors and Ergonomics Society.
Liu, Y. (1996). Queueing network modeling of elementary mental processes. *Psychological Review, 103*(1), 116-136.
Liu, Y. (1997). Queueing network modeling of human performance of concurrent spatial and verbal tasks. *IEEE Transactions on Systems Man and Cybernetics Part a-Systems and Humans, 27*(2), 195-207.
Liu, Y., Feyen, R., & Tsimhoni, O. (2004). *Queueing Network-Model Human Processor (QN-MHP): A Computational Architecture for Multitask Performance* (No. Tech Report 04-05). Ann Arbor, MI, USA: Department of Industrial & Operations Engineering, University of Michigan.
Liu, Y., Feyen, R., & Tsimhoni, O. (2005). Queueing Network-Model Human Processor (QN-MHP): A Computational Architecture for Multi-Task Performance in Human-Machine Systems. *ACM Transaction on Human Computer Interaction, In Press*.
Liu, Y. L. (1998). Queueing network modeling of reaction time and mental architecture. *Journal of Mathematical Psychology, 42*(4), 485-485.
Logothetis, N. K., & Pfeuffer, J. (2004). On the nature of the BOLD fMRI contrast mechanism. *22*(10), 1517-1531.
Luce, R. D. (1986). *Response times*. New York: Oxford University Press.
Manoach, D. S., Schlaug, G., Siewert, B., Darby, D. G., Bly, B. M., Benfield, A., et al. (1997). Prefrontal cortex fMRI signal changes are correlated with working memory load. *Neuroreport, 8*(2), 545-549.
Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences, 9*(6), 296-305.
McClelland, J. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review, 86*, 287-330.
McIntosh, A. R. (1999). Mapping cognition to the brain through neural interactions. *Memory, 7*(5-6), 523-548.
McIntosh, A. R. (2000). Towards a network theory of cognition. *Neural Networks, 13*(8-9), 861-870.
Meyer, D. E., Glass, J. M., Mueller, S. T., Seymour, T. L., & Kieras, D. E. (2001). Executive-process interactive control: A unified computational theory for answering 20 questions (and more) about cognitive ageing. *European Journal of Cognitive Psychology, 13*(1-2), 123-164.
Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance .1. Basic mechanisms. *Psychological Review, 104*(1), 3-65.
Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance .2. Accounts of psychological refractory-period phenomena. *Psychological Review, 104*(4), 749-791.
Meyer, D. E., & Kieras, D. E. (1999). Precis to a practical unified theory of cognition and action: Some lessons from EPIC computational models of human multiple-task performance. In *Attention and Performance Xvii* (Vol. 17, pp. 17-88).
Mitz, A. R., Godschalk, M., & Wise, S. P. (1991). Learning-Dependent Neuronal-Activity in the Premotor Cortex - Activity During the Acquisition of Conditional Motor Associations. *Journal of Neuroscience, 11*(6), 1855-1872.
Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature, 431*(7010), 760-767.
Mustovic, H., Scheffler, K., Di Salle, F., Esposito, F., Neuhoff, J. G., Hennig, J., et al. (2003). Temporal integration of sequential auditory events: silent period in sound pattern activates human planum temporale. *Neuroimage, 20*(1), 429-434.
Naatanen, R. (1971). Non-aging foreperiods and simple reaction time. *Acta Psychologica, 35*, 316-327.

Nakahara, H., Doya, K., & Hikosaka, O. (2001). Parallel Cortico-Basal Ganglia Mechanisms for Acquisition and Execution of Visuomotor Sequences-A Computational Approach. *Journal of Cognitive Neuroscience, 13*(5), 626-647.
Nickerson, R. S. (1965). Response time to the second of two successive signals as a function of absolute and relative duration of intersignal interval. *Perceptual and Motor Skills, 21*, 3-10.
Nickerson, R. S. (1967). Expectancy, waiting time and the psychological refractory period. *Acta Psychologica,, 27*, 23-39.
Niemi, P., & Naatanen, R. (1981). Foreperiod And Simple Reaction-Time. *89*(1), 133-162.
Nilsen, E. L. (1991). *Perceptual-motor control in human-computer integration.* University of Michigan.
Oberauer, K., & Kliegl, R. (2004). Simultaneous cognitive operations in working memory after dual-task practice. *Journal of Experimental Psychology-Human Perception and Performance, 30*(4), 689-707.
Ohbayashi, M., Ohki, K., & Miyashita, Y. (2003). Conversion of working memory to motor sequence in the monkey premotor cortex. *Science, 301*(5630), 233-236.
Pashler, H. (1984a). Processing Stages in Overlapping Tasks - Evidence for a Central Bottleneck. *Journal of Experimental Psychology-Human Perception and Performance, 10*(3), 358-377.
Pashler, H. (1984b). Processing stages in overlapping tasks: evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception & Performance, 18*, 217-232.
Pashler, H. (1989). Chronometric Evidence for Central Postponement in Temporally Overlapping Tasks. *Quarterly Journal of Experimental Psychology; Series A: Human Experimental Psychology, 41*(1), 19.
Pashler, H. (1990). Do response modality effects support multiprocessor models of divided attention? *Journal of Experimental Psychology: Human Perception & Performance, 16*(4), 826-842.
Pashler, H. (1994a). Divided Attention - Storing and Classifying Briefly Presented Objects. *Psychonomic Bulletin & Review, 1*(1), 115-118.
Pashler, H. (1994b). Dual-Task Interference in Simple Tasks - Data and Theory. *Psychological Bulletin, 116*(2), 220-244.
Pashler, H. (1994c). Overlapping Mental Operations in Serial Performance with Preview. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 47*(1), 161-191.
Pashler, H., Carrier, M., & Hoffman, J. (1993). Saccadic Eye-Movements and Dual-Task Interference. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 46*(1), 51-82.
Pashler, H., Johnston, J. C., & Ruthruff, E. (2000). Attention and performance. *Annual Review of Psychology, 52*, 629-651.
Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology, 52*, 629-651.
Pashler, H., Luck, S. J., Hillyard, S. A., Mangun, G. R., Obrien, S., & Gazzaniga, M. S. (1994). Sequential Operation of Disconnected Cerebral Hemispheres in Split-Brain Patients. *Neuroreport, 5*(17), 2381-2384.
Roland, P. E. (1993). *Brain activation*. New York: Wiley-Liss.
Rolls, E. T. (1994). Neurophysiology and Cognitive Functions of the Striatum. *Revue Neurologique, 150*(8-9), 648-660.
Rolls, E. T. (2000). Memory systems in the brain. *Annual Review of Psychology, 51*, 599-630.
Rouse, W. B. (1980). *Systems Engineering Models of Human-Machine Interaction*. New York: North Holland.
Ruthruff, E., Johnston, J. C., & Van Selst, M. (2001). Why practice reduces dual-task interference. *Journal of Experimental Psychology-Human Perception and Performance, 27*(1), 3-21.
Ruthruff, E., Johnston, J. C., Van Selst, M., Whitsell, S., & Remington, R. (2003). Vanishing dual-task interference after practice: Has the bottleneck been eliminated or is it merely latent? *Journal of Experimental Psychology-Human Perception and Performance, 29*(2), 280-289.
Ruthruff, E., Miller, J., & Lachmann, T. (1995). Does Mental Rotation Require Central Mechanisms. *Journal of Experimental Psychology-Human Perception and Performance, 21*(3), 552-570.
Ruthruff, E., Pashler, H. E., & Hazeltine, E. (2003). Dual-task interference with equal task emphasis: Graded capacity sharing or central postponement? *Perception & Psychophysics, 65*(5), 801-816.
Ruthruff, E., Pashler, H. E., & Klaassen, A. (2001). Processing bottlenecks in dual-task performance: Structural limitation or strategic postponement? *Psychonomic Bulletin & Review, 8*(1), 73-80.

Sadato, N., Ibanez, V., Campbell, G., Deiber, M. P., LeBihan, D., & Hallett, M. (1997). Frequency-dependent changes of regional cerebral blood flow during finger movements: Functional MRI compared to PET. *Journal of Cerebral Blood Flow and Metabolism, 17*(6), 670-679.
Sadato, N., Ibanez, V., Deiber, M. P., Campbell, G., Leonardo, M., & Hallett, M. (1996). Frequency-dependent changes of regional cerebral blood flow during finger movements. *Journal of Cerebral Blood Flow and Metabolism, 16*(1), 23-33.
Salthouse, T. A. (1986). Perceptual, Cognitive, and Motoric Aspects of Transcription Typing. *Psychological Bulletin, 99*(3), 303-319.
Salvucci, D. D. (2002). *Modeling driver distraction from cognitive tasks.* Paper presented at the Proceedings of the 24th Annual Conference of the Cognitive Science Society.
Salvucci, D. D., & Macuga, K. L. (2001). *Predicting the effects of cell-phone dialing on driver performance.* Paper presented at the Proceedings of the Fourth International Conference on Cognitive Modeling, Mahwah, NJ.
Schubert, T. (1996). The analysis of dual-task interference effects. *Zeitschrift Fur Experimentelle Psychologie, 43*(4), 625-656.
Schubert, T., & Szameitat, A. J. (2003). Functional neuroanatomy of interference in overlapping dual tasks: an fMRI study. *Cognitive Brain Research, 17*(3), 733-746.
Schubotz, R. I., Friederici, A. D., & von Cramon, D. Y. (2000). Time perception and motor timing: A common cortical and subcortical basis revealed by fMRI. *Neuroimage, 11*(1), 1-12.
Schubotz, R. I., & von Cramon, D. Y. (2001). Functional organization of the lateral premotor cortex: fMRI reveals different regions activated by anticipation of object properties, location and speed. *Cognitive Brain Research, 11*(1), 97-112.
Schuch, S., & Koch, I. (2004). The costs of changing the representation of action: Response repetition and response-response compatibility in dual tasks. *Journal of Experimental Psychology-Human Perception and Performance, 30*(3), 566-582.
Schumacher, E. H., Elston, P. A., & D'Esposito, M. (2003). Neural evidence for representation-specific response selection. *Journal of Cognitive Neuroscience, 15*(8), 1111-1121.
Schumacher, E. H., & Jiang, Y. H. (2003). Neural mechanisms for response selection: Representation specific or modality independent? *Journal of Cognitive Neuroscience, 15*(8), 1077-1079.
Schumacher, E. H., Lauber, E. J., Glass, J. M., Zurbriggen, E. I., Gmeindl, L., Kieras, D. E., et al. (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., et al. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological Science, 12*(2), 101-108.
Schweikert, R. (1978). A critical path generalization of the additive
factor methods: Analysis of a Stroop task. *Journal of Mathematical Psychology*(18), 105-139.
Shafritz, K. M., Kartheiser, P., & Belger, A. (2005). Dissociation of neural systems mediating shifts in behavioral response and cognitive set. *Neuroimage, 25*(2), 600-606.
Simon, S. R., Meunier, M., Piettre, L., Berardi, A. M., Segebarth, C. M., & Boussaoud, D. (2002). Spatial attention and memory versus motor preparation: Premotor cortex involvement as revealed by fMRI. *Journal of Neurophysiology, 88*(4), 2047-2057.
Slaterhammel, A. T. (1958). Psychological Refractory Period in Simple Paired Responses. *Research Quarterly, 29*(4), 468-481.
Smith, E. E. (1972). *Where is the bottleneck in information processing?* Paper presented at the American Psychological Association, Honolulu.
Smith, E. E., & Jonides, J. (1998). Neuroimaging analyses of human working memory. *Proc. Natl. Acad. Sci. USA, 95*, 12061-12068.
Smith, G. A. (1978). Comparison between Models of Psychological Refractory Period Proposed by Welford and Surwillo and Titus. *Developmental Psychobiology, 11*(2), 177-182.
Smith, M. C. (1967a). Psychological Refractory Period as a Function of Performance of a First Response. *Quarterly Journal of Experimental Psychology, 19*, 350-&.
Smith, M. C. (1967b). Theories of Psychological Refractory Period. *Psychological Bulletin, 67*(3), 202-280.
Smith, M. C. (1969). Effect of Varying Channel Capacity on Stimulus Detection and Discrimination. *Journal of Experimental Psychology, 82*(3).

Smith, M. C. (1969). Effect of Varying Information on Psychological Refractory Period. *Acta Psychologica, 30*, 220-280.
Solomons, L. M., & Stein, G. (1896). Normal motor automatism. *Psychologocial Review, 3*, 492-512.
Sommer, W., Leuthold, H., Abdel-Rahman, R., & Pfutze, E. M. (1997). Localisation of the grouping effect in overlapping tasks. *Zeitschrift Fur Experimentelle Psychologie, 44*(1), 103-116.
Sommer, W., Leuthold, H., & Schubert, T. (2001). Multiple bottlenecks in information processing? An electrophysiological examination. *Psychonomic Bulletin & Review, 8*(1), 81-88.
Sporns, O., Tononi, G., & Edelman, G. M. (2000). Connectivity and complexity: the relationship between neuroanatomy and brain dynamics. *Neural Networks, 13*(8-9), 909-922.
Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction.* Cambridge, Massachusetts: MIT Press.
Szameitat, A. J., Schubert, T., Muller, K., & von Cramon, D. Y. (2002). Localization of executive functions in dual-task performance with fMRI. *Journal of Cognitive Neuroscience, 14*(8), 1184-1199.
Tanaka, S., Honda, M., & Sadato, N. (2005). Modality-specific cognitive function of medial and lateral human Brodmann area 6. *Journal of Neuroscience, 25*(2), 496-501.
Taylor, J. (2003). Paying attention to consciousness. *Progress in Neurobiology, 71*(4), 305-335.
Taylor, J., Horwitzc, B., Shaha, N. J., Fellenzb, W. A., Mueller-Gaertnera, H.-W., & Krausee, J. B. (2000). Decomposing memory: functional assignments and brain traffic in paired word associate learning. *Neural Networks, 13*, 923-940.
Ungerleider, L., & Karni, A. (2002). maging Brain Plasticity during Motor Skill Learning. *Neurobiology of Learning and Memory, 78*, 553-564.
Van Selst, M., & Jolicoeur, P. (1997). Decision and response in dual-task interference. *Cognitive Psychology, 33*(3), 266-307.
Van Selst, M., Ruthruff, E., & Johnston, J. C. (1999). Can practice eliminate the psychological refractory period effect? *Journal of Experimental Psychology-Human Perception and Performance, 25*(5), 1268-1283.
Vartanian, O., & Goel, V. (2005). Task constraints modulate activation in right ventral lateral prefrontal cortex. *Neuroimage, 27*(4), 927-933.
Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998a). Why "Sounds Are Judged Longer Than Lights": Application of a Model of the Internal Clock in Humans. *The Quarterly Journal of Experimental Psychology, 51B*(2), 97-120.
Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998b). Why "sounds are judged longer than lights": Application of a model of the internal clock in humans. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology, 51*(2), 97-120.
Welch, J. (1898). On the measurement of mental activity through muscular activity and the determintation of a constant of attention. *American Journal of Physiology, 1*, 288-306.
Welford, A. T. (1952). The "psychological refractory period" and timing of high speed performance: A review and a theory. *British Journal of Psychology, 43*, 2-19.
Wildgruber, D., Ackermann, H., & Grodd, W. (2001). Differential contributions of motor cortex, basal ganglia, and cerebellum to speech motor control: Effects of syllable repetition rate evaluated by fMRI. *Neuroimage, 13*(1), 101-109.
Wu, C., & Liu, Y. (2004a). *Modeling Behavioral and Brain Imaging Phenomena in Transcription Typing with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA. 314-319.
Wu, C., & Liu, Y. (2004b). *Modeling human transcription typing with Queueing network-model human processor.* Paper presented at the Proceedings of the 48th Annual Meeting of Human Factors and Ergonomics Society, New Orleans, Louisiana, USA. 381-385.
Wu, C., & Liu, Y. (2004c). *Modeling Psychological Refractory Period (PRP) and Practice Effect on PRP with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA. 320-325.
Zysset, S., Muller, K., Lehmann, C., Thone-Otto, A. I. T., & von Cramon, D. Y. (2001). Retrieval of long and short lists from long term memory: a functional magnetic resonance imaging study with human subjects. *Neuroscience Letters, 314*(1-2), 1-4.

Appendix

1. Mathematical Modeling of Expected RT2 (Simple Reaction Task) in the Experiments of Karlin & Kestenbaum (1968) and Sommer et al. (2001)

Since Task 2 in Karlin & Kestenbaum (1968) and Sommer et al. (2001) are simple reaction tasks, their mathematical modeling methods can be described in the following section.

There are two conditions in modeling the expected RT2 at the simple reaction condition. At short SOA conditions (entities of T2 arrive at Server F before Server F starts its anticipation process, t_a=0, see Figure 3-30), entities of T2 have to wait until entities of T1 leave Server F; after entities of T1 leave Server F, entities of T2 enter Server F immediately. Since Server F is occupied by the entities of T2 (subjects are busy in performing judgment of T2), the anticipation process does not occur in this condition. At long SOA conditions, Server F starts its anticipation process before entities of T2 arrive at Server F. The mathematical models of RT2 (simple reaction condition) are constructed based on these two conditions in the following sections.

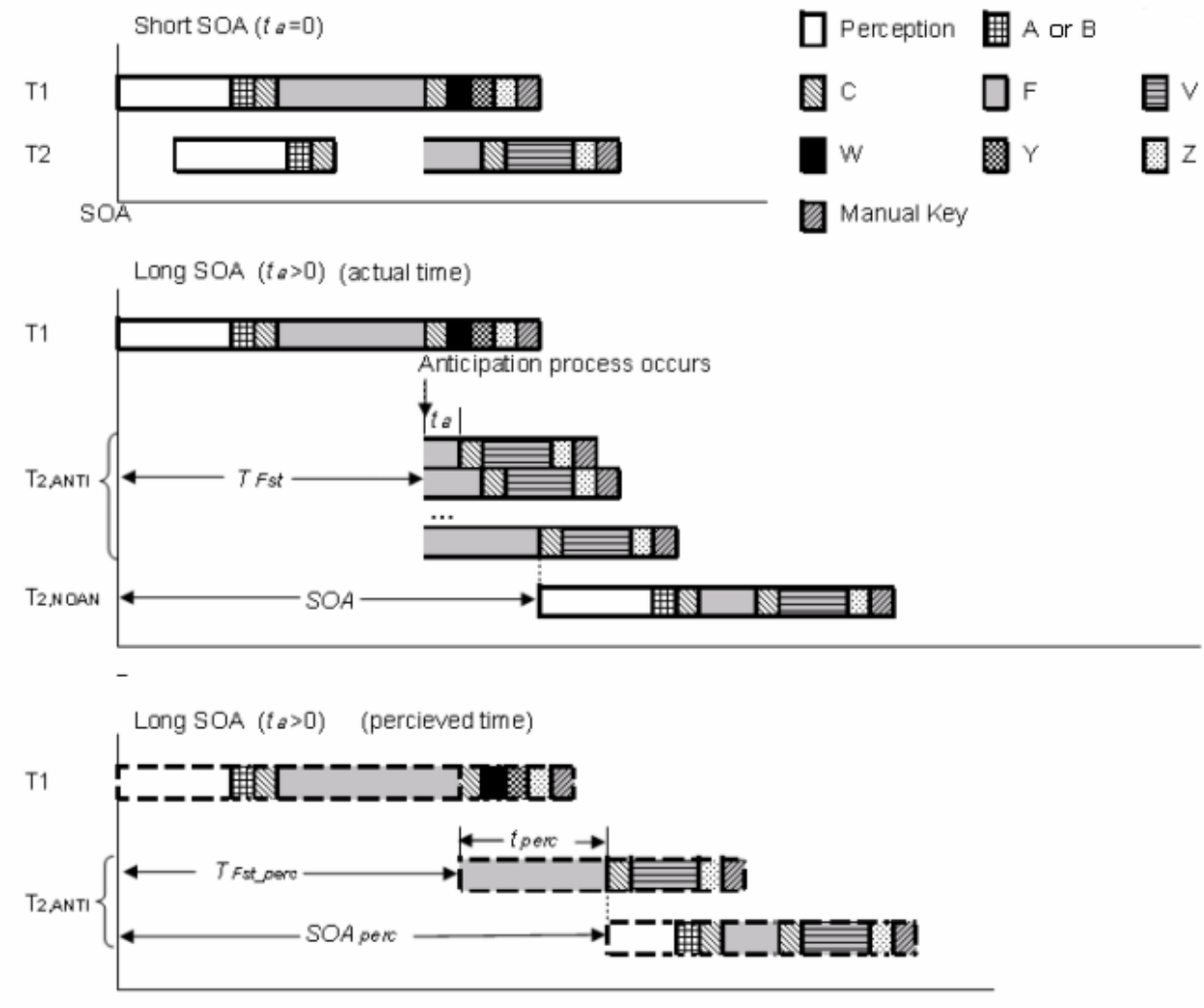


Figure 3-30 Modeling mechanisms of the expected RT2 under the simple reaction condition. t_a is the duration between when Server F starts the anticipation process and when entities of S2 arrives at the perceptual subnetwork; T_{Fst} is the time point when Server F starts its anticipation process $T_{Fst}= T_{1,AP\,|VP}+T_{1,\,B|A} +T_{1,C} +T_{1,F,}$ where $T_{1,AP\,|VP}$ is the processing time at the auditory perceptual subnetwork (Karlin & Kestenbaum's experiment) or visual perceptual subnetwork (Sommer et al.'s experiment); $T_{1,\,B|A}$ is the processing time at Server B (Karlin & Kestenbaum's experiment) or Server A (Sommer et al.'s experiment).

1.1 Short SOA Conditions ($t_a=0$)

Under short SOA conditions of T2 (simple reaction condition), the expected RT2 is also modeled with the same form of the equations as in the choice reaction condition of RT2 except the motor subnetwork's servers are replaced by the servers involved in the simple reaction time (see Equation 53 and Figure 3-30).

$E(RT_2|t_a = 0)$ (simple reaction condition)= $max(T_{1,\,AP|VP} + T_{1,B|A} + T_{1,C} + T_{1,F} - SOA,$

$T_{2,VP|VP} + T_{1,A|B} + T_{2,C}) + T_{2,F} + T_{2,C} + T_{2,V} + T_{2,Z} + T_{2,K}$ (53)

where $T_{1,AP|VP}$ is the processing time at the auditory perceptual subnetwork (Karlin & Kestenbaum's experiment) or visual perceptual subnetwork (Sommer et al.'s experiment); $T_{1,\,B|A}$ is the processing time at Server B (Karlin & Kestenbaum's experiment) or Server A (Sommer et al.'s experiment).

1.2 Long SOA conditions ($t_a > 0$)

Quantification of the Anticipation Process

The anticipation process (R2 is made without seeing S2) at Server F is quantified by the following mechanisms in time perception. According to the function of Server F in anticipating a sensory event in a simple reaction task (Schubotz, et al. 2000, 2001, see the introduction section of this article), the longer Server F anticipates S2 (defined as perceived waiting time, t_{perc}), the higher the probability (defined as p) of trigging motor response without seeing S2 (anticipation process) (see Equation 54).

$$p = \frac{t_{perc}}{T_{perc}} \tag{54}$$

where T_{perc} is the duration between when the anticipation process starts and when the probability that subjects make the motor response equal to 1.

Based on psychophysical research in studying the relationship between perceived waiting time (t_{perc}) and actual waiting time (t_a) in very short time periods, there has been considerable support for a psychophysical law for perceptual duration described by a power function following the Steven's power law (Eisler, 1975, 1976). Thus,

$$t_{perc} = k t_a^{\beta} \tag{55}$$

where t_a is the duration between when Server F starts the anticipation process and when S2 arrives at the perceptual subnetwork. k and β are the parameters in Steven's power law (Wearden, Edwards, Fakhri, & Percival, 1998). Combining Equation 54 and 55, result in:

$$p = \frac{k t_a^{\beta}}{T_{perc}} \tag{56}$$

We can also get:

$$t_a = \left(\frac{T_{perc}}{k}\, p\right)^{\frac{1}{\beta}} \tag{57}$$

Moreover, since p is defined as the probability that the response of T2 is made with the anticipation process (R2 is made without seeing S2), there are two conditions in which expected RT2 is modeled: RT2 with or without the anticipation process.

Quantification of Expected RT2 with the Anticipation Process

Based on Figure 3-30 we can develop the expected RT2 with the anticipation process ($E(RT_{2,ANTI})$) into:

$$E(RT_{2,ANTI}) = T_{Fst} + t_a + T_{2,C} + T_{2,V} + T_{2,Z} + T_{2,K} - SOA \tag{58}$$

where T_{Fst} is the time point when Server F starts its anticipation process (see Equation 59 and Figure 3-30).

$$T_{Fst} = T_{1,AP\,|\,VP} + T_{1,B\,|\,A} + T_{1,C} + T_{1,F,} \tag{59}$$

From Equation 55, t_a can be rewritten into:

$$t_a = (t_{perc} / k)^{1/\beta} \tag{60}$$

Moreover, since subjects end their waiting process of S2 when they perceive the time reaches the perceived SOA, the perceived waiting time (t_{perc}) equals the perceived SOA (SOA_{perc}) minus the perceived T_{Fst} (T_{Fst_perc}), i.e.:

$$t_{perc} = \max(SOA_{perc} - T_{Fst\_perc}, 0) \tag{61}$$

where SOA_{perc} and T_{Fst_perc} can be derived from Equation 55, i.e., $SOA_{perc}=kSOA^{\beta}$, and $T_{Fst\_perc}=k\,T_{Fst}^{\beta}$, resulting in:

$$t_{perc} = \max\left(kSOA^{\beta} - k\,T_{Fst}^{\beta}, 0\right) \tag{62}$$

Combining Equation 56, 60, and 62, results in:

$$t_a = \max[(SOA^{\beta} - T_{Fst}{}^{\beta})^{1/\beta}, 0] \quad (63)$$

$$RT_{2,ANTI} = T_{Fst} + \max[(SOA^{\beta} - T_{Fst}{}^{\beta})^{1/\beta}, 0] + T_{2,C} + T_{2,V} + T_{2,Z} + T_{2,K} - SOA \quad (64)$$

Quantification of Expected RT2 without the Anticipation Process

Under the condition that there is no anticipation, the expected RT2 ($E(RT_{2,NOAN})$) is modeled with the same form of the equation as in the choice reaction condition except the motor subnetwork's servers are replaced by the servers involved in the simple reaction task (see Equation 65 and Figure 3-30).

$$E(RT_{2,NOAN}) \text{ (simple reaction condition)} = \max(T_{1,AP|VP} + T_{1,B|A} + T_{1,C} + T_{1,F} - SOA, \quad (65)$$
$$T_{2,VP} + T_{1,A} + T_{2,C}) + T_{2,F} + T_{2,C} + T_{2,V} + T_{2,Z} + T_{2,K}$$

Hence, the expected RT2 in long SOA conditions ($t_a>0$) can be quantified by Equation 66:

$$E(RT_2 | t_a > 0) \text{ (simple reaction condition)} = pE(RT_{2,ANTI}) + (1-p)E(RT_{2,NOAN}) \quad (66)$$

In sum, the expected RT2 under the simple reaction condition is:

$$E(RT2)_{\text{(simple reaction)}} = \begin{cases} pE(RT_{2,ANTI}) + (1-p)E(RT_{2,NOAN}) & t_a>0 \text{ (long SOA conditions)} \\ \max(T_{1,AP|VP} + T_{1,B|A} + T_{1,C} + T_{1,F} - SOA,\ T_{2,VP|AP} + T_{1,A|B} + T_{2,C}) & \\ + T_{2,F} + T_{2,C} + T_{2,V} + T_{2,Z} + T_{2,K} & t_a=0 \text{ (short SOA conditions)} \end{cases} \quad (67)$$

Combing Equation 56, 63, 65, and 66, Equation 67 can be rewritten into:

$$E(RT2)=\text{(simple reaction)}\begin{cases}\left(\frac{kSOA^{\beta}-kT_{Fst}{}^{\beta}}{T_{perc}}\right)[T_{Fst}-T_{Fst}{}^{\beta}-SOA+SOA^{\beta}-(T_{2,VP|AP}+T_{,AB}+2T_{,C}+T_{2,F}+T_{2,V}+T_{2,Z}+T_{2,K})] \\ +(T_{2,VP|AP}+T_{,AB}+2T_{,C}+T_{2,F}+T_{2,V}+T_{2,Z}+T_{2,K}) \\ t_a>0 \ \text{(long SOA conditions)} \\ max(T_{1,AP|VP}+T_{1,B|A}+T_{1,C}+T_{1,F}-SOA,\ T_{2,VP|AP}+T_{1,A|B}+T_{2,C}) \\ +T_{2,F}+T_{2,C}+T_{2,V}+T_{2,Z}+T_{2,K} \qquad t_a=0 \text{ (short SOA conditions)}\end{cases} \tag{68}$$

2. Mathematical Modeling of the Percentage of Negative Responses of RT2 (P_n)

Negative response of RT2 (RT2<0) means that the R2 occurs prior to the arrival of S2 (the arrival of S2 is SOA). Based on Figure 3-30, the interval between the arrival of S1 (time=0) and the occurrence of R2 is $T_{Fst}+t_a+T_{2,C}+T_{2,V}+T_{2,Z}+T_{2,K}$.

Supposing $u=T_{2,C}+T_{2,V}+T_{2,Z}+T_{2,K}$, result in:

$$P_n = P\{RT_2 < 0\} = P\{T_{Fst} + t_a + u < SOA\} = P\{t_a < SOA - u - T_{Fst}\} \tag{69}$$

Since t_a ranges from 0 to SOA- T_{FSt} (t_a ends when S2 arrives according to its definition), the probability of the RT2<0 (P_n) is:

$$P_n = \begin{cases}\int_0^{SOA-u-T_{Fst}} \frac{1}{SOA-T_{Fst}}\,dt_a & SOA \geq u + T_{Fst} \\ 0 & SOA < u + T_{Fst}\end{cases} \tag{70}$$

Solve the integral, and then result in:

$$P_n = \begin{cases} 1 - \dfrac{T_{2,C} + T_{2,V} + T_{2,Z} + T_{2,K}}{SOA - T_{Fst}} & SOA \geq u + T_{Fst} \\ 0 & SOA < u + T_{Fst} \end{cases} \quad (71)$$

3. Derivation of Routing Probability Equation (Equation 42)

Equation 42 is derived based on Black's model (1999) as well as other neuroscience findings. Black (1999) proposed a model to explain the role of BDNF (brain derived neurotrophic factor) in its regulation of synaptic plasticity in adults: BDNF increases the activity of NMDA (N-methyl-D-aspartate) receptors, increases neuron channel open probability by increasing opening frequency, and then increases the velocity of spike trains travel (V) through these neuron channels (Black, 1999). Hence, the stronger synaptic connection strength (the amount of presynaptic transmitter released and the degree of postsynaptic responsiveness) of an individual route, the greater the probability (P_i) that spike trains (represented by entities) travel through that route (Black, 1999; Braus, 2004; Chklovskii, et al., 2004; Habib, 2003) (see Equation 72 and Figure 3-31).

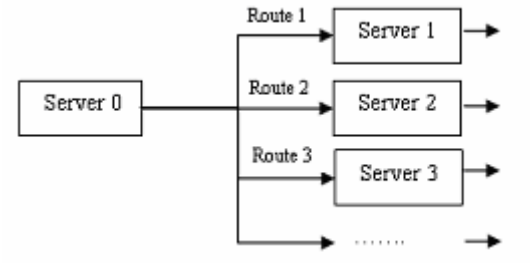


Figure 3-31Multiple routes for one location in the Queueing network (Server 0 has *U* multiple routes as output)

$$P_i = \frac{ST_i}{\sum_{j=1}^{U} ST_j} \quad (72)$$

where the numerator (ST_i) stands for the standardized synaptic connection strength of route i ($ST_i \in [0, 1]$). The denominator represents the sum of the standardized synaptic connection strength of all the multiple routes starting from the original brain region (Server 0 in Figure 3-31). Moreover, the standardized synaptic connection strength of route *i* (ST_i) is in direct proportion to the standardized velocity (V_i) that the spike trains travel through that route (Black, 1999; Bullock, 1968; Chklovskii, et al., 2004) (see Equation 73).

$$ST_i = r_0 V_i \tag{73}$$

where r_0 is a parameter stands for the ratio between ST_i and V_i.

Since the Queueing network is able to capture several properties of information processing in the human brain—spike trains carrying information (represented by entities) travel through different brain regions and form a "brain traffic" including the possible waiting of the previous information flow to be processed (see the first assumption in modeling the practice effect), the travel time of the spike trains (represented by entities) in route *i* is composed of both waiting and processing time and therefore this travel time can be regarded as the sum of waiting time (W_i) and processing time (T_i) of entities, i.e., sojourn time (S_i) in that route. Furthermore, this sojourn time or travel time (sum of waiting and processing time) is in inverse proportion to the standardized velocity (V_i) of the travel process (see Equation 74).

$$V_i = \gamma\left(\frac{1}{W_i + T_i}\right) = \frac{\gamma}{S_i} \tag{74}$$

where γ is a parameter represents the inverse ratio between W_i+T_i and Vi.

Combining Equations 72-74, Equation 75 and 76 quantify the probability (P_i) that the spike trains (entities) pass through route i in totally U multiple routes.

$$P_i = \frac{ST_i}{\sum_{j=1}^{U} ST_j} = \frac{r_0 V_i}{\sum_{j=1}^{U} r_0 V_j} = \frac{r_0\left(\frac{\gamma}{S_i}\right)}{\sum_{j=1}^{U}\left[r_0\left(\frac{\gamma}{S_j}\right)\right]} = \frac{1/S_i}{\sum_{j=1}^{U}\frac{1}{S_j}} \tag{75}$$

Thus, we have:

$$P_i = \frac{1/S_i}{\sum_{j=1}^{U}\frac{1}{S_j}} \tag{76}$$

4. Deduction of *CB(t)* in Modeling the Brain Imaging Study in PRP

Using Cohen 's model, Anderson et al. (2003, 2004) proposed that the integrated BOLD signal (*CB(t)*) in a certain brain region is mainly determined by several factors: the length of time the current buffer/server occupied throughout time *t* (*i(x):* at time *x*, if the current buffer/server is occupied, *i(x)*=1; otherwise, *i(x)*=0), latency scale *s* and magnitude scale *M* (see Equation 77) (Anderson et al., 2003; Anderson et al., 2004).

$$CB(t) = M\int_0^t i(x)B(\frac{t-x}{s})dx \tag{77}$$

where $B(T) = kT^a e^{-T/b}$ (Cohen, 1997). In the Queueing network model, assuming the length of time server *i* is being used is η, Equation 77 can be further developed into:

$$CB(t)= \begin{cases} M\int_0^{\eta} B(\frac{t-x}{s})dx & 0 \le x \le \eta \Rightarrow i(x) = 1 \\ 0 & x < 0 \text{ or } x > \eta \Rightarrow i(x) = 0 \end{cases} \tag{78}$$

Suppose $Y = \frac{t-x}{s}$ and combine Equation 78 with the Cohen's equation $B(T) = kT^a e^{-T/b}$, then Equation 78 can be rewritten into Equation 79:

$$CB(t)= \begin{cases} skM\int_{\frac{t-\eta}{s}}^{\frac{t}{s}} Y^a e^{-Y/b}dY & \frac{t-\eta}{s} \le Y \le \frac{t}{s} \\ 0 & \frac{t-\eta}{s} > Y \text{ or } Y > \frac{t}{s} \end{cases} \tag{79}$$

Solving the integer above, result in:

$$CB(t)= \begin{cases} \frac{kMsb}{a+1}\left(\begin{array}{l} e^{(-.5t/bs)}\left(t/s\right)^a \left(t/bs\right)^{-.5a} \text{WhittakerM}\left(.5a,.5a+.5,t/bs\right) - \\ e^{(-.5a\frac{t-\eta}{bs})}\left(\frac{t-\eta}{bs}\right)^{.5a} \text{WhittakerM}\left(.5a,.5a+.5,\frac{t-\eta}{bs}\right) \end{array} \right) & \frac{t-\eta}{s} \le Y \le \frac{t}{s} \\ 0 & \frac{t-\eta}{s} > Y \text{ or } Y > \frac{t}{s} \end{cases} \tag{80}$$

where the result of the Whittaker function—WhittakerM (m, n, z) can be obtained by solving the following differential equation: $y'' + (-0.25 + \frac{m}{z} + \frac{0.25 - n^2}{z^2})y = 0$.

