ABSTRACT

KIM, NAYOUNG. Effective Connectivity Based ACT-R Modeling of Workload Transition in Multitasking. (Under the direction of Dr. Chang S. Nam).

Understanding how human operators cope with unexpected workload transitions is one of the important ergonomic issues. As the level of task demand changes, the level of effort operators exerted for the current task may influence subsequent behavior (e.g., performance, perceived mental workload, or brain activity). The impact of previous demand conditions on subsequent conditions is known as 'workload transition effect', or 'hysteresis effect.'

Past research has shown inconsistent results on workload transition or Hysteresis effects at the behavioral, perceived workload, or neurophysiological level, resulting in three different hypotheses: Enhancement, Deterioration, or No change. Adaptive Control of Thought-Rational (ACT-R) is one of the cognition theories that seek to predict human performance in multitasking by utilizing a computational simulation of human cognitive processing. This study adopted ACT-R as a theoretical framework to address the challenges identified from the previous research, to better understand the underlying behavioral and neurocognitive processes that occur during workload transitions, and to predict the impacts of hysteresis effects.

This study intended to achieve the following objectives: a) Confirm workload transition effects by comparing the transitioned group and the non-transitioned group in terms of performance, mental workload, and brain activity change over time in two different settings; and b) Propose and validate a new method to quantify and adjust ACT-R's main parameters based on neural correlates of workload transition. This research consisted of two studies, each study with two phases.

This research addressed the first research question, "What is the nature of workload transition effects?", by examining the behavioral, perceived workload, and neurophysiological

measures in both a basic cognitive task (in Study 1) and a more ecologically valid task (AF-MATB) (in Study 2). The findings of the present study confirmed workload transition effects at the three measurement levels in both task settings.

The second research question, "What is the nature of the effects of workload transition variation types at the behavioral, perceived workload, and neurophysiological levels?" supported the "Enhancement" hypothesis. The cyclic transition type was shown to significantly affect task performance, perceived mental workload, and brain connectivity, as compared to other transition types. We investigated the effective connectivity and identified differentiating brain regions that coincide with the ACT-R module associated with memory retrieval.

Finally, answers to the third research question "What is a systematic way that can improve the fitness of the ACT-R model that simulated human performance during workload transitions?" (Phase II) outlined a parameter tuning framework using both behavioral measures and effectivity connectivity of human operators. We suggested an effective connectivity-based method, which significantly increases the amount of information that can be extracted from EEG data to distinguish workload profiles. By understanding neural correlates and ACT-R production rule processing, in the current study suggested a novel way to scale the significant three parameters of ACT-R; retrieval threshold, latency-scaling factor, and activation noise. Through Studies 1 and 2, we validated that the ACT-R model with the revised parameters resulted in better model fitness than the model with default parameters.

Through neural analysis, the study found that people with better performance exhibited the strong activation of the temporal regions which are involved in retrieving task-relevant information. We also found that cyclic workload profiles might better induce continued activation of the temporal regions than other workload profiles.

By combining the ACT-R modeling technique and EEG-based effective connectivity analysis, this dissertation demonstrated how researchers could identify and analyze the effects of workload transition in various work settings. Finally, the findings of the present study would pave the way towards the use of brain network patterns for automatically monitoring and predicting cognitive states, which can subsequently help us develop neuroadaptive systems.

Effective Connectivity Based ACT-R Modeling of Workload Transition in Multitasking

by
Nayoung Kim

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Industrial Engineering

Raleigh, North Carolina
2019

APPROVED BY:

_____	_____
Chang S. Nam Committee Chair	Justin Post
_____	_____
Karen Chen	Kristen A. Lindquist

DEDICATION

To my parents and husband.

BIOGRAPHY

Nayoung Kim is a Ph.D. candidate in the Edward P. Fitts Department of Industrial and Systems Engineering at North Carolina State University, with a Bachelor of Industrial Engineering from Korea Advanced Institute of Science & Technology, Daejeon, and a Master of Management Engineering from Korea Advanced Institute of Science & Technology, Seoul.

ACKNOWLEDGMENTS

I am truly thankful to my advisor, Dr. Nam. His guidance and support were invaluable during this process! I would also like to thank my committee members, Dr. Chen, Dr. Feng, Dr. Lindquist, and Dr. Post. Their encouragement and advice contributed greatly to my success in this endeavor. I truly appreciate all of the time and effort put forth by all five of my committee members.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF CONTRIBUTING PUBLICATIONS

Peer-reviewed Journal Articles

Kim, N., House, R., Yun, M. H., and Nam, C. S. (2018). Neural Correlates of Workload Transition in Multitasking: An ACT-R Model of Hysteresis Effect. *Frontiers in Human Neuroscience*, 12: 535.

Kim, N., Wittenberg, E., and Nam, C. S. (2017). Behavioral and Neural Correlates of Executive Function: Interplay between Inhibition and Updating Processes. *Frontiers in Neuroscience*, 11:378.

Peer-reviewed Conference Proceedings

Sanders, N., Choo, S., **Kim, N.**, and Nam, C. S. (2019). Neural Correlates of Trust During an Automated System Monitoring Task: Preliminary Results of An EEG Effective Connectivity Study. In *Proceedings of the Human Factors and Ergonomics Society's 63rd Annual Meeting*. Sage CA: Los Angeles, CA: SAGE Publications.

Choo, S., Sanders, N., **Kim, N.**, Kim, W., and Nam, C. S. (2019). Detecting Human Trust Calibration in Automation: A Deep Learning Approach. In *Proceedings of the Human Factors and Ergonomics Society's 63rd Annual Meeting*. Sage CA: Los Angeles, CA: SAGE Publications.

Kim, N., Kim, W., Hwan Yun, M., & Nam, C. S. (2018). Behavioral and Neural Correlates of Hysteresis Effects during Multitasking. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Sage CA: Los Angeles, CA: SAGE Publications.

Kim, N., McCune, E., Yun, M. W., and Nam, C. S. (2018). EEG-Based Neural Correlates of ACT-R Model for Multitasking. *The 2nd International Neuroergonomics Conference*.

Kim, N., Wittenberg, E., & Nam, C. S. (2017). Working Memory Capacity, Memory Load, and Cognitive Control Network. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61, No. 1, pp. 46-48). Sage CA: Los Angeles, CA: SAGE Publications.

Liu, S., Wadeson, A., **Kim, N.**, and Nam, C. S. (2016). Effects of Working Memory Capacity, Task Switching, and Task Difficulty on Multitasking Performance. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 60, No. 1, pp. 502-506). Sage CA: Los Angeles, CA: SAGE Publications.

Book Chapter

Kim, N., and Nam, C. S. (2020). Adaptive Control of Thought-Rational (ACT-R): Applying a Cognitive Architecture to Neuroergonomics. In Nam, C. S. (Ed.). *Neuroergonomics: Principles and Practices*. Springer.

LIST OF ACRONYMS

ACC	Anterior Cingulate Cortex
ACF	Auto-Correlation Function
ACT-R	Adaptive Control of Thought-Rational
AF_MATB	Air Force_Multi Attribute Task Battery
ANOVA	Analysis of Variance
BA	Brodmann Area
BOLD	Blood Oxygen Level Dependent
CLARION	Connectionist Learning with Adaptive Rule Induction On-line
dDTF	direct Directed Transfer Function
DLPFC	Dorsolateral Prefrontal Cortex
DTF	Directed Transfer Function
DWI	Diffusion Weighted Imaging
EEG	Electroencephalogram
EF	Executive Function
EOG	Electrooculographic
ERP	Event-Related Potential
FA	False Alarm
fMRI	Functional Magnetic Resonance Imaging
FR	Frustration
GCA	Granger Causality Analysis
HSD	Honestly Significant Difference
ICA	Independent Component Analysis
LMP	Li-McLeod Portmanteau
MD	Mental demand
MEG	Magnetoencephalography
MNI	Montreal Neurosciences Institute
MVAR	Multivariate Autoregressive
NASA-TLX	NASA Task Load Index
PRP	Psychological Refractory Period
QN-MHP	Queueing Network-Model Human Processor

RACE/A	Retrieval by Accumulating Evidence in an Architecture
RMSE	Root-Mean-Square Error
ROI	Regions of Interest
RSME	Rating Scale Mental Effort
RT	Reaction Time
SIFT	Source Information Flow Toolbox
SOAR	State, Operator, And Result
SWAT	Subjective Workload Assessment Technique
VAR	Vector Auto Regression
WMC	Working Memory Capacity

1. Introduction

1.1. What is the workload transition effect or hysteresis effect?

Workload transitions are common in many situations where operators have been working at a level(s) of task demand for a while and are then confronted with a substantially different level(s) of task demand while performing multiple tasks at the same time. One example is piloting an aircraft (See Figure 1.1). While flying, a pilot needs to look out for obstructions, weather, and runway conditions while monitoring various metrics, such as airspeed, altitude, fuel levels, and navigation information. At the same time, pilots must pay attention to auditory cues from air traffic controllers as well as potential warning alarms such as ground proximity warnings. Pilots are also often tasked with manually controlling acceleration and guiding the plane using steering mechanisms. Although modern airplanes are equipped with sophisticated automation systems, pilots are still required to perform multiple tasks simultaneously.

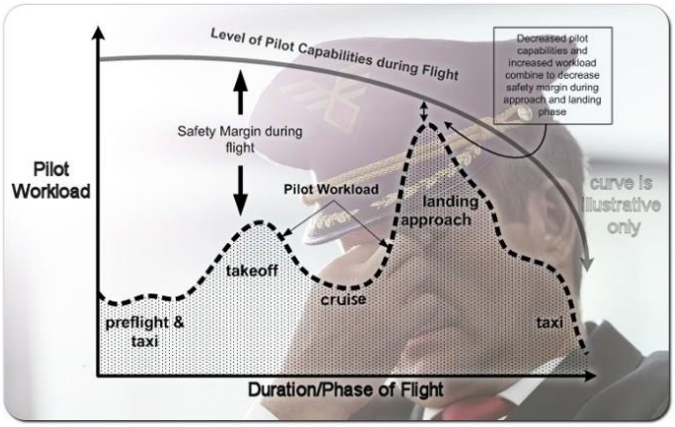


Figure 1.1. Changes in pilot workload during the flight

(adapted from Human Factors for Aviation: Basic Handbook, 2008)

In the past few years, much efforts have been made to explain how behavioral and cognitive performance are impacted in such environments of varying task demands (e.g., Lebiere et al., 2001;

Gray et al., 2005; Moon & Anderson, 2013), and develop computational models which help simulate human performance and cognitive activity (Maanen & Rijn, 2012).

Traditionally different levels of cognitive load were compared as if they are temporally independent. Specifically, it is assumed that counter-balancing or randomizing task demand levels remove any effects that one level might have on another. However, cases where workload levels are not temporally related rarely occur during real multitasking. In the aforementioned aircraft pilot example, demand transitions may occur at multiple points during a flight depending on travel conditions, even the weather regularly change, creating a temporally dependent cognitive load environment. It follows that measurement of the effects of the temporally dependent and dynamic workload is needed to improve performance prediction. Workload transition referred to as cognitive load has transitioned from one load at a specific time to a new load at a new time.

Increasing attention has been given to researching the impact of shifts in task demand on human performance and cognitive activity. An impact of previous demand conditions on current demand conditions, or hysteresis effects (Morgan & Hancock, 2011) is often the focus of such research. As such, an operator's effectiveness and accuracy may be affected by these changes in workload (Bowers et al., 2016).

Most experimental studies focusing on performance have typically maintained a constant workload factor. This method controls potential confounding variables or explores effects at a stable workload level, which allows us to observe individual responses throughout fixed workload levels (Matthews, 1986). However, the constant workload factors do not help understand how individuals respond to changing workload situations, in particular in the context of real-world situations. In fact, this focus on the constant workload has led to many advances in understanding the effects of the fixed workload, but at the same time has inadvertently limited the study of broader

workload dynamics, such as workload transition. As more emphasis is being placed on understanding the interaction between human and system in the context of functions that are shared, it should be recognized that operator workload might vary dramatically over time as systems attain a more dominant role in controlling the information rate to which she/he must respond (Matthews, 1986).

A workload transition is a shift in the difficulty of a cognitive task. This is important because a performer's effectiveness and accuracy may be affected by these transitions or workload profile (Bowers et al., 2016). For example, Morgan & Hancock (2011) used driving as an example and pointed out that the history of previously experienced events may be as influential on driver response and levels of workload as are current levels of demand. The term most applicable to the ongoing influence of such prior historical influences is "*hysteresi*s" (Morgan & Hancock, 2011). Any effects of prior task workload on post-transition performance are called hysteresis effects. For most professions, a transient change in effectiveness is not a concern; however, in some professions, there is little room for error. For instance, an investigation of air traffic control (ATC) operational errors showed that a high proportion of near misses occurred after a period of sustained high workload, suggesting that the hysteresis effect may have been a strong contributor (Farrell, 1999).

Hysteresis, a term borrowed from the physical sciences, has been applied to the study of humans for many years (Morgan, 2008). Both the behavioral and physical sciences utilize definitions of hysteresis, which roughly associate to the subject's history affecting the present experience (Morgan, 2008). In physics, hysteresis is defined as current properties being affected by forces no longer active on the object. For example, magnetically reactive materials (such as recording tapes) which continue to display a response to a magnetic field, even when the field is

removed, is said to demonstrate hysteresis. Despite the crucial impact of workload transition on behavioral performance, and cognitive states, this effect remains still unclear.

1.2. Research challenges

Several studies have described how workload transition effects or hysteresis effects develop over time by comparing aggregated performance data of multiple periods (e.g., Matthews, 1986; Gluckman et al., 1993; Ungar, 2005; Cox-Fuenzalida, 2007). Various cognitive models and approaches have been utilized to measure user performance in multitasking. Nevertheless, the previous studies on workload transition effects during multitasking still have three significant gaps.

1.2.1. Mixed findings on workload transition effects

The previous studies demonstrated that the effects of workload transitions might significantly affect performance, mental workload, and brain activity. These studies mainly examined the impacts of workload transitions on performance. Only a few studies had shown neurophysiological changes when the workload transitions occurred. Despite these attempts, the overall results of workload transition studies have been mixed and inconsistent depending on transition profiles and tasks, resulting in three different hypotheses: Enhancement, Deterioration, or No change. (See section 2.1 for more details).

1.2.2. Needs for neurocognitive model-based modeling and simulation

Various cognitive models and approaches have been utilized to measure user performance in multitasking. However, only a limited number of studies have quantitatively analyzed workload dynamics (Kim et al., 2018). Even fewer studies have shown how to predict and understand workload transition effects with multiple aspects from behavioral performance to underlying

neural mechanisms. The lack of studies on neurocognitive modeling for workload transition effects has caused inconsistent results. Therefore, we need a quantitative method including neurophysiological matrices to understand workload transition effects.

1.3. Objectives

This dissertation aims at improving the ACT-R model that can quantitatively analyze workload changes and predict operators' behavioral performance under multitasking environments. Adaptive Control of Thought-Rational (ACT-R) is a high-level computational simulation of human cognitive processing. ACT-R is one of the cognition theories that seek to predict human performance in cognitive tasks such as multitasking. ACT-R may be a useful tool for unifying multiple disciplines by incorporating behavioral, cognitive, and neurophysiological data into one computational model. In this study, we chose ACT-R as a theoretical framework and expected it to resolve the discrepancies seen in previous research and subsequently help us better understand the underlying neural mechanisms behind workload transition effects.

To address the limitations of previous research, this study addressed the following three research questions:

1) "What is the nature of workload transition effects?" by examining the behavioral, perceived workload, and neurophysiological measures in both a basic cognitive task (in Phase I of Study 1) and a more ecologically valid task (AF-MATB) (in Phase I of Study 2).
2) "What is the nature of the effects of workload transition variation types at the behavioral, perceived workload, and neurophysiological levels?", by testing which of

the three hypotheses on workload transitions effects (i.e., the "Enhancement," "Deterioration," or "No change") can be supported.

3) "What is a systematic way that can improve the fitness of the ACT-R model that simulated human performance during workload transitions?" by validating a proposed ACT-R parameter tuning framework using both behavioral measures and effectivity connectivity of human operators.

Figure 1.2 illustrates the framework of the present research that is composed of two studies, each having two phases. In Study 1, we used a simple cognitive dual-task to investigate hysteresis effects. In Study 2, we utilized the AF-MATB, which is a more ecologically valid task, to mimic the real-world multitasking settings.