Chapter 4

Queueing Network Modeling of a Real-time Psychophysiological Index of Mental Workload—P300 in Event-Related Potential (ERP)

Chapter Summary

Modeling and predicting of mental workload is one of the most important issues in studying human performance in complex systems. Ample research has shown that the amplitude of the P300 component of event related potential (ERP) is an effective real-time index of mental workload, yet no computational model exists that is able to account for the change of P300 amplitude in dual task conditions compared to single task situations. We describe the successful extension and application of a new computational modeling approach in modeling P300 and mental workload—a Queueing network approach based on the Queueing network theory of human performance (Liu, 1996, 1997) and neuroscience discoveries. Based on the neurophysiological mechanisms underlying the generation of P300, the current modeling approach accurately accounts for P300 amplitude both in temporal and intensity dimensions. This approach not only has a basis in its biological plausibility, but also has the ability to model and predict workload in real-time and can be applied to other applied domains. Further model developments in simulating other dimensions of mental workload and its potential applications in adaptive system design are discussed.

1. Introduction

Mental workload is one of the most important issues in studying human performance in complex systems (Moray, 1988; Wickens, Lee, Liu, & Gordon-Becker, 2003; Xie & Salvendy, 2000b). Overloaded operators are more likely to make errors, reducing the safety of human-machine systems (Moray, 1988). From the system engineering perspective, modeling and predicting mental workload at an early stage in system design is very helpful to reduce mental workload of operators (Moray, 1988; Xie & Salvendy, 2000a). Moreover, designing adaptive user interface in "real-time human engineering" expects real-time prediction of mental workload, so that the user interface can propose corresponding actions to keep operator mental workload at an optimal value (Tsang & Vidulich, 2003). In addition, there is a growing research field in human factors called neuroergonomics, which focuses on investigation of the neural bases of mental functions and physical performance in applied domains (Parasuraman, 2003). If a computational model can bridge the neural activities (measured by ERP or fMRI techniques) and mental workload, it might be a useful tool to assist researchers in human factors to understand the basic mechanisms of mental workload and design the interface to optimize the workload.

To measure changes in mental workload in real time, event-related brain potential (ERP) measurements stand out to be one of the most effective indexes of mental workload in comparison with some other behavioral, subjective and psychophysiological measurements (Johnson & Proctor, 2004). There are several advantages in using the ERP technique to measure mental workload. First, it provides a relatively continuous record of data over time, meeting the requirement of real-time human engineering. Second, it is not obtrusive to task performance since it does not require overt responses which are needed in measuring mental workload with secondary task measurements. Third, compared with some other physiological measurements such as the pupil diameter which is sensitive to all information stages including perceptual, cognitive and motor processing, ERP (e.g., P300 component) is diagnostic and sensitive to stimulus-evaluation process (perceptual and central processing resources) but not motor execution process (Johnson & Proctor, 2004).

Ample ERP research has shown that the amplitude of the P300 component in the ERP typically reflects the current state of mental workload (Coles, 1996; David & Friston, 2003; David, Harrison, & Friston, 2005; Kramer, Trejo, & Humphrey, 1995). P300 is a positive component characterized by a parietally maximal scalp distribution and a latency between 300 and 800 ms (Rugg & Coles, 1995). Here, latency refers to the time interval between the arrival of stimulus and the time point when the peak of the potential is observed. Because of its ease in measurement, P300 has become the most frequently measured ERP component. The most important finding of P300 related to mental workload is that the P300 amplitude (peak value) of a secondary task is reduced in dual-task conditions compared with the corresponding single task situation of performing the secondary task alone and the P300 amplitude of the secondary task decreases further when the difficulty of the primary task increases (Parasuraman, 1990; Wickens, Kramer, Vanasse, & Donchin, 1983; Wickens et al., 2003). Wickens et al.'s study (1983) (Wickens et al., 1983) is a representative study among the studies on this topic and therefore selected as the target experiment for modeling in this paper (A detailed description of their experiment is in the modeling section of this paper).

To model P300 in accordance with its biological realism—an important requirement for building cognitive models (O'Reilly, 1998), it is desirable to introduce the physiological mechanism underlying the generation of P300 discovered in neuroscience studies. Several researchers have proposed that P300 results from intracortical currents which are triggered by the release of norepinephrine (NE) (Nieuwenhuis, Aston-Jones, & Cohen, in press; Nieuwenhuis, Gilzenrat, Holmes, & Cohen, 2005). NE is a type of neurotransmitter which is responsible for synaptic transmission (Burke et al., 2006; Grid, Statements, Website, & Version, 2002; Lindquist & Rehnmark, 1998; Masur, Niggemann, Zanker, & Entschladen, 2001; Mpofu & Conyers, 2003; Nadel & Barnes, 1984; Pirke, 1996; Sanders, Happe, & Murrin, 2005; Voorhess, 1984; Xu et al., 2000)—when a spike train arrives in an axon terminal, NE is released into the synaptic cleft producing an electrical response in the postsynaptic neurons (Bear, Connors, & Paradiso, 2001; Haines, 2002) through a chain of events. First, the NE is produced by the locus coeruleus-norepinephrine (LC-NE) system (a nucleus in the pontine regions of the brain stem that consists of NE-containing cells) (Berridge & Waterhouse, 2003; Nieuwenhuis et al., in

press); the LC-NE system synthesizes the NE and then sends the neurotransmitter to the central nervous system via its efferent projections. Second, NE is released in certain brain regions (known as P300 generators), causing a change of conductivity of these regions and then producing a change of the amplitude of P300 (Nieuwenhuis et al., in press). Nieuwenhis et al. (2005) reviewed the major findings on the generators of P300 and found that P300 generators are mainly located in the prefrontal cortex, the medial temporal lobe structures (including the hippocampus), the temporal-parietal junction, and adjacent areas which are responsible for perceptual processing (Nieuwenhuis et al., in press).

Besides the experimental studies of the P300 component related to mental workload and its mechanism, it is necessary to review the related computational models of mental workload, ERP, and the LC-NE system.

In human factors engineering, several models of mental workload have been successfully developed and they can be categorized into three groups: conceptual models, mathematical and simulation models, and task-analytic models. Among the conceptual models, Wickens' resource model (1990) is one of the most influential models and it describes how the amplitude and latency of the P300 component is related to the "resource" in cognitive information processing (Wickens, 1990). In his model, the amplitude of the P300 component of the secondary task reflects the perceptual-cognitive resource which are depleted by the primary task (Wickens, 1990). Among the mathematical and simulation models, the representative models include control theory-based model (Levison, 1979), Queueing theory-based model (Moray, Dessouky, Kijowski, & Adapathya, 1991; W.B. Rouse, 1980), PROCRU (Procedure-Oriented Crew Model, (Baron & Corker, 1989)), Micro-SAINT (Chubb, Laughery, & Pritsker, 1987), HOS (Human Operator Simulator, (Harris, Glenn, Iavecchia, & Zaklad, 1986)), Rouse et al.'s mathematical model (1993) (W. B. Rouse, Edwards, & Hammer, 1993), and MHP (Model Human Processor, (Card, Moran, & Newell, 1983)). Unlike the models of ERP and the LC-NE system which focus on the biological aspect of the cognitive system, these models emphasize their engineering applications and the definition of mental workload varies based on the feature of the model itself. For example, Rouse's Queueing theory model (1980) regards the server utilization as a representation of mental workload

(W.B. Rouse, 1980). The task-analytic models include TLAP (Time Line Analysis and Prediction Model (Parks & Boucek, 1989)), TAWL (Task Analysis/Workload (Hamilton & Bierbaum, 1990)), W/INDEX (Workload Index (North & Riley, 1989)), and Bi and Salvendy's model (1994) (Bi & Salvendy, 1994) (see (Xie & Salvendy, 2000a) for a comprehensive review of these models).

Different from the models in human factors engineering, the models introduced in the following section focus on the physiological and biological mechanisms in generating ERP waves. In modeling ERP and EEG (electro-encephalography), several mathematical and simulation models have been successfully established (David & Friston, 2003; David et al., 2005; Gratton, Coles, Sirevaag, Eriksen, & Donchin, 1988; Jansen & Rit, 1995). Building on a lumped-parameter model, Jansen and Rit (1995) developed a computational model to produce EEG rhythms (Jansen & Rit, 1995). Based on Jansen and Rit's model (1995), a neural mass model proposed by David and Friston (2003) assumed that the behavior of a population of neurons (millions of interacting neurons) can be approximated using several state variables (e.g., mean membrane currents, potentials, and firing rates). The model reproduced brain signals within the oscillatory regime by simply changing the population kinetics (David & Friston, 2003).

In modeling the LC-NE system, several neural network models have been developed successfully (Nieuwenhuis et al., 2005; Usher & Davelaar, 2002). These models usually include several layers of connectionist units representing detection/input, decision, and response. These layers are connected with excitatory and inhibitory connections and the weights of these connections are updated during the learning process. Nieuwenhuis et al.'s model (2005) is able to successfully simulate LC activity and output of NE from the LC-NE system. Based on the LC activity and NE output, their model quantifies the attentional blink—a temporary deficit in processing of a target stimulus following successful processing of a previous target (Nieuwenhuis et al., 2005).

In sum, each of these models demonstrates their usefulness and ability to quantify one or several aspects of mental workload, ERP or the LC-NE system. However, none of these models quantify the major finding of P300 amplitude and latency related to mental workload based on its physiological mechanisms. As suggested by Olsen and Olsen (1990), modeling mental workload remains to be a challenge in cognitive modeling even

though overt behavior (reaction time and response accuracy) has been modeled by existing models more successfully (Olsen & Olsen, 1990).

In this paper, we describe a Queueing network modeling approach to quantify human performance and P300 as one of the most important psychophyisological indexes of mental workload, focusing on both biological realism of mental workload and its engineering application. First, we introduce the platform of this modeling approach—a simulation model of a Queueing network architecture representing information processing in the brain. Second, based on this network platform, a set of mathematical equations is developed and implemented into the simulation model to quantify the amplitude and latency of P300. Third, the modeling results are presented and validated with the results of the representative experimental study of Wickens et al. (1983). Finally, we discuss the implication of the modeling approach and its further extensions to model the experimental results of other electrophysiological and human factors studies.

2. Modeling of Human Performance and P300 in a Tracking Task

In the following section, we describe our use of the Queueing network modeling approach to model human performance and P300. First, a set of formulas are developed and implemented in the simulation model to quantify the amplitude and latency of P300. Second, a representative experiment on human performance and P300 is described, which was used to validate the modeling method proposed in this paper. In the third section, we describe how to simulate performance and P300 with QN-MHP.

2.1 Modeling the Amplitude and Latency of P300

Quantification of P300's amplitude and latency in the Queueing network model is composed of two parts: 1) modeling the entities representing the neurotransmitters in synaptic transmission; 2) based on this modeling result and existing computational models in electric fields of the brain, both the amplitude and latency of P300 are quantified by a set of formulas. All of these formulas are implemented in the simulation

model so that the model is able to generate the corresponding values for the dependant variables in real-time.

2.1.1 Modeling NE in Synaptic Transmission

As described in the introduction section of this paper, after NE is produced from the LC-NE system, NE reaches target brain regions engaged in processing the information of tasks. Based on the balance of NE before and after synaptic transmission (Bear et al., 2001; Haines, 2002), the total amount of released NE in processing the tasks (suppose there are ξ tasks which are concurrently processed) equals the difference between the amount of NE synthesized from the LC-NE system ($NE_{LC,}$ (Neff, Spano, Groppetti, Wang, & Costa, 1971; Nieuwenhuis et al., 2005)) and the amount of residual NE left (NE_0) in the presynaptic neurons (see Equation 1), where τ is a normally distributed random factor with mean equaling 0.

$$\sum_{m=1}^{\xi} NE_{rel,\,m} = NE_{LC} - NE_0 + \tau \tag{1}$$

For any one of these ξ tasks, the amount of NE released for task i ($NE_{rel,i}$) is determined by Equation 2.

$$NE_{rel,\,i} = \sum_{m=1}^{\xi} NE_{rel,\,m} - \sum_{m \neq i}^{\xi} NE_{rel,\,m} \tag{2}$$

Therefore, we have:

$$NE_{rel,\,i} = NE_{LC} - \sum_{m \neq i}^{\xi} NE_{rel,\,m} - NE_0 + \tau \tag{3}$$

Equation 3 above can be rewritten as:

$$NE_{rel,\,i} = NE_{LC} - \sum_{m \neq i}^{\xi} \sum_{all\ j} N_{m,\,j} C_{m,\,j} NE_p - NE_0 + \tau \tag{4}$$

where $N_{m,j}$ is the number of information entities of other tasks concurrently processed in server j; $C_{m,j}$ is the number of processing cycles for each of those entities at server j; NE_p is the amount of NE needed for each of those entities at each processing cycle at server j.

2.1.2 Modeling P300 Amplitude

In the computational models of brain potentials, Nunez (1981) proposed the following basic formula for quantifying the amplitude of the brain potentials (Nunez, 1981).

$$\phi = \frac{I}{4\pi r\delta} \tag{5}$$

where, ϕ is the amplitude of the ERP potential (unit: μv); r is the distance from the electrical field point (location where NE is released) to locations of the electrodes on the scalp; δ is the resistivity of the brain regions across this distance; I is the current from the electrical field point where NE is released.

Since there is an inverse proportional relation between the resistance and the amount of NE released (NE_{rel}) (Gray, Freeman, & Skinner, 1986; Nieuwenhuis et al., 2005), δ in Equation 5 can be further quantified in Equation 6 where b is a constant in this inverse relationship.

$$\delta = b / NE_{rel} \tag{6}$$

Moreover, the number of population spike trains (represented by information entities) (N) is in direct proportion to the current (Nunez, 1981), I in Equation 5 can be quantified in Equation 7 where k is a constant in this relationship.

$$I = kN \tag{7}$$

Combining Equations 5, 6, and 7, we have:

$$\phi = \frac{kN}{4\pi r(b / NE_{rel})} = (k / b)\frac{NE_{rel}N}{4\pi r} \tag{8}$$

Furthermore, since P300 comes from the generators of P300 wave in certain brain regions, Equation 4 can be further developed into:

$$NE_{rel,i} = NE_{LC} - \sum_{m \neq i}^{\xi} \sum_{all\ j'} N_{m,j'} C_{m,j'} NE_p - NE_0 + \tau \tag{9}$$

where *j'* represents the servers which can serve as the generators of P300 corresponding to the neuroscience findings (servers in the perceptual subnetwork, Servers A, B, C, E, F in the cognitive subnetwork corresponding to the P300 generators described in the introduction section).

Combining Equations 8 and 9, the P300 amplitude including its peak for task i is quantified in Equation 10.

$$\phi_i = (k / b)\frac{NE_{rel,i}N_i}{4\pi r} = (k / b)\frac{N_i}{4\pi r}(NE_{LC} - \sum_{m \neq i}^{\xi} \sum_{all\ j'} N_{m,j'} C_{m,j'} NE_p - NE_0 + \tau) \tag{10}$$

Therefore, when the amount of NE consumed by the primary task increases from 0 (single secondary task condition) to a certain value (dual task condition) the amount of NE available for the secondary tasks decreases. This decrease in the amount of NE produces an increase in the resistivity of the brain regions and then a decrease in the amplitude of P300 of the secondary task. The P300 amplitude of the secondary task reduces further as the difficulty of the primary task increases consuming a greater amount of NE.

2.1.3 Modeling P300 Latency

The latency of P300 for a certain task i (L_i) is composed of three parts: the time interval between the stimulus presentation and the arrival of stimulus information at the LC-NE system ($T_{i,P}+T_{i,A/B}+T_{i,C}+T_{i,E}$), the time interval between the arrival of stimulus information at the LC-NE system (t=0) and the time point when NE_{LC} reaches its peak (t_p), and the conduction time of NE from LC to the forebrain (NE_{cond}) which processes task information, as shown in Equation 11.

$$L_i = T_{i,P}+T_{i,A/B}+T_{i,C}+T_{i,E}+ t_p+NE_{cond} \quad (11)$$

where $T_{i,P}$, $T_{i,A/B}$, $T_{i,C}$, and $T_{i,E}$ are the processing time of task i at the perceptual subnetwork, at Server A or B, at Servers C and E, respectively.

2.2 A Representative Experiment on P300 and Human Performance

Wickens et al. (1983) measured human performance and the P300 in a concurrent task which includes a visual-manual tracking task (primary task) and an auditory probe counting task (secondary task) (Wickens et al., 1983). In the primary task, subjects manipulated a joystick and attempted to superimpose a cursor on a target which was moving in a series of discrete horizontal displacement on a visual display. There were three levels of difficulty in the primary task: i) 1st order predictable (1P): the target moved only in a left-right direction and only the magnitude of the movement/step was unpredictable; the control of the cursor with the joystick followed 1st order control—constant displacement of the joystick caused the cursor to move at a constant velocity in the movement direction of the joystick; ii) 1st order unpredictable (1U): both direction and magnitude of the movement of the target were unpredictable and the control of the

cursor with the joystick still followed 1st order control; iii) 2nd order unpredictable (2U): both direction and magnitude of the movement of the target were unpredictable and the control of the cursor with the joystick followed 2nd order control—constant displacement of the joystick accelerated the cursor's movement. Concurrently with the tracking tasks, the subjects were assigned to perform an auditory probe counting task. The subjects heard a Bernoulli series of tones of high and low pitch, occurring with equal probability and the subjects were instructed to count the number of occurrences of the low-pitched tones. They found that P300 amplitude (peak value) of the secondary task was reduced in dual-task conditions compared with a single task situation and the P300 amplitude (peak value) was decreased further when the difficulty of the primary task increased.

2.3 Simulation of Human Performance in the Target Experiment

Simulation of any human-machine interaction task requires the specification of three components: a human model, the machine or the environment with which the human model interacts, and the task input to the human model. These three components correspond to the simulation model of Queueing network (QN-MHP), a joystick, a visual display presenting the cursor and the target, and a speaker presenting the auditory stimulus, respectively, in the context of the dual task—manual tracking and auditory probe counting (see Figure 4-1).

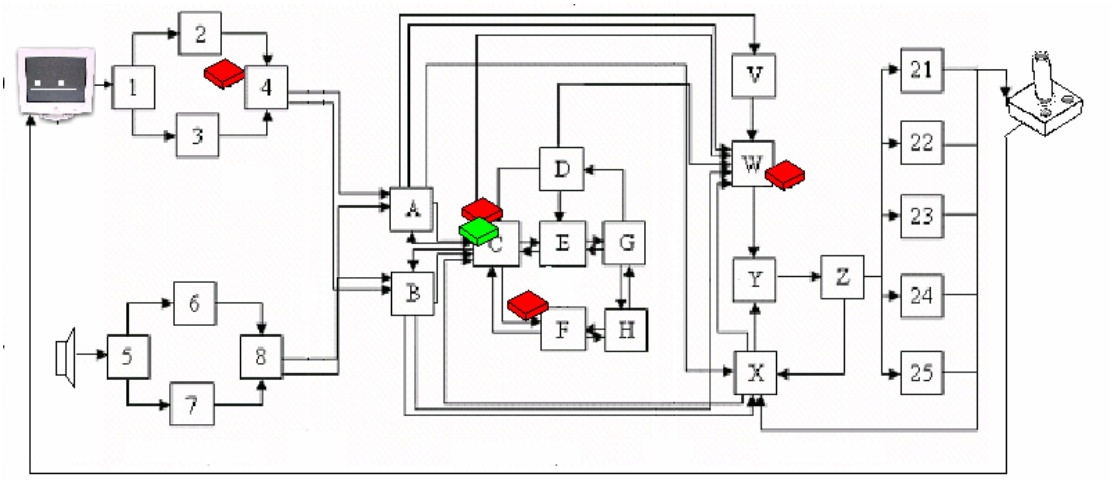


Figure 4-1Components of the simulation model (QN-MHP) in simulating the concurrent task: the manual tracking task (red entities) and the audio probe counting task (green entities)

The general human model of QN-MHP is described in the previous section. To possess the basic knowledge of tracking and counting requires the QN-MHP to have the corresponding procedure knowledge rules stored in its long-term procedure memory server. Thus, following the general method of QN-MHP simulation (Liu, Feyen, & Tsimhoni, in press), the NGOSML-style task descriptions of both the manual tracking and auditory probe counting task are developed (see Table 4-1) and stored in server D as the long-term procedure knowledge of the task in the model. For the tracking task: first, the model watches for the spatial difference between the cursor and the target; second, if there is a difference, the model computes the expected movement time (1P, 1U, and 2U conditions) and expected movement direction (1U and 2U conditions) (with an increase in tracking difficulty, the number of cycles in computation increases); third, the model executes the movement to move the joystick in the expected movement direction and time. Similarly, in the auditory probe counting task, the model increases the value of a counter if it receives a target low-pitch tone from the auditory perceptual subnetwork. All of these steps or operators are defined in a task-independent manner; task-specific information is treated as their parameters.

More importantly, one of the unique features of QN-MHP in modeling concurrent tasks is that the entities representing the information of the two tasks can be processed in the network concurrently and multitask performance emerges as the behavior of multiple streams of information flowing through a network without writing another program to either interleave two task procedures into a serial program or control the two task procedure with an executive control (Liu et al., in press).

In addition, to define the joystick with which the QN-MHP interacts, a software module called m-hJOYSTICK is implemented to represent the joystick in the tracking task. This module defines the order of control (1^{st} or 2^{nd} order), collects the movement information of the Hand server, and transmits the corresponding position of the cursor on the visual display which is implemented in a server in the model (see Figure 4-1). This module also computes and records the root-mean-square error of the tracking task. Another software module is implemented to represent the speaker which produces the entities of auditory stimulus and supplies them to the auditory perceptual subnetwork.

Human performance is generated by a natural interaction of the entities of the concurrent tasks being processed in the network following the task descriptions (see Figure 4-1).

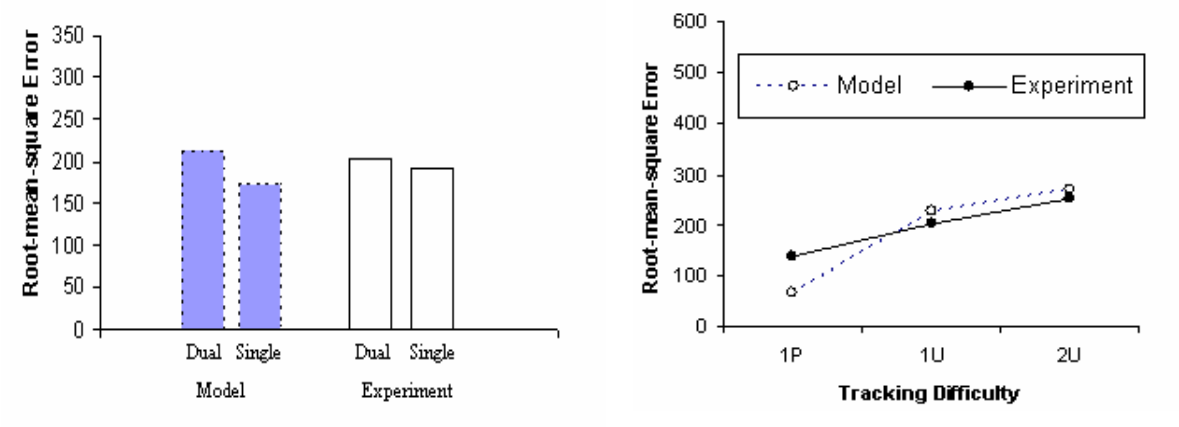
Table 4-1 NGOMSL-Style Task Description of Manual Tracking and Auditory Probe Counting Task

GOAL: Do manual tracking task	GOAL: Do auditory probe counting task
Method for GOAL: Do manual tracking task	Method for GOAL: Do auditory probe counting task
Step 1. Watch for <the spatial difference between the cursor and the target> on <the display>	Step 1. Listen to <the tone> from <the speaker>
Step 2. Retain < the spatial difference>	Step 2. Retain <the tone>
Step 3. Decide: if there is a difference then go to step 4; else go to step 7	Step 3. Compare: <the tone> with <the target tone> in memory
Step 4. Compute <the expected movement direction (θ) and time of the joystick (t)>	Step 4. Decide: If match, then go to step 5 Else move to step 1
Step 5. Move <joystick> in <direction θ> for time <t>	Step 5. Compute <increase the counter>
Step 6. Go to step 1	Step 6. Retain <the counter>
Step 7. Stop moving <joystick>	

3. Simulation Results and its Validation

By implementing the equations developed in the previous section into the simulation model (see Appendix 1 for values of several parameters in these equations), the simulation results are obtained and compared with the target experimental results of Wickens et al. (1983).

Figure 4-2 showed the simulation results of the root-mean-square error of human performance in comparison with experimental results. The R square of the model is .99 and RMS equals 13.24 (comparison between single and dual task, Figure 4-2a); the R square is .95 and RMS equals 131.6 (comparison between 3 difficulty levels, Figure 4-2b).



(a) (b)

Figure 4-2 Root-mean-square error in the study of Wickens et al. (1983) (solid lines) in comparison with the Queueing network simulation results (dashed lines) (a: comparison between single (secondary task only) and dual task; b: comparison between the 3 difficulty levels)

The latency and the amplitude of P300 (peak value) are shown in Figure 4-3 and Figure 4-4. For the latency, R square=.99 and RMS=1; for the amplitude of P300, R square=0.99; RMS=0.39. The P300 amplitudes (peak values) of the secondary task, shown in Figure 4-4, is smaller in the dual task condition than in the single task condition (R square=0.99; RMS=0.39).

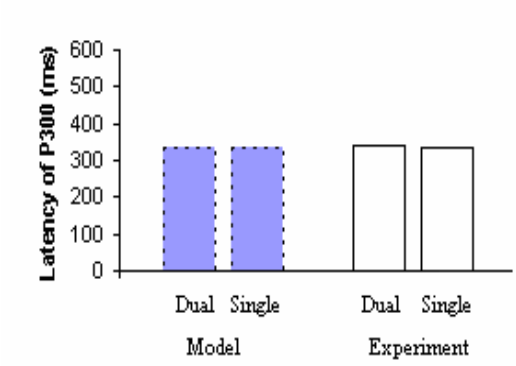


Figure 4-3 P300 latency in the study of Wickens et al. (1983) (solid lines) in comparison with the Queueing network simulation results (dashed lines) (single: secondary task only; dual: concurrent task)

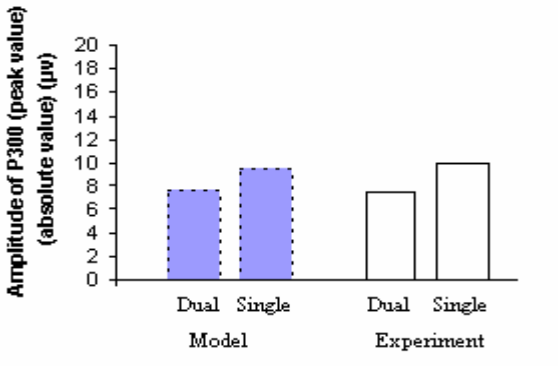
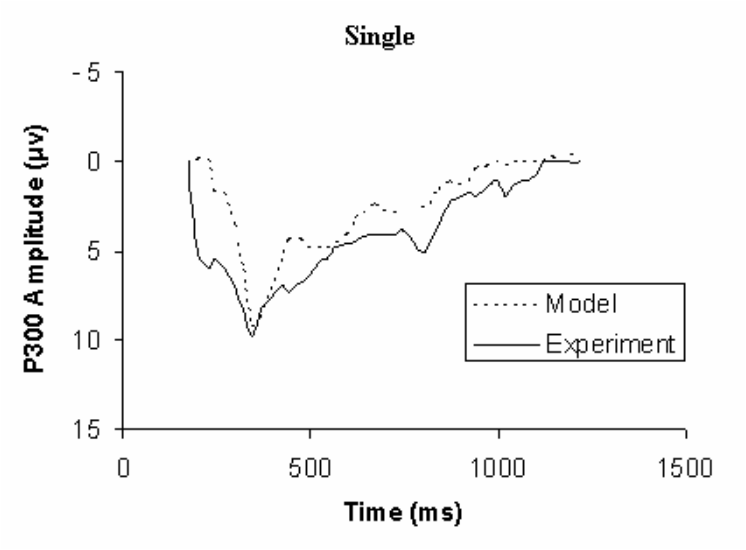


Figure 4-4 P300 amplitude (peak value) in the study of Wickens et al. (1983) (solid lines) in comparison with the Queueing network simulation results (dashed lines) (single: secondary task only; dual: concurrent task)

Figure 4-5 shows a comparison of the real-time change of the P300 amplitude of the secondary task in Wickens et al's experiment and the simulation results (secondary task

only and dual task conditions). In the single task condition (secondary task only), the R square of the model is 0.93 and RMS equals 1.63; in the concurrent task condition, the R square of the model is 0.86 and RMS equals 1.66.



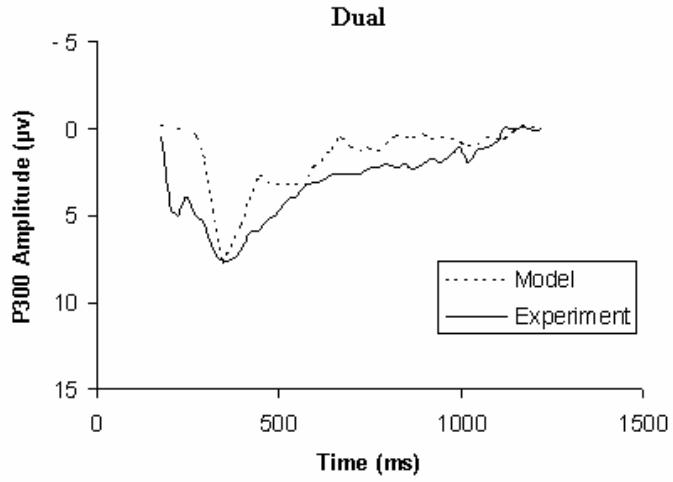


Figure 4-5 Real-time P300 amplitude in the study of Wickens et al. (1983) (solid lines) in comparison with the Queueing network simulation results (dashed lines) (single: secondary task only; dual: concurrent task)

The change of P300 amplitude (peak value) of the secondary task with an increase of tracking difficulty in the primary task is shown in Figure 4-6. R square of the model is 0.99 and RMS equals 5.86.

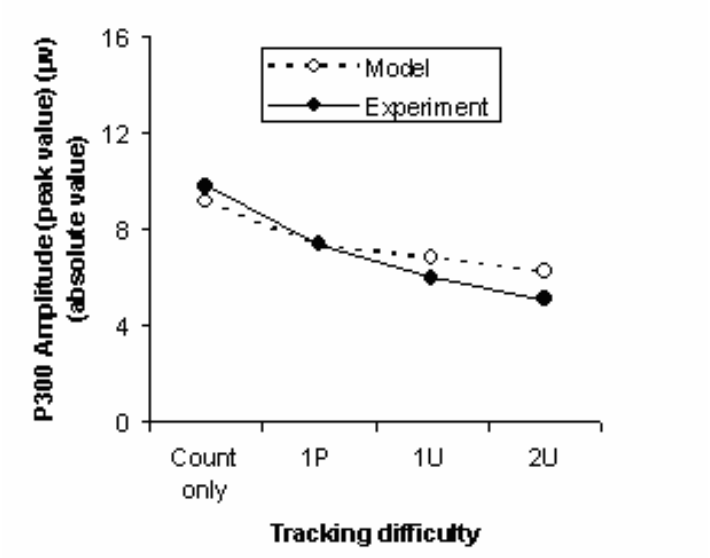


Figure 4-6 Change of P300 amplitude (peak value) with an increase of tracking difficulty in the study of Wickens et al. (1983) (solid lines) in comparison with the Queueing network simulation results (dashed lines)

4. Extension and Application of the Model

Equations developed in this work can be extended further to account for other P300 studies in multitasking and be used in designing user interface in dual tasks, including designing the stimuli or representation of multiple tasks on user interface and determining the maximal difficulty level of a task in multitasking.

Based on Equation 11, the P300 amplitude of Task 2 can be quantified into:

$$\phi_2 = (k/b)\frac{NE_{rel,i}N_i}{4\pi r} = (k/b)\frac{N_2}{4\pi r}(NE_{LC} - \sum_{m\neq i}^{\xi}\sum_{all\ j'} N_{1,j'}C_{1,j'}NE_p - NE_0 + \tau) \quad (12)$$

If there are only two tasks, Equation 12 can be simplified into:

$$\phi_2 = (k/b)\frac{N_2}{4\pi r}(NE_{LC} - \sum_{all\ j'} N_{1,j'}C_{1,j'}NE_p - NE_0 + \tau)$$

$$= \frac{(k/b)}{4\pi r}N_2(-\sum_{all\ j'} N_{1,j'}C_{1,j'}NE_p + NE_{LC} - NE_0 + \tau) \quad (13)$$

Since $\sum_{all\ j'} N_{1,j'}C_{1,j'} = NE_p\sum_{all\ j'} N_{1,j'}C_{1,j'} = NE_p n_1 \overline{N_1}\overline{C_1}$,

where $\overline{N_1}$ and $\overline{C_1}$ is the averaged number and cycle times of entities of Task 1 in the network, respectively (n_1 is the number of servers processing entities of Task 1),

$$\phi_2 = \frac{(k/b)}{4\pi r}(\alpha\overline{N_1}+\beta)(-NE_p n\overline{N_1}\overline{C_1} + NE_{LC} - NE_0 + \tau) \quad (14)$$

$$= 0.76(\alpha\overline{N_1}+\beta)(-.63n\overline{N_1}\overline{C_1} + NE_{LC} - 0.38) \text{ (assuming } \tau = 0 \text{ on average)}$$

Similarly,

$$\phi_1 = \frac{(k/b)}{4\pi r} N_1(-NE_p n_2\overline{N_2}\overline{C_2} + NE_{LC} - NE_0 + \tau) \quad (15)$$

Equations 14 and 15 can be used to optimize dual task performance in various domains. First, since P300 amplitude is one of the indexes to reflect the resource available for a task, Equations 14 and 15 can help us to identify the critical factors in maximizing the resource of Tasks 1 and 2 in dual tasks: N_2, $\overline{C_1}$, $\overline{N_1}$, n_1, and NE_{LC} to maximize the resource of Task 2; N_1, $\overline{C_2}$, $\overline{N_2}$, n_2, and NE_{LC} to maximize the resource of Task 1. Second, to maximize the resource of Task 2, N_2 and $\overline{N_1}$ can be optimized by properly designing the stimuli of the two tasks according to Equation 14; n_1 and $\overline{C_1}$ can be reduced by practicing Task 1. The same logic can also be applied to maximize the resource to Task 1. Third, Equations 14 and 15 can help us quantify the maximal difficulty level of a task in multitasking so that the resource available for the other task is maintained above the minimal level. In this paper, we focus our discussion on: 1) how to quantify the relation between N_2 and $\overline{N_1}$ with its implication in designing stimuli in multitasking; 2) how to quantify the maximal difficulty level of a task so that the resource available for the other task is above the minimal level.