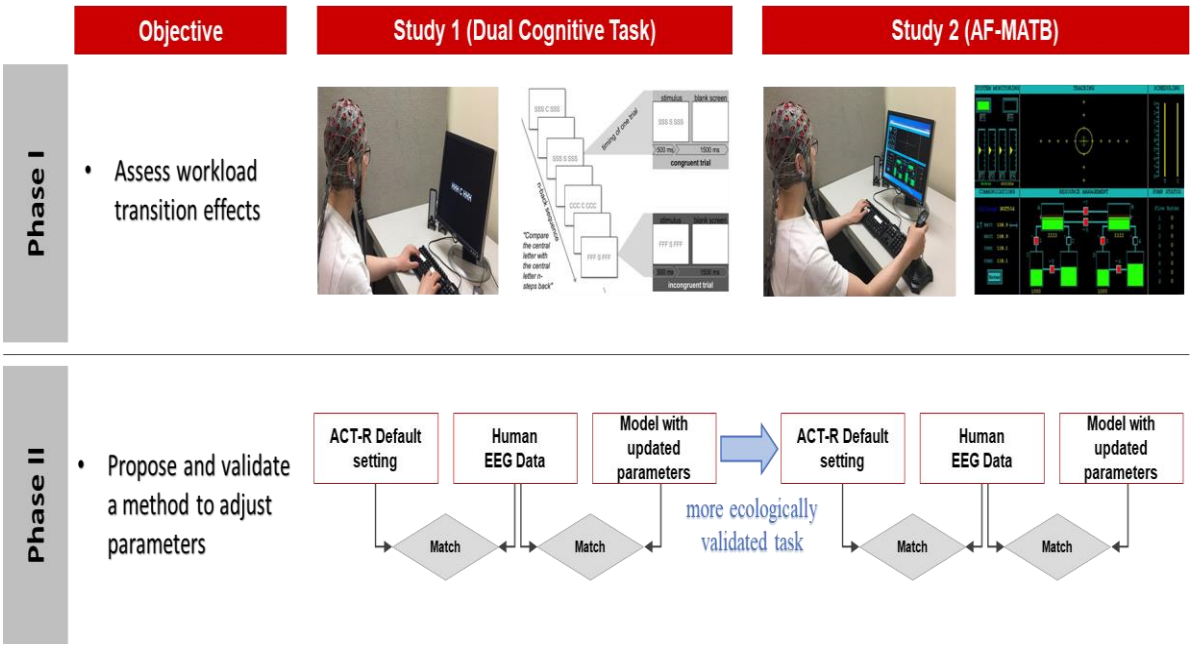


Figure 1.2. Schematic diagram of the research framework

1.4. Organization of this dissertation

The remainder of the dissertation is organized as follows:

Chapter 2 presented a comprehensive literature review related to three relevant topics of this research, including workload transition in multitasking, ACT-R, application of brain connectivity for ACT-R modeling.

Chapter 3 introduced Study 1. This chapter first explains the details of the experimental paradigm with a dual cognitive task (combined n-back and Flanker). In phase II, we proposed the method on how to modify the ACT-R parameters based on the acquired neuroimaging data, while reflecting the workload transition effect.

Chapter 4 described the research strategies of Study 2 in detail. The AF-MATB was employed as a multitasking platform, and the proposed ACT-R parameter tuning method was implied and validated on the application of AF-MATB tasks.

Chapter 5 discussed the general findings from Study 1 and 2.

Finally, Chapter 6 concluded this dissertation with a review of the research founding along with research implications.

2. Literature review

This chapter reviews a wide range of studies that investigated workload transition effects, ACT-R modeling of human performance, and brain effective connectivity. Section 2.1 covers the effects of workload transition on operators' performance, mental workload, and neurophysiological measures. Section 2.2 presents ACT-R studies on multitasking and previous ACT-R modeling approaches. Finally, section 2.3 discusses what brain effective connectivity (EC) is and how we can incorporate EC measures into ACT-R architecture.

2.1. Workload transition in multitasking

This section reviews studies that mainly examined the effects of workload transition in multitasking on performance, mental workload, and neurophysiological responses. In addition, previous studies that investigated individual differences in workload transition are reviewed.

Workload transition, or a workload shift, has significant implications for many work environments. These implications are particularly salient in occupations, where individuals are confronted with varying levels of workload demands, especially safety-sensitive occupations. The finding that demand transitions make human performance more unpredictable may have implications for human performance modelers (Wickens et al., 1998).

2.1.1. Changes in workload transition affect task performance

The directions of workload transitions (i.e., increasing or decreasing load) produced mixed evidence regarding their effects on task performance. In previous studies, a reduced performance (Krulewitz et al., 1975; Gluckman et al., 1993; Hancock et al., 1995; Moroney et al., 1995; CoxFuenzalida et al., 2006), no change (Helton et al., 2004; Cox-Fuenzalida, 2007) or an

improvement (Matthews, 1986; Ungar et al., 2005; Kim et al., 2018) were observed (see Figure 2.1); hereinafter they will be referred to as the "Deterioration" hypothesis, "No change" hypothesis, and "Enhancement" hypothesis, respectively.

First, previous studies have shown that a decrease in task demand often caused subsequent performance deterioration ((C) in Figure 2.1).

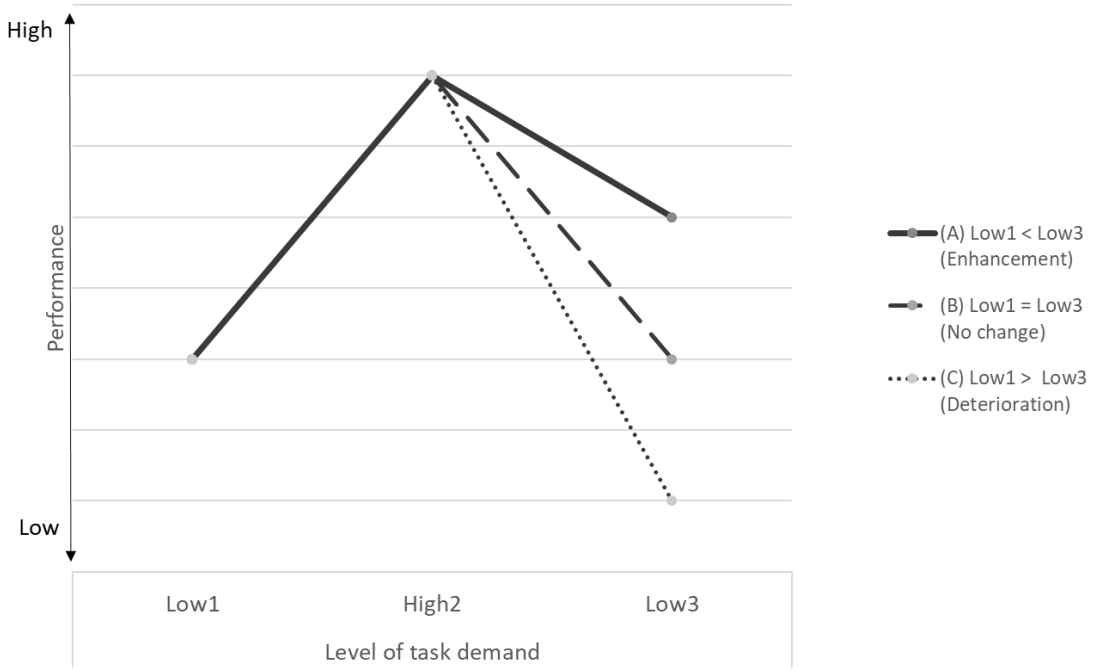


Figure 2.1. Inconsistent hysteresis effects on performance

Krulewitz et al. (1975) used a vigilance task in which participants monitored a screen where a red bar would appear. This study used twenty minutes of induction and forty minutes of post-transition performance, broken into two, twenty-minute periods. The difficulty of the task was manipulated by increasing the number of background signals presented within the timeframe. There were four conditions. The first condition was one that transitioned from a high number of background signals to a low number of background signals (high-low). The second condition

transitioned from a low number to a high number of background signals (low-high). The last two conditions had continuously high levels (high-high) or continuously low levels (low-low) of background signals. During the first twenty minutes of post-transition, significant performance differences were not found between the high-low and low-low conditions or between the low-high and high-high workload transition conditions. However, participants in the low-high condition performed significantly worse than those in the high-high condition in the second period of the post-transition session.

Gluckman et al. (1993) pointed out that demand transitions were induced by shifting from two parallel visual signal detection tasks to one signal detection task or vice versa. Pre-transition and post-transition performance were measured in two periods of ten minutes and compared against non-shifting control groups. A hysteresis effect in the form of lower performance was found only with the shift from dual-task to a single task.

Hancock et al. (1995) performed a study that focused on how the participants' perception of the current workload might be influenced by previous levels of workload. The experiment used a manual, compensatory tracking task, and each condition was made up of three, five-minute segments. The levels of difficulty for the tracking task in the first condition were medium-low-medium. In the second condition, the levels were medium-high-medium, and in the third or control condition, they were medium-medium-medium. Participants completed all three conditions, each on a separate day. The only significant result in the performance data was a decrease in the combined time lead in the last segment of the Medium-High-Medium segment compared to the last segment of the control condition. This indicated that the participants' mean response time became slower in the transition condition.

Moroney et al. (1995) conducted a study examining the effect of abrupt and gradual changes in demand on performance during low-to-high and high-to-low demand transitions. In this study, the participants performed a vigilance task in which they were required to observe two identically oriented lines. A critical event occurred when the two lines were not oriented in the same direction. Each trial contained a transition from high to low workload or from low to high workload, and those transitions could occur abruptly or gradually. The results showed that in the low-high demand group, the participants in both the gradual and the abrupt transition conditions performed significantly worse than their non-shifted controls after the transition period. The participants in the high-low demand groups, in both the abrupt and gradual transition conditions, all began the vigil with detection scores below those of the non-shifted controls, and their scores remained low for the rest of the vigil.

CoxFuenzalida et al. (2006) showed that transitions negatively affect performance, regardless of the direction of the transition. They sought to further examine those results by investigating the possibility that one type of transition (high-to-low or low-to-high) might be more detrimental than the other. To accomplish this, a visual Sternberg memory task (Sternberg, 1969) was employed. During the Sternberg memory task, participants are given a set of letters to memorize. In this case, participants had five seconds to memorize a set of six letters. Participants were assigned to either a high-medium condition or a low-medium condition. Trials were seven minutes long, and the transition in difficulty occurred after two minutes. Also, participants completed a five-minute baseline trial at each of the workload levels they experienced in the transition trial. The results of the Sternberg memory task indicated that both sudden increases and sudden decreases in workload were detrimental to performance as compared to baselines. When the mean performance from the high-medium trials (first minute after the transition) was compared

to the means of the corresponding baseline trials, the difference between the two was larger than when the low-medium trials were compared to the corresponding baseline trials. This difference led the authors to conclude that sudden decreases in demand may be more damaging than sudden increases in demand.

Next, two studies did not find hysteresis effects during workload transition ((B) in Figure 2.1). Helton et al. (2004) conducted an experiment investigating the effects of salience on performance, using three salience conditions: low-high, consistently low, and consistent high. The analysis of the results revealed no significant differences when the performance of the participants in the high-low was compared to the performance of those in the low salience control group. Similarly, no significant differences were found in the performance of low-high as compared to the high salience control group. Cox-Fuenzalida (2007) manipulated the difficulty of an auditory signal detection task and did not find a hysteretic effect in the case of a shift from low to high.

Lastly, even though observations of negative effects following workload transitions are well supported, there is a considerable body of evidence showing the positive effects of workload transitions ((A) in Figure 2.1). For example, Matthews (1986) collected performance data from a visual signal detection task in fifteen consecutive periods of ten seconds each. Task demand was manipulated by varying the number of co-occurring stimuli. Matthews (1986) designed the study to include fifteen trials in each block, and the blocks all took 2.5 minutes to complete. For all trials within a block, the workload level, determined by the number of strings in the group, remained the same. Therefore, the workload was manipulated between blocks, and each condition was made up of sixteen blocks. In one condition, the workload level increased to its peak (three strings in the first block, six strings in the second, nine strings in the third, and twelve strings in the fourth) and then systematically decreased back down to three strings. In the second condition, the level of

demand for each block was randomized. For the third condition, there was a schedule of high demand for three blocks (twelve strings each), and then there would be one block of low demand (three strings each). That pattern was repeated four times. Finally, the fourth condition was the reverse of the third condition. There were three blocks of low demand (three strings) and then one block of high demand (twelve strings), with that pattern repeated four times. Additionally, there were four control conditions in which each of the participants completed 16 blocks, each at one of the four levels of demand (3, 6, 9 or 12 strings). Matthews (1986) expected that participants who were transitioned from a high level of workload to a lower level would maintain their high level of responding despite the change; thus, transitioned participants would perform better after the transition than participants in low-workload control conditions. Except for sudden demand changes, the other results of a cyclic transition group showed that any workload transition, whether it increased or decreased the level of demand, improved performance.

Ungar et al. (2005) used a manual, compensatory tracking task that was paired with a vigilance task under dual-task performance conditions. Performance for participants who were transitioned from a dual-task condition involving concurrent difficult tracking and vigilance performance to the single-task condition requiring difficult tracking was significantly worse than that of participants who only performed the difficult tracking task in a control condition. However, the performance for the participants who transitioned from a dual-task condition requiring concurrent performance of easy tracking and vigilance tasks to a single task involving easy tracking was significantly better than that of the easy tracking controls.

Kim et al. (2018) found post-transition improvements in performance on the Airforce Multi-Attribute Task Battery (AF-AMTB) tasks, in a study which was conducted with

distinguished task demands in three stages (Low-High-Low). They found significant hysteresis effects on performance; specifically, post-transition reaction time was shorter than pre-transition.

Overall, the effects of workload transitions on performance are inconsistent, but the majority of evidence suggests they are detrimental to task performance. These studies showed that the partitioning of performance data into a sequence of post-transition periods is essential to investigate how hysteresis develops during the time frame of an experimental condition and thus potentially in real-world situations. A next question, then, is whether hysteresis also affects mental workload.

2.1.2. Variability of task demand and mental workload

Although mental workload has been the subject of increased interest and research activity in recent years, one variable that may be of some significance, workload history, has been consistently overlooked (Morgan & Hancock, 2011). Traditionally, temporal effects that are most often reported in the vigilance literature are described in terms of workload levels that are stable over time, and interest has focused largely on the performance changes associated with time on task. Several studies on demand transitions assessed mental workload in addition to performance.

The effects of workload transitions on mental work have shown inconsistent results with regards to performance data. In previous studies, both a decrease (Hancock et al., 1995; Moroney et al., 1995; Kim et al., 2018) ((C) Enhancement in Figure 2.2) and an increase in subjective mental workload (Matthews & Desmond, 2002; Morgan & Hancock, 2011) ((A) Deterioration in Figure 2.2) were observed.

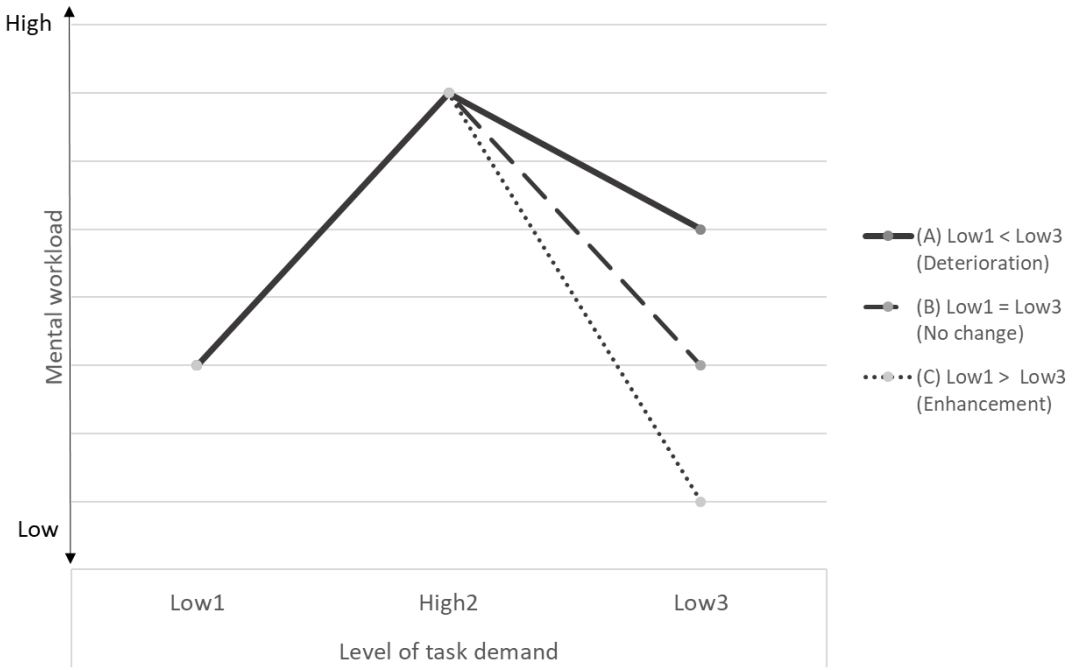


Figure 2.2. Inconsistent hysteresis effects on mental workload

Hancock et al. (1995) subjected participants to three trials on a compensatory tracking task. The first and third trials were performed at an identical difficulty level. Hancock et al. (1995) analyzed the results of the Subject Workload Assessment Technique (SWAT; Reid & Nygren, 1988) and found that participants rated the last segment in the Medium-Low-Medium condition as having a significantly higher workload than the last segment in the Medium-Medium-Medium control condition. The same pattern of results was also found with the NASA TLX. With SWAT, the participants rated the last segment in the Medium-High-Medium condition as having a significantly lower workload than the last segment in the Medium-Medium-Medium control condition. The NASA TLX results for the same comparison were not significant. These results indicated that participants' perceptions of workload were somewhat affected by previous levels of demand.