4.1 Relation between N_2 and $\overline{N_1}$ with its implication in interface design in multitasking

Suppose there is a direct proportional relation between N_2 and $\overline{N_1}$:

$$N_2 = \alpha\overline{N_1} + \beta \quad (16)$$

Then,

$$\begin{aligned}\phi_2 &= \frac{(k/b)}{4\pi r}(\alpha\overline{N_1}+\beta)(-NE_p n_1\overline{N_1}\overline{C_1}+NE_{LC}-NE_0+\tau)\\ &= 0.76(\alpha\overline{N_1}+\beta)(-.63n_1\overline{N_1}\overline{C_1}+NE_{LC}-0.38)\ \text{(assuming } \tau=0 \text{ on average)}\\ &= -.48n_1\alpha\overline{C_1}\overline{N_1}^2+(-.48n_1\beta\overline{C_1}+0.76\alpha NE_{LC}-0.29\alpha)\overline{N_1}+.76\beta(NE_{LC}-0.38)\end{aligned} \quad (17)$$

In addition, according to the properties in Queueing networks, there are an inverse proportional relation between the difficulty level of Task 1 (TD_1) and $\overline{N_1}$ as well as a direct proportional relation between TD_1 and $\overline{C_1}$:

$$TD_1 = -g\overline{N_1}+h\ (g>0; h>g\overline{N_1}) \quad (18)$$

$$TD_1 = s\overline{C_1}+t\ (s>0; t=0\ \because\ \text{when } \overline{C_1}=0,\ TD_1=0) \quad (19)$$

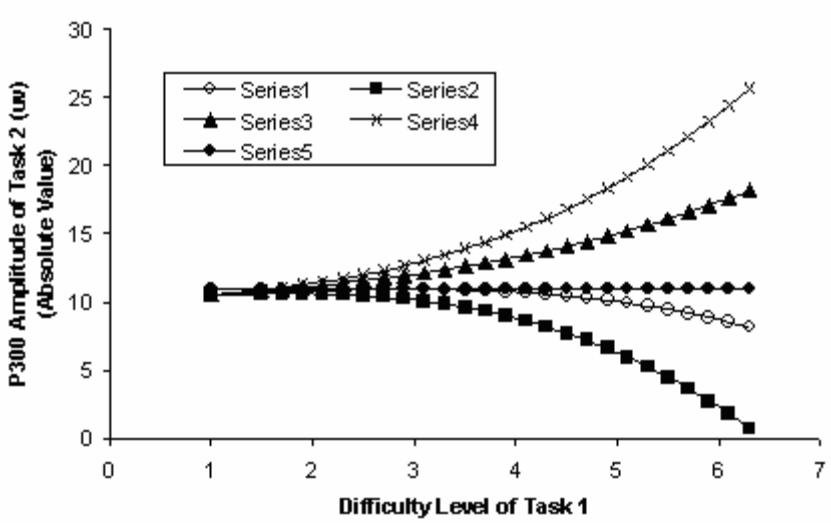
Combining Equations 17 to 19, Equation 17 can be further developed into:

$$\begin{aligned}\phi_2 &= -.48n_1\alpha s^{-1}g^{-2}TD_1^3+(.96hn_1\alpha s^{-1}g^{-2}+.48n_1\beta s^{-1}g^{-1})TD_1^2\\ &+(-.48n_1\alpha s^{-1}g^{-2}h^2-0.76\alpha g^{-1}NE_{LC}+0.29\alpha g^{-1}-.48n_1\beta hs^{-1}g^{-1})TD_1\\ &+0.76\alpha hg^{-1}NE_{LC}-0.29\alpha hg^{-1}+.76\beta(NE_{LC}-0.38)\end{aligned} \quad (20)$$

Take partial derivative:

$$\begin{aligned}\frac{\partial\phi_2}{\partial TD_1} &= -1.44n_1\alpha s^{-1}g^{-2}TD_1^2+(1.92hn_1\alpha s^{-1}g^{-2}+.96n_1\beta s^{-1}g^{-1})TD_1\\ &+(-.48n_1\alpha s^{-1}g^{-2}h^2-0.76\alpha g^{-1}NE_{LC}+0.29\alpha g^{-1}-.48n_1\beta hs^{-1}g^{-1})\end{aligned} \quad (21)$$

Based on Equation 21, depending on the value of α, β and other parameters, an increase of Task 1 difficulty level may generate an increase of ϕ_2, a decrease of ϕ_2, or no change of ϕ_2 (see Figure 4-7 and Appendix 2 for Equation *A-E* including their solutions).



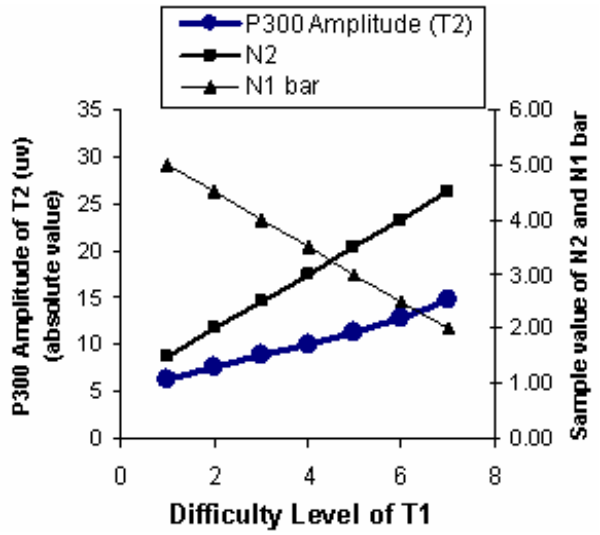
Condition 1: an increase of difficulty level of Task 1 (TD1) will increase the value of ϕ_2 (Series 3 and 4, $\alpha_4<\alpha_3<\Gamma$)

Condition 2: an increase of difficulty level of Task 1 (TD1) will decrease the value of ϕ_2 (Series 1 and 2, $\alpha_2>\alpha_1>\Gamma$)

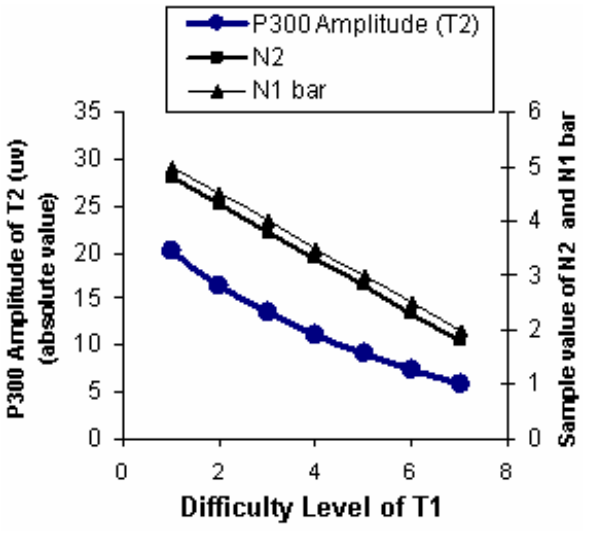
Condition 3: an increase of difficulty level of Task 1 (TD1) dose not affect the value of ϕ_2 (Series 5, $\alpha_5=\Gamma$)

Figure 4-7 The two conditions relating the difficulty level of Task 1 and the change of ϕ_2

Based on derivation results of Equation 21 in Appendix 2, Figure 4-8 provides a more intuitive illustration of Conditions 1 and 2 in Figure 4-7 connecting the relation between N_2 and $\overline{N_1}$ and the change of ϕ_2: 1) if there is an inverse proportional relation between N_2 and N1 bar ($\overline{N_1}$) (e.g., α=-1), the greater number of Task 1 ($\overline{N_1}$) on average, the less number of Task 2 (T2) entities processed in the network; Equation 21 predicts an increase of P300 amplitude (resource) of T2 (see Condition 1 in Figure 4-8); 2) if there is a direct proportional relation between N_2 and N_1 bar ($\overline{N_1}$) (e.g., α=1), the less number of Task 1 ($\overline{N_1}$) on average, the less number of Task 2 (T2) entities processed in the network; Equation 21 predicts a decrease of P300 amplitude (resource) of T2 (see Condition 2 in Figure 4-8).



Condition 1: an inverse relation between N_2 and $\overline{N_1}$ produced an increase ϕ_2 with increasing of difficulty level of Task 1 (TD1) (α= -1<Γ)



Condition 2: an direct relation between N_2 and $\overline{N_1}$ produced an decrease of ϕ_2 with increasing of difficulty level of Task 1 (TD1) (α=1> Γ)

Figure 4-8 The effect of the relationship between N2 and $\overline{N_1}$ on the change of P300 amplitude of Task 2 with an increase of difficulty level of Task 1

The predicted results in Figure 4-8 are consistent with the existing results in ERP studies (Kramer, et al., 1985). In Kramer et al.'s study (1985), when the stimuli of the primary task (T1) and the secondary task (T2) are integrated into the same stimuli (T1: tracking a moving object; T2: counting the transitional change of the same moving object) ("Dual-Task Integrality Condition" in Figure 1 in Kramer et al., 1985), the amplitude of P300 of T2 increases with an increase of difficulty level of T1. Different from the Wickens' study (1983) in which stimuli of T1 and T2 are not in the same object, "the dual-task integrality" condition in Kramer et al's study (1985) sets the stimuli of T1 and T2 into the properties of the same object. In QN-MHP, this setting of the experiment is represented as using one type of entity (called "shared entity" here) with different attributes. In other words, one entity carries the two tasks' information at the same time, generating parallel processing in the perceptual subnetwork in the model. When the difficulty level of T1 increases, the shared entities of T1 and T2 stay for a longer time at Server F to process the information of T1 (lower value of $\overline{N_1}$), increasing the number of shared entities in the other servers in the cognitive subnetwork and the perceptual subnetwork. Since the shared entities also carry the information T2, more information of

T2 get processed at the same time while the shared entities are waiting longer for the service of Server F (higher value of N_2). For explanation purposes, by plugging the value of other parameters, Equation 13 can be simplified into:

$$\begin{aligned}\phi_2 &= \frac{(k/b)}{4\pi r} N_2(-NE_p n\overline{N_1}\overline{C_1} + NE_{LC} - NE_0 + \tau) \\ &= 0.76N_2(-.75\overline{N_1}\overline{C_1} + 1.62) \\ &= 1.215N_2 - .57\overline{N_1}\overline{C_1}\end{aligned} \quad (22)$$

In Equation 22, the constant before N2 is greater than that in front of $\overline{N_1}\overline{C_1}$, therefore, the value of N2 becomes the major factor determining the value of ϕ_2. Higher value of N2 in the same object condition increases the value of ϕ_2.

In contrast, if the stimuli of T1 and T2 are not set in the same object (e.g., in Wickens et al., 1983; or the different object condition in Kramer et al, 1985), they are represented as two types of entities in the network. When the difficulty level of T1 increases, the entities of T1 spend more time at the Server F at the cognitive subnetwork (lower value of $\overline{N_1}$), decreasing number of entities of T2 receiving the service of Server F (lower value of N_2) and the value of ϕ_2 (see Equation 22).

The quantification of the relationship between N2 and $\overline{N_1}$ on the change of P300 amplitude of Task 2 above can be applied in designing user interfaces in multitasking. First, based on Figure 4-7, Figure 4-8 and Equation 13, it is recommended that the information of T1 and T2 can be encoded into the same object or stimuli, creating the inverse relation between N_2 and $\overline{N_1}$ with an increase of T1 difficulty level, and maximizing the parallel processing of information of T2 when the processing of T1 is delayed; moreover, as long as the derived α is lower than the threshold (Γ) derived in Appendix 2, an increase of T1 difficulty level will increase the value of ϕ_2 as an indication of the resource available for Task 2; the lower value of α, the more resources available for Task 2. Second, in the circumstances that stimuli of two tasks cannot be set in the same object, one of the focuses of the designer is to lower the derived α so that its value will be closer to Γ, the higher value of α, the less resources available for Task 2 when the difficulty level of Task 1 increases.

To implement this in real user interface design, first, a similar simulation process described in this paper is needed to obtain the value of α and other parameters in Appendix 2 (see the published papers (Wu & Liu, 2006b, 2006c) for a detailed description about how to use QN-MHP to simulate other tasks); second, the value of the threshold (Γ) can be obtained via Equations *A-D* and their solutions in Appendix 2; third, the original design of user information can be revised (e.g., integrating the information of two tasks in the same object or reducing the distance between the locations of the two objects belonging to T1 and T2) until the value of α_1 is lower than Γ in the same object condition (the lower, the better) or α_1 is closer to Γ in the different object condition (the closer, the better). The same logic can also be applied to the situation that the resource of Task 1 is to be maximized when the difficulty level of Task 2 increases.

4.2 Quantification of the maximal difficulty level of a task in multitasking

Equation 17 can also be used to determine the maximal difficulty level of a task in multitasking. Suppose a minimal level of resource for Task *i* is needed (say, $\phi_{i\,\min}$ found by ERP experiments), Equation 17 can be generalized into Equation 23 to quantify the maximal difficulty level of the other task *j* (TD_{jmax}) (see Figure 4-9 as an illustration of Equation 23)

$$\begin{aligned}\phi_{i\,\min} &= -.48n_j\alpha s^{-1}g^{-2}TD_{j\,\max}{}^3 + (.96hn_j\alpha s^{-1}g^{-2} + .48n_j\beta s^{-1}g^{-1})TD_{j\,\max}{}^2 \\ &+(-.48n_j\alpha s^{-1}g^{-2}h^2 - 0.76\alpha g^{-1}NE_{LC} + 0.29\alpha g^{-1} - .48n_j\beta hs^{-1}g^{-1})TD_{j\,\max} \\ &+0.76\alpha hg^{-1}NE_{LC} - 0.29\alpha hg^{-1} + .76\beta(NE_{LC} - 0.38)\end{aligned} \tag{23}$$

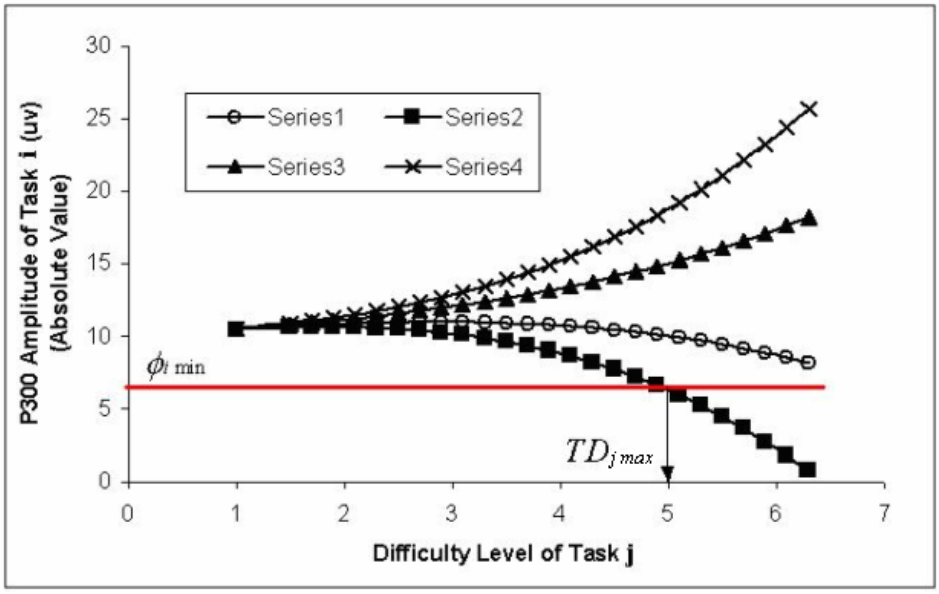


Figure 4-9 Deriving the maximal task difficulty level (TD_{jmax}) using Equation 23

For example, in a driving task, suppose T1 is steering a vehicle on a highway and T2 is operating an in-vehicle device, and we have obtained the minimal value of ϕ_1 ($\phi_{1\,min}$) based on existing ERP experiments in driving. After performing similar simulation using the current model (see the published papers (Wu & Liu, 2006b, 2006c) for a detailed description in how to use QN-MHP to model driving and other tasks), users of the model can obtain the parameter values in Equation 23 except TD_{2max}. Finally, the maximal task difficulty level of T2 (TD_{2max}) can be obtained via Equation 23; and it can be used as a guideline to design the user interface of an in-vehicle device, so that operating this device at the same time while driving will not exceed the "red-line" of resource for the primary driving task.

5. Conclusion

We described a Queueing network modeling approach to model human performance and P300 including its latency, amplitude, and real-time change of amplitude simultaneously in dual task situations. It successfully accounts for the major findings in measuring mental workload with the ERP techniques—the P300 amplitude of a secondary task decreases in dual-task conditions and this decrease is greater when the difficulty of the primary task is higher. By quantifying this major finding, QN-MHP also

offers a quantitative mechanism with corresponding neurological support to explain how these ERP phenomena are produced in the human brain as an attempt to meet requirements both for engineering applications and biological realism. This Queueing network modeling approach demonstrates its value in describing and predicting behavioral performance and the important aspects of the macroscopic electrical activities of the brain.

The simulation mechanism of the Queueing network modeling approach is consistent with existing models of mental workload and the results of other experimental studies. First, for the conceptual model proposed by Wickens (1990) (Wickens, 1990), the Queueing network modeling approach finds a potential neurological basis of the "resource" in Wickens' model. In a defined time period, the amount of NE synthesized in the brain is constrained by the amount of tyrosine, dopa, and energy (ATP, adenosine triphosphate) in the neurological system (Haines, 2002). When the amount of NE consumed by the primary task increases, the amount of NE available for the secondary task decreases naturally. This decrease in the amount of NE produces an increase in the resistivity of the brain regions and then decreases the amplitude of P300 measured by ERP techniques. Second, the modeled P300 latency (see Equation 11) is composed of the processing of entities at the perceptual and cognitive subnetwork, which is also consistent with Wickens' model that the latency of P300 results from the perceptual and cognitive processing activities before the motor response stage (Wickens, 1990).

The current modeling approach provides a useful linkage among neuron activity, mental workload, and human performance and it uses both bottom-up and top-down modeling methods: the quantification of NE in synaptic transmission in the model is a bottom-up modeling process starting from micro-activity in the brain; while the quantification of task procedure and an overall Queueing network structure of brain regions belong to a top-down modeling approach so that the model starts from a task-independent architecture to model mental workload and human performance in the target concurrent task. Moreover, the current model incorporates the NE output of the neural network model of the LC-NE system (Nieuwenhuis et al., 2005), offering a useful interface between neural network and Queueing network models.

Besides its role in connecting neural network with Queueing network models, this Queueing network model is useful in predicting mental workload in real-time in engineering applications. First, the consistency between the simulation results and experiment results suggests that this modeling approach is able to predict mental workload relatively accurately, both in the temporal dimension as reflected by P300 latency and in intensity dimension as indicated by the P300 amplitude. This relative sensitivity to the manipulation of the difficulty level of the task and the arrival patterns of task information make the model useful in engineering applications. For example, many intelligent or adaptive driver support and warning systems could benefit from computational workload models for estimating driver workload and proposing actions (e.g., redirecting messages into a voice mailbox, (Piechulla, Mayser, Gehrke, & Konig, 2003)) to prevent traffic accidents, since collecting ERP signals directly in these real world systems requires expensive devices. By implementing this computational model into these systems, driver mental workload can be estimated more accurately including the prediction when mental workload reaches "red-line" of mental workload (reflected by certain P300 amplitude) as well as by how much and for how long it exceeds that red-line.

Second, unlike the traditional models in mental workload, the current modeling approach starts from a task-independent cognitive architecture—QN-MHP which has successfully modeled human performance (e.g., reaction time, response accuracy, eye movement) of various kinds of tasks. The success of modeling mental workload significantly extends the coverage of the model in engineering applications and allows users of the model to model mental workload and human performance at the same time.

Furthermore, the mathematical models and the simulation model developed in this work can be extended to model P300 amplitude and latency in other tasks and be applied to designing user interface in multiple tasks, especially designing the representation or stimuli of multiple tasks as well as determining the maximal difficulty level of a task in multitasking.

The current modeling approach has its limitations since it is the first step to quantify P300 amplitude and latency in dual task using queueing network modeling methods. Because the biological aspects of our cognitive system are so complicated, researchers have found that other neurotransmitters, e.g., acetylcholine (Ach), may also affect P300

amplitude (Abe, Sawada, Horiuchi, & Yoshimura, 1999; Hammond, Meador, Aung-Din, & Wilder, 1987); therefore, the current model—using NE to quantify P300 amplitude—is only one of the possible approaches to quantify P300 amplitude in dual tasks.

We are extending the current modeling approach to quantify other important findings in mental workload research. For example, by quantifying subnetwork utilization, the Queueing network model is able to predict the subjective mental workload measured by various workload scales. Overall, the Queueing network modeling approach shows potential as a useful modeling method to quantify and predict mental workload, the behavioral performance, and electrophysiological phenomena of the cognitive system.

Reference

Abe, K., Sawada, T., Horiuchi, M., & Yoshimura, K. (1999). Effects of S-8510, a benzodiazepine receptor partial inverse agonist, on event related potentials (P300) in monkeys. *Psychopharmacology, 141*(1), 71-76.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience, 28*, 403-450.

Baron, S., & Corker, K. (1989). Engineering-based approaches to human performance modeling. In G. R. McMillan, D. Beevis, E. Salas, M. H. Strub, R. Sutton & L. V. Breda (Eds.), *Applications of Human Performance Models to System Design* (pp. 203-218). New York: Plenum.

Bear, M. F., Connors, B. W., & Paradiso, M. A. (2001). *Neuroscience: exploring the brain* (8th ed.). Baltimore, MD: Lippincott Williams & Wilkins,.

Bi, S. X., & Salvendy, G. (1994). A Proposed Methodology for the Prediction of Mental Workload, Based on Engineering System Parameters. *Work and Stress, 8*(4), 355-371.

Burke, W. J., Li, S. W., Gillespie, K. N., Chung, H. D., Jagadeesan, K., & Joist, J. H. (2006). Platelet Synthesis of DOPEGAL, the Free Radical Generating Metabolite of Norepinephrine: Potential Target for Protective Therapy in Arteriosclerosis. *Letters in Drug Design & Discovery, 3*(7), 481-487.

Byne, E. A., & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology, 42*, 249-268.

Card, S., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hinsdale, NJ: Lawrence Erlbaum.

Chubb, G. P., Laughery, K. R., & Pritsker, A. B. B. (1987). Simulating manned systems. In G. Salvendy (Ed.), *Handbook of Human Factors & Ergonomics*. New York: Wiley.

Coles, M. G. H. (1996). The lateralized readiness potential: Past, present, and future. *Psychophysiology, 33*, S3-S3.

Courtney, S. M., & Petit, L. (1998). An Area Specialized for Spatial Working Memory in Human Frontal Cortex. *SCIENCE 279*(27), 1347-1351.

David, O., & Friston, K. J. (2003). A neural mass model for MEG/EEG: coupling and neuronal dynamics. *Neuroimage, 20*, 1743-1755.

David, O., Harrison, L., & Friston, K. J. (2005). Modeling event-related respones in the brain. *Neuroimage, 25*, 756-770.

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

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

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11*, 355-372.

Faw, B. (2003). Pre-frontal executive committee for perception, working memory, attention, long-term memory, motor control, and thinking: A tutorial review. *Consciousness and Cognition, 12*(1), 83-139.

Gratton, G., Coles, M. G. H., Sirevaag, E. J., Eriksen, C. W., & Donchin, E. (1988). Pre- and post-stimulus activation of response channels: A psychophysiological analysis. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 331-344.

Gray, C. M., Freeman, W. J., & Skinner, J. E. (1986). Chemical dependencies of learning in the rabbit olfactory bulb: acquisition of the transient spatial pattern change depends on norepinephrine. *Behavioral Neuroscience, 100*(4), 585-596.

Grid, S., Statements, P., Website, A. S. A., & Version, P. (2002). Chronic Cocaine and Amphetamine Treatment Is Associated with Changes in Rat Spinal Norepinephrine Transporters. *Anesthesiology 96*, A794.

Haines, D. E. (2002). *Fundamental Neuroscience*. New York: Churchill Livingstone.

Hamilton, D. B., & Bierbaum, C. R. (1990). *Task analysis/workload (TAWL): AMethodology for predicting operator workload.* Paper presented at the Proceedings of the Human Factors Society 34th Annual Meeting, Santa Monica.

Hammond, E. J., Meador, K. J., Aung-Din, R., & Wilder, B. J. (1987). Role of the cholinergic system in the generation of human cognitive evoked potentials. *Neurology, 37*, 346–350.

Harris, R. M., Glenn, F., Iavecchia, H. P., & Zaklad, A. (1986). Human operator simulator. In W. Karwoski (Ed.), *Trends in Ergonomic: Human Factors III (Part A)*. Amsterdam: North-Holland.
Jansen, B. H., & Rit, V. G. (1995). Electroencephalogram and Visual-Evoked Potential Generation in a Mathematical-Model of Coupled Cortical Columns. *Biological Cybernetics, 73*(4), 357-366.
Kramer, A. F., Trejo, L., & Humphrey, D. (1995). Assessment of mental workload with task-irrelevant auditory probes. *Biological Psychology, 40*, 83-100.
Levison, W. H. (1979). A model for mental workload in tasks requiring continuous information processing. In N. Moray (Ed.), *Mental Workload: Its Theory and Measurement* (pp. 189-218). New York: Plenum.
Lim, J., & Liu, Y. (2004). *A queueing network model of menu selection and visual search.* Paper presented at the Proceedings of the 48 Annual Conference of the Human Factors and Ergonomics Society, New Orleans, Louisiana, USA.
Lindquist, J. M., & Rehnmark, S. (1998). Ambient Temperature Regulation of Apoptosis in Brown Adipose Tissue. *J Biol Chem, 273*(46), 30147-30156.
Liu, Y. (1996). Queueing network modeling of elementary mental processes. *Psychological Review, 103*(1), 116-136.
Liu, Y. (1997). Queueing network modeling of human performance of concurrent spatial and verbal tasks. *IEEE Transactions on Systems Man and Cybernetics Part a-Systems and Humans, 27*(2), 195-207.
Liu, Y. (2005). *Queueing Network Modeling of Mental Architecture, Response Time, and Response Accuracy: Reflected Multi-dimensional Diffusions.* Paper presented at the Annual Meeting of Mathematical Psychology Society, Memphis, TN.
Liu, Y., Feyen, R., & Tsimhoni, O. (in press). Queueing Network-Model Human Processor (QN-MHP): A Computational Architecture for Multitask Performance. *ACM Transaction on Human Computer Interaction*.
Lytton, W. W., & Lipton, P. (1999). Can the hippocampus tell time? The temporo-septal engram shift model. *NeuroReport*(10), 2301-2306.
Masur, K., Niggemann, B., Zanker, K. S., & Entschladen, F. (2001). Norepinephrine-induced Migration of SW 480 Colon Carcinoma Cells Is Inhibited by beta Blockers. *Cancer Research*(61), 2866-2869.
Moray, N. (1988). Mental workload since 1979. *International Review of Ergonomics, 2*, 123-150.
Moray, N., Dessouky, M. I., Kijowski, B. A., & Adapathya, R. (1991). Strategic Behavior, Workload, and Performance in Task-Scheduling. *Human Factors, 33*(6), 607-629.
Mpofu, E., & Conyers, L. M. (2003). Neurochemistry in the comorbidity of conduct disorder with other disorders of childhood and adolescence: implications for counselling. *Counseling Psychology Quarterly, 16*(1), 37-41.
Nadel, J. A., & Barnes, P. J. (1984). Autonomic Regulation of the Airways. *Annual Review of Medicine, 35*(1), 451-467.
Neff, N. H., Spano, P. F., Groppetti, A., Wang, C. T., & Costa, E. (1971). A simple procedure for calculating the synthesis rate of norepinephrine, dopamine and serotonin in rat brain. *The Journal of Pharmacology and Experimental Therapeutics, 176*, 701-709.
Nellgard, B. M. G., Miura, Y., Mackensen, G. B., Pearlstein, R. D., & Warner, D. S. (1999). Effect of intracerebral norepinephrine depletion on outcome from severe forebrain ischemia in the rat. *Brain Research, 847*(2), 262-269.
Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (in press). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin*.
Nieuwenhuis, S., Gilzenrat, M. S., Holmes, B. D., & Cohen, J. D. (2005). The role of the locus coeruleus in mediating the attentional blink: A neurocomputational theory. *Journal of Experimental Psychology-General, 134*(3), 291-307.
North, R. A., & Riley, V. A. (1989). W/INDEX: A predictive model of operator workload. In G. R. McMillan, D. Beevis, E. Salas, M. H. Strub, R. Sutton & L. V. Breda (Eds.), *Applications of Human Performance Models to System Design* (pp. 203-218). New York: Plenum.
Nunez, P. L. (1981). *Electric Fields of the Brain*. New York: Oxford University Press.
Olsen, G., & Olsen, J. R. (1990). The Growth of Cognitive Modeling in Human-Computer Interaction Since GOMS. *Human-Computer Interaction, 5*(2&3), 221-265.

Parasuraman, R. (1990). Event-related brain potentials and human factors research. In J. W. Rohrbaugh, R. Parasuraman & R. Johnston, Jr. (Eds.), *Event-related brain potentials*. London, UK: Oxford University Press.

Parasuraman, R. (2003). Neuroergonomics: research and practice. *Theoretical Issues in Ergonomics, 4*(1-2), 5-20.

Parks, D. L., & Boucek, G. P., Jr. (1989). Applications of human performance models to system design. In G. R. McMillan, D. Beevis, E. Salas, M. H. Strub, R. Sutton & L. V. Breda (Eds.), *Applications of Human Performance Models to System Design* (pp. 203-218). New York: Plenum.

Paton, J. F. R., Foster, W. R., & Schwaber, J. S. (1993). Characteristic Firing Behavior of Cell-Types in the Cardiorespiratory Region of the Nucleus-Tractus-Solitarii of the Rat. *Brain Research, 604*(1-2), 112-125.

Piechulla, W., Mayser, C., Gehrke, H., & Konig, W. (2003). Reducing drivers' mental workload by means of an adaptive man-machine interface. *Transportation Research, Part F, 6*, 233-248.

Pirke, K. M. (1996). Central and peripheral noradrenalin regulation in eating disorders. *Psychiatry Research 62*(1), 43-49.

Rieke, F., Warland, D., R.S., R., & Bialek, W. (1997). *Spikes: Exploring the Neural Code (Computational Neuroscience)*: MIT Press.

Roland, P. E. (1993). *Brain activation*. New York, NY: Wiley-Liss.

Rouse, W. B. (1980). *Systems Engineering Models of Human-Machine Interaction*. New York: North Holland.

Rouse, W. B., Edwards, S. L., & Hammer, J. M. (1993). Modeling the Dynamics of Mental Workload and Human-Performance in Complex-Systems. *IEEE Transactions on Systems Man and Cybernetics, 23*(6), 1662-1671.

Rugg, M. D., & Coles, M. G. H. (1995). *Electrophysiology of Mind*. New York: Oxford University Press.

Sanders, J. D., Happe, H. K., & Murrin, L. C. (2005). A transient expression of functional alpha2-adrenergic receptors in white matter of the developing brain. *Synapse, 57*, 213-222.

Schluter, N. D., & Krams, M. (2001). Cerebral dominance for action in the human brain: the selection of actions. *Neuropsychologia*(39), 105-113.

Smith, E. E., & Jonides, J. (1998). Neuroimaging analyses of human working memory. *Proc. Natl. Acad. Sci. USA, 95*, 12061-12068.

Taylor, J., Horwitzc, B., Shaha, N. J., Fellenzb, W. A., Mueller-Gaertnera, H.-W., & Krausee, J. B. (2000). Decomposing memory: functional assignments and brain traffic in paired word associate learning. *Neural Networks, 13*, 923-940.

Tsang, P. S., & Vidulich, M. A. (2003). *Principles and practice of aviation psychology*. Mahwah, NJ.: Lawrence Erlbaum Associates.

Voorhess, M. L. (1984). Low plasma norepinephrine responses to acute hypoglycemia in children with isolated growth hormone deficiency. *Journal of Clinical Endocrinology & Metabolism,, 59*(4), 790-793.

Wickens, C. (1990). Applications of event-related potentials to problems in human factors. In J. W. Rohrbaugh, R. Parasuraman & R. Johnston, Jr. (Eds.), *Event-related brain potentials*. London, UK: Oxford University Press.

Wickens, C., Kramer, A., Vanasse, L., & Donchin, E. (1983). Performance of Concurrent Tasks - a Psychophysiological Analysis of the Reciprocity of Information-Processing Resources. *Science, 221*(4615), 1080-1082.

Wu, C., & Liu, Y. (2004a). *Modeling Behavioral and Brain Imaging Phenomena in Transcription Typing with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA.

Wu, C., & Liu, Y. (2004). *Modeling human transcription typing with Queueing network-model human processor.* Paper presented at the Proceedings of the 48th Annual Meeting of Human Factors and Ergonomics Society, New Orleans, Louisiana, USA.

Wu, C., & Liu, Y. (2004b). *Modeling Psychological Refractory Period (PRP) and Practice Effect on PRP with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA.

Wu, C., & Liu, Y. (2006a). *Queueing Network Modeling of a Real-time Psychophysiological Index of Mental Workload—P300 Amplitude in Event-Related Potential (ERP).* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.

Wu, C., & Liu, Y. (2006b). *Queueing Network Modeling of Driver Workload and Performance.* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.

Wu, C., & Liu, Y. (2006c). *Queueing Network Modeling of Age Differences in Driver Mental Workload and Performance.* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.

Wu, C., & Liu, Y. (2006d). *Queueing Network Modeling of Reaction time, Response Accuracy, and Stimulus-Lateralized Readiness Potential Onset Time in a Dual Task.* Paper presented at the 28th Annual Conference of the Cognitive Science Society, Vancouver, BC, Canada.

Wu, C., & Liu, Y. (2006e). *Modeling fMRI BOLD Signal and Reaction Time of a Dual Task with a Queueing Network Modeling Approach.* Paper presented at the 28th Annual Conference of the Cognitive Science Society, Vancouver, BC, Canada.

Xie, B., & Salvendy, G. (2000). Review and reappraisal of modeling and predicting mental workload in single-and multi-task environments. *Work & Stress, 14*(1), 74-99.

Xu, F., Gainetdinov, R. R., Wetsel, W. C., Jones, S. R., Bohn, L. M., Miller, G. W., et al. (2000). Mice lacking the norepinephrine transporter are supersensitive to psychostimulants. *Nature Neuroscience, 3*, 465-471.

Appendix 1. Setting of Parameters

Parameter setting in simulation: no free parameter is used during the simulation process and all of the values of parameters come from original setting of QN-MHP (Liu et al., in press) and the existing neuroscience studies (see Table 4-2) (free parameter refers to parameters whose value is adjusted by researchers so that the modeling results fit the experimental results).

Table 4-2 Parameters used in Simulation

Parameter	Value	Description	Source
$T_{i,AP,}$ $T_{i,VP}$	126 ms	Time for auditory or visual perception (1 cycle is 42 ms at each server in perceptual subnetwork)	(Liu et al., in press)
$T_{i,A,}$ $T_{i,B,}$ $T_{i,C,}$ $T_{i,E,}$ $T_{i,F,}$	18 ms	1 processing cycle time at servers in cognitive subnetwork	(Liu et al., in press)
$T_{i,W,}$ $T_{i,Y,}$ $T_{i,Z,}$ $T_{i,X,}$	24 ms	1 processing cycle time at servers in motor subnetwork	(Liu et al., in press)
NE_p	0.01 μmol	Amount of NE needed for each of those entities at each processing cycle at server *j*	(Paton, Foster, & Schwaber, 1993)
NE_0	0.166 μmol	Amount of residual NE left (NE_0) in presynaptic neurons	(Nellgard, Miura, Mackensen, Pearlstein, & Warner, 1999)
k/b	18.0	Parameter in Nunez's equation	(Gray et al., 1986)
r	5.8 cm	Average distance from the servers as P300 generators to the scalp	(Bear et al., 2001; Haines, 2002)
t_p	100 ms	Duration that NE_{LC} reaches its peak	(Nieuwenhuis et al., in press)
NE_{cond}	65 ms	Average to conduction time from the LC to forebrain	(Aston-Jones & Cohen, 2005)

Appendix 2. The Three Conditions in Relating the Difficulty Level of Task 1 and the Change of ϕ_2

1) Condition 1: $\frac{\partial \phi_2}{\partial TD_1} > 0$:

$$-1.44m\alpha s^{-1}g^{-2}TD_1^2 + (1.92hm\alpha s^{-1}g^{-2} + .96m_1\beta s^{-1}g^{-1})TD_1$$
$$+(-.48m\alpha s^{-1}g^{-2}h^2 - 0.76\alpha g^{-1}NE_{LC} + 0.29\alpha g^{-1} - .48m_1\beta hs^{-1}g^{-1}) \geq 0$$
$$\Rightarrow \frac{(1.92hm\alpha s^{-1}g^{-2} + .96m_1\beta s^{-1}g^{-1}) + \sqrt{(1.92hm\alpha s^{-1}g^{-2} + .96m_1\beta s^{-1}g^{-1})^2 + 5.76m\alpha s^{-1}g^{-2}(-.48m\alpha s^{-1}g^{-2}h^2 - 0.76\alpha g^{-1}NE_{LC} + 0.29\alpha g^{-1} - .48m_1\beta hs^{-1}g^{-1})}}{2.88m\alpha s^{-1}g^{-2}}$$
$$> TD_1 >$$
$$\frac{-(1.92hm\alpha s^{-1}g^{-2} + .96m_1\beta s^{-1}g^{-1}) + \sqrt{(1.92hm\alpha s^{-1}g^{-2} + .96m_1\beta s^{-1}g^{-1})^2 + 5.76m\alpha s^{-1}g^{-2}(-.48m\alpha s^{-1}g^{-2}h^2 - 0.76\alpha g^{-1}NE_{LC} + 0.29\alpha g^{-1} - .48m_1\beta hs^{-1}g^{-1})}}{-2.88m\alpha s^{-1}g^{-2}}$$
$$\Rightarrow$$
$$A) -\sqrt{(1.92hm\alpha s^{-1}g^{-2} + .96m_1\beta s^{-1}g^{-1})^2 + 5.76m\alpha s^{-1}g^{-2}(-.48m\alpha s^{-1}g^{-2}h^2 - 0.76\alpha g^{-1}NE_{LC} + 0.29\alpha g^{-1} - .48m_1\beta hs^{-1}g^{-1})}$$
$$> (1.92hm\alpha s^{-1}g^{-2} + .96m_1\beta s^{-1}g^{-1}) - 2.88m\alpha s^{-1}g^{-2}TD_1$$
and
$$B) (1.92hm\alpha s^{-1}g^{-2} + .96m_1\beta s^{-1}g^{-1}) - 2.88m\alpha s^{-1}g^{-2}TD_1$$
$$< \sqrt{(1.92hm\alpha s^{-1}g^{-2} + .96m_1\beta s^{-1}g^{-1})^2 + 5.76m\alpha s^{-1}g^{-2}(-.48m\alpha s^{-1}g^{-2}h^2 - 0.76\alpha g^{-1}NE_{LC} + 0.29\alpha g^{-1} - .48m_1\beta hs^{-1}g^{-1})}$$

Solving Equations A and B, we have:

$$\alpha < \frac{\beta(2.77h - 5.53TD_1)}{-2.77g^{-1}h^2 - 4.38sn_1^{-1}NE_{LC} + 1.67sn_1^{-1} - 8.3g^{-1}TD_1^2 - 11.06TD_1hg^{-1}} = \Gamma \text{ (and } \alpha < 0) \quad (24)$$

If α and β satisfy Equation 24 (i.e., Equation A and B) at the same time, i.e. α is lower than the threshold (Γ), an increase of Task 1 difficulty level will increase the value of ϕ_2. Similarly,

2) Condition 2: $\frac{\partial \phi_2}{\partial TD_1} < 0$

$$-1.44 m \alpha s^{-1} g^{-2} TD_1^2 + (1.92 h m \alpha s^{-1} g^{-2} + .96 m \beta s^{-1} g^{-1}) TD_1$$
$$+(-.48 m \alpha s^{-1} g^{-2} h^2 - 0.76 \alpha g^{-1} NE_{LC} + 0.29 \alpha g^{-1} - .48 m \beta h s^{-1} g^{-1}) < 0$$
$$\Rightarrow TD_1 > \frac{(1.92 h m \alpha s^{-1} g^{-2} + .96 m \beta s^{-1} g^{-1}) + \sqrt{(1.92 h m \alpha s^{-1} g^{-2} + .96 m \beta s^{-1} g^{-1})^2 + 5.76 m \alpha s^{-1} g^{-2} (-.48 m \alpha s^{-1} g^{-2} h^2 - 0.76 \alpha g^{-1} NE_{LC} + 0.29 \alpha g^{-1} - .48 m \beta h s^{-1} g^{-1})}}{2.88 m \alpha s^{-1} g^{-2}}$$

or

$$TD_1 < \frac{-(1.92 h m \alpha s^{-1} g^{-2} + .96 m \beta s^{-1} g^{-1}) + \sqrt{(1.92 h m \alpha s^{-1} g^{-2} + .96 m \beta s^{-1} g^{-1})^2 + 5.76 m \alpha s^{-1} g^{-2} (-.48 m \alpha s^{-1} g^{-2} h^2 - 0.76 \alpha g^{-1} NE_{LC} + 0.29 \alpha g^{-1} - .48 m \beta h s^{-1} g^{-1})}}{-2.88 m \alpha s^{-1} g^{-2}}$$

$\Rightarrow$

$$C) \sqrt{(1.92 h m \alpha s^{-1} g^{-2} + .96 m \beta s^{-1} g^{-1})^2 + 5.76 m \alpha s^{-1} g^{-2} (-.48 m \alpha s^{-1} g^{-2} h^2 - 0.76 \alpha g^{-1} NE_{LC} + 0.29 \alpha g^{-1} - .48 m \beta h s^{-1} g^{-1})}$$
$$< 2.88 m \alpha s^{-1} g^{-2} TD_1 - (1.92 h m \alpha s^{-1} g^{-2} + .96 m \beta s^{-1} g^{-1})$$

or

$$D) (1.92 h m \alpha s^{-1} g^{-2} + .96 m \beta s^{-1} g^{-1}) - 2.88 m \alpha s^{-1} g^{-2} TD_1$$
$$> \sqrt{(1.92 h m \alpha s^{-1} g^{-2} + .96 m \beta s^{-1} g^{-1})^2 + 5.76 m \alpha s^{-1} g^{-2} (-.48 m \alpha s^{-1} g^{-2} h^2 - 0.76 \alpha g^{-1} NE_{LC} + 0.29 \alpha g^{-1} - .48 m \beta h s^{-1} g^{-1})}$$

$$\alpha > \frac{\beta(2.77h - 5.53TD_1)}{-2.77 g^{-1} h^2 - 4.38 s n_1^{-1} NE_{LC} + 1.67 s n_1^{-1} - 8.3 g^{-1} TD_1^2 - 11.06 TD_1 h g^{-1}} = \Gamma \text{ (and } \alpha > 0) \qquad (25)$$

If α and β satisfy Equation 25, i.e., α is higher than the threshold (Γ), an increase of Task 1 difficulty level will decrease the value of ϕ_2.

3) Condition 3: $\frac{\partial \phi_2}{\partial TD_1} = 0$:

$$-1.44sn\alpha s^{-1}g^{-2}TD_1^2 + (1.92hsn\alpha s^{-1}g^{-2} + .96sn\beta s^{-1}g^{-1})TD_1$$
$$+(-.48sn\alpha s^{-1}g^{-2}h^2 - 0.76\alpha g^{-1}NE_{LC} + 0.29\alpha g^{-1} - .48sn\beta hs^{-1}g^{-1}) = 0$$
$$\Rightarrow TD_1 = \frac{-(1.92hsn\alpha s^{-1}g^{-2} + .96sn\beta s^{-1}g^{-1}) \pm \sqrt{(1.92hsn\alpha s^{-1}g^{-2} + .96sn\beta s^{-1}g^{-1})^2 + 5.76sn\alpha s^{-1}g^{-2}(-.48sn\alpha s^{-1}g^{-2}h^2 - 0.76\alpha g^{-1}NE_{LC} + 0.29\alpha g^{-1} - .48sn\beta hs^{-1}g^{-1})}}{-2.88sn\alpha s^{-1}g^{-2}}$$
$$\Rightarrow$$
$$E)\ (1.92hsn\alpha s^{-1}g^{-2} + .96sn\beta s^{-1}g^{-1})^2 + 5.76sn\alpha s^{-1}g^{-2}(-.48sn\alpha s^{-1}g^{-2}h^2 - 0.76\alpha g^{-1}NE_{LC} + 0.29\alpha g^{-1} - .48sn\beta hs^{-1}g^{-1})$$
$$= [2.88sn\alpha s^{-1}g^{-2}TD_1 - (1.92hsn\alpha s^{-1}g^{-2} + .96sn\beta s^{-1}g^{-1})]^2$$

Thus,

$$\alpha = \frac{\beta(2.77h - 5.53TD_1)}{-2.77g^{-1}h^2 - 4.38sn_1^{-1}NE_{LC} + 1.67sn_1^{-1} - 8.3g^{-1}TD_1^2 - 11.06TD_1hg^{-1}} = \Gamma \qquad (26)$$

If α and β satisfy Equation 26, an increase of Task 1 difficulty level will not affect the value of ϕ_2.

Chapter 5

Queueing Network Modeling of Driver Workload and Performance

Chapter Summary

Drivers overloaded with information significantly increase the chance of vehicle collisions. Driver workload, a multi-dimensional variable, is measured by both performance-based and subjective measurements and affected by driver age differences. Few existing computational models are able to cover these major properties of driver workload or simulate subjective mental workload and human performance at the same time. We describe a new computational approach for modeling driver performance and workload—a Queueing network approach based on Queueing network theory of human performance (Liu, 1996, 1997) and neuroscience discoveries. This modeling approach not only successfully models mental workload measured by the six NASA-TLX workload scales in terms of subnetwork utilization, but also simulates driving performance, reflecting mental workload from both subjective and performance-based measurements. In addition, it models age differences in workload and performance and allows us to visualize driver mental workload in real-time. Further usage and implementation of the model in designing intelligent and adaptive in-vehicle systems are discussed.

1. Introduction

The expanding usage of in-vehicle systems increases the chance that drivers perform dual tasks in driving, e.g., driving and using a mobile phone concurrently. These dual tasks may impose high information load on drivers, increasing driver mental workload

(Alm & Nilsson, 1995; Wagner, Vercruyssen, & Hancock, 1997; Wickens, Kramer, Vanasse, & Donchin, 1983) which in turn may increase the chance of vehicle collisions by about 4 times compared to a single task condition (Alm & Nilsson, 1995; Redeleier & Tibshirani, 1997; Violanti & Marshall, 1996). Moreover, it is reported that older drivers' crash rates were higher than young drivers (McKelvey & Stamatiadis, 1989) and using in-vehicle systems is one of the main causes of this increase in crash rates since older drivers' information processing efficiency decreases with an increase in age (Hing, Stamatiadis, & Aultman-Hall, 2003). In practice, modeling and predicting driver workload and performance is very useful in designing in-vehicle systems to prevent drivers (especially older drivers) from being overloaded with information (Piechulla, Mayser, Gehrke, & Konig, 2003). Significant costs of implementation and modification can be saved if driver mental workload can be predicted at an early stage of vehicle design.

Several decades of research on mental workload has shown that mental workload has three important properties. First, it is a multidimensional variable (perceptual, cognitive, and motor dimensions) and operators are often capable of reporting the demands on separate workload dimensions (Annett, 2002; Card, Moran, & Newell, 1983; Hendy, Hamilton, & Landry, 1993; Lee & Liu, 2003; Ohsuga, Shimono, & Genno, 2001; Rubio, Diaz, Martin, & Puente, 2004; Tsang & Velaquez, 1996; Xie & Salvendy, 2000). Second, age differences are one of the most important factors in affecting driver workload (Verwey, 2000). Aging causes the slowing of older drivers' information processing in perceptual, cognitive and psychomotor aspects (Hing et al., 2003; Kirby & Nettelbeck, 1991; Salthouse, 1982, 1985). For the same amount of information being processed in the same time period, older drivers usually perceive higher levels of mental workload than young drivers (Feyen & Liu, 1998; Tomporowski, 2003). Third, performance-based measurements alone may not fully reflect mental workload because of the potential dissociation of performance and mental workload (Vidulich & Wickens, 1986). Thus, subjective or physiological measurements of mental workload should be applied in addition to performance-based measurements (Johnson & Proctor, 2004). In this regard, subjective measurements are relatively easy to implement, nonintrusive and inexpensive, and have a high face validity (Johnson & Proctor, 2004; Wickens et al.,

1983). For example, NASA-TLX (National Aeronautic and Space Administration Task Load Index, (Hart & Staveland, 1988)) is one of the most frequently used subjective mental workload scales which reflect the multidimensional property of mental workload (Tsang & Velaquez, 1996). It measures mental workload with six rating scales: mental demand, physical demand, temporal demand, performance, effort, and frustration levels. NASA-TLX has been successfully applied in a number of multi-task system environments (Mcdowd, Vercruyessen, & Birren, 1991).

In accordance with the three properties of mental workload discussed above, a computational model of mental workload is expected to capture the multidimensional property of mental workload and to account for its age differences; it should also model mental workload from both performance-based and subjective measurements. Several computational models have been developed to model mental workload in driving (see Table 5-1). Using control theory, Horiuchi and Yuhara (2000) modeled drivers' mental and physical workload based on lead time constraints and steering wheel angle (Horiuchi & Yuhara, 2000). Lin et al. (2005) modeled driver performance using artificial neural network methods including counter propagation network, the radial basis function network and the back propagation network (Lin, Tang, Zhang, & Yu, 2005). A statistical model was applied to model visual workload/demand in the driving context by Easa and Ganguly (2005) (Easa & Ganguly, 2005): regression analysis was used to determine the best regression model of visual demand with independent variables (e.g., lane width). Assuming the driver as a semiotic system, Goodrich and Boer (1998) modeled mental workload by interactions of several mental model agents (Goodrich & Boer, 1998). Piechulla et al. (2003) estimated driver mental workload by multiplying a weight factor with a basic estimated workload (*w)* based on the road information (e.g., intersection ahead) (Piechulla et al., 2003). Based on the production-rule architecture—ACT-R (Anderson & Lebiere, 1998), Salvucci et al. (2001) developed a model of driving behavior to simulate driver performance in a dual task situation (Salvucci, Boer, & Liu, In press). However, as shown in Table 5-1, few models are able to simulate human performance and mental workload in dual tasks while reflecting the multidimensional property of mental workload. None of these models takes into account the effect of age

differences on driver workload or visualizes mental workload, an important feature for enhanced usability and applicability (Koshman, 2004; Trickett, Trafton, & Schunn, 2000).

Table 5-1 Coverage of driver mental workload in computational models

	Multi-dimensional	Task	Subjective Measurement	Performance-based Measurement	Age Difference	Visual-ization
Queueing Network Model (Wu and Liu, this paper)	Yes	Dual	Yes	Yes	Yes	Yes
Control Theory (Horiuchi et al., 2000)	Mental and physical	Single	Yes	Yes	-	-
Neural Network (Lin, et al., 2005)	Mental only	Single	-	Yes	-	-
Semiotics model (Goodrich et al., 1998)	Mental only	Single	Yes	-	-	-
Statistic Model (Easa et al, 2005)	Visual only	Single	Yes	-	-	-
Engineering Model (Piechulla, et al., 2003)	Mental only	Single	Yes	-	-	-
Rule-based Model (Salvucci, et al., 2001)	Visual/Cog-nitive/Motor	Dual	-	Yes	-	-

-: not covered

In this paper, to model the major properties of driver mental workload summarized in Table 5-1, we describe how to model driver mental workload and performance using a new computational modeling approach—the Queueing network modeling approach (Liu, 1996, 1997). First, we describe the Queueing network mental architecture representing information processing in the mental system and how the model was used to account for subjective mental workload including its age differences. Then, we describe how the model was validated with an experimental study on driver performance and workload.

2. Modeling Mental Workload in Driving

Since subjective mental workload reflects the perception of information processing throughout each trial in a task, the average utilization of a subnetwork ($\bar{\rho}_i$)—the average

utilization of subnetwork *i* in total task time of each trial (*T*)—is regarded as a natural index of subjective mental workload in QN-MHP (see Equation 1). In computational modeling of mental workload, Rouse (1980) modeled mental workload in a single task situation using server utilization as an index of the workload (Rouse, 1980); Just et al. (2003) also regard the capacity utilization as a typical representation of mental workload (Just, Carpenter, & Miyake, 2003). In terms of the physiological mechanism of mental workload, it is also reasonable to use utilization as the index of mental workload: increasing utilization of certain brain regions causes the consumption of more neurotransmitters (e.g., amino acids, norepinephrine/NE, 5-hydroxytryptamine/5-HT) in synaptic transmissions, which in turn increase the perception of mental fatigue (Blomstrand, 2001, 2006; Davis, 1995; Newsholme & Blomstrand, 1996, 2006).

$$\bar{\rho}_i = (\int_0^T \rho_i dt) / T \qquad (0 \le \bar{\rho}_i \le 1) \tag{1}$$

where $\bar{\rho}_i$ can represent the average utilization of visual perceptual subnetwork ($\bar{\rho}_{vp}$), auditory perceptual subnetwork ($\bar{\rho}_{ap}$), cognitive subnetwork ($\bar{\rho}_c$), and motor subnetwork ($\bar{\rho}_m$), respectively. Moreover, based on the definition of each scale in NASA-TLX (Hart & Staveland, 1988), the score of physical demand (*PD*) reflects workload at the motor component, and therefore it is in direct proportion to the average utilization of motor subnetwork ($\bar{\rho}_m$) (see Equation 2); the scores of temporal demand (*TD*), frustration (*FR*), performance (*PE*) and effort (*EF*) represent the overall workload in the system, which is reflected by the average utilization of all the subnetworks (see Equations 3-6); the score of mental demand (*MD*) is judged based on the perceptual and cognitive demands (how much perceptual and mental activities were required, (Hart & Staveland, 1988)), and therefore it is in direct proportion to the average utilization of perceptual and cognitive subnetworks (see Equation 7).

$$PD = a\bar{\rho}_m + b \qquad (0 \le PD \le 100) \tag{2}$$

$$TD = a(\sum_{All\ i} \bar{\rho}_i) / 4 + b \qquad (0 \le TD \le 100) \tag{3}$$

$$EF = a(\sum_{All\ i} \bar{\rho}_i) / 4 + b \qquad (0 \le EF \le 100) \tag{4}$$

$$PE = a(\sum_{All\ i} \bar{\rho}_i) / 4 + b \qquad (0 \le PE \le 100) \tag{5}$$

$$FR = a(\sum_{All\ i} \bar{\rho}_i)/4 + b \qquad (0 \le FR \le 100) \tag{6}$$

$$MD = a(\sum_{i=ap,vp,c} \bar{\rho}_i)/3 + b \qquad (0 \le MD \le 100) \tag{7}$$

where parameters *a* and *b* are constants in representing the direct proportional relation between the averaged utilizations and subjective responses (*a>0*). Equations 2-7 are implemented in the simulation model to generate subjective workload responses (See (Liu, Feyen, & Tsimhoni, in press) for descriptions of how QN-MHP is able to simulate driver performance).

In addition, research evidence suggests that the major difference in information processing between the older and young adults is a generalized slowing in information processing speed for older adults (Bunce & Macready, 2005; Hing et al., 2003; Mcdowd et al., 1991); therefore, considering age differences, the information processing speed at server *j* (μ_j) in the network is:

$$\mu_j = (\frac{1}{A})\mu_{0,j} \tag{8}$$

where *A* is a factor of aging ($A \ge 1$): the value of *A* is directly proportional to the driver's age; $\mu_{0,j}$ is the original processing speed of server *j* for young adults in QN-MHP (Liu et al., in press). Moreover, according to the traffic intensity function in Queueing network theory (Gross, 1998), utilization of a certain subnetwork *i* (ρ_i) (the fraction of time the subnetwork *i* is processing entities in a defined time period) is in inverse proportion to the average processing speed of all the servers in the subnetwork ($\bar{\mu}_i$) (see Equation 9).

$$\rho_i \equiv \frac{\lambda_i}{C_i \bar{\mu}_i} = \frac{\lambda_i}{C_i[(\sum_{j=1}^{Ci} \mu_j)/C_i]} = \frac{\lambda_i}{\sum_{j=1}^{Ci} \mu_j} \qquad (0 \le \rho_i \le 1) \tag{9}$$

where λ is the arrival rate of the subnetwork and C_i is the total number of servers in the subnetwork.

Mathematically, we can derive that the expected subjective mental workload of older drivers is equal to or greater than young drivers from the equations above. Combining Equations 1-9 above, we have:

$$PD = a(\int_0^T \rho_m dt)/T + b = a(\int_0^T \frac{\lambda_m}{\sum_{j=1}^{Cm} \mu_j} dt)/T + b \tag{10}$$

$$= a(\int_0^T \frac{\lambda_m}{\sum_{j=1}^{Cm} (\mu_{0,j}/A)} dt)/T + b = a(\int_0^T \frac{\lambda_m}{(\frac{1}{A})\sum_{j=1}^{Cm} \mu_{0,j}} dt)/T + b$$

$$= Aa(\int_0^T \frac{\lambda_m}{\sum_{j=1}^{Cm} \mu_{0,j}} dt)/T + b$$

Similarly, we can derive:

$$TD = EF = PE = FR = Aa(\int_0^T \frac{\lambda_{all\ i}}{\sum_{j=1}^{Call\ i} \mu_{0,j}} dt)/4T + b \tag{11}$$

$$MD = Aa(\int_0^T \frac{\lambda_{i=vp,ap,c}}{\sum_{j=1}^{Ci=vp,ap,c} \mu_{0,j}} dt)/3T + b \tag{12}$$

If the arrival rate (λ_m, $\lambda_{all\ i}$, $\lambda_{i=vp,ap,c}$) and the total task time of each trial (T) remain the same in different age groups, (13)

$\because A \geq 1$

$\therefore PD_{old} \geq PD_{young}$, $TD_{old} \geq TD_{young}$, $EF_{old} \geq EF_{young}$, $PE_{old} \geq PE_{young}$, $FR_{old} \geq FR_{young}$, and $MD_{old} \geq M$

Similarly, age differences in driving performance can also be quantified. In Queueing network theory, the performance of a network (HP) is in direct proportion to its servers' processing speeds (see Equation 14, (Gross, 1998)).

$$HP = (1/A) \underset{all\ j}{\Omega} (\mu_j) \tag{14}$$

where Ω is a function describing a negative relationship between human performance and all of the servers' processing times as variables. Since $A \geq 1$, the expected performance of older drivers is equal to or lower than young drivers.

3. An Experiment on Driver Workload and Performance

Feyen and Liu (1998) conducted an experimental study in which drivers of two age groups performed a dual task of vehicle steering and button-pressing in a simulator (see Figure 5-1) (Feyen & Liu, 1998). In the primary vehicle steering task, subjects were asked to keep the vehicle in control by maintaining the lane position and the same driving

speed (45 miles/hour). In the secondary button-pressing task, subjects were instructed to press one of the buttons on a panel mounted on the right side of the steering wheel when they saw a command presented on the display.

Figure 5-1 A subject responded to a command prompt during driving (Feyen & Liu, 1998)

The independent variables were: 1) the age group of the subjects (four young drivers, 17-30 years old; four older drivers, 61-75 years old); 2) the number of buttons on the panel with 3 difficulty levels (2, 4, or 6 buttons). The dependent variables included: 1) the lane position deviation difference from the baseline (LPDDB) and it was calculated by subtracting a baseline lane position standard deviation from the lane position standard deviation during the task time segment (a negative value indicated a more stable lane positioning while a positive value indicated a less stable lane positioning); 2) reaction time of the button-pressing task as a performance-based mental workload measurement: the time interval between the command presentation and pressing of a button; 3) subjective ratings on the 6 scales of NASA-TLX after each trial. Since overall mental workload calculated by weighting the scales does not appear to add to the sensitivity of the NASA-TLX (Eggemeier & Wilson, 1991; Johnson & Proctor, 2004), the overall mental workload was not collected in this experimental study (Feyen & Liu, 1998).

4. Simulation Results and Validation

By implementing Equations 2-7 described in the previous section in the Queueing network simulation model, the simulation results are obtained and then compared with

the experimental results (see Appendix for the method of setting parameters in these equations).

4.1 Driver Workload

Figure 5-2 shows the comparison between the simulation results and experimental results for each of the scales of NASA-TLX. Table 5-2 summarizes the R square and RMS of the model for each scale.

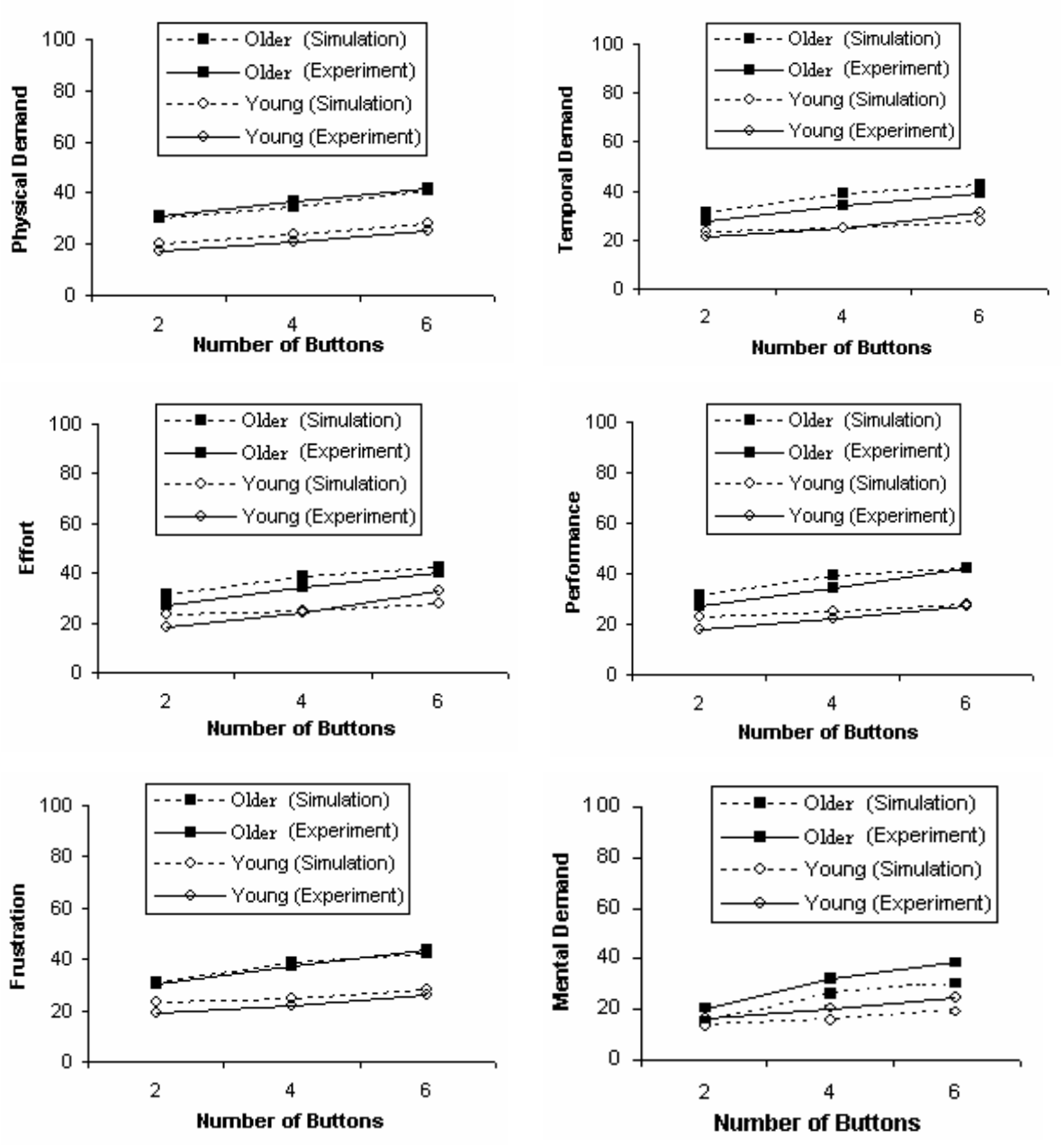


Figure 5-2 Subjective mental workload in the experimental study of Feyen and Liu (1998) (solid lines) in comparison with the Queueing network simulation results (dashed lines).

Table 5-2 R Square and RMS of the Model for Each Scale

Scales	Young Drivers		Older Drivers	
	R Square	RMS	R Square	RMS
Physical Demand	.99	1.83	.95	1.74
Temporal Demand	.99	1.26	.97	3.92
Effort	.99	2.43	.97	4.01
Performance	.97	2.37	.93	3.79
Frustration	.99	1.78	.95	1.69
Mental Demand	.99	1.52	.99	6.56
Average	.99	2.11	.96	3.62

4.2 Driver Performance

Figure 5-3 and Figure 5-4 show the simulation results of driver performance in comparison with the experimental results (LPDDB: R square=.98, RMS=.03; RT to the secondary task: R square=.94, RMS=50.4).

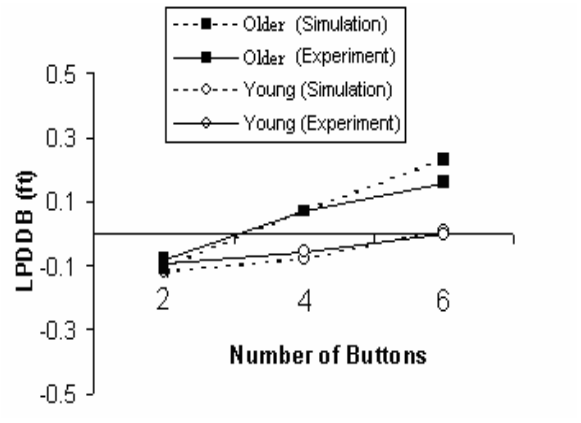


Figure 5-3 LPDDB in the experimental study (solid lines) in comparison with simulation results (dashed lines)

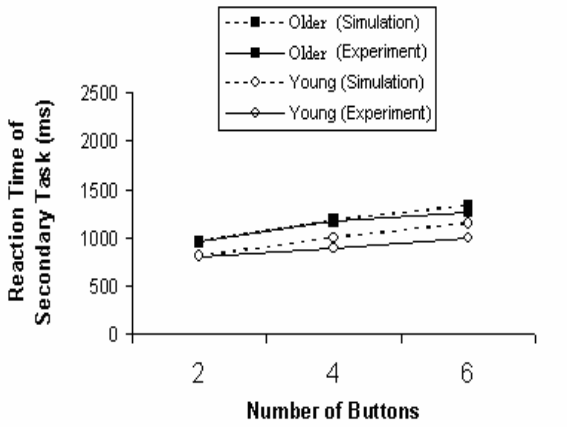
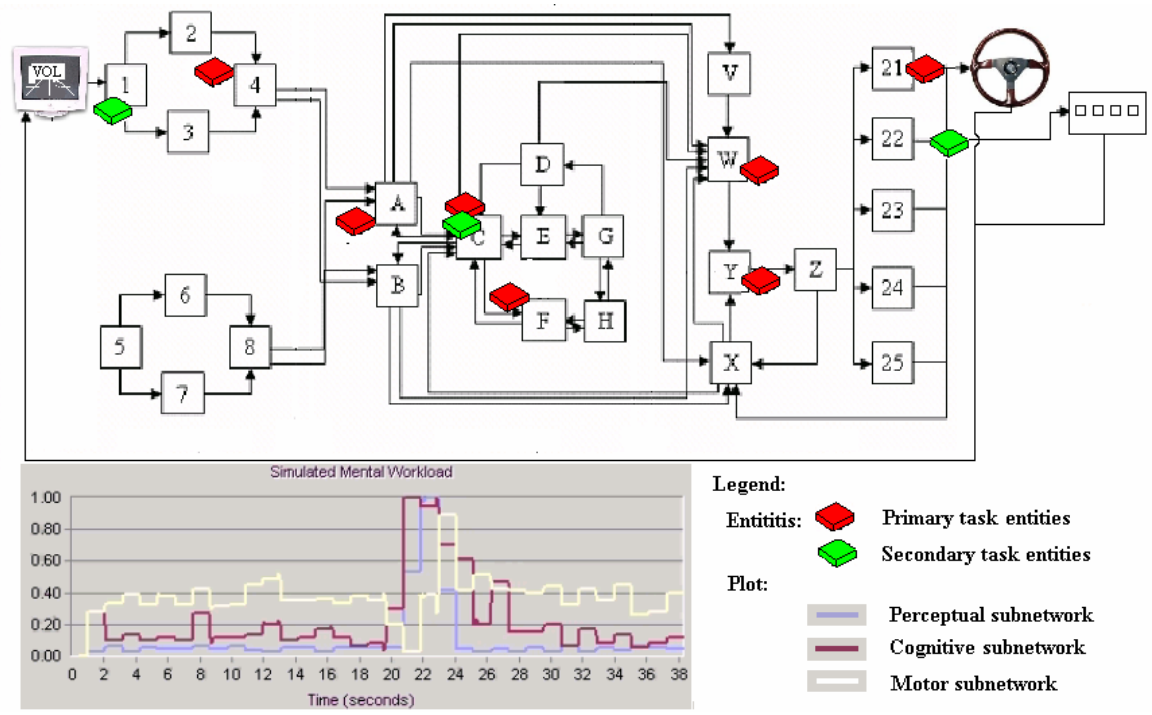


Figure 5-4 Reaction time to the secondary task in the experimental study (solid lines) in comparison with simulation results (dashed lines)

4.3 Workload Visualization

As shown in Figure 5-5 and Figure 5-5, the model allows a modeler to visualize the overall and the subnetwork mental workload by observing the entity activities and the network flow patterns during the simulation. Dynamic values of subnetwork utilizations are also shown in the simulation so that the user of the model can observe the dynamic changes of mental workload in real-time.

(a) High mental workload condition



(b) Low mental workload condition

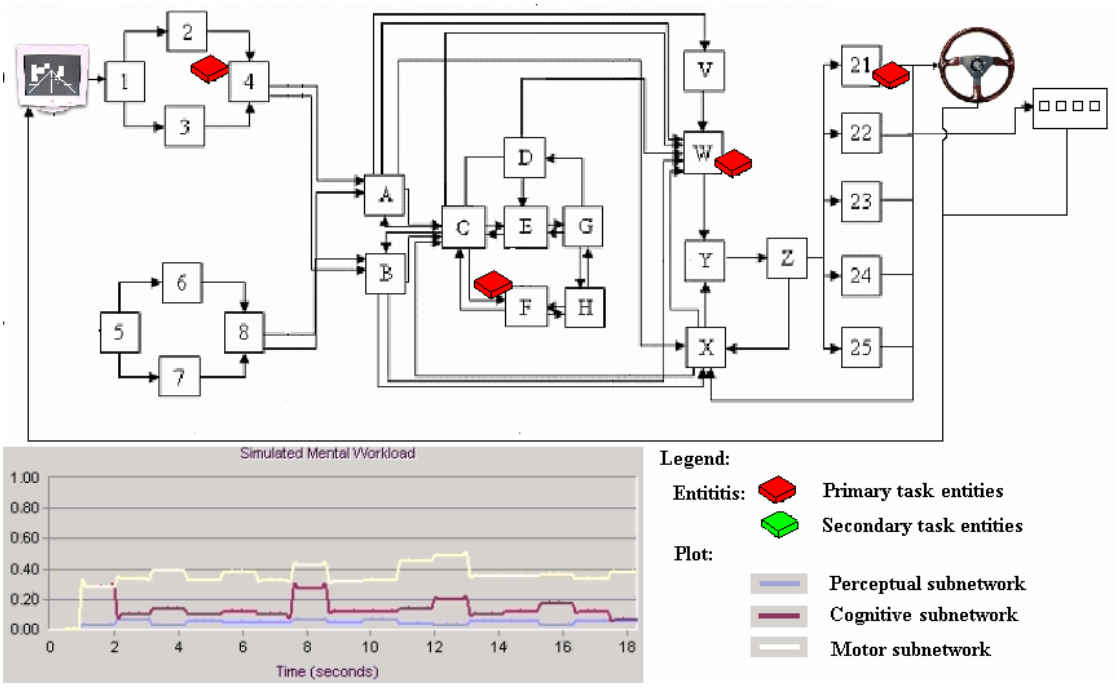


Figure 5-5 Visualizing mental workload in QN-MHP during the simulation

A short movie clip can be seen on the website:
http://www.umich.edu/~yililiu/ or http://www.acsu.buffalo.edu/~changxu/

5. Conclusions

We described a Queueing network modeling approach to model subjective mental workload and multitask performance including their age differences in a driving context, reflecting the multidimensional nature of mental workload from both subjective and performance-based measurements. Few existing computational models are able to simulate all of these major properties of driver workload at the same time in dual task situations. This modeling work offers a natural quantification of subjective mental workload with subnetwork utilization and initiates a step in connecting the output of an engineering model with the measurement of the subjective mental workload.