As in the previously described studies in Section 2.1.1, in Moroney et al. (1995) study, the NASA TLX was administered to each participant at the conclusion of the vigilance task. Even though participants in the Low-High and High-Low conditions all experienced the same levels of demand during the vigil, there were some significant differences in the participants' ratings on the mental demand subscale of the TLX. In the abrupt shift condition, those shifted from High-to-Low reported significantly higher mental demand than those who were transitioned from Low-to-High. Interestingly, the participants in the gradual transition condition reported the opposite effect. Those in the Low-to-High group reported significantly higher mental demand than those in the High-to-Low demand group. Thus, the results of the NASA TLX provide evidence that participants found the High-Low and Low-High conditions to be quite different in terms of the mental demand placed upon them and depending on the other conditions experienced at the time (but only for the mental demand scale).

Kim et al. (2018) used the unidimensional Instantaneous Self-Assessment (ISA) scale (Tattersall & Foord, 1996) to assess mental workload during an AF-MATB experiment with three stages (Low-High-Low). ISA questionnaires (i.e., "Report how much mental workload the task just required") were collected verbally two times during each condition. The results showed a lower score in the post-transition period from high to low, but there was no significant difference between the first low condition and the third low condition.

On the other hand, two studies have shown increasing mental workload after shifts of task-load. Matthews & Desmond (2002) induced a demand transition by shifting from a dual-task driving trial (i.e., with a signal detection task) to a single-task driving trial. Shifted drivers reported higher mental workload on the NASA-TLX compared to non-shifted drivers.

Morgan & Hancock (2011) conducted a study using a driving simulation task to explore the effects of a sudden increase in workload on performance and subjective workload ratings. The workload was assessed by using a simplified version of the SWAT technique (S-SWAT; Luximon & Goonetilleke, 2001). Participants in this study completed four, 5-minute driving trials. During each trial, the heads-up navigation display would malfunction at approximately 3 minutes and 20 seconds into the drive. To correct the malfunction, the drivers would read a 10-digit alphanumeric code aloud to the experimenter. This was the only significant change in workload during the drive. Morgan & Hancock (2011) examined performance and subjective workload at three points during the drive: once at one minute and twenty seconds into the drive (low workload), again after the navigation error had been corrected (high workload), and finally, at the end of the trial (low workload). Since the S-SWAT assesses three factors (i.e. time, mental effort, and psychological stress) participants verbally reported their ratings (scale of 0-100) to the experimenter at those points during the drive.

Morgan & Hancock (2011) calculated the overall mental workload score for the S-SWAT by taking an unweighted average of the ratings on the subscales of time, mental effort, and psychological stress. These S-SWAT scores indicated that workload within trials significantly increased from the first, pre-transition measurement period to the second, high workload period. There was no significant difference between the second and the third measurement periods, although there was a slight decrease in the reported workload. For the individual S-SWAT subscales, there was a significant increase in perceived time demand between all of the measurement periods. Similarly, the mental effort increased significantly from the first period to the second and then dropped significantly between the second and third. However, the third-period score was still significantly higher than that of the first period. Therefore, the significant

differences between the first period and the third period in the S-SWAT overall workload scores and in each of the subscales scores all demonstrate hysteresis effects.

The aforementioned studies demonstrate that hysteresis is also manifested in mental workload. However, none of them partitioned the data in a sequence of post-transition periods because subjective workload ratings were collected only once after each experimental condition. Consequently, the development of mental workload during experimental conditions could not be investigated. Neuroimaging techniques can address this issue because they allow us to continuously measure the operator's cognitive states during variable task load. Therefore, the next section covers workload transition studies using neuroimaging techniques.

2.1.3. Workload transition effects on neurophysiological responses

Studies have been conducted to investigate the relationship between the hysteresis effect and the brain activity caused by the change of cognitive states. This section reviews studies that used three different neuroimaging techniques; functional magnetic resonance imaging (fMRI), functional near-infrared spectroscopy (fNIRS) and electroencephalography (EEG).

To the best of my knowledge, there is no research on workload transition effects by using fMRI; only a few studies have shown different brain activities related to workload. Schweizer et al. (2013) used fMRI to identify differences in brain activation by the level of workload in distracted driving. Deprez et al. (2013) found that there is a difference in brain activation during single, dual, and multiple tasks in visual and auditory domains.

One of the fNIRS studies investigated the effects of workload transition (McKendrick, 2016). A spatial memory task with an event-related design was used as a proxy for showcasing a design matrix regression on fNIRS data for workload transition effects. The analysis identified a

region of the ventrolateral prefrontal cortex (VLPFC) in which oxygenated hemoglobin (HbO2) was increased during post transition period from steady-state low load to transient high load. Task performance negatively modulated HbO2 in this brain region during that period. The results show neural effects specifically within the period of post-transition.

Two studies on workload transition (within the AF-MATB environment) have assessed the hysteresis effect by using EEG data. In the first of these, Bowers et al. (2014) reported there was a hysteretic effect in gamma activity of the EEG signal following a high- to low-demand transition as observed in a delay of gamma power's changes. They found that temporal gamma oscillations changed rapidly following a transition and settled after task difficulty changed from easy to hard, but settled more slowly when the difficulty changed from hard to easy. Frontal theta oscillations, in contrast, exhibited consistently rapid settling, which may indicate rapid changes in working memory utilization and conflict resolution (Gevins et al., 1997). Secondly, Kim et al. (2018) found significant hysteresis effects on brain network measures such as prefrontal cortex outflow and connectivity magnitude, which allowed us to clarify the direction and strength of Granger causality flow under workload transitions. They showed that after completing the high-demand condition, the prefrontal area remained to activate, so the direction and magnitude of connectivity within the neural network might differ from the pre-transition period.

Observation of this neuroimaging data would allow for objective classification of different cognitive load states. These states could then be identified in individual performers and used to study the effects of workload transitions to different cognitive load states. The following section reviews previous studies on individual differences accounting for workload transition.

2.1.4. Individual differences and workload transition

Except in one case (Bowers et al., 2014), workload transition studies have made no attempt to adapt task demand manipulations to individuals. Without adaptation, individual differences in cognitive capacity can result in different levels of cognitive load in different individuals even when task demands are identical. When workload transitions to different levels of an 'optimal' cognitive load state have minimal confounding effects. However, if workload transitions cause some individuals to transition out of an 'optimal' cognitive load state and into an overload or underload state a lack of individual adaptation is problematic for assessing workload transition effects. The examination of overload and underload is important because these workload states have the greatest effect on task performance. While different levels of an 'optimal' cognitive load state should elicit positive task performance, both overload and underload are likely to elicit inferior task performance. The overload state should result in errors related to an inability to process more, or new information, leading to an inability to cope with the cognitive demands of current or novel tasks (Parasuraman et al., 2008). The underload state which is believed to be as detrimental (Hancock & Parasuraman, 1992) and more difficult to measure than overload (Hancock & Verwey, 1997) appears to induce its own unique task decrements. In order to address issues of workload transitions into overload and underload it is necessary to adapt task demands on an individual basis. Furthermore, given that cognitive load and task performance are not equivalent (Parasuraman et al., 2008) other objective measures of cognitive load are required.

Baddeley (2012) defines working memory as "the system or systems that are assumed to be necessary in order to keep things in mind while performing complex tasks such as reasoning, comprehension, and learning" (p. 136). As a construct for studying workload transitions, working memory offers a number of advantages over complex tasks and signal detection tasks. Working

memory refers to a limited capacity store (be it a unique buffer or part of long term memory) (Logie, 2011; Baddeley, 2012) that works in conjunction with a cohort of executive functions (Unsworth & Engle, 2007). Working memory capacity (WMC) is predictive of performance in several complex cognitive tasks. Specifically, individuals with high WMC exhibit superior visual attention (Engle, 2002), inhibition of irrelevant representations (Unsworth & Engle, 2007) and improved time-critical decision making (Endsley, 1995). Individuals with a greater working memory capacity also store more task-relevant information and recall that information more quickly. High working memory individuals also update information more efficiently, shift and maintain task goals with less error, and better cope with distractions (Unsworth & Engel, 2007). The theoretical underpinnings of working memory, specifically its predictive power for basic and complex tasks and its property of instantly affecting cognitive load make it a prime concept for studying workload transitions.

2.1.5. Limitations of current studies on workload transition in multitasking

From a systematic literature review of the workload transition in current studies in Section 2.1, two important issues were identified as follows:

(1) The overall results have been mixed, and workload transition effects are still poorly understood. Although the previous studies demonstrated that the effects of workload transitions may significantly impact performance, mental workload, and brain activity, those outcomes resulted in three different hypotheses: Enhancement, Deterioration, or No change.

(2) Lack of a theoretical framework for workload transition effects has caused inconsistency in results across previous studies (McKendrick, 2016). Therefore, it is necessary to

develop a computational model to quantitatively analyze workload transition effects and multitasking performance.

2.2. Adaptive Control of Thought-Rational (ACT-R): A cognitive architecture

In order to overcome the limitations of the studies on the workload transition effects reviewed in the previous section, we need to consider various empirical data. Logical interpretation of the results of workload transition requires incorporating behavioral, cognitive, and neural outcomes. Specifically, brain activity measures would enable us to explain underlying neural mechanism, which has not been uncovered by previous studies. ACT-R is an effective model that can incorporate the various types of data (behavioral, cognitive, and neurophysiological).

This section reviews one of the neurocognitive architectures, ACT-R that can unify different scientific disciplines: cognitive psychology, cognitive modeling, and traditional neurosciences.

Figure 2.3 depicts a neurocognitive modeling paradigm that helps to explain how ACT-R may be a useful tool for unifying multiple disciplines by incorporating behavioral data, brain data, and cognitive data into one model. ACT-R, a model that systematically incorporates data from these three approaches, helps us to resolve the discrepancies seen in previous research and allows us to better understand the underlying physical and cognitive processes that occur within workload transitions. This section reviews the related research on ACT-R, including basic knowledge of modeling as well as research on multitasking and discusses how ACT-R can be used as a neurocognitive architecture that can unify the three scientific disciplines.

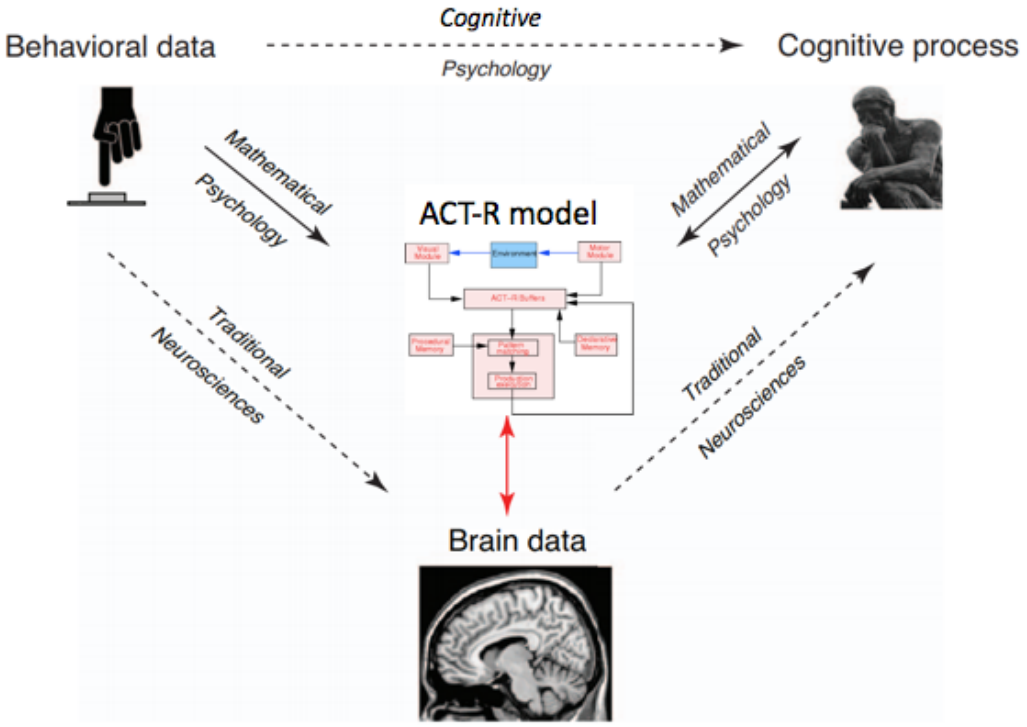


Figure 2.3. Neurocognitive modeling paradigm (modified from Forstmann et al., 2011)

2.2.1. Overview of ACT-R

In 1983, Anderson published his seminal work *The Architecture of Cognition* (Anderson, 1983). Cognitive architecture refers to both a unified theory of cognition - the outline of the structure of the various parts of the mind - and computational implementation of the theory with specified rules and associative networks. The symbiotic relationship between cognitive modeling and cognitive neuroscience results in palpable progress towards the shared goal of better understanding the functional architecture of human cognition. Specifically, cognitive architectures are frameworks that can be used to develop computational models of human cognitive processes (Langley, Laird, & Rogers, 2009; Taatgen & Anderson, 2010; Smart et al., 2016). Cognitive architectures help us understand human cognition in specific task environments, and they are used to develop various intelligent systems and agents (e.g., cognitive robots). Although multiple cognitive architectures

exist, such as State, Operator, And Result (SOAR) (Laird, 2012), ACT-R (Anderson et al., 2004; Anderson, 2007), and Connectionist Learning with Adaptive Rule Induction On-line (CLARION) (Sun, 2006; Sun, 2007), the current dissertation is focused on ACT-R.

Adaptive Control of Thought-Rational (ACT-R) is a rule-based system that has been widely used by cognitive scientists to model human cognitive performance. It is also one of the few cognitive architectures that have an explicit link to research in the neurocognitive domain: the structural elements of the core ACT-R architecture (i.e., its modules and buffers) map onto different regions of the human brain (Anderson, 2007). This enables cognitive modelers to predict the activity of different brain regions at specific junctures in a cognitive task (Anderson, 2007).

ACT-R is a formalized, integrated cognitive architecture that combines the *Spreading Activation Memory* theory with a production system to model the high level of cognitive tasks. Like many successful architectures, ACT-R is a modular theory of mind that aims to provide an integrated account of human cognition. Figure 2.4 shows an overview of ACT-R modules (Anderson et al., 2004).

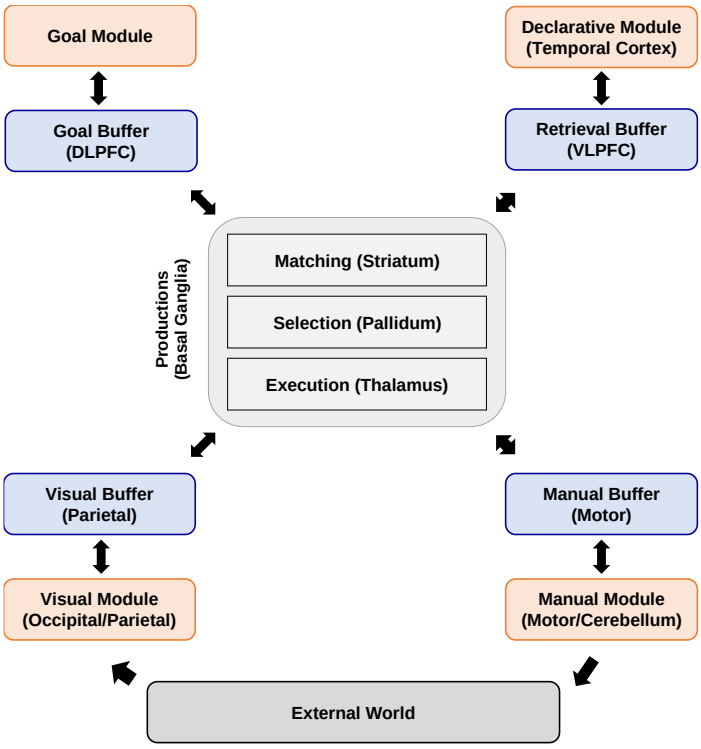


Figure 2.4. Overview of ACT-R modules (adapted from Kim et al. 2018)

It treats the mind as being composed of distinct modules that exist for particular functions. The ACT-R modules represent the functions of the brain as well as how these functions are mapped to different parts of the brain.