In practice, this modeling approach has several significant values for user interface design of in-vehicle systems. First, the Queueing network simulation model is able to predict and visualize where workload is concentrated in the perceptual (auditory or visual), cognitive or motor subnetworks. For example, if the visual perceptual workload predicted by the model is heavy in certain circumstances, interface designers can design the user interface to present auditory information and use the model to test whether driver's visual perceptual workload can be reduced and whether the design creates other workload and performance problems.

Second, an accurate estimation of mental workload is vital for the design of intelligent or adaptive driver support and warning systems. Typically, these systems rely on computational models to estimate driver workload and propose actions to prevent traffic accidents (e.g., redirecting messages into a voice mailbox, see related papers (Jennifer & Picard, 2005; Piechulla et al., 2003; Vahidi & Eskandarian, 2003). By implementing this computational model into these systems, driver mental workload in different information processing components can be estimated more accurately.

Third, the capability of mental workload visualization is unique feature of the current modeling approach. Information visualization is an important step to increase the usability and face validity of a model (Trickett et al., 2000) and allows users of the model to view the input, processing activities, and output of the model intuitively. Moreover, the dynamic change of mental workload in perceptual (auditory or visual), cognitive and motor subnetworks can be viewed and estimated directly in real time. This may help users of the model predict when mental workload reaches "red-line" (reflected by a

certain level of the average subnetwork utilization) as well as by how much and for how long it exceeds that red-line (see Figure 5-5).

In addition, the current modeling work accounts for age differences by simply considering an aging factor in servers' processing times, and it is consistent with findings in empirical studies on age differences. For example, Salthouse (1982, 1985) suggested that age differences are simply a function of a generalized slowing of information processing in older adults (Salthouse, 1982, 1985). Moreover, the current modeling work simplifies the estimation of 4 scales of NASA-TLX (TD, EF, PE and FR) by using the same index in the network (averaged utilization of all the subnetworks). This simplification is supported by empirical studies developing and using NASA-TLX in dual tasks. Hart and Staveland (1988) found that there is a high correlation among TD, EF, PE and FR (correlation efficient >.65) when NASA-TLX is used to measure the subjective mental workload in a dual task (Hart & Staveland, 1988).

Even though the current modeling approach demonstrates its effectiveness and simplicity in accounting for the six mental workload scales in NASA-TLX in the driving context, several important topics need to be investigated in future research. The model in the future may need to differentiate the workload scores in the four scales (*PE, EF, FR, TD*) since they may stem from different psychological mechanisms. For example, frustration (*FR*) may be not only related to the utilization of resources or capacities in the system, but also affected by a person's subjective sensitivity to temporal pressure. Compared with mental workload measured by the other scales related to the utilization of resources or capacities, mental workload measured by the performance (*PE*) scale may result from a complex subjective self-evaluation of one's performance including his or her prior experience in performing the same or relevant tasks, self-confidence, and self-evaluation strategies. This is also relevant to the modeling of individual differences, which is a very important topic to be covered in our future research and development of the Queueing network model. In addition, even though the overall mental workload calculated by weighting the scales does not appear to add to the sensitivity of the NASA-TLX (Eggemeier & Wilson, 1991; Johnson & Proctor, 2004), its value is another important topic to be investigated in the future because evaluation of some systems only need one index to represent mental workload.

We are extending the current modeling approach to other related mental workload research including modeling physiological measurements of mental workload. Overall, our current work demonstrates the value of the Queueing network modeling approach in modeling and quantifying driver subjective mental workload and performance.

Acknowledgments This article is based upon work supported by the National Science Foundation under Grant No. NSF 0308000. However, any opinions, findings and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation (NSF).

Reference

Alm, H., & Nilsson, L. (1995). The Effects of a Mobile Telephone Task on Driver Behavior in a Car Following Situation. *Accident Analysis and Prevention, 27*(5), 707-715.

Anderson, J. R., & Lebiere, C. (1998). *The Atomic Components of Thought*: Lawrence Erlbaum Associates.

Annett, J. (2002). Subjective rating scales: science or art? *Ergonomics, 45*(14), 966-987.

Bear, M. F., Connors, B. W., & Paradiso, M. A. (2001). *Neuroscience: exploring the brain* (8th ed.). Baltimore, MD: Lippincott Williams & Wilkins,.

Blomstrand, E. (2001). Amino acids and central fatigue. *Amino Acids, 20*(1), 25-34.

Blomstrand, E. (2006). A role for branched-chain amino acids in reducing central fatigue. *Journal of Nutrition, 136*(2), 544S-547S.

Bunce, D., & Macready, A. (2005). Processing speed, executive function, and age differences in remembering and knowing. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 58*(1), 155-168.

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

Card, S., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hinsdale, NJ: Lawrence Erlbaum.

Cook, A. S., & Woollacott, M. H. (1995). *Motor Control: Theory and Practical Applications*: Williams& Wilkins.

Davis, J. M. (1995). Carbohydrates, Branched-Chain Amino-Acids, and Endurance - the Central Fatigue Hypothesis. *International Journal of Sport Nutrition, 5*, S29-S38.

Easa, S., & Ganguly, C. (2005). Modeling driver visual demand on complex horizontal alignments. *Journal of Transportation Engineering-ASCE, 131*(8), 583-590.

Eggemeier, F. T., & Wilson, G. F. (1991). Performance-based and subjective assessment of workload in multi-task environment. In D. L. Damos (Ed.), *Multiple Task Performance*: Taylor & Francis.

Faw, B. (2003). Pre-frontal executive committee for perception, working memory, attention, long-term memory, motor control, and thinking: A tutorial review. *Consciousness and Cognition, 12*(1), 83-139.

Feyen, R. (2002). *Modeling Human Performance using the Queueing Network-Model Human Processor (QN-MHP)*. University of Michigan, Doctor Thesis, Ann Arbor.

Feyen, R., & Liu, Y. (1998). *Driver Performance Model for Instrument Panel Control Use, Technical Report: 99-03*: Department of Industrial and Operations Engineering, University of Michigan.

Fletcher, P. C., & Henson, R. N. A. (2001). Frontal lobes and human memory - Insights from functional neuroimaging. *Brain, 124*, 849-881.

Gilbert, P. F. C. (2001). An outline of brain function. *Cognitive Brain Research, 12*(1), 61-74.

Goodrich, M. A., & Boer, E. R. (1998). *Semiotics and mental modes: modeling automobile driver behavior.* Paper presented at the IEEE ISIC/CIRA/ISAS Joint Conference, Gaithersburg, MD.

Gordon, A. M., & Soechting, J. F. (1995). Use of tactile afferent information in sequential finger movements. *Experimental Brain Research, 107*, 281-292.

Gross, D. (1998). *Fundamentals of queueing theory*: John Wiley & Sons, Inc.

Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload,* . Amsterdam: North-Holland.

Hendy, K. C., Hamilton, K. M., & Landry, L. N. (1993). Measuring Subjective Workload - When Is One Scale Better Than Many. *Human Factors, 35*(4), 579-601.

Hing, J. Y. C., Stamatiadis, N., & Aultman-Hall, L. (2003). Evaluating the impact of passengers on the safety of older drivers. *Journal of Safety Research, 34*(4), 343-351.

Horiuchi, S., & Yuhara, N. (2000). An analytical approach to the prediction of handling qualities of vehicles with advanced steering control system using multi-input driver model. *Journal of Dynamic Systems Measurement and Control-Transactions of the ASME, 122*(3), 490-497.

Jennifer, H. A., & Picard, R. W. (2005). Detecting Stress During Real-World Driving Tasks Using Physiological Sensors. *IEEE Transactions on Intelligent Transportation Systems, 6*(2), 156-166.

Johnson, A., & Proctor, R. W. (2004). *Attention*. London: SAGE Publications.

Just, M. A., Carpenter, P. A., & Miyake, A. (2003). Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work. *Theoretical Issues in Ergonomics Science, 4*(1), 56-88.
Kansaku, K., Hanakawa, T., Wu, T., & Hallett, M. (2004). A shared neural network for simple reaction time. *Neuroimage, 22*(2), 904-911.
Kirby, N. H., & Nettelbeck, T. (1991). Speed of information processing and age. *Personality & Individual Differences*(12), 183-188.
Koshman, S. (2004). Comparing usability between a visualization and text-based system for information retrieval. *Journal of Documentation, 60*(5), 565-580.
Lee, Y. H., & Liu, B. S. (2003). Inflight workload assessment: Comparison of subjective and physiological measurements. *Aviation Space and Environmental Medicine, 74*(10), 1078-1084.
Lim, J., & Liu, Y. (2004). *A queueing network model of menu selection and visual search.* Paper presented at the Proceedings of the 48 Annual Conference of the Human Factors and Ergonomics Society, New Orleans, Louisiana, USA.
Lin, Y., Tang, P., Zhang, W. J., & Yu, Q. (2005). Artificial neural network modeling of driver handling behaviour in a driver-vehicle-environment system. *International Journal of Vehicle Design, 37*(1), 24-45.
Liu, Y. (1996). Queueing network modeling of elementary mental processes. *Psychological Review, 103*(1), 116-136.
Liu, Y. (1997). Queueing network modeling of human performance of concurrent spatial and verbal tasks. *IEEE Transactions on Systems Man and Cybernetics Part a-Systems and Humans, 27*(2), 195-207.
Liu, Y., Feyen, R., & Tsimhoni, O. (2006). Queueing Network-Model Human Processor (QN-MHP): A Computational Architecture for Multitask Performance. *ACM Transaction on Human Computer Interaction, 13*(1), 37-70.
Manoach, D. S., Schlaug, G., Siewert, B., Darby, D. G., Bly, B. M., Benfield, A., et al. (1997). Prefrontal cortex fMRI signal changes are correlated with working memory load. *Neuroreport, 8*(2), 545-549.
Mcdowd, J., Vercruyessen, M., & Birren, J. E. (1991). Aging, divided attention, and dual-task performance. In D. L. Damos (Ed.), *Multiple Task Performance*: Taylor & Francis.
McKelvey, F. X., & Stamatiadis, N. (1989). Highway accident patterns in Michigan related to older drivers. *Transportation Research Record, 1210*, 53- 57.
Mitz, A. R., Godschalk, M., & Wise, S. P. (1991). Learning-Dependent Neuronal-Activity in the Premotor Cortex - Activity During the Acquisition of Conditional Motor Associations. *Journal of Neuroscience, 11*(6), 1855-1872.
Mustovic, H., Scheffler, K., Di Salle, F., Esposito, F., Neuhoff, J. G., Hennig, J., et al. (2003). Temporal integration of sequential auditory events: silent period in sound pattern activates human planum temporale. *Neuroimage, 20*(1), 429-434.
Newsholme, E. A., & Blomstrand, E. (1996). The plasma level of some amino acids and physical and mental fatigue. *Experientia, 52*(5), 413-415.
Newsholme, E. A., & Blomstrand, E. (2006). Branched-chain amino acids and central fatigue. *Journal of Nutrition, 136*(1), 274S-276S.
Ohsuga, M., Shimono, F., & Genno, H. (2001). Assessment of phasic work stress using autonomic indices. *International Journal of Psychophysiology, 40*(3), 211-220.
Piechulla, W., Mayser, C., Gehrke, H., & Konig, W. (2003). Reducing drivers' mental workload by means of an adaptive man-machine interface. *Transportation Research, Part F, 6*, 233-248.
Proctor, R. W., Vu, K. L., & Pick, D. F. (2005). Aging and response selection in spatial choice task. *Human Factors, 47*(2), 250-270.
Redeleier, D. A., & Tibshirani, R. J. (1997). Associatation between Cellular-Telephone Calls and Motor Vehicle Collisions. *New England Journal of Medicine, 336*(7), 453-458.
Rieke, F., Warland, D., R.S., R., & Bialek, W. (1997). *Spikes: Exploring the Neural Code (Computational Neuroscience)*: MIT Press.
Roland, P. E. (1993). *Brain activation*. New York, NY: Wiley-Liss.
Rolls, E. T. (2000). Memory systems in the brain. *Annual Review of Psychology, 51*, 599-630.
Rouse, W. B. (1980). *Systems Engineering Models of Human-Machine Interaction*. New York: North Holland.

Rubio, S., Diaz, E., Martin, J., & Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Applied Psychology-an International Review-Psychologie Appliquee-Revue Internationale, 53*(1), 61-86.

Salthouse, T. A. (1982). *Adult Cognition. An Experimental Psychology of Human Aging*. New York: Springer-Verlag.

Salthouse, T. A. (1985). *Speed of behavior and its implications for cognition*. New York: Van Nostrand Reinhold.

Salvucci, D. D., Boer, E. R., & Liu, A. (2001). Toward an Integrated Model of Driver Behavior in a Cognitive Architecture. *Transportation Research Record, 1779*, 9-16.

Simon, S. R., Meunier, M., Piettre, L., Berardi, A. M., Segebarth, C. M., & Boussaoud, D. (2002). Spatial attention and memory versus motor preparation: Premotor cortex involvement as revealed by fMRI. *Journal of Neurophysiology, 88*(4), 2047-2057.

Smith, E. E., & Jonides, J. (1998). Neuroimaging analyses of human working memory. *Proc. Natl. Acad. Sci. USA, 95*, 12061-12068.

Taylor, J., Horwitzc, B., Shaha, N. J., Fellenzb, W. A., Mueller-Gaertnera, H.-W., & Krausee, J. B. (2000). Decomposing memory: functional assignments and brain traffic in paired word associate learning. *Neural Networks, 13*, 923-940.

Tomporowski, P. D. (2003). Performance and perceptions of workload among young and older adults: effects of practice during cognitively demanding tasks. *Educational Gerontology, 29*, 447-466.

Trickett, S. B., Trafton, J. G., & Schunn, C. D. (2000). *Blobs, dipsy-doodles and other funky things: Framework anomalies in exploratory data analysis.* Paper presented at the Proceedings of the 22nd Annual Conference of the Cognitive Science Society, Mahwah, NJ.

Tsang, P. S., & Velaquez, V. L. (1996). Diagnosticity and multidimensional subjective workload ratings. *Ergonomics, 39*(3), 358-381.

Vahidi, A., & Eskandarian, A. (2003). Research Advances in Intelligent Collision Avoidance and Adaptive Cruise Control. *IEEE Transactions on Intelligent Transportation Systems, 4*(3), 143-153.

Verwey, W. B. (2000). On-line driver workload estimation. Effects of road situation and age on secondary task measures. *Ergonomics, 43*(2), 187-209.

Vidulich, M. A., & Wickens, C. (1986). Causes of dissociation between subjective workload measures and performance: Caveats in the use of subjective assessments. *Applied Ergonomics, 17*, 291-296.

Violanti, J. M., & Marshall, J. R. (1996). Cellular phones and traffic accidents: an epidemiological approach. *Accident Analysis & Prevention, 28*(2), 265-270.

Wagner, D., Vercruyssen, M., & Hancock, P. (1997). A computer-based methodology for evaluating the content of variable message signage. *ITS Journal, 3*(4), 353-373.

Wickens, C., Kramer, A., Vanasse, L., & Donchin, E. (1983). Performance of Concurrent Tasks - a Psychophysiological Analysis of the Reciprocity of Information-Processing Resources. *Science, 221*(4615), 1080-1082.

Wu, C., & Liu, Y. (2004a). *Modeling Behavioral and Brain Imaging Phenomena in Transcription Typing with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA.

Wu, C., & Liu, Y. (2004). *Modeling human transcription typing with Queueing network-model human processor.* Paper presented at the Proceedings of the 48th Annual Meeting of Human Factors and Ergonomics Society, New Orleans, Louisiana, USA.

Wu, C., & Liu, Y. (2004b). *Modeling Psychological Refractory Period (PRP) and Practice Effect on PRP with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA.

Xie, B., & Salvendy, G. (2000). Review and reappraisal of modeling and predicting mental workload in single-and multi-task environments. *Work & Stress, 14*(1), 74-99.

Appendix

In simulating subjective mental workload, the values of parameters *a* and *b* in Equations 2-7 are estimated based on the parameter setting method in a classic cognitive modeling work (Byrne & Anderson, 2001)—*a* and *b* are estimated only for the physical demand scale (change the value of these two parameters to generate the maximum fitness between the modeling results and experimental results), and then the same value is used to estimate subjective responses on the other 5 scales. Therefore, no free parameter is used in estimating the subjective responses in all the other 5 scales for young and older drivers (free parameter refers to parameters whose value is adjusted by researchers so that the modeling results fit the experimental results). Moreover, based on the method in calculating R square, the high R square values indicate that without using parameter *a* and *b* in Equations 2-7, the average subnetwork utilizations in Equations 2-7 are able to predict the variance of subjective mental workload accurately. In addition, no free parameter is used in predicting human performance.

The aging factor, *A*, is set according to a review of Proctor et al. (2005): Proctor et al. reviewed seven experimental studies and found that the mean reaction time for younger adults (24 years old on average) in spatial-visual choice RT task is 417 ms (compatible condition: 369 ms; incompatible condition: 465 ms); mean RT for older adults (70 years old on average) is 527 ms (compatible condition: 457 ms; incompatible condition: 597 ms) (see Table 1 in the empirical study (Proctor, Vu, & Pick, 2005)). Therefore, A=1.26 (527/417=1.26) is selected for older drivers; and A=1 for young drivers (Proctor et al., 2005).

Chapter 6
Development of an Adaptive Workload Management System using Queueing Network-Model of Human Processor (QN-MHP)

Chapter Summary

Drivers overloaded with information from in-vehicle systems significantly increase the chance of vehicle collisions. Developing adaptive workload management systems (AWMS) to dynamically control the rate of messages from these in-vehicle systems is one of the solutions to this problem. However, existing AWMS do not use a model of driver to estimate workload and only suppress or redirect messages without changing the rate of messages from the in-vehicle systems. In this work, we propose a prototype of a new adaptive workload management system (QN-MHP AWMS) which includes: a model of driver workload based on queueing network theory of human performance (Liu, 1996, 1997; Liu, Feyen, and Tsimhoni, 2006) estimating driver workload in different driving situations, and a message controller dynamically controlling the rate of messages presented to drivers. QN-MHP AWMS was able to adapt the rate of messages to the properties of the secondary task, driving conditions (speeds and curvatures) and characteristics of drivers (age). A corresponding experimental study was conducted to validate the potential effectiveness of this system in reducing driver workload and improving driver performance.

1. Introduction

With the development of in-vehicle system technology, more and more in-vehicle information and entertainment systems (e.g., navigation aides, mobile phones, email, web browsers, vehicle-to-vehicle communication systems, and traffic information displays) have been installed in vehicles. Multitasking between driving and using these systems may impose high information load on drivers, increasing drivers' mental workload (Alm & Nilsson, 1995; Wagner, Vercruyssen, & Hancock, 1997; Wickens, Kramer, Vanasse, & Donchin, 1983) which in turn increases the chance of vehicle collisions comparing to a single driving condition (Violanti and Marshall, 1996; Alm & Nilsson, 1995). In addition, besides multitasking in the vehicles of regular drivers, multitasking in driving becomes more common for drivers with special duties. For example, police officers need to drive, communicate with other police officers, and monitor the speed of other cars via radar systems at the same time; ambulance vehicles drivers need to steer vehicles, navigate vehicles to patients' locations, and communicate with dispatches and hospitals at the same time; fire fighting vehicles drivers may also need to steer and navigate vehicles to target locations and communicate with headquarters at the same time to know situations of target locations.

Recently, several adaptive workload management systems have been developed as one of the possible solutions to reduce driver mental workload via the design and use of adaptive workload management systems (AWMS) (Piechulla, Mayser, Gehrke, & Konig, 2003). Recently, several adaptive workload management systems have been developed (See Table 6-1). Some available systems include BMW's phone adaptive system (Piechulla, et al., 2003) and Toyota's voice adaptive system (Uchiyama et al., 2004) (see reviews in Green, 2004 and Kantowitz, 2004). There are two important components in these systems. First, to estimate driver workload, these adaptive systems are able to collect current driving information such as steering wheel angle and lane position and then use computational algorithms to directly estimate the current workload of the driver. Second, based on these estimations of driver workload, the systems propose corresponding actions to reduce driver workload, i.e., suppressing messages from in-vehicle systems (Uchiyama et al., 2004) or redirecting messages into a voice mailbox when the driver's estimated mental workload is high (Piechulla, et al., 2003).

Table 6-1 Summary of four major adaptive workload management systems (AWMS)

	Workload Estimation (Human)		Message Management (System)
Existing AWMS	Human Model Used	Dual Task (2^{nd} Task's Property)	
1. Phone Adaptive System (BMW, Germany) (Piechulla, et al., 2003)	No	Single (Not considered)	Redirect to a phone mailbox only
2. Voice Adaptive System (Toyota, Japan) (Uchiyama et al., 2004)	No	Single (Not considered)	Suppress only
3. SaveIt Adaptive System (Delphi, 2004-2007)	Under Development		Under Development
4. In-Vehicle Message System (Leeds, UK) (Jamson, et al., 2004)	No	Single (Not considered)	Wait for driver's response only

There are two important human factors aspects of these adaptive workload management systems which need further improvements: first, at the human end, a model of driver cognitive system might be useful in these systems so that drivers' workload can be estimated in multitasking situation, especially quantification of the effects of the secondary task on driver workload (e.g., the processing time of the secondary task in different components—perceptual, cognitive and motor—of the cognitive system); second, at the system end, the all-or-none solution (suppressing or redirecting messages from the in-vehicle systems) might be too simple and a more general solution might to treat the delay between messages as a continuous variable (range from 0 to several seconds) whose value is set depending on different driving situations. In addition, there are two potential problems if the in-vehicle messages are controlled by a driver's response: the drivers need additional actions to turn on (or off) the device and drivers may not be able to manage or prioritize messages from the in-vehicle and the primary task (see a review of Haigney & Westerman 2001, discussing effects of concurrent mobile phone use on driving).

Among various models in quantifying mental workload, QN-MHP is able to cover many important features of driver workload including its multi-dimensional properties, workload in single and dual task, age differences, prediction of subjective and physiological workload, as well as workload visualization (See Table 6-2).

In this work, first, we propose a new adaptive workload management system (QN-MHP AWMS) which includes: a) a model of driver workload to estimate driver workload based on research in cognitive modeling; b) a message controller which dynamically controls the rate of messages in various driving situations (Section 2). Second, we describe how QN-MHP can be used to simulate driver workload and performance in a typical multitasking situation in driving (Sections 3 and 4). Third, a corresponding experimental study to validate the potential effectiveness of this system in reducing the driver workload is described (Section 5).

Table 6-2 Coverage of the existing models in modeling driver workload

Models	Multi-dimensional	Dual Task	Task Covered	Age Differences	Subjective Workload Prediction	Performance Prediction	Physiological Prediction	Dynamic	Workload Visualization
Control Theory (Horiuchi et al., 2000)	Mental and physical	-	Steering	-	Yes	Yes	-	Yes	-
Neural Network (Lin, et al., 2005)	-	-	Steering	-	Yes	Yes	-	-	-
Rule-based Model (Salvucci, et al., in press)	Visual /Cognitive/ Motor	Yes	Steering & phone dialing	-	-	Yes	-	Yes	-
Statistic Model (Easa and Ganguly, 2005)	Visual	-	Steering	-	Yes	-	-	Yes	-
Probability Model (Vadeby, 2004)	-	-	Steering	Yes	-	Yes	-	-	-
Piechulla 's Engineering Model (Piechulla, et al., 2003)	-	-	Steering	-	Yes	-	-	Yes	-
QN-MHP (Liu, Feyen, & Tsimhoni, 2006) (Wu & Liu, 2006a,2006b,2006c)	Visual /Auditory /Cognitive/ Motor	Yes	Steering & operating in-vehicle system	Yes	Yes	Yes	Yes	Yes	Yes

-: not covered

2. Designing the Prototype of QN-MHP Adaptive Workload Management System (QN-MHP AWMS)

The purpose of QN-MHP adaptive workload management system (QN-MHP AWMS) is to regulate the rate of messages from the in-vehicle based on driving condition and properties of the secondary task, so that the drivers' workload will be reduced effectively. Figure 6-1 shows the prototype of the adaptive system which is composed of two parts: QN-MHP and a message controller (MC). QN-MHP AWMS receives three types of information: 1) driving conditions (e.g., current driving speed and curvatures); 2) properties of messages from in-vehicle systems (processing time at the perception, cognitive and motor part); 3) properties of driver (e.g., age and level of driving experience). QN-MHP simulates the driver workload and performance, and the message controller regulates the rate of messages in real time and outputs the messages to the driver based on optimal rate derived from the simulation results.

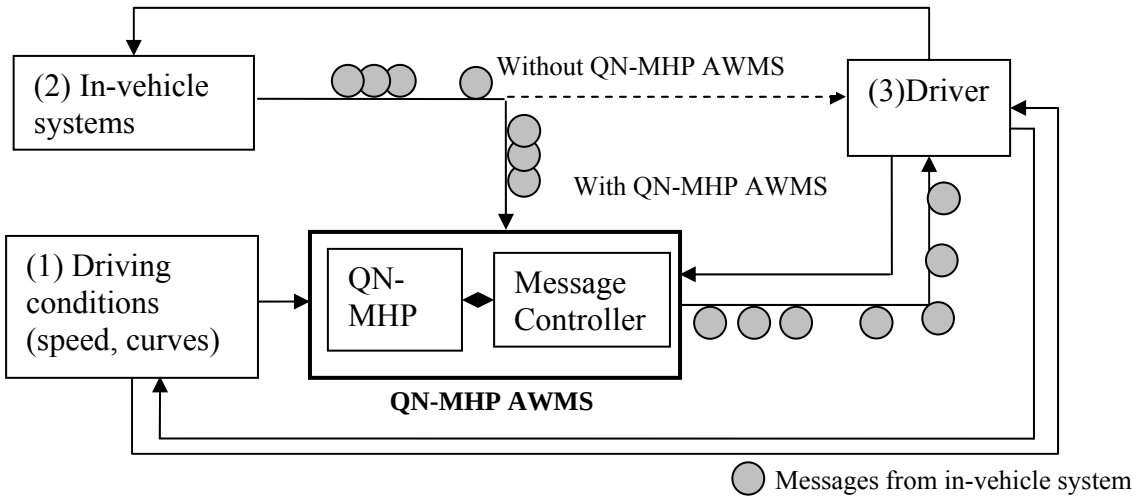


Figure 6-1 Illustration of the prototype of the QN-MHP adaptive workload management system (QN-MHP AWMS)

In this current study, we focus on the evaluating the potential effectiveness of AWMS in reducing driver workload and improving performance when driving conditions (speed and curves) and one of characteristics of drivers (age) change. The optimal delays of messages were obtained by running the simulation model of QN-MHP offline (see Section 4); based on these optimal delays, a simulated message controller dynamically sets the delay in real time according to the current driving condition (see Section 5).

3. A Sample Multitasking in Driving with Practical Importance

Speeding is one of the main causes of traffic accidents (Ewing, 1999) and traffic law enforcement of police officers to detect speeding and issue speeding tickets is one of the most critical measures to prevent speeding. However, besides detecting speeding, police officers also have to perform other tasks at the same time, e.g., communicating with dispatches, navigate the vehicle to a target location etc. Based on an informal interview with four police officers at the public safety service center at University of Michigan, a typical multitasking scenario of police officers was obtained: 1) Speeding detection or judgment task (called "radar-judgment task"): Officers need to read two numbers on a display of an in-vehicle radar system mounted on dashboards of the police vehicles: The first number was the speed of a target vehicle measured by the radar system; the second number was the distance from the police vehicle to the target vehicle. Whether the target vehicle is speeding or not is determined by the speed and the distance together, for example, "on a road with speed limit 55 miles/hr, a) if the speed is between 56 and 64 miles/hr and the distance is below 100 yards, it is speeding; if the distance is above 100 yards, it is not speeding; b) If the speed is above 65 miles/hr, it is speeding. c) If the speed is below 55 miles/hr, and it is not speeding." 2) Radio message response task (called "message-response task"): Messages received by the offices usually came multiple resources (headquarters, other police officers, and maintenance, and the officers need respond to higher priority messages (e.g., headquarters) by pressing a button on the radio. The design of this task was also inspired by the ALERT project in the Texas Transportation Institute and that project focused on the development of an integrated interface of various devices (radar detection system, radio, video recording systems etc) for police officers to improve their performance and safety (Hoelscher, 2007).

Even though the multiple task scenario described above is mainly focused on drivers in the police cars, it can be generalized into other common multitasking situations of drivers because the three aspects considered in the QN-MHP AWMS (driving conditions, properties of the 2nd tasks, and characteristics of drivers) are common in all of multitasking situations using the in-vehicle systems.

In order to design the adaptive workload management system which is able to manage the rates of messages in these in-vehicle systems presented to a driver, it is necessary to obtain the simulation or modeling results of mental workload and human performance when driving conditions, properties of the secondary task as well as the characteristics of drivers change.

4. Simulation of Multiple Tasks in Driving with QN-MHP

4.1 Simulation using QN-MHP

Following the steps described in simulating human performance and workload using QN-MHP (see Liu et al., 2006), the multiple tasks in driving were simulated with the following steps.

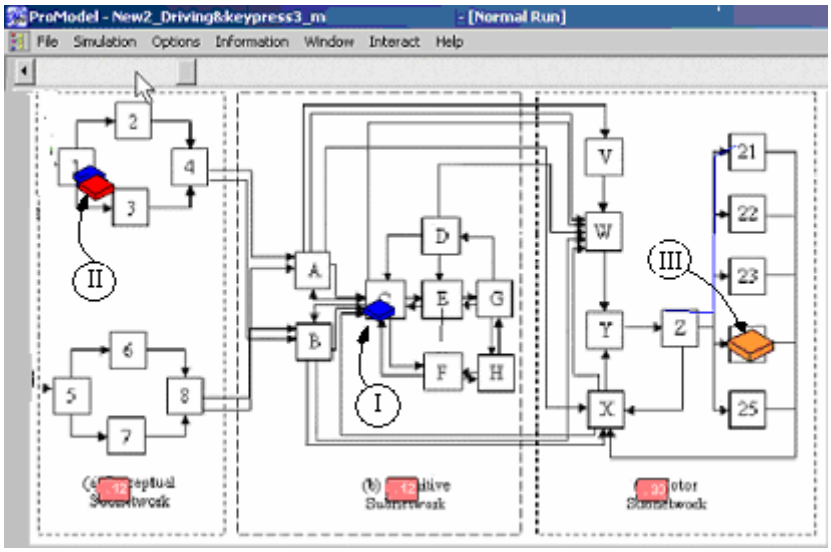
To model driver workload and performance, the input to the model was modified so that it can represent: 1) a road with two levels of curvature (straight and curves of 250 meter radius, a column of digits in the input file changed their values from 0 to positive values to represent the increase of road curvatures); 2) driving speed (45 and 65 miles/hr, a variable in the arrival table of the model changed from 0 to positive values to represents the increase of driving speed). The task analysis of a driving task was described in the work of Liu et al., (2006) in detail.

To model the secondary task, a new input to the model was added so that it can represent the stimuli of the secondary task based on its arrival interval (a delay between stimuli/messages). NGOMSL-Style task analysis was performed so that the model was able to know how to route and process entities (information) among different servers in the network (See Table 6-3) (each step in the NGOMSL-Style corresponded to an operator in the model and the operators determined the processing of entities in the model). In addition, the physical distance from the steering wheel to the target buttons on an in-vehicle user interface as well as sizes of the buttons were also input to the model so that the implemented Fitts' law in the model was able to simulate the motor execution time of in-vehicle messages.

Table 6-3 NGOMSL-Style Task Description of the Secondary Task

GOAL: Do visual speeding judgment task Method for GOAL: Do visual speeding judgment task Step 1. Watch for <the first number> on <the display> Step 2. Retain < the first number> Step 3. Retrieve from LTM (lower and upper boundary) Step 4. Decide: if <the first number> is <less than> <the lower boundary> than go to step 6; else go to step 5; Step 5. Decide: if <the first number> is <less than> <the upper boundary> than go to step 7; else go to step 11; Step 6. Current task completes, go to step 1; Step 7. Watch for the <the second number> on <the display> Step 8. Retain <the second number> Step 9. Retrieve from LTM (detection limit) Step 10. Decide: if the second number is <less than> <detection limit> then go to step 11; else go to step 6; Step 11. Press <SPEEDING button>	GOAL: Do auditory response task Method for GOAL: Do auditory response task Step 1. Listen to <the voice > from <the speaker> Step 2. Retain <the voice> Step 3. Retrieve from LTM <target voice> Step 4. Compare: <the voice> with <the target voice> in memory Step 5. Decide: If match, then go to step 6; else go to step 7 Step 6. Press <H button> Step 7. Current task completes, go to step 1;

Figure 6-2 shows a snapshot of the simulation model, when it was simulating the multitask situation in driving. The total length of road driven by the model was 5,000 meters in each run (the model performed six replications with different sets of random numbers).



Blue entity (I): entity (information) of the driving task
Red entity (II): entity of the speeding judgment task (subtask 2 of the 2nd task)
Orange entity (III): entity of the auditory message response task (subtask 2 of the 2nd task)

Figure 6-2 A snapshot of the simulation when it was simulating the multiple tasks in driving

4.2 Simulation Results (Young Group)

By changing the delay (inter-arrival time between the messages in the secondary task), simulation results of workload, standard deviation of lane positions and reaction time of the secondary task were obtained and plotted. Figure 6-3 and Figure 6-4 show the simulation results of overall workload, and delta overall workload ($\Delta Workload = Workload_{delay\ i} - Workload_{delayi-1}$, $delay_1=3$; $delay_2=5$, $delay_3=10$, $delay_4=15$, $delay_5=20$, and $delay_6=30$), which represent the change of subjective workload when the delay time increase.