ACT-R incorporates both declarative knowledge (e.g., addition facts) and procedural knowledge (e.g., rules for solving multi-column addition) into a production system where procedural rules act on declarative chunks (Anderson et al., 2004). In ACT-R, the structures of declarative knowledge are called “chunks” and held in the Declarative module, whereas those of procedural knowledge is called “rules” and held in the Procedural module. The rules also have access to other modules, including the Visual module for perception, the Manual module for action, the Imaginal module for storing visual problem representation, and the Goal module for keeping track of current intentions. These modules are linked to specific areas of the brain: Manual in the

motor cortex (BA 3/4), Imaginal in the parietal cortex (BA 39/40), Declarative in the dorsolateral prefrontal cortex (DLPFC) (BA 45/46), Goal in the anterior cingulate cortex (ACC) (BA 24/32), Visual in the fusiform gyrus (BA 37), and Procedural in the caudate of the basal ganglia (Matessa, 2008) (Refer to Figure 2.4). Each module of ACT-R has its buffer that can store only one chunk of information extracted from the corresponding module. In ACT-R, the condition statements of all production rules are compared with the current contents of buffers every 50 ms (Anderson, 2007).

When there are one or more matching productions, the matching one or the one with the highest expected value among them, respectively, is selected to fire, and the action statement of the selected rule then makes each module exchange chunks of information or creates new chunks in the particular modules (Taatgen & Lee, 2003). At this time, only one chunk can be processed through each buffer at a time, while all modules can be executed simultaneously. For example, the declarative module cannot retrieve the additional memory chunk until it has completed the previous retrieval. In contrast, the manual module can perform the physical movement, while the visual module looks at a new object. Moreover, there is a limit to the number of objects and a time limit in the visual and retrieval buffer. These are related to the working memory limitation of human cognitive processing (Anderson et al., 2004).

2.2.2. Cognitive modeling of multitasking using ACT-R

Salvucci et al. (2004) maintained that human multitasking arises in many real-world situations, from mundane everyday tasks to the most complex, demanding work environments. Computational cognitive models such as ACT-R are formal theories of cognition that provide predictions of human performance in cognitive tasks such as multitasking. These models generally

include simulated representations of human cognitive processes such as memory, attention, visual and motor processing, problem-solving, learning, and other related phenomena (Anderson et al., 2004). Important insights about the cognitive mechanisms behind multitasking and how brain networks are affected by changes in internal cognitive strategies, external interface properties, and task demands arise by comparing simulated human cognitive data from the ACT-R architecture to real human data (Kim et al., 2018).

Previous studies conducted within the framework of cognitive architectures have focused on multitasking in small-scale (e.g., Psychological Refractory Period (PRP)) tasks (Byrne & Anderson, 2001; Salvucci et al., 2004). These models tend to have 'customized executives' (Kieras et al., 2000) that are limited to a particular task. Other modeling efforts have focused on the general characteristics of domain-independent multitasking by integrating smaller task models (Kieras et al., 2000). For example, Salvucci (2005) has described a general executive module for the ACT-R architecture, and Taatgen (2005) has explored a general way in which this architecture can account for multiple concurrent tasks.

Salvucci & Taatgen (2008) extended ACT-R with threaded cognition to explain how a serial architecture can account for multitasking. Instead of an explicit control strategy, threaded cognition adds a simple interleaving task-scheduling method to ACT-R. Nijboer et al. (2016) have shown that an ACT-R model that assumed a distributed version of working memory accounts for both behavioral and fMRI data better than a model that takes a more centralized approach. Furthermore, their proposed modeling idea took into consideration working memory components within ACT-R such as attentional focus, declarative memory, and a subvocalized rehearsal mechanism. Thus, the data and model favor an account where working memory interference in

dual-tasking is the result of interactions between different resources that together form a working memory system.

2.2.3. ACT-R modeling approaches

ACT-R has a subsymbolic level in which continuously varying quantities are processed in parallel to produce much of the qualitative structure of human cognition. These subsymbolic quantities participate in neural-like activation processes that determine the speed and success of access to chunks in declarative memory as well as the conflict resolution among production rules. The core of ACT-R memory modeling can be summarized in the following equations (Anderson et al., 2004).

In this study, we focused on three parameters for reflection of hysteresis effects based on Nijboer et al. (2015) study. The first is a *retrieval threshold* (τ) (see equation (3)), which determines the minimal amount of activation required to retrieve a chunk. Next, the *activation noise* (ε) (see equation (1)) is the amount of random activation added or subtracted to individual chunk activations. Lastly, the *latency-scaling factor* (***F***) (see equation (4)) scales the retrieval times of chunks based on their activation.

The activation A(t) of a particular chunk is defined as where B(t) is the base activation of the knowledge chunk at time t, and the summation is the associative strength of the chunk dependent on related chunks, defined as “the elements in the current goal chunk and the elements currently being processed in the perceptual field." (Anderson, 1993, p 51.)

$$A(t) = B(t) + \sum_j w_j \, s_{ij} + \varepsilon \qquad (1)$$

where each $w_j \in [0,1]$ reflects the salience or validity of [related] element j.

The $s_{ij} \in [-\infty + \infty]$ are the strengths of association to i from elements j in the current context.

ε: The noise value, as described in the last unit.

The base activation B (t) is increased whenever it is used, either through practice (learning events) or when matched to a production rule.

$$B(t) = ln \sum_k t_k^{-d} + B \qquad (2)$$

The decay rate, d, is defined for the particular learning event k.

If we make a retrieval request and there is a matching chunk, that chunk will only be retrieved if it exceeds the retrieval activation threshold, τ. The probability of this happening depends on the expected activation, A_i, and the amount of noise in the system which is controlled by the parameter s:

$$recall\ probability_i = \frac{1}{1 + e^{\frac{\tau - A_i}{s}}} \qquad (3)$$

Inspection of that formula shows that as A_i tends higher, the probability of recall approaches 1, whereas as *t* tends higher, the probability decreases. When $\tau = A_i$, the probability of recall is 0.5. The *s* parameter controls the sensitivity of recall to changes in activation. If *s* is close to 0, the transition from near 0% recall to near 100% will be abrupt, whereas when *s* is larger, the transition will be a slow sigmoidal curve.

The activation of a chunk also determines how quickly it can be retrieved. When a retrieval request is made the time it takes until the chunk that is retrieved is available in the **retrieval** buffer is given by this equation:

$$Time = Fe^{-A} \qquad (4)$$

A: The activation of the chunk which is retrieved.

F: The latency factor parameter.

If no chunk matches the retrieval request, or no chunk has an activation which is greater than the retrieval threshold, then a retrieval failure will occur (Taatgen et al., 2006). These equations provide an integrated foundation for theory development in memory encoding, storage, retrieval, and by extension to Workload Transition Effect analysis, which is critical for the explanation of hysteresis effects.

To evaluate models to be developed in the present study, the basics of ACT-R memory modeling steps are reviewed. Those previously well-defined structures allow us to model our paradigm by using a cognition framework that has been extensively tested. As stated before, current multitasking models do not incorporate the workload transition effect. The next section presents previous modeling approaches to workload based on behavioral data and neural data.

2.2.3.1. Behavioral data-based modeling

Lebiere et al. (2001) proposed a workload prediction method based on ACT-R to predict the workload in air traffic control simulations. The workload of an ACT-R model was assumed as the scaled ratio between the times spent on critical unit tasks to the total time of the task. However, he considered only time as the resource related to mental workload. Gray et al. (2005) suggested cognitive-metrics profiling, also based on ACT-R, as a theory-based prediction of transient changes in the workload demanded by dynamic task environments, but they did not introduce the detailed procedure or mathematical representation to predict the mental workload over time.

Retrieval by Accumulating Evidence in an Architecture (RACE/A) is an extension of the ACT-R that models the process of retrieving memory (Van Maanen et al., 2012). In RACE/A, retrieval of a chunk from declarative memory is a process in which evidence is accumulated for

the likelihood that a chunk will be needed, like sequential sampling models of cognitive behavior (Ratcliff, 1978; Usher & McClelland, 2001). The dynamics of the competition between memory chunks governs the retrieval time from declarative memory. Because RACE/A dynamically updates the activation values of chunks, it can explain the interference effects observed in semantic memory retrieval better than existing ACT-R models. RACE/A differs from default ACT-R in a number of ways. However, these adaptations do not change the central assumptions (Cooper, 2007) that underlie ACT-R. The first obvious deviation from the default architecture is the inclusion of an accumulative process for memory retrievals. The new model increases the explanatory power of the architecture. However, in the absence of competition between chunks RACE/A makes the same predictions as ACT-R.

In order to assess what changes were required to the model to better fit the data, Nijboer et al. (2016) made several small changes that together resulted in the updated model: To accurately capture tone-counting performance, they introduced the possibility of making a mistake during counting. Anecdotal evidence suggested that participants were sometimes mistaking low tones for high tones. Therefore, model parameters were adjusted to optimize the quantitative fit: The retrieval threshold determines the minimal amount of activation required to retrieve a chunk. The activation noise is the amount of random activation added or subtracted to individual chunk activations. The latency-scaling factor scales the retrieval times of chunks based on their activation. Finally, the sound decay time determines how long sounds observed by the aural module are available for further processing before they are discarded.

While ACT-R has no dedicated WM system, it has two modules that can be used as part of a WM strategy: declarative (long-term) memory and the problem state. Although the capacity of declarative memory is essentially unlimited, the chance of being able to retrieve an item

decreases over time. This is implemented by giving each item an activation value, which is a numerical expression of its strength in memory. This activation value decays over time (Anderson, 2007), but increases when the associated fact is retrieved: Items that have been used more often or more recently will have a higher activation. The activation value is used during the retrieval phase, where it determines how long the retrieval will take (a higher activation value results in shorter retrievals), and whether retrieval is possible at all (a chunk can only be retrieved if its activation value exceeds a predefined retrieval threshold value).

2.2.3.2. Neurological data-based ACT-R modeling

Borst et al. (2016) maintained that cognitive models are hard to validate and estimate their quality based on behavioral measures alone. Neuroimaging data can provide additional constraints, but this requires a mapping from model components to brain regions. Several mappings between ACT-R modules and brain regions exist (Anderson, 2007; Anderson et al., 2008; Nijboer et al., 2016). One of the current research areas with ACT-R is utilizing the buffers to track the activity of their associated modules and then comparing that activity to data from neuroimaging studies (e.g., fMRI, MEG, or EEG) in order to identify correlations between regions of the brain and particular buffer/module activity in ACT-R models.

ACT-R models can be tested by comparing predicted response times and accuracies to those observed in humans (Anderson et al., 2004; Salvucci, 2006). More recently, ACT-R modules have been linked to specific brain regions (Anderson et al., 2008; Borst & Anderson, 2015). Because of the advancement of neuroimaging techniques, it is possible to compare activity in ACT-R's modules with activity in these brain regions. When brain regions are active, their metabolic demands increase (Anderson, 2007). As a result, the oxygen flow in the blood towards

these regions increases. fMRI allows us to track this using the blood oxygen level-dependent (BOLD) signal. As such, the BOLD signal is often used as a proxy of brain activity in an area, though this proxy is delayed, as the flow is not instantaneous. (Rosenberg-Lee et al., 2009).

The ACT-R architecture is also flexible enough that innovations made in neuro-computational models can be implemented. Originally, the existence of a connectionist implementation of ACT-R was sufficient proof of the neural plausibility of the architecture (Lebiere & Anderson, 1993). But with the increased focus on mapping functionality to specific brain regions of interest (ROIs), the question of how the ROIs implement their respective modules' functionality has risen. This is especially so for modules that are part of the basic information processing circuit (Anderson et al., 2008), which function independently from specific inputs or outputs (Anderson et al., 2007). Explaining how they are connected to other brain regions is a crucial part of understanding their functionality.

Research using fMRI has been able to develop an understanding of the roles of various brain regions and assign the cognitive functionality of those regions to ACT-R modules and buffers (Cassenti et al., 2011). Despite the strong correlation between fMRI data and ACT-R, it is unknown whether ACT-R also holds a strong correlation with EEG data. Recently, interest in this area has piqued with researchers, who have been outlining methods to study the neural correlates of ACT-R in electrophysiological data and have been demonstrating how different ACT-R modules can be associated with observable EEG data (Cassenti et al., 2011; Vugt, 2012; Vugt, 2014). Despite initial attempts to correlate ACT-R and EEG data, more investigation is required in this area.

This dissertation extended these initial findings by further exploring EEG correlates of ACT-R modules and discussed the broader implications of this approach for both cognitive

neuroscience and cognitive modeling with ACT-R. Specifically, this study aimed to develop models of multitasking behavior in a realistically complex workspace, incorporating a wide range of cognitive processes that are affected by task demand transitions. Multitasking environments can often give rise to such transitions in task demand and subjective mental workload, as the operator is actively monitoring and engaging in several different tasks at once.

In this dissertation, we also proposed a method to create data-driven parameter mappings from components of cognitive models to brain regions. The new ACT-R model parameters mapping/tuning method would be at least as powerful as an existing parameter adjustment that was based on the literature and indicated where the models were supported by the data and where they have to be improved. In the next section, we reviewed previous literatures regarding how to adjust ACT-R model parameters in order to improve models.

2.2.3.3. How to tune ACT-R model parameters

There have been some attempts to optimize parameters for ACT-R models. However, previous studies set their model parameters to the same values directly from the literature without a stated reason or found them through trial and error. For example, Faubel & Schöner (2008) found all of the model's parameters through trial and error. Samuelson et al. (2011) set six parameters for their word-learning model through trial and error, while the rest of the parameters were set to value directly from an earlier study. Many ACT-R models (Altmann & Trafton, 2002; Gonzalez et al., 2003; Lewis & Vasishth, 2005) were developed with parameters either directly from earlier literature or derived manually. Chen et al. (2015) presented a cognitive model of visual search that used reinforcement learning to discover bounded optimal strategies of eye movements. Parameters of the model were set manually based on existing literature. For example, Kangasrääsiö et al.

(2017) optimized the value of the same parameter in a later study using automatic methods, ending up with a smaller fixation duration around 250 ms, which resulted in a better model fit to observation data.

Their fit and generalizability often evaluate computational cognitive models. These properties of a model are related to two aspects of model complexity: (1) the number of parameters and (2) the functional forms of computation. In part, such evaluations seek to evaluate the extent to which noise is unnecessarily captured (Pitt et al., 2002; Oaksford, 2002). Using cross-validation, Taatgen et al. (2007) estimated parameters of a base-model once and then made use of these estimated values throughout subsequent models. To provide an alternative to estimate parameters, Wong et al. (2010) developed a database by collecting estimated and modified ACT-R parameters from the ACT-R modeling community. However, no systemically assessment has validated whether there is any sustained regularity of these estimated parameters for ACT-R models across other published studies.

2.2.4. Limitations of current studies on ACT-R modeling of multitasking

Previous studies have demonstrated different ACT-R modeling approaches for evaluating and understanding user performance and cognition processes in multitasking environments. However, these studies still present weakness: (1) few studies have quantitatively analyzed the workload transition effects, (2) lack of neurocognitive approaches to improve the current ACT-R model. Many recent studies (Salvucci, 2006; Anderson et al., 2008; Borst & Anderson, 2015) have compared ACT-R module activity with brain activity measured in fMRI studies. However, the exact timing of cognitive processes cannot be obtained by fMRI. EEG, which is brain's direct electrical activity, has a superior temporal resolution (i.e., on the order of milliseconds) than fMRI and thus could provide insights into the differences in the time course of ACT-R module activation

as well as the modules' interaction (van Vugt, 2014; Kim et al., 2018). (3) The ACT-R architecture requires multiple parameters (Bothell, 2010) which affect the model's performance. However, as previously mentioned many researchers have used the default values for all parameters (Baker, 1963; Gartenberg et al., 2014) or found optimized parameters through trial and error (Wong et al., 2010).

This study proposed a new method to quantitatively calibrate ACT-R parameters based on brain connectivity measures. Many ACT-R studies reported their parameter modifications or estimations, with the average 5.93 parameters modified for each study (Wong et al., 2010). Among the ACT-R parameters that were modified in these studies, the three most frequently modified ones were retrieval threshold, module activation noise, and latency-scaling factor. The next section presented how our method calibrates these three parameters by using EEG data and brain effective connectivity.

2.3. Application of brain connectivity for ACT-R modeling

As previously mentioned, one of the ACT-R model's advantages is its ability to control the activation of modules by a millisecond. For example, time duration to activate the production rule is set to 50ms. One of the best ways to fully utilize this advantage is comparing EEG data and time-series data of module activation (Perzenski, van Vugt). Unlike fMRI, EEG data can collect data with superior temporal resolution, thus contributing more to ACT-R modeling adjustments. Specifically, ACT-R model is suitable to quantify neural mechanism behind workload transition effects that change over time.

To measure temporal changes in neural dynamics and information flows that index and predict task-relevant changes in cognitive states, this study incorporated brain connectivity

measures into ACT-R modeling as a major model component. This allows us to evaluate the interactions between modules and causal relationships based on brain effective connectivity. Therefore, based on several connectivity measures, we proposed a new way to adjust three major ACT-R model parameters mentioned in the previous section. This section reviews previous studies on brain connectivity.

2.3.1. Analyzing effective connectivity with Granger causality

Classic physiological trends relating to workload Electroencephalography (EEG) has been employed in many studies to explore changes in cognitive workload during a variety of tasks including memory, pilot simulations and operational pilot flight (Gevins & Smith, 2005; Berka, et al., 2007). The recorded signal potential is representative of the superposition of field potentials produced by many simultaneously active neuronal components. Because EEG normally includes a composite of signals that make up the total signal from 0.5 to 100Hz, many analyses are conducted following the extraction of the Fourier components which may be further broken into the traditional frequency bands. The power spectral density (PSD) for each channel was further divided into traditional spectral frequency bands (Delta [0-3 Hz], Theta [4-7 Hz], Alpha [8-12 Hz], Beta [13-30 Hz] and Gamma [31-42 Hz]). Published research on EEG indices of workload has shown that the Alpha and Theta bands exhibit workload-dependent changes that are likely reflective of changes in cortical activity (Gevins, et al., 1979). However, these traditional measures focus on frequency and temporal correlation within the EEG signal. This only provides insights into relationships between the stimulus and activity in specified brain regions. Conversely, effective connectivity measures, such as Granger Causality, can be used to analyze time-varying interactions between brain areas in order to draw conclusions regarding causal relationships.

Studies of human brain connectivity generally fall under three categories: structural, functional, and effective connectivity (Bullmore & Sporns, 2009). Structural connectivity denotes networks of anatomical (e.g., axonal) links. The primary goal is to understand the influences of brain structures via direct or indirect axonal connections. This might be studied in vivo using invasive axonal labeling techniques or noninvasive MRI-based diffusion-weighted imaging (DWI/DTI) methods (Bullmore & Sporns, 2009). Functional connectivity denotes (symmetrical) correlations in activity among brain regions during information processing. Here the primary goal is to understand which regions are functionally related through correlations, as measured by some imaging technique. fMRI is a popular tool to analyze functional connectivity. The analysis has computed the pairwise correlation (or partial correlation) in BOLD activity for a large number of voxels or regions of interest within the brain volume. In contrast to the symmetric nature of functional connectivity, effective connectivity denotes asymmetric or causal dependencies between brain regions. Here the primary goal is to identify which brain structures in a functional network are (causally) influencing other elements of the network during information processing. Often the term "information flow" is used to indicate directionally specific (although not necessarily causal) effective connectivity between neuronal structures. Popular effective connectivity methods applied to fMRI and/or electrophysiological (EEG, iEEG, or MEG) imaging data, include dynamic causal modeling, structural equation modeling, transfer entropy, and Granger-causal methods (Mullen, 2010).

Granger causality (GC) is a method of inferring certain types of causal dependency between stochastic variables. The method is based on the reduction in prediction error of a putative effect when past observations of a putative cause and effect are used to predict the effect. The

concept was first introduced by Wiener (1958) and later reformulated by Granger (1969). In the context of linear stochastic autoregressive models, it relies on two assumptions:

1. Causes should precede their effects in time (temporal precedence)

2. Information in a cause's past should improve the prediction of the effect, above and beyond the information contained in past of the effect (and other measured variables)

Granger causality is based on the principle that directional causal influence from time series A to time series B can be inferred if past values of time series A help predict the present and future values of the time series B (Granger, 1969).

Given k time series $X(t) = [\, x_1(t) \times x_2(t) \ldots\ x_k(t)]$,

The multivariate vector autoregressive (MVAR) of order p is

$$X_t = \sum_{k=1}^{p} A_k x_{t-k} + E(t) \qquad (1)$$

Where A(n) is the coefficient of the model, as shown in Eq. (2) and E(t) is the model error.

$$A_k = \begin{bmatrix} {a_{11}}^{(n)} & \cdots & {a_{1k}}^{(n)} \\ \vdots & \ddots & \vdots \\ {a_{k1}}^{(n)} & \cdots & {a_{kk}}^{(n)} \end{bmatrix}$$

Multivariate Granger causality between the time series can be inferred from the model coefficients (Kamiński et al., 2001). The estimation of the coefficients can also be achieved in the frequency domain. That gives rise to several metrics, including directed transfer function and partial directed coherence (Kus et al., 2004). Bivariate and conditional Granger causality between time series can also be obtained using Geweke's formulation. Accordingly, the total linear dependence between time series can be split into instantaneous and causal parts, which can be estimated using either model coefficients or the error variances of the restricted and unrestricted VAR models (Geweke, 1982). The direct Directed Transfer Function (dDTF), a measure of frequency-domain conditional Granger causality, was estimated from the fitted model coefficients.

The Directed Transfer Function (DTF) allows for analysis of short epochs of EEG activity to analyze information flow between different brain structures while making it possible to determine the spectral content of the signal (Kamiński et al., 2001). Below formula indicates normalization of the DTF, where A(f) = $\sum_{k=0}^{p} \widehat{A_k} e^{-i2\pi fk}$, H(f) is the transfer matrix of A(f).

$$\eta^2{}_{ij}(f) = \frac{\left|H_{ij}(f)\right|^2}{\sum_f \sum_{k=1}^{M} \left|H_{ij}(f)\right|^2}$$

However, DTF is limited by its ability to differentiate between direct and indirect connections. By combining DTF and partial coherence (pCoh) measures, dDTF quantifies conditional, directionally-specific information transfer between sources over the trial time period at each frequency (Korzeniewska et al., 2003).

The partial coherence between i and j is the remaining coherence which cannot be explained by a linear combination of coherence between i and j and other measured variables.

$$P_{ij}(f) = \frac{\hat{S}_{ij}(f)}{\sqrt{\hat{S}_{ii}(f)\hat{S}_{jj}(f)}}$$

Thus, $P_{ij}(f)$ can regards as the conditional coherence between i and j with respect to all other measured variables, where spectral density matrix $S(f) = H(f)\Sigma H(f)-1$, $\hat{S} = S-1$ (Brillinger, 2001).

The direct Directed Transfer Function (dDTF) from process *j* to *i* is obtained from a fitted VAR[*p*] model by

$$\delta^2{}_{ij}(f) = \eta^2{}_{ij}(f)\mathrm{P}^2{}_j(f)$$

It can be shown that $\eta^2{}_{ij}(f,t)$ will be nonzero if and only if there exists a direct (multivariate) Granger-causal influence from *X j* to *Xi* at time *t* and frequency *f*.

2.3.2. Measures of brain network

A network is a mathematical representation of a real-world complex system and is defined by a collection of nodes (vertices) and links (edges) between pairs of nodes. Nodes in large-scale brain networks usually represent brain regions, while links represent anatomical, functional, or effective connections (Friston, 1994), depending on the dataset. Anatomical connections typically correspond to white matter tracts between pairs of brain regions.

Functional connections correspond to magnitudes of temporal correlations in activity and may occur between pairs of anatomically unconnected regions. Depending on the measure, functional connectivity may reflect linear and nonlinear interactions, as well as interactions at different time scales (Zhou et al., 2009). Effective connections represent direct or indirect causal influences of one region on another and may be estimated from observed perturbations (Friston et al., 2013).

The nature of nodes and links in individual brain networks is determined by combinations of brain mapping methods and measures of connectivity. Many combinations are used for experimental settings (Horwitz et al., 2004). The choice of a given combination must be carefully selected, as the nature of nodes and links largely determines the neurobiological interpretation of network topology (Butts, 2009).

The interpretation of the results obtained with graph analysis is mediated by choice of the effective connectivity measure and by the used neuroimaging technique. A node interaction analysis was performed to better extract information on the temporal relations among these brain networks obtained from GCA. Edges in a graph represent significant Granger causality interactions on the difference of influence term ($F_{X\to Y|Z} - F_{Y\to X|Z}$), allowing the application of graph-theoretic techniques.

We chose to base our graph theory analyses on the difference of influences, as opposed to the influences *per se* because these were conserved more consistently over subjects. Since Granger causality is in general not symmetric, these edges are directional. A general definition of the GCA graph-theoretic properties, as provided in previous studies (Seth, 2005; Sridharan et al., 2008; Stevens et al., 2009) is listed in the following:

2.3.2.1. Graph-theoretic measures

The causal participation of process *j* within the rest of the system can be represented by graph-theoretic measures. Outflow characterizes the causal influence of a node on the rest of the system, while the degree to which a node is causally driven by other elements of the system is represented by the inflow. This casual flow profile identifies nodes that differentially affect, or are affected by, the others. A node with a highly positive causal flow exerts a strong causal influence. On the contrary, a node with negative casual flow can be considered to be largely affected by the other ones.

- **Outflow ($\boldsymbol{\Omega_j}$)**: Sum of connectivity strengths over outgoing connections; $\Omega_i = \sum_{i=1}^{n} \delta^2{}_{ij}(f)$
- **Inflow**: Sum of connectivity strengths over incoming connections; $\Upsilon_i = \sum_{j=1}^{n} \delta^2{}_{ij}(f)$
- **Causal flow**: Difference between the outflow and inflow as a measure of the causal flow associated with a node; $F_i = \Omega_i - \Upsilon_i$
- **Causal asymmetry ratio**: The causal asymmetry ratios represent asymmetry in the causal influence of a given node. Large positive values of *Fi* (or *Ri*= 1) indicate a causal source (a driving process) while large negative values (or *Ri* = −1) indicate a causal

sink. Values near zero indicate balanced inflow and outflow or nonsignificant flow; $R_i = \frac{\Omega_i - \Upsilon_i}{\Omega_i + \Upsilon_i}$

In this study, we constructed and analyzed the distribution of these graph-theoretic properties for each node of the Granger causality network, and also calculate the mean value and standard error across all the subjects (Sridharan et al., 2008).

2.3.3. Limitations of current studies on effective connectivity of workload transition

Majority of neuroimaging studies have traditionally used neurophysiological measurements (e.g., P300 amplitude or alpha band power, etc.) to evaluate the user's cognitive state. However, these traditional measures focus on frequency and temporal correlations within the EEG signal (Kim et al., 2017). This only provides insights into overall relationships between the stimulus and activity in specific brain regions. In order to understand a human operator's cognitive state as well as the level of mental workload, we must fully understand the underlying neural mechanisms. As mentioned earlier, brain effective connectivity analysis allows us to understand causal relationships between brain regions beyond traditional methodologies. To my best knowledge, few studies have evaluated effective connectivity accounting for workload transition effects.

There exist few brain imaging studies to discover effective connectivity between brain regions matched ACT-R modules. We report some preliminary results of its application and illustrate some of the expected outcomes of this study. Borst et al. (2015) used model-based fMRI analysis to create a data-driven model-brain mapping for five modules of the ACT-R cognitive architecture. However, they used the only BOLD signal to match ACT-R modules' activations and did not find any causality or direction of information flow between ACT-R modules and brain

regions. Kim et al. (2017) investigated the interaction between two executive function processes - inhibition and updating - through analyses of behavioral, neurophysiological, and effective connectivity metrics. Using independent component analysis, source localization (DIPFIT), and Granger Causality analysis of the EEG time-series data, they demonstrated that manipulation of cognitive demand in a dual executive function task influenced the causal neural network. They compared connectivity across three updating loads (n-back levels) and found that experimental manipulation of working memory load enhanced the causal connectivity of a large-scale neurocognitive network. This network contains the prefrontal and parietal cortices, which are associated with inhibition and updating executive function processes.

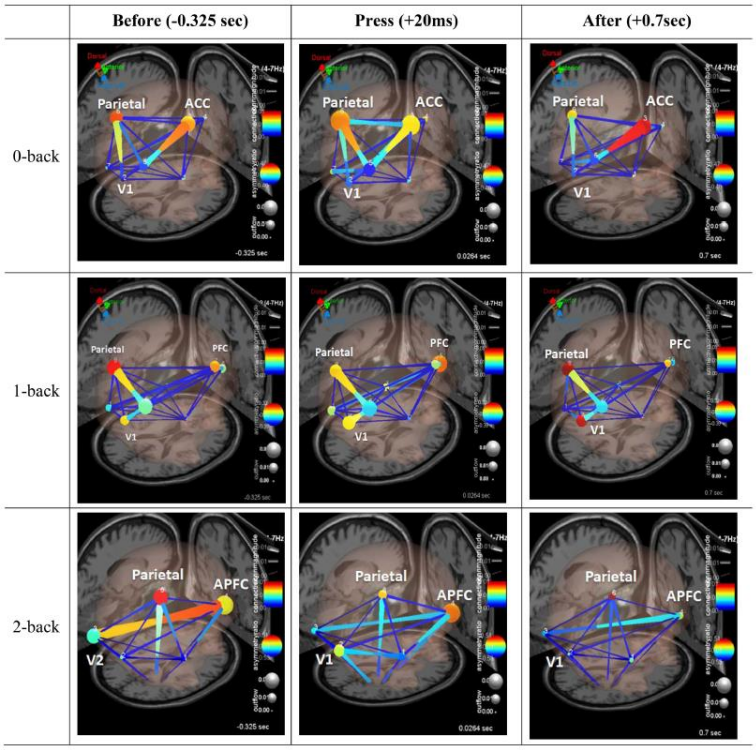
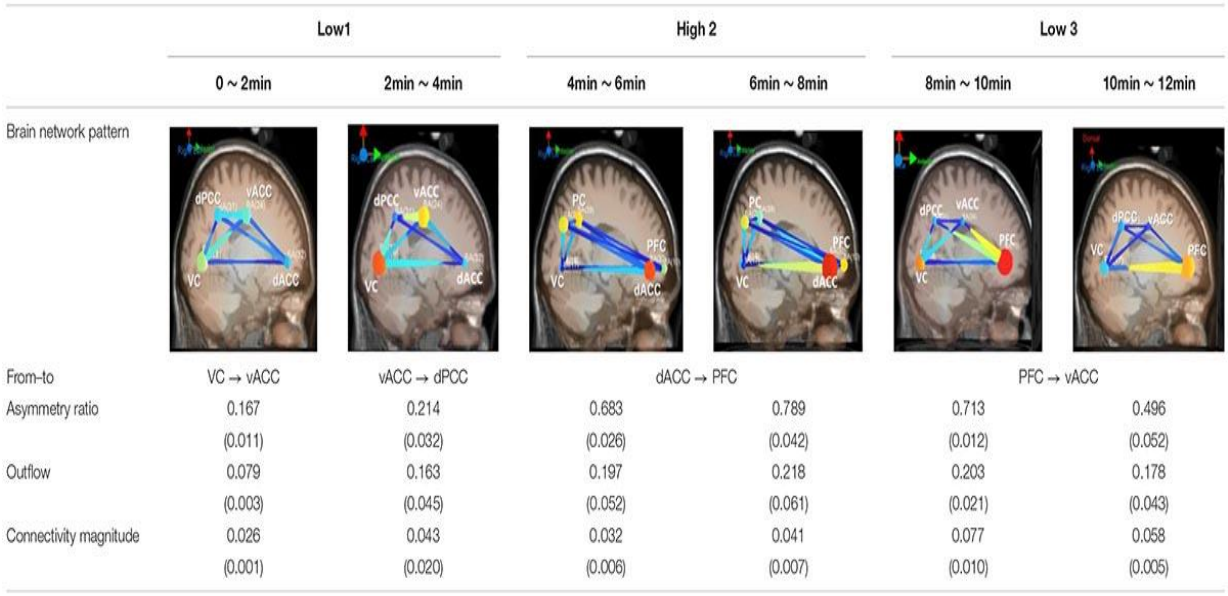


Figure 2.5. Effective connectivity linked to executive functions during n-back tasks (Adapted from Kim et al., 2017)

Figure 2.5 depicts the outcome of effective connectivity analysis. Three frames of a causal brain maps show transient information flow during the correct commission of a dual-task: The

frame corresponds to −325ms (left), 20ms (center), and 70ms (right) relative to the button press (0ms). The color of the edges represents connectivity strength (i.e., amount of information flow along that edge). Red = high connectivity, green = low connectivity. The size of edges of the graph represents connectivity magnitude (absolute value of connectivity strength). The color of node represents the asymmetry ratio of connectivity for that source. Red = causal source, blue = causal sink, green = balanced flow. The size of a node represents the amount of information outflow from the source. However, Kim et al. (2017) didn't compare EEG data to the ACT-R model. Therefore, we extended this previous founding to identify causal relationship between ACT-R modules' activation.