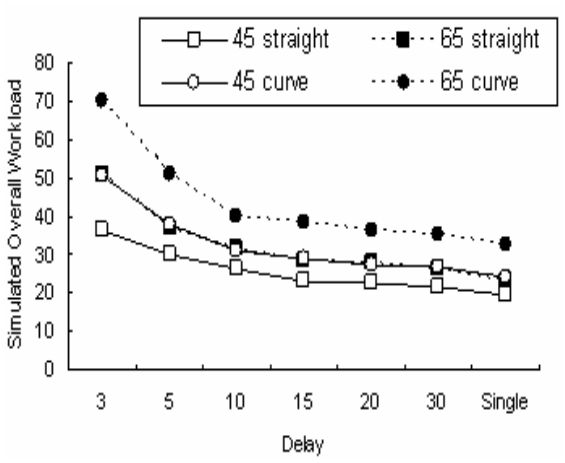


Figure 6-3 Simulated overall workload using QN-MHP (Young Group)

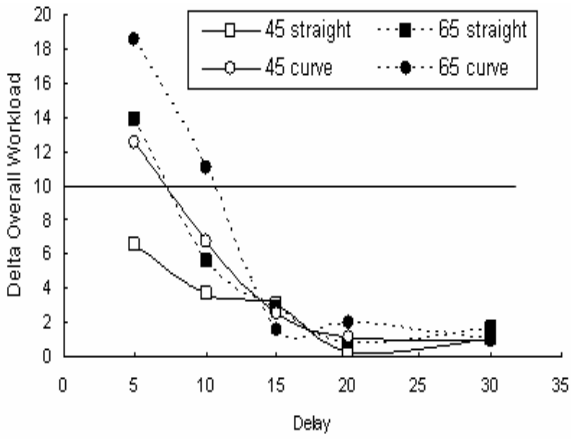


Figure 6-4 Simulated delta overall workload ($Workload_{delay\ i} - Workload_{delay\ i-1}$) (Young Group)

The optimal delay of messages was defined as a delay time that an increase of delay produces at least a decrease of one major unit in the workload scale (i.e., 10, in the 0-100 workload rating)[10]. Accordingly, based on Figure 6-3 and Figure 6-4, the differential threshold in decreasing workload is set at 10 (See the straight line in Figure 6-4); therefore, the following minimal delay times are obtained for young drivers (25-35 years old) in the four driving conditions: 65 curve: Delay>=15 sec; 65 straight: Delay>=10 sec; 45 curve: Delay >=10 sec; 45 straight: Delay >=5 sec.

[10] This is an arbitrary definition of optimal delay which can be changed by users of the system depending on different workload situations (see Discussion section).

The SD of simulated lane positions and its delta values are shown in Figure 6-5 and Figure 6-6.

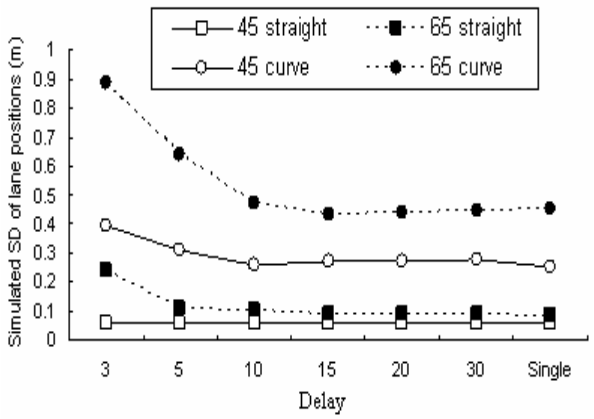


Figure 6-5 Simulated SD of lane positions using QN-MHP (Young Group)

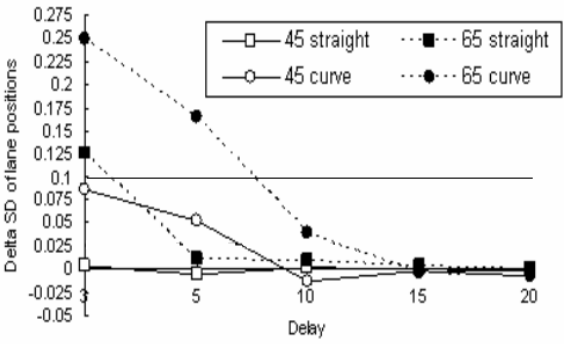


Figure 6-6 Simulated Delta SD of lane positions ($SD_{delay\ i}$-$SD_{delay\ i-1}$) (Young Group)

According to Figure 6-5 and Figure 6-6, similar to the logic in the delta overall workload in determining the delay time, the differential threshold in reducing SD of lane positions is set at .1 (a safe standard deviation of lane positions (SDLP) set based on a driving safety study of SDLP (Green, 1994; Zhang & Smith, 2004) and the following minimal delay times are obtained for young drivers (25-35 years old) in the four driving conditions: 65 curve: Delay >=10 sec; 65 straight: Delay >=5 sec; 45 curve: Delay >=3 sec; 45 straight: Delay >=3 sec.

The simulated average reaction time of the secondary task is presented in Figure 6-7, which suggests that the minimal delay of messages in the secondary task for young drivers (25-35 years old) should be at least greater than 5 sec for the 65 mile/hr condition including straight and curve conditions.

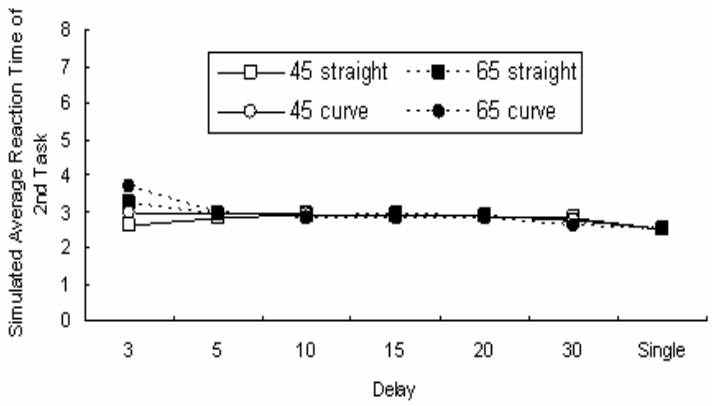


Figure 6-7 Simulated average reaction time of the secondary task (Young Group)

Combining the three lists of suggestions from the simulated overall workload, driving performance and performance of secondary task, we can derive the following suggestions of the minimal delays for the four driving conditions when a young driver (25-35 years old) is performing the secondary task: 65 curve: Delay>=15 sec; 65 straight: Delay>=10 sec; 45 curve: Delay >=10 sec; 45 straight: Delay >=5 sec. In other words, in the adaptive workload management system, the rates of messages presented to a driver need to follow the final suggestion list above to reduce drivers' overall workload and improve driving performance and performance of the secondary task. The same simulation model can be used to model driver workload and performance when the properties of the secondary task changes as well as driving conditions change.

4.3 Simulation Results (Old Group)

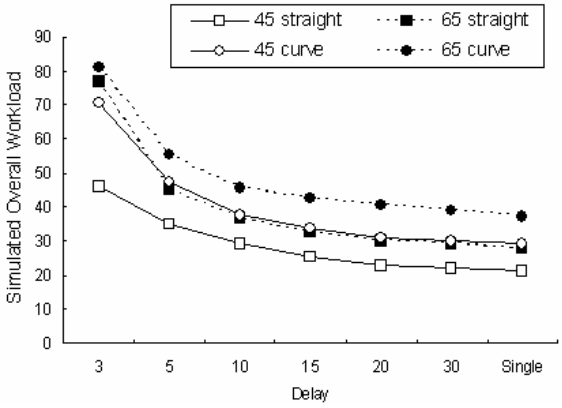


Figure 6-8 Simulated overall workload using QN-MHP (Old Group)

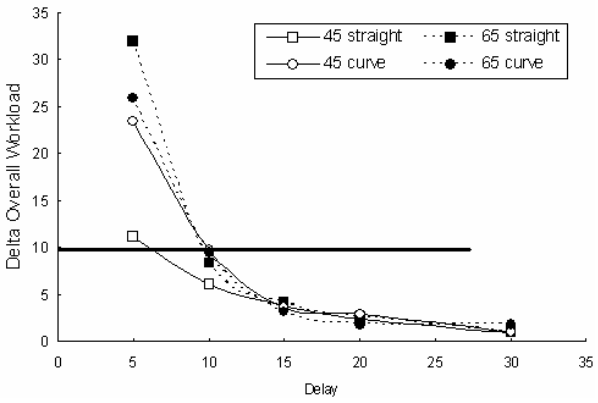


Figure 6-9 Simulated delta overall workload ($Workload_{delay\ i}$-$Workload_{delay\ i-1}$) (Old Group)

Based on Figure 6-8 and Figure 6-9, the following minimal delay times are obtained for older drivers (60-75 years old) in the four driving conditions: 65 curve: Delay>=15 sec; 65 straight: Delay>=10 sec; 45 curve: Delay >=15 sec; 45 straight: Delay >=10 sec.

The SD of simulated lane positions and its delta values are shown in Figure 6-10 and Figure 6-11.

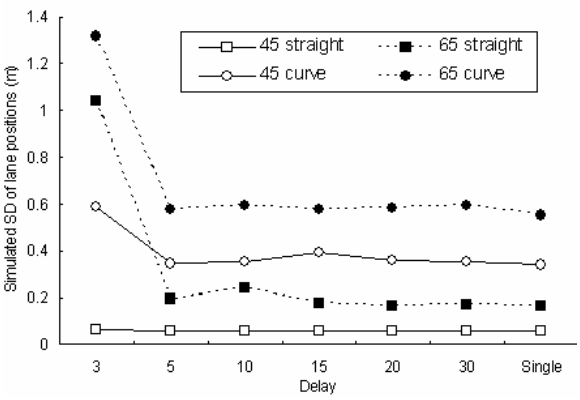


Figure 6-10 Simulated SD of lane positions using QN-MHP (Older Group)

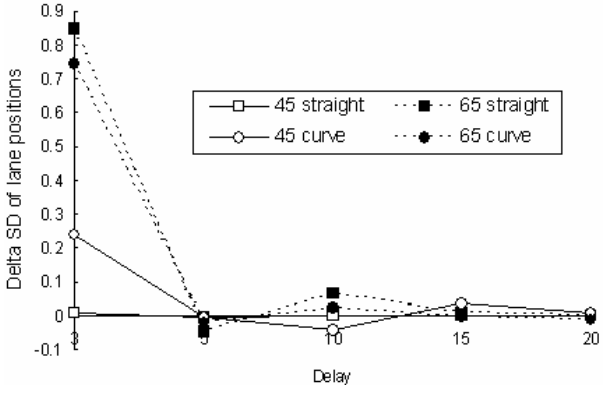


Figure 6-11 Simulated Delta SD of lane positions ($SD_{delay\ i}$-$SD_{delay\ i-1}$) (Older Group)

According to Figure 6-10 and Figure 6-11, similar to the logic in the delta overall workload in determining the delay time, the differential threshold in reducing SD of lane positions is set at .1 and the following minimal delay times are obtained for older drivers (60-75 years old) in the four driving conditions: 65 curve: Delay >=5 sec; 65 straight: Delay >=5 sec; 45 curve: Delay >=5 sec; 45 straight: Delay >=3 sec.

The simulated average reaction time of the secondary task is presented in Figure 6-12, which suggests that the minimal delay of messages in the secondary task for older drivers (60-75 years old) should be at least greater than 5 sec for the 65 mile/hr condition including straight and curve conditions.

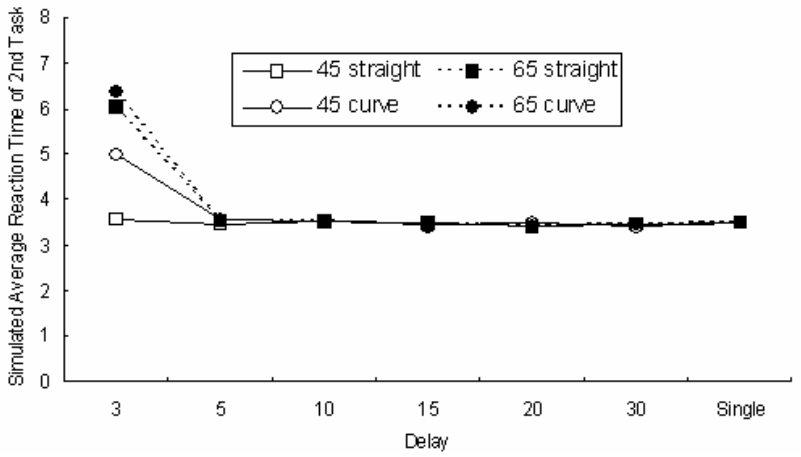


Figure 6-12 Simulated average reaction time of the secondary task (Older Group)

Combining the three lists of suggestions from the simulated overall workload, driving performance and performance of secondary task, we can derive the following suggestions of the minimal delays for the four driving conditions when an older driver (60-75 years old) is performing the secondary task: 65 curve: Delay>=15 sec; 65 straight: Delay>=10 sec; 45 curve: Delay >=15 sec; 45 straight: Delay >=10 sec. In other words, in the adaptive workload management system, the rates of messages presented to a driver need to follow the final suggestion list above to reduce drivers' overall workload and improve driving performance and performance of the secondary task. The same simulation model can be used to model driver workload and performance when the properties of the secondary task changes as well as driving conditions change.

5. Experimental Exploration of the Prototype of QN-MHP AWMS

The effectiveness of the adaptive system was tested using two conditions. 1) Adaptive condition: between the four driving conditions (two speeds × two curvatures), delay time (duration between stimuli of the secondary task) was set based on the optimal rates derived from simulation results of QN-MHP (See section 4); within each driving

condition, this delay time was kept constant. 2) Random condition: between the four driving conditions, the delay time was random (range: 3-30 sec); this delay time was constant within each driving condition. The total amount of stimuli, i.e., the total number of radio messages and amount of radar detection results, is the same in both adaptive and random conditions. There are three reasons to select this random condition: first, if a standard system for comparison is not available, a design of the same system with random configuration is usually used to validate the effectiveness of a new design (e.g., Mou & Zhang, 2001); second, compared to a fixed rate condition, the random condition might be a closer simulation of the real world situation without adaptive workload management system; third, this design of random condition is able to prevent the two extreme situations: fixed rate (a non-realistic condition) and random rate in both between and within four driving conditions. If both between and within driving conditions are random (called "complete random" condition here), it brings a new confounding factor—degree of randomness to the experiment because the degree of randomness of the delay in this complete random condition is higher than that in adaptive condition: whether the adaptive condition is better than the complete random condition or not may stem from the degree of randomness rather than the manipulation of the independent variable.

5.1 Experimental Design

A 2×2 two-factor mixed subject design was used in this experiment. The independent variables were: 1) the within-subject variable was the two conditions of the system (random vs. adaptive); 2) the between-subject variable was the age of drivers, i.e., young (25-35 years old) vs. older.(25-35 years old). The dependent variables were the driver workload measured by NASA-TLX, driving performance measured by standard deviation of lane position and performance of the secondary task. Each participant experienced two conditions of the system (adaptive and random) combined with four levels of driving conditions (straight with speed 45 miles/hr, straight with speed 65 miles/hr, curve with speed 45 miles/hr, and curve with speed 65 miles/hr). Participants were randomly assigned to one of two groups, which performed the experimental task either first in the adaptive condition followed by the random condition or vice versa. Within each of these

groups, the order of the four levels of driving conditions was also randomized and each of these driving conditions appeared once for each participant.

5.2 Participants

Sixteen licensed drivers were paid to participate in this experiment, including a group of young subjects (Young Group) (age 25-35 years, mean=30, SD=2.9) and a group of older subjects (age 60-75 years, mean=65, SD=3.8). All participants had corrected far visual acuity of 20/40 or better. All had midrange (80 cm) visual acuity of 20/70 or better. Prescreening of all participants ensured they had good driving records and were physically healthy.

5.3 Equipment and Test Materials

Driving Simulator. The simulator consisted of a full size cab, computers, video projectors, cameras, audio equipment, and other items (Figure 6-13). The simulator has a forward field of view of 120 degrees (3 channels) and a rear field of view of 40 degrees (1 channel). The forward screen was approximately 16-17 feet (4.9-5.2 m) from the driver's eyes, close to the 20-foot (6 m) distance often approximating optical infinity in accommodation studies. The vehicle mockup consisted of the A-to-B pillar section of a 1985 Chrysler Laser with a custom-made hood and back end. Mounted in the mockup was a torque motor connected to the steering wheel (to provide steering feedback), an LCD projector under the hood (to show the speedometer/tachometer cluster), a touch screen monitor in the center console (for in-vehicle tasks), a 10-speaker sound system (for auditory warnings), a sub-bass sound system (to provide vertical vibration), and a 5-speaker surround system (to provide simulated background road noise). The 10-speaker sound system was from a 2002 Nissan Altima and was installed in the A-pillars and lower door panel, and behind each of the two front seats. The stock amplifier (from the 2002 Nissan Altima) drove the speakers. The main simulator hardware and software was a DriveSafety Vection simulator running version 1.6.1 of the software. The display cards, GeForce3's, did not support anti-aliasing (Cullinane & Green, 2006).

Figure 6-13 UMTRI Driving Simulator

Simulated Roads. The simulated roads had two levels of road curvature (straight sections, curves of 250m radius sections) which was the same with the input to QN-MHP. Both lanes of the two-lane road were 3.66 meter (12 feet wide). Speed-limit signs (45 and 65 miles/hr) were placed in each section (straight and curve). The length of each road section was 5,000 meters, consistent with the input to QN-MHP.

Touch Screen. An IBM laptop X60 with a 12'' touch screen was mounted on the right of the driver at arm's length. This touch screen was located in the center console of the vehicle, 23º ± 3º below the horizontal line of sight and 30º ± 3º to the right of the center (see Figure 6-14). To allow easy reading, numbers on the display were relatively large (digit height = 11 mm, 1º at 63 cm).

Figure 6-14 Driver's view of the road and the touch screen

Message Controller in the Prototype of QN-MHP AWMS. The Wizard of OZ method (e.g., Tsimhoni, et al., 2004; Green & Wei-Hass, 1985) was used to simulate the message controller in the prototype of QN-MHP AWMS. The experimenter, acted as the message controller to read the current speed and curvatures from a) the screen of the simulator and b) speedometers of the simulator shown on a computer connected with the simulator. When participants started to drive in one of four driving conditions, the Wizard (the experimenter) input the speed and curvature information immediately into the prototype by pressing a button on a self-developed Visual Basic program; and this program transmitted the speed and curvature information to the prototype in real time and then the prototype selected the delay time based on the simulation results (see Section 4) (the experimenter had more than 100 practices on this selection process prior to running subjects. the duration of this selection process was less than 2 sec with no selection errors observed).

5.4 Experimental Task and Procedure

Driving Task. Participants were instructed to drive in the right lane and maintain a speed following the speed-limit signs on the simulated roads. To maintain the driving speed of each participant, if they drove 5 miles/hr above or below the speed shown on the speed-limit signs, they heard a computer-generated voice "too fast" or "too slow".

Secondary Task. The secondary task was composed of two subtasks simulating a typical multitasking scenario when a police officer was patrolling on a road (e.g., police officers respond to messages from dispatches by turning on a radio and monitor the speed of other car detected by an in-vehicle radar system).

The first subtask was a radio-message response task: participants were instructed to press the button with caption "H" on the touch screen (See Figure 6-15) as quickly as possible and then speak aloud "In route" once they hear the word "Headquarter" from the speakers. If they heard "Maintenance", they did not need to make a response.

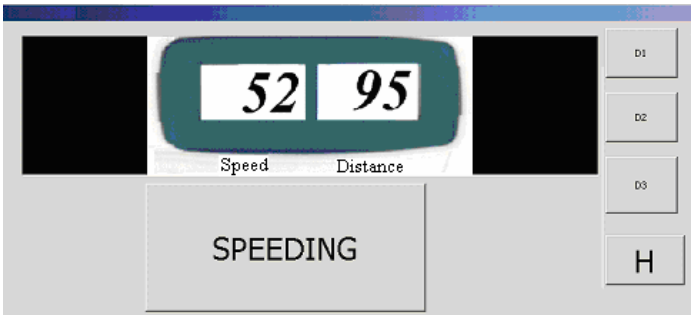


Figure 6-15 A screenshot of the touch screen of the secondary task

The second subtask was a speeding judgment task: participates were asked to judge whether other vehicles were speeding or not only based on the two numbers (the number on the left side is the detected speed, the number on the right side is the distance from the participant's car to the other car) by a radar system (see Figure 6-15), following the three rules: a) If the speed is above 65 miles/hr (including 65), it is speeding; b) If the speed is below 55 miles/hr (including 55), it is not speeding; c) If the speed is between 56 and 64 miles/hr (including 56 and 64) and the distance is below 100 yards, then it is speeding; if the distance is above 100, it is not speeding.

If participants judged that the other car was speeding based on the numbers on the screen, they were instructed to press the "SPEEDING" button on the touch screen as quickly as possible. Each time right before the numbers of the second subtask shown on the screen, a 50 ms high-pitch tone was presented to subjects as a cue for the visual stimuli. All of the buttons on the touch-screen produced an auditory feedback (a beep with length 100 ms) if they were pressed by participants.

During the experiment, the stimuli of the two subtasks in the secondary task were presented to a participant in a serial order (e.g., a radio-message followed by numbers of the radar system or another radio-message). The duration between stimuli was called delay time, manipulated in the adaptive and random condition. The ratio of each subtask in the secondary task throughout the experiment was 50%.

After filling in the pretest forms and vision tests, in the practice section, participants first practiced single task situation: driving (straight and curve) alone without the secondary task; and the secondary task alone while the simulator was in parked condition. Then, participants practiced dual task situation: driving while performing the secondary task at the same time. In the test section, participants were instructed to drive with System A (random condition) and System B (adaptive condition) according to which group they were assigned. After participants finished all of the driving conditions (two speeds and two curvatures) in the random or adaptive condition, they were asked to complete the NASA-TLX form to report their subjective workload.

5.5 Experimental Results

5.5.1 Subjective Workload

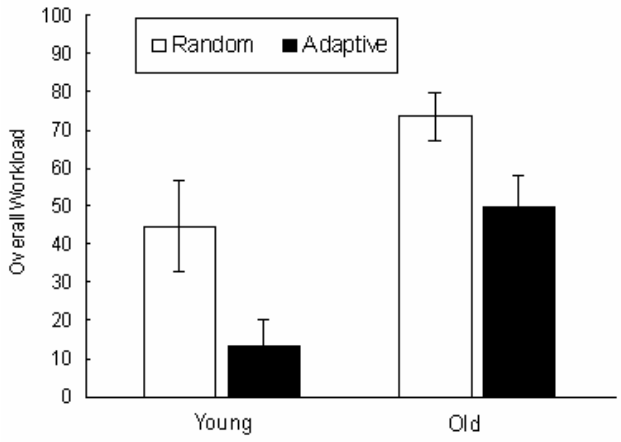


Figure 6-16 Comparison of the overall workload between the random and adaptive condition (error bar shows 1± SD of overall workload rating across subjects)

Figure 6-16 shows the comparison of overall workload ratings measured in the NASA-TLX between the random and adaptive conditions. Mixed-factor (between and within-subject) ANOVA showed that the main effect of system (random vs. adaptive) on

the overall workload was significant ($F(1,14)=30.6\ 1,\ p<.01$). In addition, the main effect of age on the overall workload was significant ($F(1,\ 14)=21.09,\ p<.01$); but the Age × System interaction was not significant ($F(1,\ 14=0.35$). Within each age group, the one-factor MANOVA found that there is a significant difference of overall workload between the random and adaptive condition (Young Group: $F(1,7)=26.57,\ p<.01$; Old Group: $F(1,\ 7)=4.67,\ p<.05$).

The comparison of workload ratings in the six subscales between the random and adaptive conditions is presented inFigure 6-17. The main effect of the system was significant at workload rating on each of the six subscale/dimensions of NASA-TLX (MD (Mental Demand) using, mixed-factor ANOVA: $F(1,14)=18.01,\ p<.01$; PH (Physical Demand): $F(1,14)=6.95,\ p<.05$; TD (Temporal Demand): $F(1,\ 14)=30.21,\ p<.01$; PE (Performance): $F(1,14)=8.73,\ p<.01$; EF (Effort): $F(1,\ 14)=30.97,\ p<.01$; FR (Frustration): $F(1,\ 14)=28.30,\ p<.01$). In addition, the main effect of age on each was also significant at each of these dimensions (MD: $F(1,\ 14)=15.28,\ p<.01$; PH: $F(1,\ 14)=12.07,\ p<.01$; TD: $F(1,14)=11.09,\ p<.01$; PE: $F(1,14)=17.52,\ p<.01$; EF: $F(1,\ 14)=27.26,\ p<.01$; FR: $F(1,14)=43.97,\ p<.01$). The Age×System interaction was not significant (MD: $F(1,14)=.96,\ p>.05$; PH: $F(1,14)=.01,\ p>.05$; TD: $F(1,14)=.70,\ p>.05$; PE: $F(1,14)=.15,\ p>.05$; EF: $F(1,14)=.96,\ p>.05$; FR: $F(1,14)=.003,\ p>.05$). In the young group, one-factor MANOVA found that there is a significant difference of workload rating between the random and adaptive condition on the TD ($F(1,7)=24.93,\ p<.01$), PE ($F(1,7)=6.36,\ p<.05$), EF ($F(1,\ 7)=5.79,\ p<.05$), and FR ($F(1,7)=21.81,\ p<.01$) subscales. In the older group, one-factor MANOVA found that there is a significant difference of workload rating between the random and adaptive condition on the EF subscale ($F(1,\ 7)=7.50,\ p<.05$).

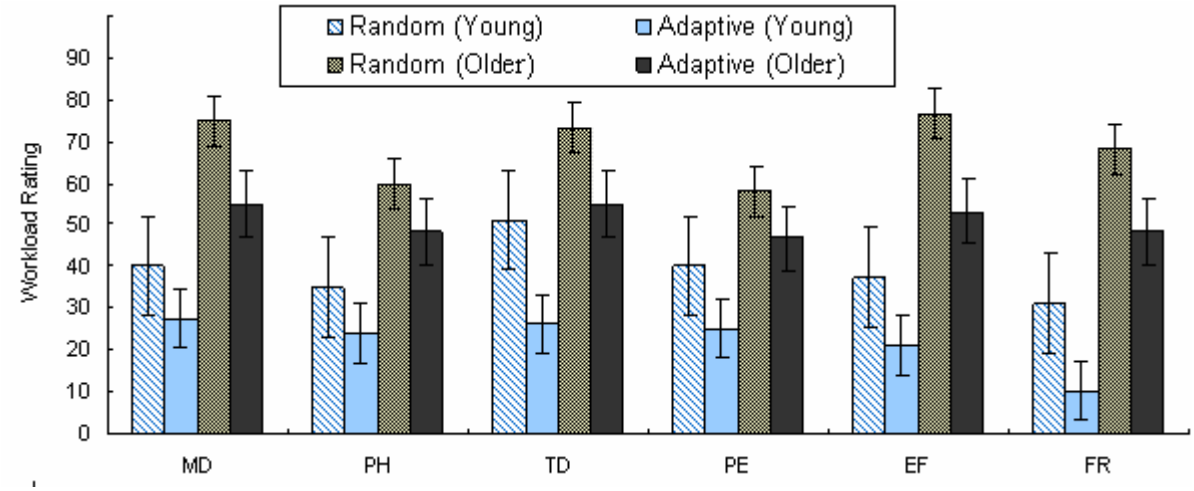


Figure 6-17 Comparison of the six workload rating in NASA-TLX between the random and adaptive condition (error bar shows 1± SD of workload rating across subjects)

In other words, the adaptive system significantly reduced the subjective workload in both young and older age groups; and the subjective workload was reflected in overall workload and the six subscales in NASA-TLX.

5.5.2 Performance in Driving and Secondary Task

In terms of driving performance, the main effect of the system on the standard deviation of lane positions was also significant (Mixed-factor ANOVA, $F(1,14)=33.37$, $p<.01$). The main effect of age was not significant ($F(1,14)=.012$). The system×age interaction was significant ($F(1,14)=7.3$, $p<.05$). In the young group, the adaptive condition significantly reduced the SD of lane positions ($F(1,7)=20.50$, $p<.01$); in the older group, the SD of lane positions in the adaptive condition was also significantly reduced compared to the random condition ($F(1,7)=5.91$, $p<.05$) (See Figure 6-18).

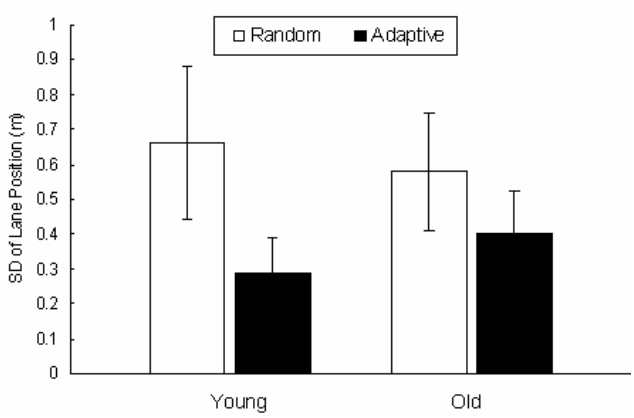


Figure 6-18 Comparison of the standard deviation of lane positions between the random and adaptive condition (error bar shows 1± SD of standard deviation of lane positions across subjects)

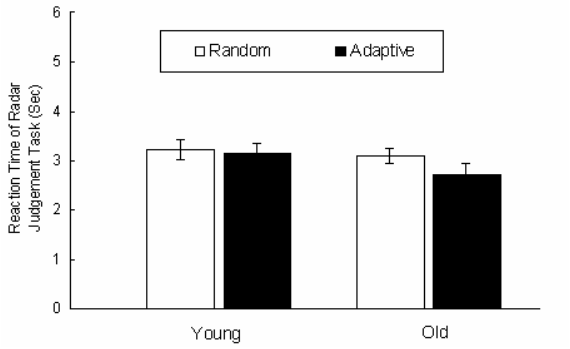


Figure 6-19 Comparison of the radar-judgment task between the random and adaptive condition (error bar shows 1± SD of this RT across subjects)

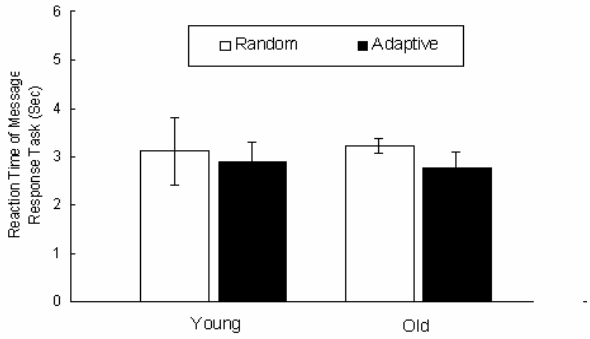


Figure 6-20 Comparison of the radar-judgment task between the random and adaptive condition (error bar shows 1± SD of this RT across subjects)

Figure 6-19 and Figure 6-20 showed the comparison of reaction time of the radar-judgment task and the message-response task between the random and adaptive condition (error rate of the both the secondary task is less than 1% in both conditions). For the reaction time of the radar-judgment task, the main effect of both system and age on the reaction time of the radar-judgment task was significant (system: $F(1,14)=10.47$, $p<.01$; age: $F(1,14)=13.48$, $p<.01$); The system×age interaction was also significant ($F(1,14)=4.95$, $p<.05$). The simulated adaptive system significantly reduced the reaction time of the radar-judgment task in the older group but not in the young group (older group: $F(1,7)=17.44$, $p<.01$; young group: $F(1,7)=.53$, $p>.05$).

For the reaction time of the message-response task, main effect of the system on the reaction time of the message-response task was not significant ($F(1,14)=4.23$, $p>.05$); and main effect of age as well as the interaction system×age were not significant (age: $F(1,14)=0.016$, $p>.05$; interaction: $F(1,14)=.48$, $p>.05$). The simulated adaptive system significantly reduced the reaction time of the message-response task in the older group but not in the young group (older group: $F(1,7)=12.62$, $p<.01$; young group: $F(1,7)=.64$, $p>.05$).

6. Discussion

To reduce driver workload in multitasking, a prototype of a new adaptive workload management system (QN-MHP AWMS) was developed in this work. QN-MHP AWMS was composed of two components: QN-MHP estimating driver workload in different driving situations and a simulated message controller to change the rate of messages from the in-vehicle systems. QN-MHP AWMS adaptively changed the rate of messages based on the three types of information: 1) driving conditions (e.g., current driving speed and curvatures); 2) properties of messages from in-vehicle systems (processing time at the perception, cognitive and motor part); 3) properties of driver (e.g., age). The experimental study validated the potential effectiveness of the system in reducing the workload measured by NASA-TLX in terms of overall workload as well as the workload rating at the temporal demand, performance, effort, frustration and effort subscales. Driving performance was also improved using this adaptive workload management system.

There are two possible applications for the proposed system. First, to reduce driver workload, design engineers of in-vehicle system can use QN-MHP AWMS to modify their design at the early stage of development of various in-vehicle systems. QN-MHP AWMS lets the user estimate driver workload when drivers are manipulating different user interfaces of in-vehicle systems. Engineers can estimate the level of driver workload and performance when the properties of messages from the in-vehicle systems change in terms of modalities (processing time at perceptual part), difficulty of messages (processing time at the cognitive part), and motor execution time. Engineers can also set the absolute threshold (workload "redline") and differential threshold of the simulated workload to determine the optimal design of the messages as well as whether the current design may produce workload higher than the "redline" or not.

Second, QN-MHP AWMS can be implemented into vehicles with the development of computer technologies. Even though current QN-MHP AWMS needs simulation software installed on a computer, the simulation results of QN-MHP and the suggested optimal message rates can be approximated by relatively simple algorithms; these algorithms can be implemented into microcomputers in vehicles, especially vehicles with special duties (police vehicles, ambulance vehicles etc.). The message controller simulated in the

experiment in the current work can also be easily replaced by software in the in-vehicle microcomputers because it only needs to read information of the vehicle speed and angles of the steering wheel from the bus line (a parallel circuit that connects the major components and sensors in a vehicle). Global Positioning System (GPS) can also be used to measure road curvatures and speed on the next road section so that QN-MHP AWMS can estimate driver workload a few seconds in advance.

There are several limitations of the current work that need to be examined in future research. First, because the focus of QN-MHP AWMS is to reduce driver workload, it is only suitable for non-urgent messages of in-vehicle systems (when delaying messages for a few seconds is allowable). For urgent messages which need immediate response of drivers, e.g., forward collision warning messages, the extra delay created by the adaptive system will prolong drivers' reaction time in these urgent situations. Actually, this limitation applies to many adaptive workload systems because of the extra delay or suppression of messages will delay drivers' responses to these non-urgent messages (however, it is possible to add an option in the QN-MHP AWMS so that users can disable the message delay function). Second, the current adaptive system developed in this work only focuses on the rate of two types of messages with equal priority. New algorithms are needed to manage messages with different priorities, including the order and length of these messages; but QNMHP AWMS may still serve as a platform for designing and optimizing the other properties of information presented to drivers. Third, the current work only tested the adaptive part of QN-MHP AWMS depending on four driving conditions (speed ×curvature) and one property of drivers (age), and future simulation and experimental studies are expected to add more driving conditions (e.g., intensity of traffics, weather condition etc.) and drivers' properties (e.g., driving experience, gender etc.) into the simulation and empirical validations of the system. Previous published work of QN-MHP has considered aging factor (variable *A* (age factor) in Wu & Liu, 2006c, 2006d) as one of the major factors in predicting driver workload, and this already builds a foundation for testing the adaptive system incorporating three sources of information (driving conditions, information from the in-vehicle systems and drivers' properties) at the same time.

In summary, we are extending the current approach both in modeling different tasks in driving and applying the model in designing intelligent in-vehicle systems to improve transportation safety. Our comprehensive computational model of driver workload offers not only theoretical insights into driving workload, but is a step toward developing proactive ergonomic design and multi-purpose analysis tools for tasks in transportation.

Acknowledgement

This article is based upon work supported by the Doctoral Studies Program (DSP) at the University of Michigan Transportation Research Institute (UMTRI). We would like to thank Mr. Jacob Mouro, a public safety officer at University of Michigan Security Service who gave us tremendous help in describing police vehicles and multitasking scenarios in their daily work.

Reference

Alm, H., & Nilsson, L. (1995). The Effects of a Mobile Telephone Task on Driver Behavior in a Car Following Situation. *Accident Analysis and Prevention, 27*(5), 707-715.