Kim et al. (2018) set task demands at three levels in sequential stages (Low-High-Low) using Air Force Multi-Attribute Task Battery (AF-MATB) tasks and identified differences in neural connectivity between task demand stages. Granger Causality makes it possible to conduct an effective connectivity and brain network analysis from EEG data. Figure 2.6 describes the effective connectivity analysis of each task demand stage. After completing the high demand (High 2) condition, the prefrontal area remained to activate under the Low 3 stage, the direction and magnitude of connectivity within the neural network might differ from the first demand condition (Low 1) due to hysteresis effects. Based on these studies, this dissertation analyzed such brain networks using network theoretic measures and links to module activations of ACT-R data.



From-to	VC → vACC	vACC → dPCC	dACC → PFC		PFC → vACC	
Asymmetry ratio	0.167	0.214	0.683	0.789	0.713	0.496
	(0.011)	(0.032)	(0.026)	(0.042)	(0.012)	(0.052)
Outflow	0.079	0.163	0.197	0.218	0.203	0.178
	(0.003)	(0.045)	(0.052)	(0.061)	(0.021)	(0.043)
Connectivity magnitude	0.026	0.043	0.032	0.041	0.077	0.058
	(0.001)	(0.020)	(0.006)	(0.007)	(0.010)	(0.005)

Figure 2.6. Average and standard deviation of network measures
(Adapted from Kim et al., 2018)

The extraction of the salient characteristics from brain connectivity patterns is an open challenging topic since often the estimated cerebral networks have relatively large size and complex structures. Since a graph is a mathematical representation of a network, which is essentially reduced to nodes and connections between them, the use of a theoretical graph approach would extract significant information from the brain networks estimated through different neuroimaging techniques. The present work intends to support the development of the "brain network analysis:" a mathematical tool consisting of a body of indexes based on the graph theory able to improve the comprehension of the complex interactions within the brain. In the present work, we applied for demonstrative purposes some graph indexes to the time-varying networks estimated from a set of high-resolution EEG data in a group of healthy subjects during the dual-, and multi-task. Altogether, our findings aim at proving how the brain network analysis could reveal important information about the time-frequency dynamics of effective cortical networks.

As effective connectivity can be interpreted as directional effects between cortical regions, effective connectivity can be directly linked to the nature of ACT-R computations (Ketola et al., 2019). As discussed above, ACT-R functions by firing one production at a time during its cognitive cycle; this production, in turn, changes the state of the system by modifying or copying information from one buffer to the other. For example, in what is perhaps the most common operation in ACT-R models, a production rule extracts values from the slots of chunks placed in either the imaginal or the visual buffer (to extract contextual task information) and places them in the retrieval buffer, so that they function as cues for retrieving relevant information from long-term memory. In fact, production rules are the only way information is exchanged between modules. Given their role in coordinating module-to-module communication, we made the assumption that patterns of effective connectivity follow the sequence of production rules firing.

There are multiple measures of effective connectivity. This dissertation utilizes the following three measures - connectivity magnitude (Lin et al., 2016), causal flow, and asymmetry ratio (Kim et al. 2018). Kim et al. (2018) found significant hysteresis effects on brain network measures such as prefrontal causal outflow and asymmetry, which allowed us to clarify the direction and strength of Granger causality flow under workload transitions. They showed that after completing the high-demand condition, the prefrontal area remained activated, so the direction and magnitude of connectivity within the neural network might differ from the pre-transition period.

3. STUDY 1: ACT-R modeling of workload transition in dual cognitive tasks

From a comprehensive review of previous studies, two important research questions were identified: (1) mixed findings on workload transition effects and (2) lack of a systematic method to adjust major parameters of ACT-R in a way that increases the model fit. To address the two main research issues, Study 1 was conducted in two phases (see Figure 3.1).

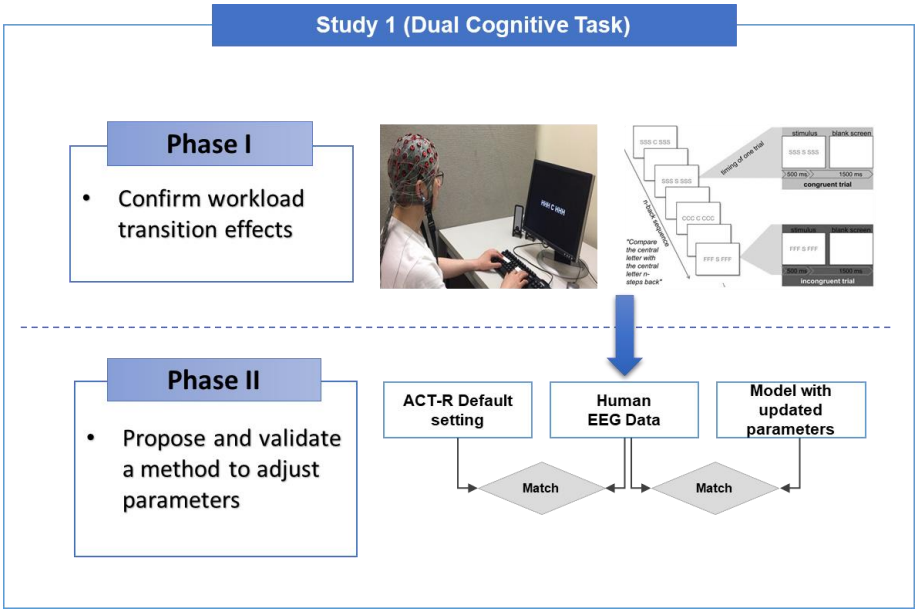


Figure 3.1. Framework of Study 1 with two phases

Phase I was designed to assess workload transition effects in a dual cognitive task. In Phase II, a method was proposed, in which three main ACT-R parameters can be effectively adjusted based on effective connectivity measured through a comparison of a priori ACT-R model built from the conventional modeling method with the empirical data collected in Phase I.

3.1. Phase I: Workload transition affecting human performance and cognition

3.1.1. Objectives and hypotheses

Phase I of Study 1 aimed to investigate the effects of workload transitions on performance, mental workload, and neurophysiological measures overtime during a dual cognitive task.

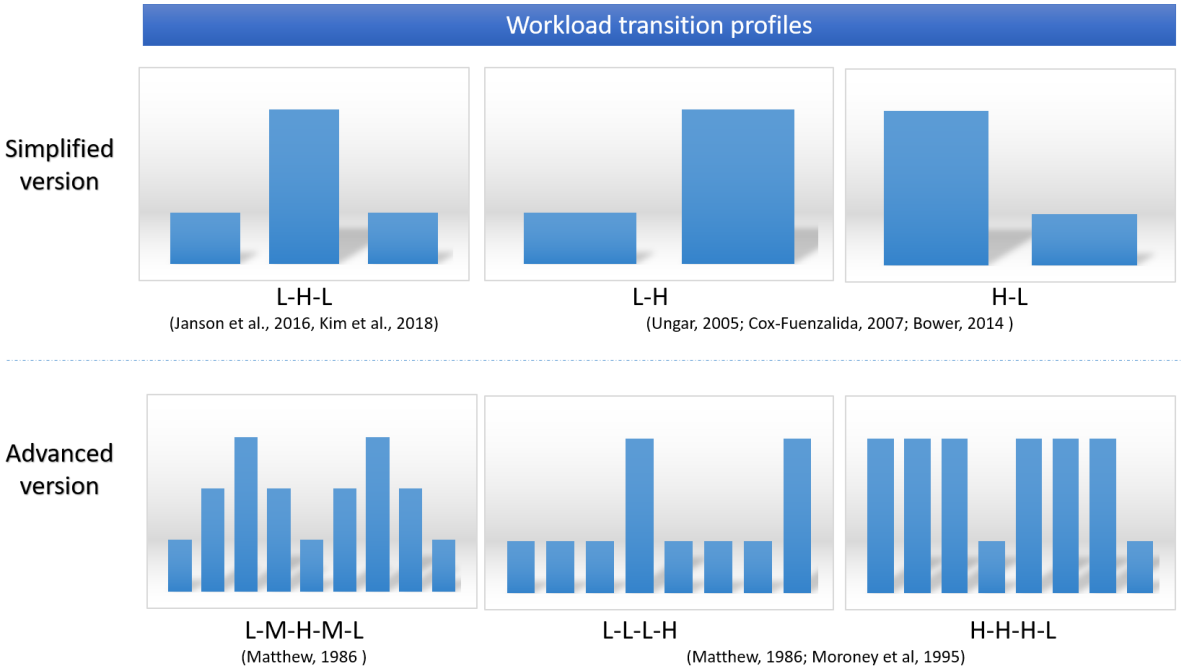


Figure 3.2. Workload transition profiles

As reviewed previously, several transition profiles have been studied in the literature, but they have shown mixed findings. In this dissertation, we utilized the advanced version of workload profile (Matthew,1986, Moroney et al., 1995) instead of a simplified version (Ungar, 2005, Cox-Fuenzalida, 2007, Bower et al., 2014, Jansen et al., 2016, Kim et al., 2018). See Figure 3.2.

In case of the advanced version, we identified three types of workload transition profiles; one case of cyclic variation and two extreme cases of workload variations that can occur; ramp-down and ramp-up because of following strong points:

(a) It allows us to collect more time points to investigate workload transition effects than the simplified version.

(b) This dissertation focused on the frequency and magnitude of workload transitions. Workload transition profiles consist of three components; (1) transition direction (increase or decrease), (2) degree of transition (moderate change (L-M-H) or extreme change (L-H)), and (3) time-oriented- the distinction of change (gradual change or abrupt change).

Based on these components, the experiment utilized three different types of workload profiles; one case of "cyclic variation" (increase & decrease / moderate change / gradual) and two extreme cases of workload variations: "ramp-down" (decrease / extreme change / abrupt) and "ramp-up" (increase / extreme change / abrupt). With 3 types of workload transition profiles, we can investigate multiple aspects of workload transition effects. In phase I, three sets of hypotheses associated with performance, mental workload and neurophysiology were tested for each transition profile group.

Performance-related hypotheses:

H3.1: The post-transition performance of a V1 group (cyclic variation) would be better than that of a non-transitioned group.

H3.1 (a): Participants' averaged reaction time would be shorter under the cyclic workload profiles than under the non-transitioned conditions.

H3.1 (b): Participants' averaged accuracy would be higher under the cyclic workload profiles than under the non-transitioned conditions.

H3.2: The post-transition performance of a V2 group (Ramp-up) will be lower than that of a non-transitioned group.

H3.2 (a): Participants' averaged reaction time would be longer under the ramp-up workload profiles than under the non-transitioned conditions.

H3.2 (b): Participants' averaged accuracy would be lower under the ramp-up workload profiles than under the non-transitioned conditions.

H3.3: The post-transition performance of a V3 group (Ramp-down) would be better than that of a non-transitioned group.

H3.3 (a): Participants' averaged reaction time would be longer under the ramp-down workload profiles than under the non-transitioned conditions.

H3.3 (b): Participants' averaged accuracy would be lower under the ramp-down workload profiles than under the non-transitioned conditions.

Mental workload-related hypotheses:

H3.4: The post-transition mental workload of a V1 group (cyclic variation) should be lower than a non-transitioned group.

H3.4 (a): Participants ' averaged ISA would be lower under the cyclic workload profiles than under the non-transitioned conditions.

H3.5: The post-transition mental workload of a V2 group (Ramp-up) should be higher than a non-transitioned group.

H3.5 (a): Participants ' averaged ISA would be higher under the ramp-up workload profiles than under the non-transitioned conditions.

H3.6: The post-transition mental workload of a V3 group (Ramp-down) should be higher than a non-transitioned group.

H3.6 (a): Participants ' averaged ISA would be higher under the ramp-up workload profiles than under the non-transitioned conditions.

Neurophysiology-related hypotheses:

H3.7: The post-transition information flow of the cortical hub node of a V1 group (cyclic variation) would be stronger than a non-transitioned group.

H3.7 (a): Participants' averaged connectivity magnitude of the hub node would be higher under the cyclic workload profiles than under the non-transitioned conditions.

H3.7 (b): Participants' averaged outflow of the hub node would be higher under the cyclic workload profiles than under the non-transitioned conditions.

H3.7 (c): Participant's averaged asymmetry ratio of the hub node would be higher under the cyclic workload profiles than under the non-transitioned conditions.

H3.8: The post-transition information flow of the cortical hub node of a V2 group (ramp-up) should be weaker than a non-transitioned group.

H3.8 (a): Participants' averaged connectivity magnitude of the hub node would be lower under the ramp-up workload profiles than under the non-transitioned conditions.

H3.8 (b): Participants' averaged outflow of the hub node would be lower under the ramp-up workload profiles than under the non-transitioned conditions.

H3.8 (c): Participants ' averaged asymmetry ratio of the hub node would be lower under the ramp-up workload profiles than under the non-transitioned conditions.

H3.9: The post-transition information flow of the cortical hub node of a V3 group (ramp-down) should be weaker than a non-transitioned group.

H3.9 (a): Participants' averaged connectivity magnitude of the hub node would be lower under the ramp-down workload profiles than under the non-transitioned conditions.

H3.9 (b): Participants' averaged outflow of the hub node would be lower under the ramp-down workload profiles than under the non-transitioned conditions.

H3.9 (c): Participant's averaged asymmetry ratio of the hub node would be lower under the ramp-down workload profiles than under the non-transitioned conditions.

3.1.2. Methods

3.1.2.1. Participants

A total of 60 right-handed participants (30 male; 30 female) with normal hearing and normal or corrected-to-normal vision was recruited from a local University. Table 3.1 presents the overall demographics of participants who took part in the experiments, including gender, age, and OSPAN score (for details, see Table 3.1).

Table 3.1. Demographics of participants (N=60)

		Gender (Number)		Age (Year)		OSPAN (Absolute score)	
Group		**M**	**F**	**Mean**	**SD**	**Mean**	**SD**
Workload transition profile	V1	5	5	21.8	2.34	48.5	12.73
	V2	5	5	19.7	1.05	46.1	17.10
	V3	5	5	21.7	1.76	49.4	17.68
Constant workload	C1	5	5	21.9	2.46	49.9	10.89
	C2	5	5	21.1	2.51	45.1	16.01
	C3	5	5	20.1	0.73	51.1	13.94

Note: V1: Cyclic variation; V2: Ramp-up; V3: Ramp-down; Control: C1, C2, and C3; OSPAN- Operation SPAN

Only participants whose first language is English were invited to the present study to avoid the influence of language on letter stimuli of the n-back task and AF-MATB operation (e.g.,

respond to English audio commands). The inclusion and exclusion criteria used during the screening process were:

1) participants should have 20/20 or corrected vision and no color-impairment due to the study requirement for visual perception of colored items/text on a display screen;
2) participants should have a general familiarity with the use of computers as the study involves using a simulation program running on a laptop computer; and
3) participants should have little to no experience with AF-MATB tasks (so as not to bias performance in control tasks).

3.1.2.2. Apparatus and materials

Operation span task (OSPAN): This study used the OSPAN scores as individual differences in working memory capacity to ensure group comparability (i.e., No difference in individual differences within each group). As a predictor of multitasking performance, WMC has been often measured by various working memory span tasks such as reading span task (Daneman & Carpenter, 1980), and counting span task (Case et al., 1982). In particular, the operation span task (OSPAN) (Turner & Engle, 1989) has been widely used as a reliable and valid marker of WMC (e.g., Kane & Engle, 2003; Konig et al., 2005; Redick & Engle, 2006). An automated version of OSPAN was employed in this study for two main reasons: (1) it is easy-to-administer and (2) it requires little intervention from the experimenter because participants can complete it online by themselves (Unsworth et al., 2005).