Cullinane, B., & Green, P. (2006). *SAfety VEhicles using adaptive Interface: Technology Visual Demand of Curves and Fog-Limited Sight Distance and Its Relationship to Brake Response Time*: University of Michigan Transportation Research Institute (UMTRI).

Ewing, R. (1999). Traffic Calming: State of the Practice: ITE/FHWA Washington (DC): Institute of Transportation Engineers; Washington (DC): U.S. Federal Highway Administration.

Feyen, R. (2002). *Modeling Human Performance using the Queueing Network---Model Human Processor (QN-MHP).* University of Michigan, Ann Arbor.

Feyen, R., & Liu, Y. (1998). *Driver Performance Model for Instrument Panel Control Use, Technical Report: 99-03*: Department of Industrial and Operations Engineering, University of Michigan.

Green, P. (1994). *Measures and Methods Used to Assess the Safety and Usability of Driver Information Systems, UMTRI-93-12*: University of Michigan Transportation Research Institute (UMTRI).

Green, P. (2004). *Driver Distraction, Telematics Design, and Workload Managers: Safety Issues and Solutions.* Paper presented at the Proceedings of the 2004 International Congress on Transportation Electronics, Warrendale, PA.

Green, P., & Wei-Haas, L. (1985). The Rapid Development of User Interfaces: Experience with the Wizard of Oz Method, *Proceedings of the Human Factors Society 29th Annual Meeting* (pp. 470-474). Santa Monica, CA: HFES Society.

Haigney, D., & Westerman, S. J. (2001). Mobile (cellular) phone use and driving: a critical review of research methodology *ERGONOMICS, 44*(2), 132 - 143.

Hancock, P. A., & Chignell, M. H. (1988). Mental workload dynamics in adaptive interface design. *Systems, Man and Cybernetics, Part A, IEEE Transactions on, 18*(4), 647-658.

Hoelscher, M. (2007). *High-tech police car builds bridges between agencies*. Retrieved 2/2, 2007, from http://tti.tamu.edu/publications/researcher/newsletter.htm?vol=36&issue=1&article=4

Horiuchi, S., & Yuhara, N. (2000). An analytical approach to the prediction of handling qualities of vehicles with advanced steering control system using multi-input driver model. *Journal of Dynamic Systems Measurement and Control-Transactions of the Asme, 122*(3), 490-497.

Jamson, A. H., Westerman, S. J., Hockey, G. R. J., & Carsten, O. M. J. (2004). Speech-based E-mail and driver behavior: Effects of an in-vehicle message system interface. *Human Factors, 46*(4), 625-639.

Kantowitz, B. (2004). *Safety vehicles using adaptive Interface technology, Technique Report of SAVEIT project*.

Lim, J., & Liu, Y. (2004). *A queueing network model of menu selection and visual search.* Paper presented at the Proceedings of the 48 Annual Conference of the Human Factors and Ergonomics Society, New Orleans, Louisiana, USA.

Lin, Y., Tang, P., Zhang, W. J., & Yu, Q. (2005). Artificial neural network modelling of driver handling behaviour in a driver-vehicle-environment system. *International Journal of Vehicle Design, 37*(1), 24-45.

Liu, Y. (1996). Queueing network modeling of elementary mental processes. *Psychological Review, 103*(1), 116-136.

Liu, Y. (1997). Queueing network modeling of human performance of concurrent spatial and verbal tasks. *IEEE Transactions on Systems Man and Cybernetics Part a-Systems and Humans, 27*(2), 195-207.

Liu, Y. (2005). *Queueing Network Modeling of Mental Architecture, Response Time, and Response Accuracy: Reflected Multi-dimensional Diffusions.* Paper presented at the Annual Meeting of Mathematical Psychology Society, Memphis, TN.

Liu, Y., Feyen, R., & Tsimhoni, O. (2006). Queueing Network-Model Human Processor (QN-MHP): A Computational Architecture for Multitask Performance. *ACM Transaction on Human Computer Interaction, 13*(1), 37-70.

Moray, N., Dessouky, M. I., Kijowski, B. A., & Adapathya, R. (1991). Strategic Behavior, Workload, and Performance in Task-Scheduling. *Human Factors, 33*(6), 607-629.

Mou, W., & Zhang, K. (2001). A compatible chord code for inputting elements of Chinese characters. *Applied Ergonomics, 32* 293-297.

Olson, A., & Green, P. (1997). Description of the UMTRI Driving Simulator Architecture and Alternatives, (Technical Report UMTRI-97-15). In T. U. o. M. T. R. Institute (Ed.).

Piechulla, W., Mayser, C., Gehrke, H., & Konig, W. (2003). Reducing drivers' mental workload by means of an adaptive man-machine interface. *Transportation Research, Part F, 6*, 233-248.
Pindo. (2002). *Scheduling: Theory, Algorithms, and Systems.* Upper Saddle River, NJ,: Precentice Hall.
Redeleier, D. A., & Tibshirani, R. J. (1997). Associatation between Cellular-Telephone Calls and Motor Vehicle Collisions. *New England Journal of Medicine, 336*(7), 453-458.
Rouse, W. B. (1980). *Systems Engineering Models of Human-Machine Interaction*. New York: North Holland.
Salvucci, D. D., Boer, E. R., & Liu, A. (In press). Toward an Integrated Model of Driver Behavior in a Cognitive Architecture. *Transportation Research Record*.
Sarter, N. B. (2001). Multimodal Communication In Support of Coordinative Functions In Human-Machine Teams. *Journal of Human Performance in Extreme Environments, 5*(2), 50-54.
Sklar, A. E., & Sarter, N. B. (1999). Good Vibrations? The Use of Tactile Feedback In Support of Mode Awareness on Advanced Technology Aircraft. *Human Factors, 41*(4), 543-552.
Tsang, P. S., & Velaquez, V. L. (1996). Diagnosticity and multidimensional subjective workload ratings. *Ergonomics, 39*(3), 358-381.
Tsimhoni, O., Smith, D., & Green, P. (2004). Address entry while driving: speech recognition versus a touch-screen keyboard. *Human Factors, 46*(4), 600-610.
Uchiyama, Y., Kojima, S., Hongo, T., Terashima, R., & Toshihiro, W. (2004). Voice information system that adapts to driver's mental workload. *R&D Review of Toyota CRDL, 39*(1), 16-22.
Vadeby, A. M. (2004). Modeling of relative collision safety including driver characteristics. *Accident Analysis & Prevention, 36*(5), 909-917.
Verwey, W. B. (1993). How can we predict overload of driver? In A. M. Parkes & S. Franzen (Eds.), *Driving Future Vehicles*. London: Taylor & Francis.
Violanti, J. M., & Marshall, J. R. (1996). Cellular phones and traffic accidents: an epidemiological approach. *Accident Analysis & Prevention, 28*(2), 265-270.
Wagner, D., Vercruyssen, M., & Hancock, P. (1997). A computer-based methodology for evaluating the content of variable message signage. *ITS Journal, 3*(4), 353-373.
Wickens, C., Kramer, A., Vanasse, L., & Donchin, E. (1983). Performance of Concurrent Tasks - a Psychophysiological Analysis of the Reciprocity of Information-Processing Resources. *Science, 221*(4615), 1080-1082.
Wu, C., & Liu, Y. (2004a). *Modeling Behavioral and Brain Imaging Phenomena in Transcription Typing with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA.
Wu, C., & Liu, Y. (2004). *Modeling human transcription typing with Queueing network-model human processor.* Paper presented at the Proceedings of the 48th Annual Meeting of Human Factors and Ergonomics Society, New Orleans, Louisiana, USA.
Wu, C., & Liu, Y. (2004b). *Modeling Psychological Refractory Period (PRP) and Practice Effect on PRP with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA. 320-325.
Wu, C., & Liu, Y. (2004c). *Modeling Psychological Refractory Period (PRP) and Practice Effect on PRP with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA.
Wu, C., & Liu, Y. (2006a). *Modeling fMRI BOLD Signal and Reaction Time of a Dual Task with a Queueing Network Modeling Approach.* Paper presented at the 28th Annual Conference of the Cognitive Science Society, Vancouver, BC, Canada.
Wu, C., & Liu, Y. (2006). *Queueing Network Modeling of a Real-time Psychophysiological Index of Mental Workload—P300 Amplitude in Event-Related Potential (ERP).* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.
Wu, C., & Liu, Y. (2006b). *Queueing Network Modeling of Age Differences in Driver Mental Workload and Performance.* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.
Wu, C., & Liu, Y. (2006c). *Queueing Network Modeling of Driver Workload and Performance.* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.

Wu, C., & Liu, Y. (2006d). *Queueing Network Modeling of Reaction time, Response Accuracy, and Stimulus-Lateralized Readiness Potential Onset Time in a Dual Task.* Paper presented at the 28th Annual Conference of the Cognitive Science Society, Vancouver, BC, Canada.

Wu, C., Zhang, K., & Hu, Y. (2003). Human performance modeling in temporary segmentation Chinese character handwriting recognizers. *International Journal of Human Computer Studies, 58*, 483-508.

Xie, B., & Salvendy, G. (2000). Review and reappriaisal of modelling and prediting mental workload in single-and multi-task environments. *Work & Stress, 14*(1), 74-99.

Zhang, H., & Smith, M. (2004). *A Final Report of SAfety VEhicles using adaptive Interface Technology (Phase I: Task 7): Visual Distraction Research*.

Chapter 7
Application of Scheduling and the Queueing Network Modeling Methods in Designing Multimodal In-vehicle Systems

Chapter Summary

Usage of multimodal in-vehicle system is one of ways to improved time sharing performance of drivers. However, few computational methods has been developed to assist designers of multimodal in-vehicle systems (MIVS) to select appropriate modalities and determine the order of messages in the system based on the properties of the in-vehicle tasks. This paper proposed a general procedure to select several scheduling methods and used them to schedule two tasks in an example multimodal in-vehicle system. An empirical study was conducted and validated the scheduling results including the optimal modality and order of these tasks. Further extensions of the current methodology and usage of this general procedure in selecting other scheduling methods were discussed.

1. Introduction

1.1 Importance of multimodal in-vehicle system design

With the development of technology, there is an increased usage of many vehicle information systems (e.g., road guidance and directions; vehicle status information), vehicle safety/warning systems (e.g., lane departure warning, collision warning system, curve speed warning etc., Gupta, et al., 2002), as well as vehicle communication system (e.g., vehicle-to-vehicle communication, usage of cellular phone while driving), Multitasking between driving and using these systems may impose high information load

on drivers, increasing drivers' mental workload (Alm & Nilsson, 1995; Wagner, Vercruyssen, & Hancock, 1997; Wickens, Kramer, Vanasse, & Donchin, 1983) which in turn increases the chance of vehicle collisions comparing to a single driving condition (Violanti and Marshall, 1996; Alm & Nilsson, 1995). This introduced a very import topic in in-vehicle system design and transportation safety—how to present information from these in-vehicle systems to drivers properly to improve driver performance and reduce driver workload.

Multimodal studies in multimodal user interface (MUI) suggest that presentation of concurrent tasks via different sensory channels leads to improved time sharing performance (Sarter, 2001). For in-vehicle systems, multimodal communication between drivers and in-vehicle systems might be an effective way to improve driver performance and reduce information overloading in visual modality (Cellario, 2001; Mariani, 2002; Gupta, et al., 2002; Siewiorek, et al., 2002). Several important qualitative guidelines how to design multimodal user interface have been summarized in Sarter (2001); however, computational methods in analyzing the multimodal information processing have received only scant attention in this field, especially how to assist designers of multimodal in-vehicle systems (MIVS) to select appropriate modalities and determine the order of messages in the system based on the properties of the in-vehicle tasks. In other words, if some basic and/or quantitative information of these tasks are given (e.g., their difficulty levels in cognitive process, their response modalities (hand or body parts), distance from the body parts to the in-vehicle devices etc.), the important question becomes how to assist designers of MIVS so that they can follow a list of algorithms to calculate and select the optimal modality and decide the order of these messages to be presented to drivers.

At an abstract level of analysis, there are two important dimensions to analyze MIVS as a subset of MUI: 1) at the spatial dimension, designers of a MIVS need to arrange which input modalities of users to receive these information and which output modalities to execute the control actions; 2) At the temporal dimension, users of a MIVS perceive sequence(s) of messages/information come from MIVS and execute sequence(s) of control movements to manipulate the interface of a MIVS. If information presented to subjects can be regarded as "jobs" and the cognitive system can be treated as a system

which is composed of several "processors" or "machines" processing these jobs, scheduling—a group of computational methods which deal with how to arrange the order and assignment of the jobs to machines—can be used as one of methods to quantitatively analyze human information processing in multimodal man-machine interaction.

1.2 Review of Scheduling Concepts and Methods with their Application in MUI

Starting from 1950s, scheduling has become one of major branches in industrial engineering and many scheduling methods and algorithms have been developed depending on number of machines in a system as well as measurements of the system performance. Table 7-1 summarized several important concepts in scheduling theory (Pinedo, 2002; French, 1982) with their corresponding meaning in MIVS.

Table 7-1 Concepts scheduling theory with their corresponding meaning in MIVS

Term	Meaning in Scheduling Theory	Corresponding meaning in multimodal in-vehicle systems
Machine /Serial Processor/Bottleneck	A processing unit or server which can only process one job at one time	A serial processor in the cognitive system
Working Station /Stage	A processing unit which is composed of several machines in parallel	A perceptual/cognitive/motor stage which consists of several serial processor arranged in a parallel manner

One of the most commonly used performance measurements in scheduling is makespan (*Cmax*), defined as the duration between the arrival of the first job and the time when the last job leaves the system (Pinedo, 2002; French, 1982). *Cmax* might be the performance measurement most relevant to human performance because it is equivalent to the total task completion time in human performance. Table 7-2 summarized several scheduling methods depending on number of machines or processors in a system to minimize makespan (*Cmax*) of a system. In the situation that machines in a system are arranged in a serial manner (jobs need to go from one processor first and then go to another processor): a) in the single machine condition, since the exchange of order of jobs does not affect the makespan, there is no scheduling method developed to minimize *Cmax;* b) when there are two machines or processors arranged in a serial manner,

Johnson's rule is a classic scheduling method to minimize makespan (see description of this rule in detail in the following section of this chapter); c) when there are three or more machines, this scheduling problem becomes *NP-hard* which means that it does not have a polynomial time algorithm to solve it (Pinedo, 2002). In the parallel arrangement situation, depending on the number of processors in a system, parallel scheduling method (French, 1982) and critical path analysis have been proposed to solve the scheduling problem (see a review of critical path analysis method in Harold, 2001) . In the field of multimodal user interface, it seemed that only critical path analysis method has been used to design multimodal user interface (Baber and Mellor, 2001) while the other simple but effective scheduling methods including Johnson's Rule, have not been applied to design MUI, especially in in-vehicle systems.

Table 7-2 Summary of scheduling methods to minimize makespan (*Cmax*) of a system

Configuration	Number of Machines	Scheduling Methods	Application in MUI
Serial	Single machine	-	-
	Two machines	**Johnson's Rule**	**Not yet**
	Three or more machine	*NP-Hard*	-
Parallel	**Two or more machines (machine number< job number)**	**Non-identical Parallel Machine Scheduling Method**	**Not yet**
	Infinite number of machines (machine number> job number)	Critical Path Analysis	Baber and Mellor (2001)

In the following section, the scheduling methods which have not been used in designing a MUI including MIVS are described in detail (see Baber and Mellor (2001) for a review of critical path analysis method in designing MUI).

1) Johnson's Rule

Johnson (1954) proposed an optimal scheduling method to arrange the order/sequence of jobs entering a system in which two machines arranged in a serial order: the optimal sequence can be obtained by partitioning the jobs into two sets, with Set I containing all the jobs with $p_{1j} < p_{2j}$ and Set II all the jobs with $p_{1j} \geq p_{2j}$. The jobs in Set I go first and they go in increasing order of p_{1j}; the jobs in Set II follow in decreasing order of p_{2j}.

2) Non-Identical Parallel Machine Scheduling Method

Sule (1996) proposed a scheduling method to assign jobs to machines in parallel with different processing time as well as arrange the order of jobs entering these machines. This method includes three steps.

Step 1. Rank the *m* parallel machines such that the most efficient machine (the one taking the least amount of time to process) is machine 1, the next efficient machine is machine 2, and so on. We can also rank the jobs in descending order of the processing times, indicating the job with the largest processing time as job1, the job with the next longest processing time as job2, and so on.

Step 2. Add the processing times of all jobs on machine 1. This is the current value of TT_1 (total of processing times assign to machine 1). TT_i for i=2,3,…m is 0, because no jobs are assigned to machine 2 through m. TT_1 is the present value of the makespan.

Step 3. Examine the feasibility of reassignment of jobs starting with the first job and proceeding toward job *n*. To do so, first select the candidate job. Temporarily remove it from machine 1 and assign it to all remaining machines. Reduce the value of TT_1 by processing time of the candidate job. Increase TT_i for machine *i* by the associated processing of the job on that machine and determine the minimum value of TT_i for $i=2,...m$ (except ignore TT_i =0). The associated processor is where the job should be assigned if it is to be moved from machine 1. Compare the minimum of TT_i with TT_1 and determine the least value between the two. If the new value of the makespan is less than the present value of the makespan, make the new assignment permanent and assign the makespan the new value. If the new makespan is not less than the present makespan, the reassignment of this job is rejected. Select the next job in the sequence and repeat the step. If all jobs are examined, stop; we have the best assignment.

1.3 Relationship between *Cmax* (Performance) and Subjective Driver Workload

The specific usage of multimodal systems in vehicles and the measurement of subjective driver workload suggest a specific relation between subjective mental workload and the total in-vehicle task completion time. In driving experiments, the data of subjective workload can only be collected after a driver drives for a certain amount of

time with at least several trials of an in-vehicle task; otherwise he or she will not have enough time to experience the workload in using the in-vehicle system (see Figure 7-1).

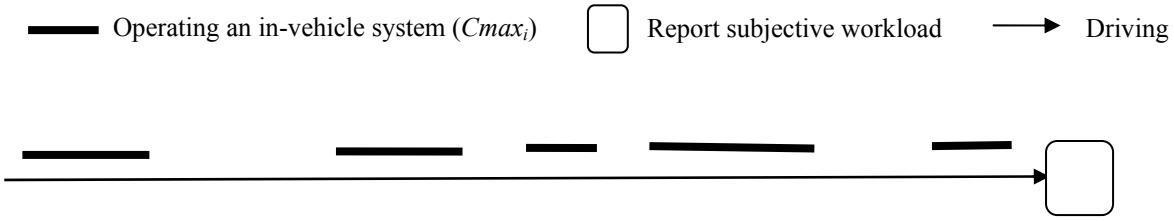


Figure 7-1 An illustration of operating an in-vehicle system while driving

If we regard the whole cognitive system as a server which process information from both a road and an in-vehicle system, there is a direct proportional relation between utilization of this server (ρ) and subjective driver workload (*WL)* (Wu & Liu, 2006a, 2006b, 2006c, 2006d):

$$WL=a\rho+b \quad (1)$$

where a and b are constants depending on different driving situations and in-vehicle systems ($a>0$). In Queueing network theory, utilization (ρ) of a single server can be quantified using the following equation:

$$\rho=\lambda/\mu \quad (2)$$

where λ is the arrival rate of information and μ is the processing speed of the server. Since *Cmax* is the total task completion time in each trial using the in-vehicle system, it is in inverse proportion to the processing speed of the cognitive system, i.e.:

$$\mu=1/Cmax \quad (3)$$

Combining equations above, we can easily have:

$$WL=a\lambda Cmax+b \quad (a>0) \quad (4)$$

which indicates a direct proportional relation between makespan (*Cmax*) and subjective driver workload. In other words, scheduling algorithms which minimize *Cmax* can also be used to reduce the subjective driver workload under the condition that the arrival rate of information remains the same[11].

[11] The arrival rate λ in this equation can also explain the situation that in-vehicle Tasks A and B with different arrival rates (e.g.., $\lambda_A > \lambda_B$) and same Cmax (CmaxA=CmaxB) will produce different subjective workload ($WL_A > WL_B$).

2. A General Procedure to Select the Scheduling Methods in MIVS

Based on the Theory of Constraints (TOC) and general procedure in application of scheduling methods in practice (McMullen, 1998; Pinedo, 2002), a general procedure was proposed to select and use the scheduling methods in designing MIVS. In addition, in order to illustrate the procedure clearly, the following definitions are proposed:

1) Bottleneck Stage (BS*i*): a stage (workstation) *i* with only one serial processor (machine)

2) Parallel Stage (PS*i*): a stage (workstation) *i* with more than one serial processor (machine) arranged in parallel

Step 1. Identify serial processors and status of stages

Qualitative method: a series of rules can be used to determine the serial or parallel processing at each processing stage (perceptual, cognitive and motor stages). For each job/task to be analyzed, first, it needs to be decomposed into (one, two or three) stages in the three major stages in processing information. Second, in each of these stages, it is necessary to identify how many serial processors/machines in that stage. For example, in one of the three stages, if there are two serial processors which can process information at the same time (e.g., the right hand is operating an in-vehicle device while the right foot is pressing a break), this stage can be regarded as a parallel stage (PS). On the other hand, if in a stage has only one serial processor processing the information/jobs one by one, this stage is regarded as a bottleneck stage (BS) (e.g., in the cognitive stage, subjects can only perform one arithmetic problem at one time).

Quantitative Method using QN-MHP: Based on the bottleneck identification methods in simulation and scheduling (Roser et al., 2002; Bank, 2004), the following bottleneck identification method is proposed to allocate a bottleneck stage: in the simulation results of the Queueing network model, if the lower bound of 95% confidence interval (CI) of the utilization of a server is greater than 0, this server will be regarded as a machine or a serial processor. Accordingly, each subnetwork can be analyzed in the following way:

1) Entities enter only one server in a subnetwork:
If that server is a serial processor → status of the subnetwork: BS
Else → Ignore that subnetwork in further analysis (I)
2) Entities enter more than one server in a subnetwork:
Case 1: If these servers are in parallel configuration→ status of the subnetwork: PS
Case 2: If these servers are in a serial configuration:
If that server is a serial processor → status of the subnetwork: BS
Else → Ignore that subnetwork in further analysis
Case 3: If these servers are in a network configuration:
If only one serial processor is found → status of the subnetwork: BS
If there are more than one serial processor is found:
Subcase 1: If these serial processors in parallel configuration→ status of the subnetwork: PS
Subase 2: If these serial processors are in a serial configuration→ status of the subnetwork: BS
Subase 2: If these serial processors are in a network configuration:
If there are branches in the network configurations→ status of the subnetwork: PS
Else→ status of the subnetwork: BS
3) No entities enter in a subnetwork→ Ignore that subnetwork in further analysis

Step 2. Choose the corresponding scheduling methods and schedule the job at each stage

Once the statuses of subnetworks are identified (BS, PS, or Ignored (I)), the second step is to select the scheduling methods to arrange the order of jobs based on following rules:

1) 1 BS:
1BS/1 PS/1 I: 1 BS→ Use Non-identical Parallel Machine Scheduling Method
1 BS/2PS → Use Non-identical Parallel Machine Scheduling Method to schedule each PS
1 BS/2 I: No scheduling methods is recommended for this situation since change of jobs orders will not affect *Cmax* and workload
2) 2BS:
2BS/1 PS: Use Non-identical Parallel Machine Scheduling Method to schedule the PS
If the 2 BS are connected directly → Use Johnson's Rule
Else→ only schedule the PS
3) 3BS: Simulation of performance and workload all of the combinations of modalities and orders
4) 0 BS: No scheduling methods is recommended for this situation since change of jobs orders will not affect *Cmax* and workload

Step 3. Rearrange the order of jobs or/and reassign the jobs to avoid inconsistency

Step 2 may generate different job orders or different assignment of jobs in in different stages. If this happens, considering the nature of human information processing—change the order of jobs within the cognitive system may cause extra load on the cognitive system, it is recommended to use the scheduling results (job orders and assignment of jobs) of the subnetwork/stage with higher utilization.

Step 4. Validate the scheduling results with simulation or experiment and design the MIVS based on the validated scheduling results

The following section described a case study in using the scheduling methods and procedure to select optimal modality and job orders when drivers are operating a multimodal in-vehicle system.

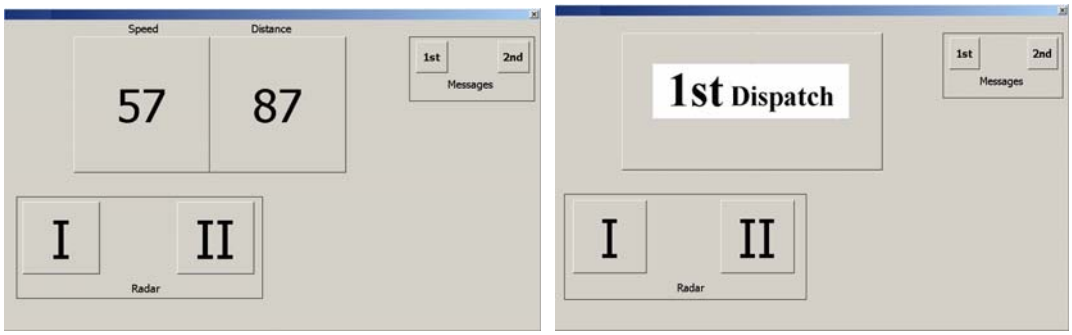
3. A Case Study

3.1 A Sample Multimodal In-Vehicle System with Practical Importance

Similar multitasking scenario of police officers described in Chapter 6 was used in this case study: 1) Speeding/Radar detection or judgment task (called Radar Judgment Task): Officers need to read two numbers on a display of an in-vehicle radar system mounted on dashboards of the police vehicles: the first number was the speed of a target vehicle measured by the radar system; the second number was the distance from the police vehicle to the target vehicle. The level of speeding of a target vehicle is determined by the speed and the distance together; 2) Message response task: Messages received by the offices usually come from multiple dispatches (first dispatch: headquarters; second dispatch: other police officers, etc.) by pressing a button on the radio.

The usage of the scheduling methods was able to assist designers of these multimodal in-vehicle systems to select the optimal modality to present the information of these tasks and determine which task is to be presented to drivers earlier. In this specific scenario, there are four possible combinations of modality and order of tasks: a) The message-response task was presented in the auditory modality and it was earlier than radar judgment task shown in visual modality (Mesg_AUD condition); b) The message-response task was shown in visual modality and it was earlier than the radar judgment task presented in the auditory modality (Mesg_VIS condition); c) The radar judgment task was presented in the auditory modality and it was earlier than message-response task shown in visual modality (Radar_AUD condition); b) The radar judgment task was shown in visual modality and it was earlier than the message-response task presented in the auditory modality (Radar_VIS condition). Figure 7-2 shows the user interface of the

multimodal system which includes the two pairs of response keys for the radar judgment and message response task (the response keys of message-response task were located 13 cm away from the response keys of the radar judgment task).



Mesg_AUD and Radar_VIS Conditions Mesg_VIS and Radar_AUD Conditions

Figure 7-2 The user interface of the multimodal in-vehicle system

For the message response task, whenever subjects heard or saw the word "first dispatches" (the presentation duration of the word "first" was 300 ms in the auditory modality, and 5 seconds in the visual modality) from the speakers or the touch screen, they were asked to double click on the "1st" button on the touch screen with their right index fingers; if they heard or saw "second dispatches" (the presentation duration of the word "second" was the same with that of word "first") , they were instructed to double click on the "2nd" button on the touch screen with the same fingers.

For the radar judgment task, subjects were asked to judge the level of speeding of another vehicle based on speed and distance information from the speakers or the touch screen the using following rules (the presentation duration of the speed and distance information were 850 ms in the auditory modality, and 5 seconds in the visual modality): a) If the speed is within the range from 55 to 60 (including 55 and 60), they need to see the distance: if the distance is beyond 65 yards (including 65), they were asked to press "II" button because it is a moderate speeding (level II); if the distance is below 65 yards, they were instructed to press "I" button since it is severe speeding (level I). b) If the speed is above 61 (including 61), you need to see the distance, if the distance is beyond 105 yards (including 105), it is moderate speeding (level II) and subjects were asked to press "II" button; if the distance is below 105 yards, subjects were instructed to press "I" button because it is severe speeding (level I).

3.2 Arrangement of Modality and Job Orders based on the General Procedure

Step 1: Identify serial processors and status of stages

Qualitative Method: Since both message-response task and radar judgment task involved decision making or judgment process in the cognitive stage, the cognitive stages was regarded as a bottleneck stage (BS). Because the information of the two tasks was processed through different sensory modalities, the perceptual stage was regarded as a parallel stage (PS). In addition, pressing the touch screen using the same finger implied a strictly serial processing at the motor processing stage for the two tasks (BS).

Quantitative Method using QN-MHP: Table 7-3 showed the simulation results of QN-MHP (5 replications, each replication includes 265 trials of the secondary task in operating the multimodal in-vehicle system), includings the CI of all servers and subnetwork in of the dual task (driving while perform the two tasks at the same time). The statuses on each stage were also listed in this table based on the rules in Step 1 in the general procedure.

Table 7-3 The 95% Confiendce Interval (CI) of all servers and subnetwork in QN-MHP (Simulation Results)

Subnetwork	Server/Subnetwork	95% CI of UTL			
		Lower Bound	Upper Bound	Mean	Subnetwork Status
Perceptual	Server 1	0.33	0.39	0.36	
	Server 2	0	0.08	0.04	
	Server 3	0	0.08	0.04	
	Server 4	-0.01	0.02	0.005	PS
	Server 5	-0.02	0.05	0.015	
	Server 6	-0.05	0.1	0.025	
	Server 7	-0.05	0.1	0.025	
	Server 8	0	0	0	
	Subnetwork	.18	.24	.21	
Central	Server A	-0.01	0.02	0.005	
	Server B	0	0.001	0.0005	
	Server C	0.17	0.25	0.21	
	Server D	0	0	0	SB
	Server E	0	0	0	
	Server F	0.12	0.53	0.325	
	Server G	0	0	0	
	Server H	0	0	0	
	Subnetwork	0.07	0.18	0.13	
Motor	Server V	0	0	0	
	Server W	-0.05	0.1	0.025	
	Server Y	-0.02	0.05	0.015	
	Server Z	-0.08	0.33	0.125	
	Server X	0	0	0	
	Server 21 (eye)	0.84	0.87	0.855	PB
	Server 22 (mouth)	0	0	0	
	Server 23 (left hand)	0	0	0	
	Server 24 (right hand)	0.23	1.07	0.65	
	Server 25 (left foot)	0	0	0	
	Server 26 (right foot)	0	0	0	
	Subnetwork	0.05	.36	.18	

Based on both qualitative and quantitative analysis above, a simplified configuration of the cognitive system in this specific scenario is summarized in Figure 7-3:

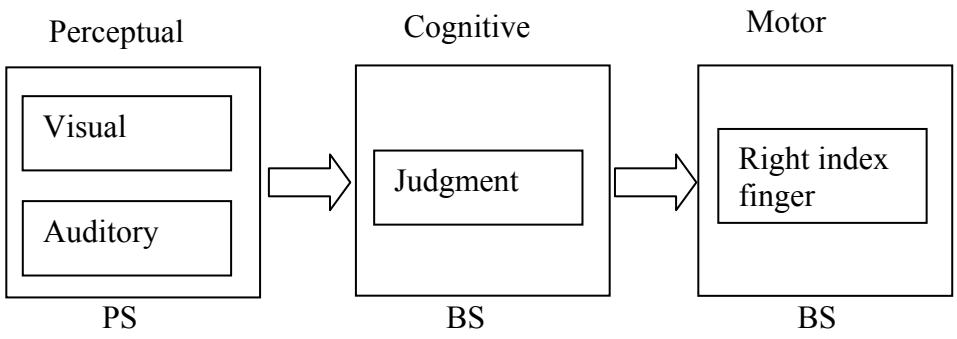


Figure 7-3 Status of stages in the cognitive system in performing the tasks in the case study

Step 2: Choose the corresponding scheduling methods and schedule the job at each stage

Based on the results of step 1,

1) Scheduling 2-Bottleneck (Cognitive and Motor Stage) Using Jonson's Rule

Based on the current design of the experimental task, $F_{mesg}<F_{radar}$ (complex rules operations of the radar task compared with the simple choice reaction task of the message task in Server F) and $M_{mesg}>M_{radar}$ (greater movement distance and double click movement in the message task compared to the smaller movement distance and single click in the radar task). According to the scheduling algorithm of two-machine, the message-response task (called "Message Job/Task", *Jm*) was assigned to Set I and the radar judgment task based on the radar's detection results (called "Radar Job/Task", *Jr*) was assigned to Set II. The order of these jobs enter Server F is *Jm* and then *Jr*. Accordingly, the *Jm* should be presented to subjects earlier than *Jr* (order of tasks). In order to guarantee that *Jm* arrives at Server F earlier than *Jr*, *Jm* should preferably be presented at a faster modality (modality of tasks). In the current experiment setting, auditory modality is the faster modality compared with the visual modality due to the following reasons: in the driving condition, it took driver at least one glance to shift their fixation from the road with curvatures to the visual stimuli of the in-vehicle task (Tsimhoni et al, 1999) compared with the condition when these information were presented in the auditory modality without eye movements. Therefore, *Jm* is assigned to the auditory modality so that the chance that *Jr* catches *Jm* and arrives at Server F is lower in the Mesg_AUD condition compared with Mesg_VIS condition. This can be described by the following simple mathematical equations:

$$A\_Jm=Per\_Jm+Cog\_Jm$$

A_Jr=Delay+Per_Jr+ Cog_Jr

Where *A_Ji* (Travel/Processing time of a task starting when it reaches the sensory modality to when it reaches the Server F), *Per_Ji* (perception time of a task in the perceptual subnetwork), *Cog_Ji* (processing time of a task in the cognitive subnetwork before they reaches Server F which is responsible for decision making and judgment), and *Delay* is the delay time between the stimuli of *Jm* and *Jr*.

If *Delay>0* and *Jm* is presented at auditory modality while *Jr* is presented at the visual modality in the current experimental setting (See Table 7-4 for detailed estimation of the processing time in each modality), then

Per_Jr> Per_Jm

Since both types go through the phonological server (Server B) and central executive server (Server C) in working memory,

Cog_Jr=Cog_Jm

Combing equations above, we can easily have *A_Jr>A_Jm,* so that the order of arrival of these two types of tasks in Server F in the Mesg_AUD condition.

While in the other conditions, since either *Delay<0* or *Per_Jr<Per_Jm*, inequality in Equation may not become true. Therefore, we can derive that the optimal order and modality for the current experiment is Mesg_AUD. And it is predicted that this optimal combination of order and modality produces the shortest makespan comparing with other combinations including Mesg_VIS, Radar_VIS and Radar_AUD.

2) Scheduling Perceptual Stage with non-identical parallel machine method

At the perceptual stage (subnetwork), since it is composed of multiple sensory modalities (processors) arranged in a parallel manner, the parallel non-identical machine scheduling method can be applied to arrange jobs/tasks in different modalities. Table 7-4 summarized the estimation of the processing time at the auditory and visual modalities based on current scenario.

Table 7-4 Estimated processing time of the two tasks in the auditory and visual modality

	Estimated Processing Time of the Two Tasks		
Modality/Processor	Message (Jm)	Radar (Jr)	Average
Auditory (P1)	300 ms (experiment setting)	800 ms (experiment setting)	550 ms
Visual (P2)	676[12] ms	676 ms	676 ms
Average	488 ms	738 ms	

Step 1. Rank the processing time of processors and jobs

Efficiency of processors: auditory modality (P1) faster than visual modality (P2)

Processing time of jobs: Jr>Jm

Step 2. Put all of the jobs at P1 with descending order of processing time

Jr (Job 1), Jm (job 2)→P1 (Auditory)

Step 3. Move longest job (job1) from P1 to other processors

Jm→ P1 (Auditory), Jr→ P2 (Visual)

=>*Cmax*=max(300, 676)=676 ms

Step 4. Move job 2 from P1 to other processors

Jr→ P1 (Auditory), Jm→ P2 (Visual)

=>*Cmax*=max(800, 676)=800 >676 → Reject

Therefore, the *Cmax* can be reduced if we assign Jm to auditory modality and Jr to visual modality[13].

Step 3. Rearrange the order of jobs or/and reassign the jobs to avoid inconsistency

[12] This processing time was estimated based on number of glances multiplied by glance duration looking at an in-vehicle system: 1.9 glances for an in-vehicle task with similar level of task difficulty (Tsimhoni et al, 1999). The duration of each glance was estimated based on MHP (Card, et al., 1983) and QN-MHP (Liu, et al., 2006): 230 ms (average eye movement time) +126 ms (42*3=126: servers' processing time at the visual subnetwork. Therefore, 1.9*(230+126)=676 ms

[13] The analysis using non-identical parallel scheduling method is assumed the two jobs arrive at the P1 and P2 at the same time, however, this method is still valid in *DELAY*>0 conditions:
Cmax (Move Jr to visual modality)=*max(Jm at AUD, Jr at VIS+DELAY)=max(300,DELAY+676)*
Cmax(Move Jm to visual modality)=*max(Jm at VIS, Jr at AUD+DELAY)=max(676,DELAY+800)*
If the selection of modality still works, then *Cmax*(Move Jr to visual modality)<=*Cmax*(Move Jm to visual modality)
i.e., max(150,DELAY+676)<=max(676,DELAY+800). If DELAY>0, then this equation can be simplified to *DELAY+676<=DELAY+800* Since *DELAY>0*, *DELAY+676<=DELAY+800* is always true.

Since the application of scheduling methods in the three stages in the step 2 produced the same scheduling results (message response task was to be presented in the auditory modality and it was earlier than the radar judgment task shown in the visual modality), it was not necessary to rearrange the order of jobs to avoid inconsistency

Step 4. Validate the scheduling results with simulation or experiment and design the MIVS based on the validated scheduling results (see the following section for validation of the scheduling results)

3.3 Experimental Validation

As described in the first section of the case study, it is predicted that the MESG_AUD condition selected by the scheduling methods should produce the minimal total task completion time and lowest subjective workload. An experiment was conducted to validate this prediction as described in the following section.

3.3.1 Experimental Design

A one-factor within-subject design was used in this experiment. The independent variable was the four combinations of modality and order of tasks as described in the first section of the case study: Mesg_AUG, Mesg_VIS, Radar_AUD, and Radar_VIS. The dependent variables were the makespan (total task completion time) of the secondary task (the in-vehicle task composed of message-response and radar judgment tasks), error rate of the secondary task, subjective workload measured by NASA-TLX, and driving performance measured by standard deviation of lane position. Each participant used the in-vehicle system in all of the four combinations of modality and order of tasks. The order of the four combinations in each participant was arranged following a Latin Square design so that the four combinations appeared first, second, third or fourth for exactly 1 participant.

3.3.2 Participants

Sixteen licensed drivers were paid to participate in this experiment (age 25-34 years, mean=31, SD=2.5; 8 male and 8 female). All participants were right-handed and had corrected far visual acuity of 20/40 or better. All had midrange (80 cm) visual acuity of

20/70 or better. Prescreening of all participants ensured they had good driving records and were physically healthy.

3.3.3 Equipment and Test Materials

Driving Simulator (see Chapter 6 for the description of UMTRI driving simulator).

Simulated Roads. The simulated road was a road with 250 m radius curvature. Both lanes of the two-lane road were 3.66 meter (12 feet wide). The length of the road in each condition of the in-vehicle system was 5,000 meters with 4 speed-limit signs (65 miles/hr) placed in the road in every 1,250 meters.

Touch Screen. An IBM laptop X60 with a 12'' touch screen was mounted on the right of the driver at arm's length. This touch screen was located in the center console of the vehicle, 23º ± 3º below the horizontal line of sight and 30º ± 3º to the right of the center (see Figure 7-4). To allow easy reading, numbers on the display were relatively large (digit height = 11 mm, 1º at 63 cm).

Figure 7-4 Driver's view of the road and the touch screen

3.3.4 Experimental Task and Procedure

Driving Task. Participants were instructed to drive in the right lane and maintain a speed following the speed-limit signs on the simulated roads. To maintain the driving speed of each participant, if they drove 5 miles/hr over or below the speed shown on the speed-limit signs, they heard a computer-generated voice "too fast" or "too slow".

Secondary Task. The secondary task was composed of two tasks (message-response and radar judgment) as described in the first section of this case study. Participants were

asked to complete the tasks as quick and accurate as possible (during the experiment, the delay between stimuli of the two task was set at five hundred million seconds).

Experimental Procedure. After filling in the pretest forms and vision test, in the practice section, participants first practiced single task situation: driving along without the secondary task; and the secondary task along while the simulator was in parked condition. Then, participants practiced dual task situation: driving while performing the secondary task at the same time. In the test section, participants were instructed to drive with the multimodal system in its four conditions (participants drove 5,000 meters in each condition). After participants finished each condition, they were asked to complete the NASA-TLX form to report their subjective workload.

3.3.5 Experimental Results

1) Performance of the Secondary Task

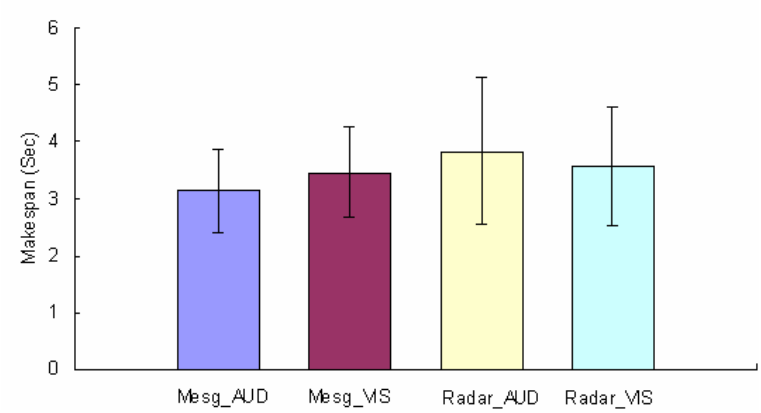


Figure 7-5 The average makespan in the four combinations of modalities and orders (Error bars represent ±1 SD of Cmax)

Figure 7-5 shows the average makespan in the four combinations of modalities and order in driving condition. The main effect of the four combinations of modality and order on makespan was significant ($F(3,45)=14.46$, $p<.001$). The tests of one-factor within-subject contrasts (treating the 4 combinations of modalities and order as one within-subject variable) found that significant difference between the Mesg_AUD with other conditions (Mesg_AUD vs. Mesg_VIS: $F(1,15)=16.61$, $p<.001$; Mesg_AUD vs. Radar_AUD: $F(1,15)=62.85$, $p<.001$; Mesg_AUD vs. Radar_VIS: $F(1,15)=49.96$, $p<.05$).

In addition, the main effect of the four combinations of the modality and order on error rate of the secondary task was not significant ($F(3, 45)=1.64$, $p>.05$). Furthermore, the error rates at Mesg_AUD condition was not significantly different from other three combinations of the modality and order (Mesg_AUD vs. Mesg_VIS: $F(1,15)=1.59$, $p>.05$; Mesg_AUD vs. Radar_AUD: $F(1,15)=1.56$, $p>.05$; Mesg_AUD vs. Radar_VIS: $F(1,15)=.65$, $p>.05$).

2) Mental Workload

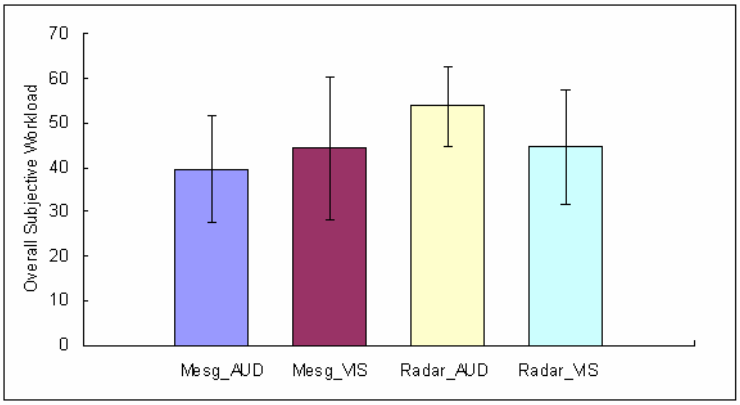
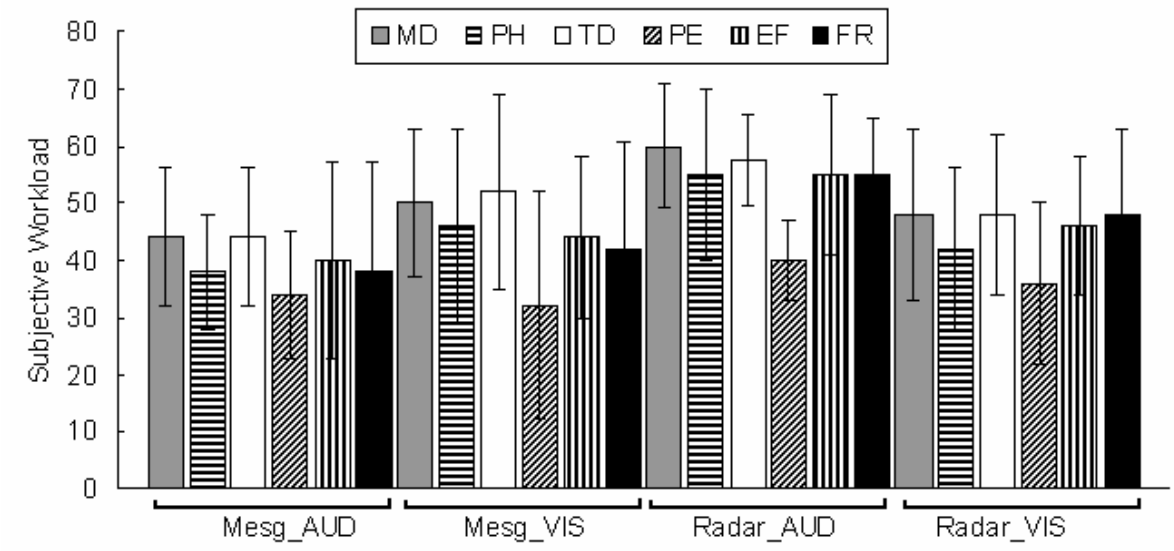


Figure 7-6 Overall subjective workload in the four combinations of modalities and orders (Error bars represent ±1 SD of the overall subjective workload)

Figure 7-6 showed the overall subjected workload measured by NASA-TLX in the four combinations of modality and order. The main effect of the four combinations of modality and order on mental workload was significant ($F(3, 45)=19.98$, $p<.001$). The overall workload at the Mesg_AUD condition was significantly less than the other conditions (Mesg_AUD vs. Mesg Vis: $F(1,15)=6.05$, $p<.05$; Mesg_AUD vs. Radar_AUD $F(1,15)=43.75$, $p<.001$; Mesg_AUD vs. Radar_VIS: $F(1,15)=6.51$, $p<.05$).



Mental Demand (MD), Physical Demand (PH), Temporal Demand (TD),Performance(PE), Effort (EF), Frustration (FR)

Figure 7-7 The six dimensions of subjective NASA-TLX workload in the four combinations of modality and order (Error bars represent ±1 SD of the subjective workload)

Figure 7-7 presented the comparison of subjective workload between Mesg_AUD condition and the other 3 conditions in the six dimensions measured by NASA-TLX. In mental demand (MD), physical demand (PD), and temporal demand (TD) dimensions, the subjective workload of Mesg_AUD at Mesg_AUD was significantly lower than Mesg_VIS and Radar_AUD conditions; in the effort (EF) and frustration (FR) dimensions, the subjective workload at Mesg_VIS condition was significantly lower than all of the other conditions; however, in the performance (PE) dimension, subjective workload at Mesg_AUD condition was only significantly lower than the Mesg_AUD condition (See Table 7-5). In addition, the main effect of combinations on the subjective workload of the six dimensions of NASA-TLX was significant except the PE dimension (MD: $F(3, 45)=18.01$, $p<.001$; PH: $F(3, 45)=27.30$, $p<.001$; TD: $F(3, 45)=10.06$, $p<.001$; PE: $F(3, 45)=1.92$, $p>.05$; EF: $F(3, 45)=20.60$, $p<.001$; FR: $F(3, 45)=21.97$, $p<.001$).

Table 7-5 Comparison of Mesg_AUD condition with the other conditions in the six dimensions of NASA-TLX

Dimensions	Comparison	F(1,15)	Sig	Dimensions	Comparison	F(1,15)	Sig
Mental	Mesg_AUD vs. Mesg_VIS	5.95	*	Performance	Mesg_AUD vs. Mesg_VIS	.32	
Demand	Mesg_AUD vs. Radar_AUD	22.84	**	(PE)	Mesg_AUD vs. Radar_AUD	5.95	*
(MD)	Mesg_AUD vs. Radar_VIS	2.14			Mesg_AUD vs. Radar_VIS	.17	
Physical	Mesg_AUD vs. Mesg_VIS	15.00	**	Effort	Mesg_AUD vs. Mesg_VIS	11.67	**
Demand	Mesg_AUD vs. Radar_AUD	55.51	**	(EF)	Mesg_AUD vs. Radar_AUD	33.99	*
(PH)	Mesg_AUD vs. Radar_VIS	2.049			Mesg_AUD vs. Radar_VIS	7.64	**
Temporal	Mesg_AUD vs. Mesg_VIS	5.99	*	Frustration	Mesg_AUD vs. Mesg_VIS	11.67	**
Demand	Mesg_AUD vs. Radar_AUD	23.50	**	(FR)	Mesg_AUD vs. Radar_AUD	34.61	**
(TD)	Mesg_AUD vs. Radar_VIS	2.14			Mesg_AUD vs. Radar_VIS	40.00	**

*: $p<.05$; **: $p<.01$

3) Driving Performance

The main effect of the four combinations of the modality and order on the standard deviation of lateral lane position was not significant ($F(3,45)=1.05$, $p>.05$). Furthermore, standard deviation of lateral lane position at Mesg_AUD condition was not significantly different from the other three combinations of the modality and order (Mesg_AUD vs. Mesg_VIS: $F(1,15)=1.07$, $p>.05$; Mesg_AUD vs. Radar_AUD: $F(1,15)=.18$, $p>.05$; Mesg_AUD vs. Radar_VIS: $F(1,15)=2.04$, $p>.05$).

4. Discussion

This study proposed a general procedure to apply several scheduling methods in designing multimodal in-vehicle systems including how to select the modalities and arrange the order of tasks. Theoretically, it introduced two new scheduling methods—Johnson's Rule and non-identical parallel machine scheduling method from scheduling theory to human factor research in transportation. Practically, the general procedure and scheduling methods described in this study can also be applied to design the multimodal user interface in other man-machine systems.

The case study in the current work used a small number of tasks and considered the two most commonly used modalities (visual and auditory) in man-machine interaction, however, when the number of tasks or modalities increases because of increased usage of

in-vehicle information/warning/security systems, it becomes more effective to use these scheduling methods to design multimodal in-vehicle systems. For example, if there are four messages to be processed by a driver in visual, auditory and tactile modalities (e.g., Message 1 from road guidance system, Message 2 from vehicle status monitoring system, Message 3 from vehicle-to-vehicle communication system, and Message 4 from cellular phone), the minimal number of full combination of modality and order is [14]: 3×2×1×3×2=36. In practice, it might be very time consuming to test all of the possibilities of modalities and orders; while using the scheduling methods described in this chapter can save part of the effort and select the optimal combinations following some algorithms.

More importantly, the current general procedure can be a platform for human factors researchers to select other scheduling methods which can consider other aspects of jobs (e.g., the priority of jobs, number of tardy jobs etc.). For example, if it is selected that there are two serial stages connected directly, even though it is difficulty to use Johnson's Rule to arrange the jobs with priority, we can use the general procedure to select scheduling methods which can handle this problem because the taxonomy of scheduling methods are organized in this manner (starting from single machine, multiple machines, and parallel machine etc.). Users can easily access these scheduling methods via the major reviews and textual book in scheduling theory (e.g., Sule, 1996; Pinedo, 2002; French, 1986) and even use the free scheduling software (e.g., LEKIN® developed by School of Business at New York University). In scheduling, many scheduling algorithms are very complex including using dynamic programming and artificial intelligence techniques which are far beyond the scope of human factors and transportation safety research, therefore, the critical thing becomes how to define a human-machine problem into a scheduling problem and select a proper scheduling method to solve this problem because the algorithms themselves have been coded in these scheduling software. Accordingly, before researchers in human factors and transportation safety use these

[14] 3×2×1×3 (the first message can be assigned to one of the three modalities; the second message can be assigned to the two modalities left; the third message is assigned the last modality; the fourth message restarts this process) ×2 (the order of 4th message and one of the previous messages also need to be considered)=36.

scheduling methods, the general procedure proposed in this chapter can be one of bases to assist them to select and use these complex scheduling methods.

There are several limitations of the current work that need to be examined in future research. First, the current scheduling methods and general procedure introduced are not able to predict the makespan (total task completion time) and workload of drivers; in other words, they can suggest modalities and order only at an ordinal scale. In many cases, these ordinal results can satisfy the purposes in designing in-vehicle systems, however, new algorithms or simulation models are needed if a designer hopes to compare the makespan and workload at the interval or ratio scale. Second, step 1 in the current general procedure used on a complex quantitative method to allocate the bottleneck stages and serial processors, and designers who do not have experience in using simulation model of human performance may feel reluctant to use this quantitative method in step 1; therefore, future research may need to develop a relatively easy-to-use quantitative method or algorithm to identify the bottleneck stages and serial processors. Third, the modality shifting effect (Spencer & Driver, 1997) was not considered in the current work because the order of tasks within each condition of the in-vehicle system in the case study was fixed while modality shifting effect is mainly related to a shift of modalities in an unexpected condition. Future research which can predict the makespan and workload need to consider this important effect in multimodal research, either considering it as part of delay time of the second task/job entering the cognitive system or prolonging the perception time of the second task.

Acknowledgement

This article is based upon work supported by the Doctoral Studies Program (DSP) at the University of Michigan Transportation Research Institute (UMTRI). We would like to thank Mr. Jacob Mouro, a public safety officer at University of Michigan Security Service who gave us tremendous help in describing police vehicles and multitasking scenarios in their daily work.

Reference

Banks, J., Carson, J., Barry, L. N., & Nicol, D. (2002). Discrete-Event System Simulation. Upper Saddle River, N.J. : Prentice Hall,.

Cellario, M. (2001). Human-centered intelligent vehicles: toward multimodal interface integration. IEEE Intelligent Transportation System, 16(4), 78-81.

Gupta, N., Bisantz, A. M., & Singh, T. (2002). The effects of adverse condition warning system characteristics on driver performance: an investigation of alarm signal type and threshold level. Behaviour & Information Technology, 21(4), 235-248.

Harold, K. (2001). Project Management: A Systems Approach to Planning, Scheduling & Controlling. New York John Wiley.

Johnson, S. M. (1954). Optimal two-and three-machine production schedules with setup times included. Naval Research Logistics Quarterly, 1(61-68).

Mariani, M. (2002). COMUNICAR: Designing multimodal interaction for advanced in-vehicle interfaces. In Human Factors in Transportations, Communication, and the Workplace (pp. 113–120): Netherlands: Shaker Publishing.

McMullen, T. B. (1998). Introduction to the Theory of Constraints (TOC) Management System: St. Lucie Press.

Roser, C., Nakano, M., & Tanaka, M. (2002). SHIFTING BOTTLENECK DETECTION Proceedings of the 2002 Winter Simulation Conference.

Sarter, N. B. (2001). Multimodal Communication In Support of Coordinative Functions In Human-Machine Teams. Journal of Human Performance in Extreme Environments, 5(2), 50-54.

Siewiorek, D., Smailagic, A., & Hornyak, M. (2002). Multimodal contextual car-driver interface. Paper presented at the Proceedings of the Fourth IEEE International Conference on Multimodal Interfaces.

Spence, D., & Driver, J. (1997). Cross-Modal Links in Attention Between Audition, Vision, and Touch: Implications for Interface Design. International Journal of Cognitive Ergonomics, 1(4), 351-373.

Sule, D. R. (1996). Industrial Scheduling: PWS Publishing Company.

Tsimhoni, O., Yoo, H., & Green, P. (1999). Effects of Visual Demand and In-Vehicle Task Complexity on Driving and Task Performance as Assessed by Visual Occlusion, Technical Report UMTRI-99-37: UMTRI.

Wu, C., & Liu, Y. (2004a). *Modeling human transcription typing with Queueing network-model human processor.* Paper presented at the Proceedings of the 48th Annual Meeting of Human Factors and Ergonomics Society, New Orleans, Louisiana, USA.

Wu, C., & Liu, Y. (2004b). *Modeling Psychological Refractory Period (PRP) and Practice Effect on PRP with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA. 320-325.

Wu, C., & Liu, Y. (2006a). *Modeling fMRI BOLD Signal and Reaction Time of a Dual Task with a Queueing Network Modeling Approach.* Paper presented at the 28th Annual Conference of the Cognitive Science Society, Vancouver, BC, Canada.

Wu, C., & Liu, Y. (2006b). *Queueing Network Modeling of a Real-time Psychophysiological Index of Mental Workload—P300 Amplitude in Event-Related Potential (ERP).* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.

Wu, C., & Liu, Y. (2006c). *Queueing Network Modeling of Age Differences in Driver Mental Workload and Performance.* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.

Wu, C., & Liu, Y. (2006d). *Queueing Network Modeling of Driver Workload and Performance.* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.

Wu, C., & Liu, Y. (2006e). *Queueing Network Modeling of Reaction time, Response Accuracy, and Stimulus-Lateralized Readiness Potential Onset Time in a Dual Task.* Paper presented at the 28th Annual Conference of the Cognitive Science Society, Vancouver, BC, Canada.

Chapter 8
Conclusions and Future Research

Chapter Summary

This chapter summarizes the major results of computational modeling in the previous chapters. It also describes the properties of Queueing networks which are able to quantify various phenomena and aspects of the perceptual-motor tasks. The limitations of the current modeling approaches in theory and practice and the corresponding focuses in future research are also discussed.

1. Summary of the Thesis

This thesis used QN-MHP as a platform to model four representative tasks (transcription typing, psychological refractory period, visual-manual tracking and steering) in the taxonomy of perceptual-motor tasks; then QN-MHP is applied into the design of an adaptive workload management system in vehicles and multimodal in-vehicle systems. The modeled dependent variables include: a) human performance measured by reaction time and response errors; b) fixation duration of and saccade size of eyes measured by eye trackers; and c) mental workload measured by P300 amplitude and NASA-TLX.

In modeling the discrete perceptual-motor task in a single task situation (transcription typing), QN-MHP quantifies and unifies 32 transcription typing phenomena involving many aspects of human performance—interkey time, typing units and spans, typing errors, concurrent task performance, eye movements, and skill effects, providing not only

an alternative way to model this basic and common activities in human-machine interaction, but also a multi-purpose analysis tools for textual data-entry tasks in human-computer interaction.

In quantifying the discrete perceptual-motor task in dual task situation (psychological refractory period, PRP), the Queueing network model is able to account for various experimental findings in PRP including all of these major counterexamples of existing models with less or equal number of free parameters and no need to use task-specific lock/unlock assumptions required by both EPIC and ACT-R/PM, thus demonstrating its unique advantages in modeling dual-task performance of discrete perceptual-motor tasks.

In modeling human performance and mental workload in the continuous perceptual-motor tasks (visual-manual tracking and car steering), we used QN-MHP as the simulation platform and then developed a set of equations to establish the quantitative relationships between Queueing networks (e.g., subnetwork's utilization and arrival rate) and P300 amplitude measured by ERP techniques and subjective mental workload measured by NASA-TLX. This modeling approach not only has a basis in its biological plausibility, but also has the ability to model and predict workload in real-time and allows researchers to visualize driver mental workload in real-time.

Extending mental workload modeling approach, this thesis also developed a prototype of adaptive workload management systems (AWMS) to dynamically control the rate of messages from these in-vehicle systems based on the properties of the secondary task, driving conditions (speeds and curvatures) and characteristics of drivers (age). A corresponding experimental study was conducted to validate the potential effectiveness of this system in reducing driver workload and improving driver performance.

In addition, based on the simulation results of QN-MHP, this thesis also proposed a general procedure to select several scheduling methods and use them to schedule two tasks (modality assignment and order of presentation) in an example multimodal in-vehicle system. An empirical study was conducted and validated the scheduling results including the optimal modality and order of these tasks.

2. Properties of Queueing Networks in Modeling Perceptual-Motor Tasks

The major properties of Queueing networks allow QN-MHP to unify several important aspects of the perceptual-motor tasks into one cognitive architecture. First, in modeling the discrete and continuous perceptual-motor behavior in single task situations (e.g., phenomena 1, 7, and all of the three eye-movements phenomena in transcription typing and visual-manual tracking), the number of entities being sampled by the visual subnetwork at one time is determined by how fast the entities are processed at the cognitive and motor subnetwork. This information processing property of Queueing/waiting is a unique feature of Queueing networks to quantify coordination and interaction among different components in the cognitive system.

Second, the entity-based network structure/arrangement in Queueing networks permits QN-MHP to quantify the utilizations of different components in the cognitive system and visualize the congestions of information processing in the cognitive system naturally; in addition, the routing probability of entities in the network during the learning process of the perceptual-motor tasks and the various units of visual-motor coordination (e.g., copy span, eye-hand span etc.) can also be predicted using this feature of Queueing networks. These two points above are consistent with the findings in neuroscience (Bullock, 1968; Eagleman, Jacobson, & Sejnowski, 2004; E. E. Smith & Jonides, 1998; Taylor et al., 2000; Braus, 2004; Chklovskii, et al., 2004; Habib, 2003).

Third, both serial and parallel information processing capability in the network provides QN-MHP a mathematical framework to quantify constraints of the cognitive system in performing the discrete and continuous perceptual-motor task in dual task situations (psychological refractory period, visual-manual tracking and steering with secondary tasks). For example, the network structure with both serial and parallel information processing capability allows QN-MHP to quantify the experimental results of Schumacher et al. (1999) and Hawkins et al. (1979) in the subadditive difficulty effect in PRP; the serial information processing at Server F allows QN-MHP to model the basic PRP including its brain imaging patterns, the response grouping effect, and also account for the reason why it is extremely difficult for people to perform two complex mental operations (e.g., mathematical calculations) simultaneously without extensive practice;

moreover, the parallel information processing in the other servers in the cognitive subnetwork enables QN-MHP to model the disappearance of the PRP effect after extensive practice.

Fourth, compared with other computational architectures, another important feature of QN-MHP is its overall mathematical structure which allows researchers to model simple and complex perceptual-motor tasks with both mathematical and computer simulation methods: when the perceptual-motor tasks are relatively simple (e.g., PRP), human performance can be quantified directly using mathematical equations in the Queueing network; when the perceptual-motor tasks are complex (e.g., visual-manual tracking and driving), the mathematical properties of Queueing networks (e.g., arrival rate and utilization) provide ideal connection points between the simulation results of the network and the other dependent variables (e.g., subjective workload, P300 amplitude). As long as a quantitative relationship between a dependent variable and one or several mathematical properties of Queueing networks can be built, the Queueing network may be used to quantify or predict the dependent variable. This can also explain the modeling methodology used in this thesis—integration of simulation approach and mathematical modeling approach in modeling the perceptual-motor tasks.

In summary, even though all of the four representative perceputual-motor tasks cover the four different categories of perceptual-motor tasks, this thesis mainly used the properties of queueing network theory to quantify the four tasks without making post-hoc assumptions. The properties of queuieng network modeling approach can be used to model other perceptual-motor tasks, reflecting the characterisitics of perceptual-motor tasks with its unique features.

3. Limitations of Current Modeling Approach and Future Research

There are several limitations of the current work that need to be examined in future research. Theoretically, the Queueing network approaches only quantified four representative perceptual-motor tasks, while high-level cognitive phenomena including problem solving, reading comprehension, spatial cognition, and complex reasoning are not covered in this thesis work. This is also the reason why QN-MHP is not able to quantify the two typing phenomena in transcription typing related to reading

comprehension. Moreover, even within perceptual-motor tasks, this thesis mainly covers the coordination and interactions among different components of the cognitive systems in the four representative perceptual-motor tasks. The other characteristics (e.g., stimulus-response compatibility, Chua, Weeks, & Goodman, 2002) of perceptual-motor tasks need to be modeled or incorporated into the modeling mechanisms in future research. These high-level cognitive tasks and other characteristics of perceptual-motor tasks can be modeled by improving the processing logic at the servers' level and the subnetwork's level.

Based on the similar mechanisms in modeling PRP and other dual tasks, further modeling work with QN-MHP is needed to explain and quantify how different brain regions were activated in various kinds of dual-task situations summarized by Collette & Linden (2002). Because the functions of the cognitive system in QN-MHP are distributed among the servers in the network and QN-MHP does not rely on a certain "conductor" server in coordinating or controlling the processing of other servers, it is possible to model these fMRI studies' results with the natural interaction among servers without an executive control mechanism.

QN-MHP currently uses the NGOMSL method to analyze each task before a model simulation; if the strategies of subjects change, the NGOMSL description needs to be changed, thus constraining the model's ability to quantify individual differences as a function of task strategies. Future research is needed to generate NGOMSL task description by the model itself via learning and practice processes and quantify individual differences in planning and executing actions in various tasks.

In practice, one of the major directions of QN-MHP is to develop a computational model of a driver. At the current stage, QN-MHP only quantifies driver performance and workload in steering with a secondary task while a driving task can involve speed control (car following and responding to road events), route planning and navigation, and lane changing. A relatively complete model of drivers can help people in the transportation research area understand the mechanisms in driving, quantify/predict the driver behavior and workload, and design the corresponding in-vehicle systems to improve transportation safety.

Unlike the established cognitive architectures including ACT-R, SOAR and EPIC, QN-MHP at the current stage is not a widely used cognitive model in the human factors area. Accordingly, we plan to upload the whole model including its documentations on the INTERNET so that researchers in the cognitive modeling and human factors area are able to download and use the model; the corresponding tutorials and workshop will also be held to introduce the model to people who are interested in using this model to quantify human performance and mental workload in various tasks. In addition, it is important to create a relatively easy-to-use user interface for QN-MHP so that the users of the model can easily quantify human performance and workload in various tasks. We are currently working on this solution using Visual Basic Application in Excel to add a user interface to QN-MHP so that researchers and even user interface designers can use this model with minimal efforts to learn programming and mathematics (Wu & Liu, In Press).

In summary, the Queueing network modeling approaches offer alternative methods in modeling and quantifying perceptual-motor tasks. We are systematically extending these modeling approaches to cover a broader range of tasks. Our comprehensive computational model of perceptual-motor tasks (QN-MHP) offers not only theoretical insights into human performance and mental workload, but is a step toward developing proactive ergonomic design and multi-purpose analysis tools for tasks in human-computer interaction.

Reference

Brass, M., & Cramon, D. (2002). The role of the frontal cortex in the task preparation. *Cerebral Cortex, 12*, 908-914.

Braus, D. F. (2004). Neurobiology of learning - The basis of an alteration process. *Psychiatrische Praxis, 31*, S215-S223.

Bullock, T. (1968). Representation of information in neurons and sites for molecular participation. *Proc Natl Acad Sci U S A, 60*(4), 1058-1068.

Chklovskii, D. B., Mel, B. W., & Svoboda, K. (2004). Cortical rewiring and information storage. *Nature, 431*(7010), 782-788.

Chua, R., Weeks, D. J., & Goodman, D. (2002). Perceptual-Motor Interaction: Some Implications for Human-Computer Interaction. In J. A. Jacko & A. Sears (Eds.), *Human-Computer Interaction Handbook* (pp. 23-34): Lawrence Erlbaum Associates, INC.

Collette, F., & Linden, M. V. (2002). Brain Imaging of the Central Executive Component of Working Memory. *Neuroscience and Biobehavioral Review, 26*, 105-125.

Eagleman, D., Jacobson, J., & Sejnowski, T. (2004). Perceived luminance depends on temporal context. *Nature, 428*(6985), 854-856.
Eagleman, D. M., & Churchland, P. S. (2005). *Ten Unsolved Questions of Neuroscience.*: MIT Press.
Habib, M. (2003). Rewiring the dyslexic brain. *Trends in Cognitive Sciences, 7*(8), 330-333.
Hawkins, H. L., Rodriguez, E., & Reicher, G. M. (1979). *Is time-sharing a general ability?* (No. Rep. No.3). Eugene: University of Oregon.
Schumacher, E. H., Elston, P. A., & D'Esposito, M. (2003). Neural evidence for representation-specific response selection. *Journal of Cognitive Neuroscience, 15*(8), 1111-1121.
Schumacher, E. H., & Jiang, Y. H. (2003). Neural mechanisms for response selection: Representation specific or modality independent? *Journal of Cognitive Neuroscience, 15*(8), 1077-1079.
Schumacher, E. H., Lauber, E. J., Glass, J. M., Zurbriggen, E. I., Gmeindl, L., Kieras, D. E., et al. (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., et al. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological Science, 12*(2), 101-108.
Smith, E. E., & Jonides, J. (1998). Neuroimaging analyses of human working memory. *Proc. Natl. Acad. Sci. USA, 95*, 12061-12068.
Taylor, J. (2003). Paying attention to consciousness. *Progress in Neurobiology, 71*(4), 305-335.
Taylor, J., Horwitzc, B., Shaha, N. J., Fellenzb, W. A., Mueller-Gaertnera, H.-W., & Krausee, J. B. (2000). Decomposing memory: functional assignments and brain traffic in paired word associate learning. *Neural Networks, 13*, 923-940.
Wu, C., Li, H., Yang, Q., & Zhang, K. (2003b). Human performance modeling in spatial segmentation Chinese handwriting recognizer. *Acta Psychologica Sinica, 35*(6), 767-776.
Wu, C., & Liu, Y. (2004a). *Modeling Behavioral and Brain Imaging Phenomena in Transcription Typing with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA. 314-319.
Wu, C., & Liu, Y. (2004b). *Modeling Psychological Refractory Period (PRP) and Practice Effect on PRP with Queueing Networks and Reinforcement Learning Algorithms.* Paper presented at the Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA. 320-325.
Wu, C., & Liu, Y. (2004c). *Modeling human transcription typing with Queueing network-model human processors.* Paper presented at the Proceedings of the 48th Annual Meeting of Human Factors and Ergonomics Society, New Orleans, Louisiana, USA. 381-385.
Wu, C., & Liu, Y. (2006a). *Queueing Network Modeling of a Real-time Psychophysiological Index of Mental Workload—P300 Amplitude in Event-Related Potential (ERP).* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.
Wu, C., & Liu, Y. (2006b). *Queueing Network Modeling of Driver Workload and Performance.* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.
Wu, C., & Liu, Y. (2006c). *Queueing Network Modeling of Age Differences in Driver Mental Workload and Performance.* Paper presented at the 50th Annual Conference of the Human Factors and Ergonomics Society, San Francisco, CA, USA.
Wu, C., & Liu, Y. (2006d). *Queueing Network Modeling of Reaction time, Response Accuracy, and Stimulus-Lateralized Readiness Potential Onset Time in a Dual Task.* Paper presented at the 28th Annual Conference of the Cognitive Science Society, Vancouver, BC, Canada.
Wu, C., & Liu, Y. (2006e). *Modeling fMRI BOLD Signal and Reaction Time of a Dual Task with a Queueing Network Modeling Approach.* Paper presented at the 28th Annual Conference of the Cognitive Science Society, Vancouver, BC, Canada.
Wu, C., & Liu, Y. (In Press). A New Software Tool for Modeling Human Performance and Mental Workload. *Ergonomics in Design*.