The OSPAN test includes items (letters) to remember and math problems to solve. There are three practice sessions and one experimental session. Appendix C includes instruction of the automated OSPAN task which was provided to the participants. The score range of this automated

OSPAN task is 0 – 75. At the end of the task, the program reports five scores: OSPAN score, total number correct, math errors, speed errors, and accuracy errors (Unsworth et al., 2005). The task requires approximately 20–25 minutes.

Table 3.1 presents the OSPAN scores of the participants by control and experimental groups. According to previous studies (Kane & Eagle, 2003; Goldinger et al., 2003; Watson et al., 2005), the cut-off scores of the OSPAN test in the present study was 19 and 30 (25% and 40%) for low- and high-span groups, respectively.

Edinburgh Handedness Inventory: Only right-handed participants were recruited in the present study because handedness is related to the dominance of the cerebral hemispheres (Witkowski et al., 2019). The Edinburgh Handedness Inventory (see Appendix B), developed by Oldfield (1971), was used to avoid the influence of handedness on neuropsychological performances.

3.1.2.3. Experimental task

A combined n-back and Flanker task was employed as a dual cognitive task in the present study (See Figure 3.3). The flanker task, also known as the Eriksen flanker task, was designed where the letters S, H, C, and F were used as target stimuli (Scharinger et al., 2015). For each trial, one of the four target letters, which was randomly chosen (e.g., H), was flanked by non-target or irrelevant distractor letters that are either congruent (e.g., HHH H HHH) or by incongruent (e.g., FFF H FFF). The n-back task was designed for the participants to decide if the current stimulus is the same as the one presented n trials ago.

In the combined n-back and Flanker task, participants were asked to indicate as quickly and accurately as possible if the randomly chosen target letter (i.e., S, H, C, or F) flanked by either

same letters (i.e., congruent) or different letters (i.e., incongruent) is the letter that matches a pre-specified letter (0-back), was displayed one trial ago (1-back) or two trials ago (2-back).

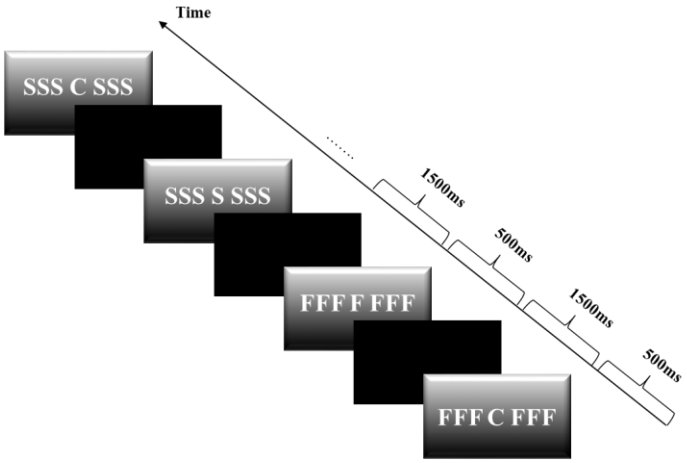


Figure 3.3. Illustration of cognitive dual-task: A combined n-back and flanker task used in the study as a dual cognitive task that requires participants to perform two tasks simultaneously

In the present study, the n-back stimuli were presented in blocks, which means that participants performed Flanker tasks under each of the three n-back conditions. Under an n-back condition, for example, participants were asked to press the "L" key with their right index finger or "A" key with their left index finger on a standard QWERTY keyboard if the currently presented central letter is or is not the n-back target, respectively. Each block consisted of 40 trials. Half of the trials were targets; another half of the trials were non-targets. About one-third of the stimuli of each response category were incongruent (e.g., FFF H FFF); two-thirds were congruent (e.g., FFF F FFF).

All letters were presented in white on black backgrounds in Arial at 25-point font size. Each stimulus was shown for 500 ms, followed by a black screen for 1500 ms. Thus, each trial lasted 2,000 ms. Stimuli were presented using the BCI2000 presentation tool with a list of predefined stimuli. The trial sequences within the block were pseudorandomized. To avoid attenuation of the interference effect for incongruent stimuli due to conflict adaptation processes

(i.e., the so-called Gratton effect; Gratton, Coles, & Donchin, 1992), incongruent-incongruent stimuli sequences were not used in this. To further avoid any Gratton-like effects, congruent trials following incongruent trials were also excluded from any further data analyses. Apart from these constraints, the stimulus sequence (i.e., the letters were chosen as stimuli as well as their assignment as target/non-target or congruent/incongruent) was randomly generated for each block and each participant (Scharinger, Soutschek, Schubert, & Gerjets, 2015).

3.1.2.4. EEG signal acquisition and processing

EEG signals were recorded using an EEG cap (Electro-Cap International, Inc.) embedded with 64 active electrodes, based on the modified 10–20 system of the International Federation (Sharbrough et al., 1991). Figure 3.4 shows the electrode montage used in the study. Recordings were referenced to the left ear lobe and grounded between AFz and FPz. EEG signals were amplified with a g.USBamp amplifier (g.tec Medical Engineering). The EEG sampling rate was 256 Hz and down-sampled to 128Hz for the connectivity analysis. EEG signal was bandpass filtered between 0.01 Hz and 75 Hz and notch-filtered at 60 Hz.

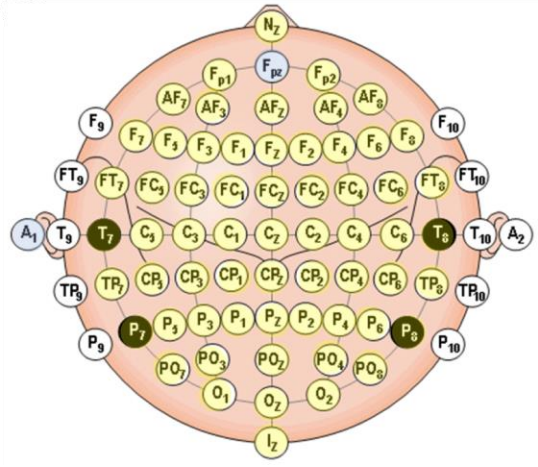


Figure 3.4. EEG electrode montage used

Epochs were determined using event-locked time windows, ranging from 500 ms before the subject response (button press) to 1250 ms after the subject response. Independent component analysis (ICA) was applied to decompose the EEG signal into independent components (ICs). Any ICs that resemble electrooculographic (EOG) activity were rejected from further analysis through visual inspection. Signal acquisition and processing were conducted using BCI2000 (Schalk et al., 2004), MATLAB (The MathWorks), and EEGLAB (Delorme et al., 2011).

3.1.2.5. Effectivity connectivity analysis

The Source Information Flow Toolbox (SIFT) for EEGLAB was used to evaluate effective connectivity – the causal flow of information between brain sources (Delorme et al., 2011).

- **Brain connectivity analysis procedures.** Following ICA analysis and artifact rejection procedures, all retained ICs were localized using the EEGLAB Dipole Fitting plug-in (DIPFIT) (see Figure 3.6). SIFT was used to evaluate effective connectivity. A multivariate autoregressive model (MVAR) was fit to the ensemble-normalized ICs using the Vieira-Morf algorithm with a 350 ms window length, step size of 30, and model order was determined. Model order was optimized from 1 to 40, such that the Hannan-Quinn criterion for each participant was minimized. The optimized model order values were averaged across all participants.

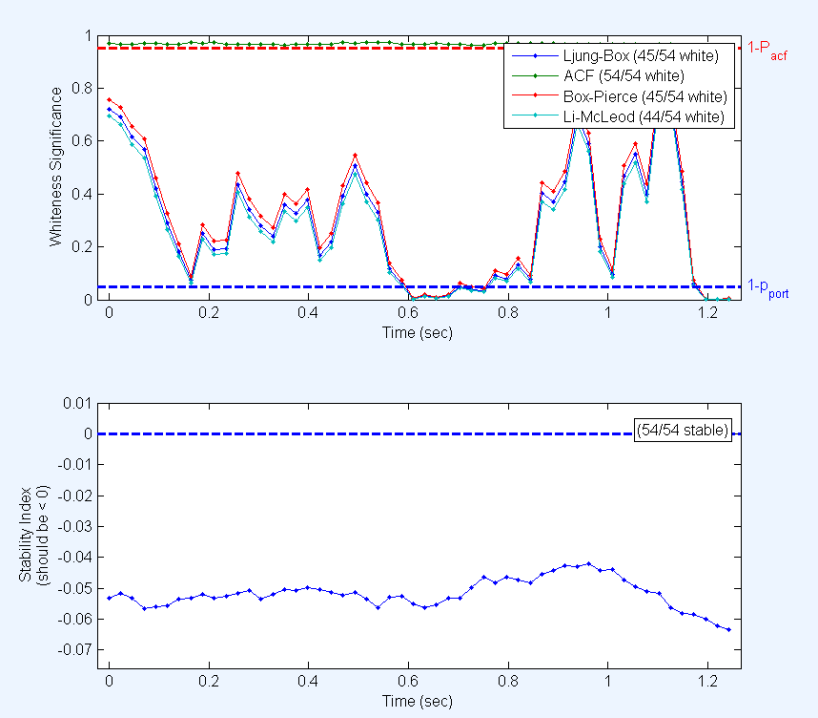


Figure 3.5. Validation of MVAR model using ACF, Ljung-box and Box-pierce

To validate the MVAR model, the whiteness of the residuals, model stability, and percent consistency were determined for each trial. See Figure 3.5. The autocorrelation function (ACF) and the Li-McLeod Portmanteau (LMP) tests were used as whiteness test criteria. The LMP test was used due to its improved small-sample properties and lack of variance inflation compared to other available Portmanteau tests (Mullen, 2010). In addition to meeting the ACF and LMP criterion, the model stability index was less than zero, and percent consistency was above 85% for each trial, indicating a validated model. The direct Directed Transfer Function (dDTF), a measure of frequency-domain conditional Granger causality, was estimated from the fitted model coefficients. The Directed Transfer Function (DTF) allows for analysis of short epochs of EEG activity to analyze information flow between different brain structures while making it possible to determine the spectral content of the signal (Kamiński et al., 2001). However, DTF is limited by

its ability to differentiate between direct and indirect connections. By combining DTF and partial coherence measures, dDTF quantifies conditional, directionally-specific information transfer between sources over the trial time period at each frequency (Korzeniewska et al., 2003). In this study, dDTF was determined over the 1-40 Hz frequency range. A percentile threshold of 97.5% is used for each frequency to visualize relevant directional connections between brain sources (Kim et al., 2018).

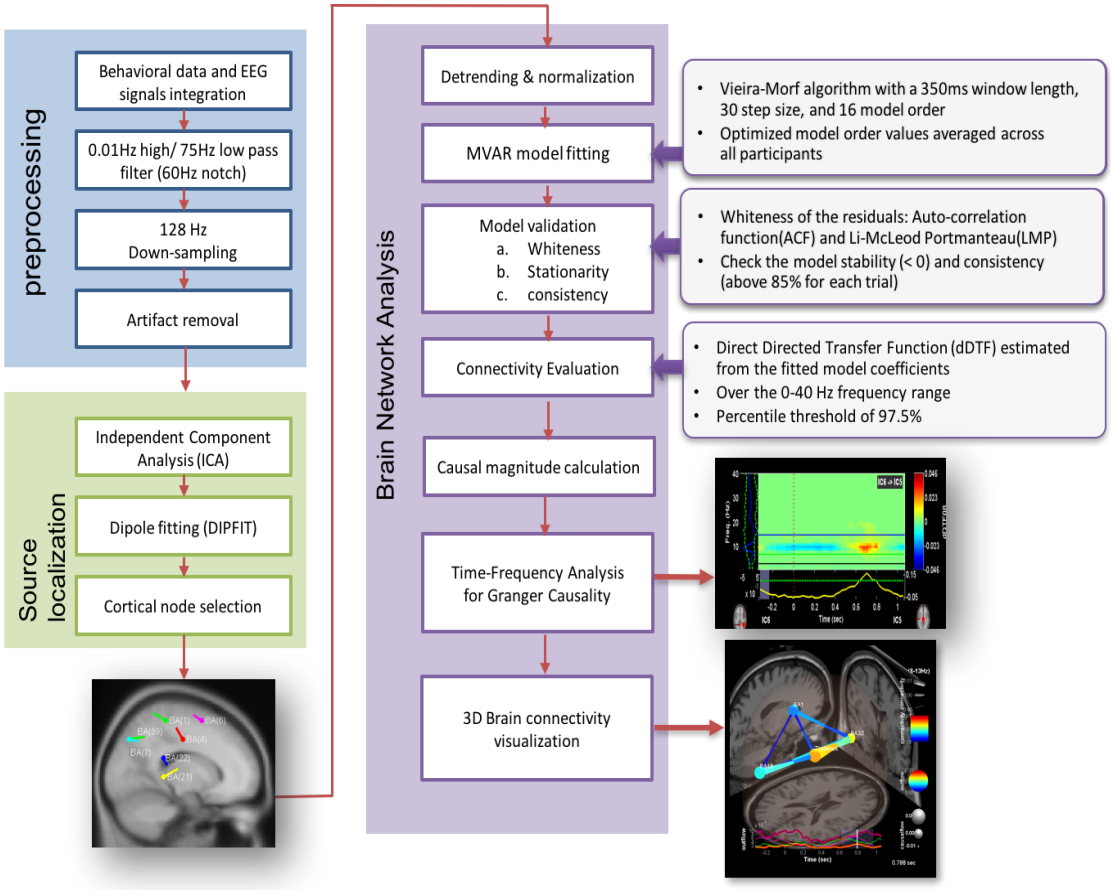


Figure 3.6. Flowchart of the connectivity analysis procedure

- **Effective connectivity measures**

When the connectivity evaluation is finished, then we can calculate graph-theoretic measures such as inflow/outflow, causal flow (refer to section 2.3.2) and asymmetry ratio.

N is the set of all nodes in the network, and n is the number of nodes. a_{ij} is the connection status between i and j: a_{ij}=1 when link i, exists (when i and j are neighbors); aij=0 otherwise (aii=0 for all i). Degree of node i, can be calculated by this formula: $k_i = \sum_{j \in N} a_{ij}$

- **Outflow:** Sum of connectivity strengths over outgoing connections
- **Inflow:** Sum of connectivity strengths over incoming connections
- **Causal flow:** Difference between the outflow and inflow as a measure of the causal flow associated with a node.
- **Asymmetry ratio:** Indicates whether a cortical hub node is a causal sink or source.

$$AR = \frac{Inflow - Outflow}{Inflow + Outflow}$$

$$-1 \leq AR \leq 1$$

AR = -1 indicates all connectivity related to that node is inflowing (a causal sink)

AR = +1 indicates all connectivity related to that node is outflowing (a causal source)

AR = 0 indicates either a balanced flow or no significant flow

3.1.2.6. Experimental design and independent variables

The experimental design was a one-way factorial design with a workload profile as a between-subject factor with six levels (V1, V2, V3, C1, C2, and C3) as shown in Table 3.2. The six workload profile types were created to simulate multitasking environments where people perform multiple tasks with varied task demand levels. Simulated variations have been used in workload transition research (e.g., Matthews, 1986).

The workload was manipulated by three n-back levels (0-back, 1-back, and 2-back) that represent varying levels of demand on human operators during multitasking. The variations can be categorized into two groups, experimental and control. First, the experimental group was

created with three types of n-backload variations during a session (Matthew, 1986): 1) **Cyclic** variation (V1): a cyclic variation of the load across successive trial blocks (e.g., 0 → 1 → 2 → 2 → 1 → 0), 2) **Ramp-up** variation (V2): a significant increase in the levels of workload from the lowest (0-back) to the highest level (2-back) on the fourth trial block during each session, and 3) **Ramp-down** variation (V3): a significant decrease in the levels of workload from the highest (2-back) to the lowest level (0-back) on the fourth trial block during each session, which is the reversed V2 pattern. Three control groups performed the entire session at a single load level (i.e. 0-back, 1-back, and 2-back, respectively).

Table 3.2. The sequence of the n-back levels by six variation groups during sessions

		Session 1				Session 2				Session 3			
Block		1	2	3	4	5	6	7	8	9	10	11	12
Experimental Variation	V1	0	1	1	2	2	1	1	0	0	1	1	2
	V2	0	0	0	2	0	0	0	2	0	0	0	2
	V3	2	2	2	0	2	2	2	0	2	2	2	0
Control Variation	C1	0	0	0	0	2	2	2	2	1	1	1	1
	C2	1	1	1	1	0	0	0	0	2	2	2	2
	C3	2	2	2	2	1	1	1	1	0	0	0	0

3.1.2.7. Dependent variables

To investigate the effect of task demand transitions or hysteresis effect, this study measured several dependent variables which can be categorized into three types of variables at the behavioral,

perceptual, and neurophysiological levels, respectively: (1) task performance, (2) mental workload, and (3) brain activity.

(1) **Task Performance**: Task performance was measured by accuracy and reaction time on the n-back tasks flanked by either congruent or incongruent stimuli.

- ***Accuracy (%):*** Accuracy was measured as the hit rate, which was calculated as the proportion of correct responses at each n-back level.
- ***Response time (RT, millisecond):*** Response time was calculated as the time between the beginning of a stimulus display and until the time of button pressing at each n-back level. Only the trials with correct responses were included for calculating this variable.

(2) **Mental Workload**: To measure perceived workload or mental workload during a block, this study used the unidimensional Instantaneous Self-Assessment (ISA) scale that measures how much mental workload the task just required. The ISA is one of the most frequently used measures of mental workload (Tattersall et al., 2007). Participants were required to verbally report their perceived mental workload on a 7-point Likert scale (1 corresponding with a very low mental workload and 7 with a very high mental workload) after completing each block.

(3) **Brain Activity**: Neurophysiological measures included Effective connectivity metrics, calculated as explained in Section 3.1.2.5.

3.1.2.8. Procedure

Prior to the experiment, participants were asked to complete a demographic questionnaire and a handedness survey (see Appendix A and B, respectively). We confirmed that all participants had

no experience with the dual cognitive task and AF-MATB system before the main experiments. At the beginning of the experiment, all participants first performed a practice session under each static-load condition in which they became familiar with the stimuli and the task. The practice session was repeated until participants reach an accuracy of at least 60 percent correct responses. Any subjects with an accuracy below 60 percent were excluded from the analysis. Figure 3.7 presents an experimental set-up that housed an EEG acquisition system and a computer system that displays a dual cognitive task and measures task performance.

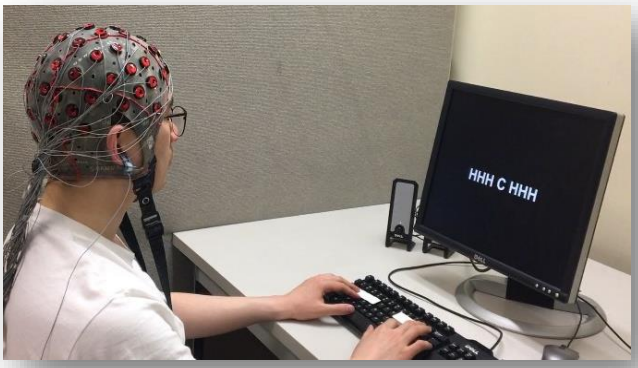


Figure 3.7. Experimental set-up in Study 1

After the practice session, participants performance the main task, a set of dual cognitive tasks that combined the Flanker and N-back tasks, that was composed of three sessions, each session having four blocks of 40 trials (see Table 3.2 for the sequence of the n-back levels by six variation groups during sessions). Figure 3.8 presents the procedure for the experiment. After each block, participants were asked to complete the Instantaneous Self-Assessment (ISA) to measure their perceived mental workload. The total experiment, including EEG preparation, task training, task run, and breaks took about 1.5 hours.

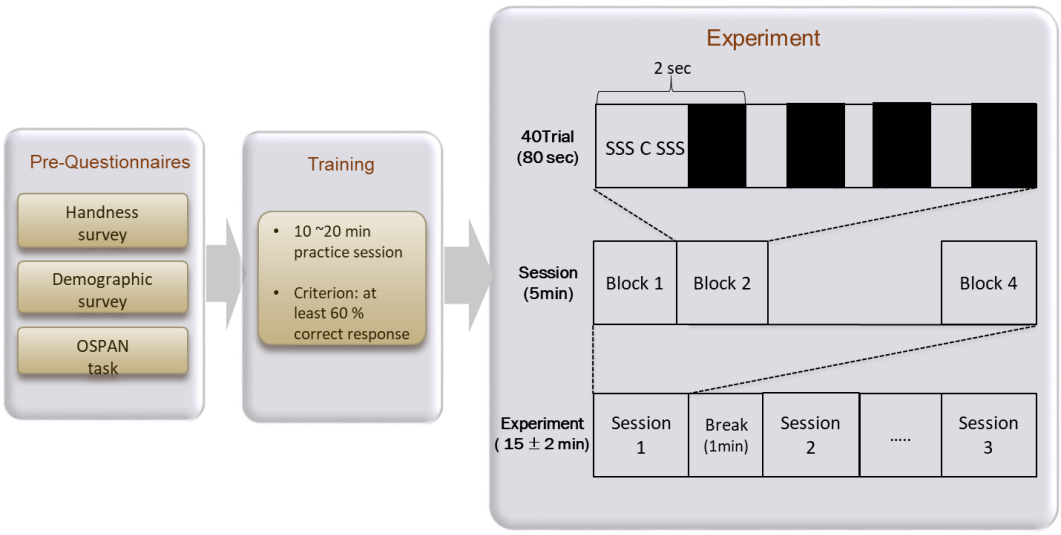


Figure 3.8. Schematic illustration of the procedure used in Phase 1

3.1.2.9. Statistical analyses

A repeated measure ANOVA without interaction was conducted for each dependent measure to evaluate the main effects of workload profile (V1, V2, V3, and constant 0-back, 1-back and 2-back) and time block effect (from 5th block to 12th block). Data from the first four blocks (Session 1) were not included in this analysis because subjects had been exposed to too few trials to establish a workload history until the first four blocks.

$$Y_{ij} = \mu + \alpha_i + \beta_j + B_k + \varepsilon_{ijk}$$

where

α_i = main effect of the workload transition profile with six levels (i = 1 to 6)

β_j = main effect of the time block (2-min) with eight levels (j = 1 to 8)

B_k = blocking factors of subject k (k = 60)

$\varepsilon_{(ijk)}$ = error term

Before performing ANOVAs, the assumptions of homoscedasticity and residual normality were examined using Bartlett's test (Garson, 2012) and the Shapiro-Wilk normality test (Shapiro and Wilk, 1965), respectively.

Power analysis was conducted to calculate the actual power for the experiment result. For all analyses, the level of significance was set to a = 0.05 and eta square (η^2) was reported. The eta squared (η^2) was reported to measure the effect sizes, the proportions of the variability associated with an effect compared to the total variance, characterized by small, medium, and large effect with values of .01, .06, and .14, respectively (Cohen, 1988; Richardson, 2011).

The paired t-tests were used to determine which conditions had significantly different means from other conditions. In order to compare the performance of the variable-load groups with the constant-load controls at a macro level, each of the load levels was contrasted with ordinally equivalent trial blocks. For example, in order to compare the performance at the 0-back level for the variable-load groups with the performance of controls, the mean trial blocks on which the 0-back level occurred in these groups were contrasted with the mean of the trial blocks in the 0-back control group at the same ordinal position. See Table 3.3 (b). For example, in the 0-back condition, we compared V1 and control the averaged performance 1 group's 8^{th} and 9^{th} blocks and control based on corresponding trial blocks with 0-back, which is the C2 group's 8^{th} block.

In addition, in this study, the workload profile was implemented in multiple blocks with varying n-back levels. Thus workload profile effects were assessed by comparing dependent measures between the constant n-back groups (0-back, 1-back, and 2-back), and transitioned groups during the blocks with the same n-back levels in the simple t-test (See Table 3.3 (a) and (b)).

Table 3.3. Comparison template for time bins between the control and transitioned groups

(a) The sequence of the n-back levels by six variation groups during sessions

		Session 1				Session 2				Session 3			
Block		1	2	3	4	5	6	7	8	9	10	11	12
Variable-load group	V1	0	1	1	2	2	1	1	0	0	1	1	2
	V2	0	0	0	2	0	0	0	2	0	0	0	2
	V3	2	2	2	0	2	2	2	0	2	2	2	0
Constant-load group	0-back	0	0	0	0	0	0	0	0	0	0	0	0
	1-back	1	1	1	1	1	1	1	1	1	1	1	1
	2-back	2	2	2	2	2	2	2	2	2	2	2	2

(b) Blocks with same n-back levels

	Comparison	**Corresponding time blocks**
0-back	Control vs. V1	8^{th}, 9^{th}
	Control vs. V2	5^{th}, 6^{th}, 7^{th}, 9^{th}, 10^{th}, 11^{th}
	Control vs. V3	8^{th}, 12^{th}
1-back	Control vs. V1	6^{th}, 7^{th}, 10^{th}, 11^{th}
2-back	Control vs. V1	5^{th}, 12^{th}
	Control vs. V2	8^{th}, 12^{th}
	Control vs. V3	5^{th}, 6^{th}, 7^{th}, 9^{th}, 10^{th}, 11^{th}

3.1.3. Results and discussions

3.1.3.1. Behavioral measures

In this study, the workload transition profile was implemented in multiple blocks with varying n-back levels. The separate repeated measure ANOVAs were conducted for each behavioral measure; reaction time and accuracy. The analysis yielded statistically significant differences. The significant effects were investigated using multiple contrasts. The comparison of baseline scores

and the first test trial provided the basis for investigating the nature of the immediate decrement following

3.1.3.2.1. Reaction time (RT)

Table 3.4 Mean and standard deviation of RTs (sec)

	Mean	**SD**
V1	0.509	0.023
V2	0.513	0.016
V3	0.608	0.022
0-back	0.459	0.013
1-back	0.522	0.020
2-back	0.627	0.017

The results of ANOVA on reaction time showed significant main effects of workload profile, $F(5, 103.2) = 4.079$, $p < .001$, $\eta^2 = .23$. RTs increased with increasing n-back levels within control group (0-back: M= 0.459, SD = 0.013, 1-back: M = 0.522, SD = 0.020, 2-back: M = 0.627, SD = 0.107). In experimental variation groups, participants performed with following RTs (V1: M= 0.509, SD = 0.023, V2: M = 0.513, SD = 0.016, V3: M = 0.608, SD = 0.022). Tukey's HSD test showed that 2-back control group was significantly different from V1, and V2 group.

Table 3.5. ANOVA results on RT

	Effect	**F value**	***p* levels**	**Eta square**
RT	Workload profile	F(5, 103.2) = 4.079	0.0001**	0.23
	Time block	F(7, 378) = 5.709	0.0001**	0.12

Note: * significance level : p <0.05, ** significance level: p < 0.001

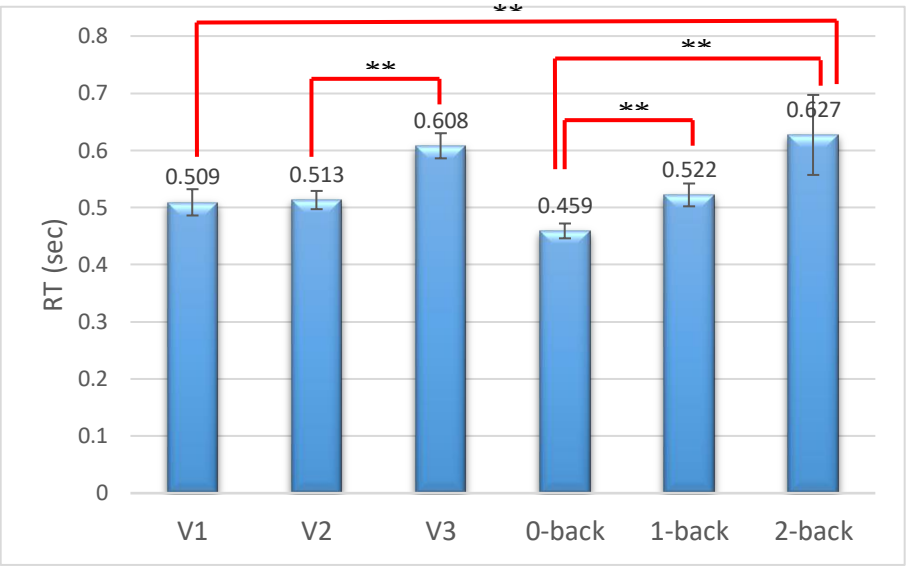


Figure 3.9. Result of Tukey's HSD test on RT

The main effect of time block was also found, $F(7, 378) = 5.709$ $p < 0.001$, $\eta^2 = .08$. For further analysis, a paired t-test was performed to test the significance of the differences between the control groups and variation groups with the same n-back level (Table 3.6). The result of the paired t-test showed that the average RT of the V1 (0.582 sec; $p < 0.001$) was significantly shorter than the constant 2-back control group (0.618 sec). This result might indicate that the cyclic workload profile positively affected their correct RT. However, there were no significant differences between sudden change groups (V2 and V3) and non-transitioned group.

Table 3.6. Mean reaction time: Control vs. Transitioned group on RT

		Reaction times (sec)			
Workload profile		Control	V1	V2	V3
N-back level	0-back	0.433	0.428		
		0.452		0.460	
		0.445			0.433
	1-back	0.532	0.496		
	2-back	0.618	0.557**		
		0.648		0.659	
		0.656			0.677

Note: Means for the variable load conditions with an asterisk are significantly different ($p < 0.05$) form their appropriate controls, which appear to their left on the same line of the table.

3.1.3.1.2. Accuracy

The results of ANOVA on reaction time showed significant main effects of workload profile, $F(5, 103.2) = 6.436$, $p < .001$, $\eta^2 = .13$. The mean accuracy was decreased with increasing n-back levels within the control group (See Table 3.7).

Table 3.7. Mean and standard deviation of accuracy

	Mean	**SD**
V1	0.811	0.020
V2	0.834	0.022
V3	0.754	0.019
0-back	0.849	0.015
1-back	0.820	0.017
2-back	0.750	0.023

The main effects of time block was found, $F(7, 378) = 8.273$, $p < 0.001$, $\eta^2 = .11$. For further analysis, a simple t-test was performed to test the significance of the differences between the control groups and variation groups with the same n-back level (Table 3.9). The result of the t-test showed that the average accuracy of the V1 (0.779 sec; $p = 0.002$) was significantly higher than the constant 2-back control group (0.738 sec). This result supports the H3.1 (b) that the cyclic workload profile affects performance. However, there were no significant differences between abrupt change groups (V2 and V3) and non-transitioned group.

Table 3.8. Significant ANOVA results on accuracy

	Effect	**F value**	***p* levels**	**Eta square**
Accuracy	Workload profile	$F(5, 103.2) = 6.436$	0.0001**	0.13
	Time block	$F(7, 378) = 8.273$	0.0001**	0.11

Note : * significance level : $p < 0.05$, ** significance level: $p < 0.001$

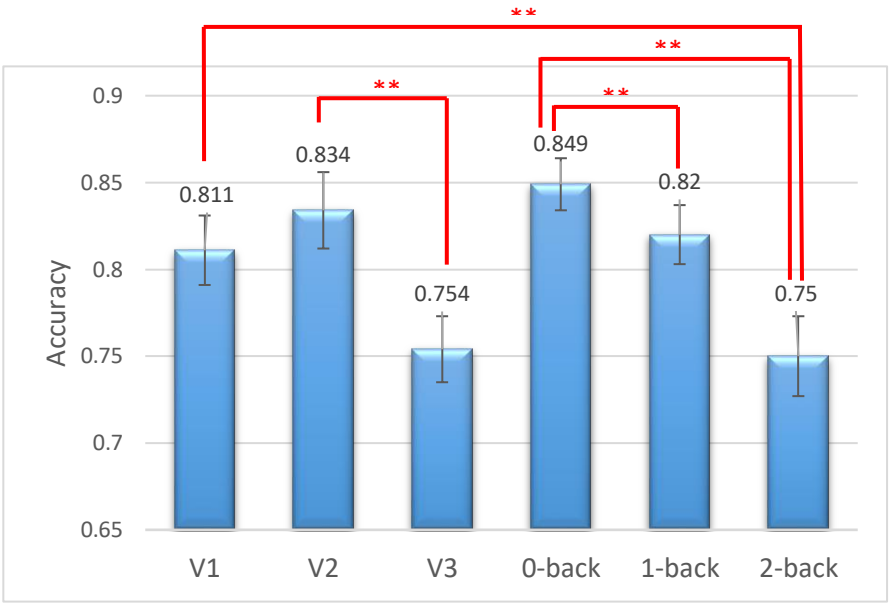


Figure 3.10. Result of Tukey's HSD test on accuracy

Table 3.9. Mean accuracy: Control vs. Transitioned group on accuracy

		Mean accuracy			
Workload transition		Control	V1	V2	V3
N-back level	0-back	0.806	0.786		
		0.813		0.807	
		0.815			0.794
	1-back	0.729	0.732		
	2-back	0.738	0.779**		
		0.736		0.758	
		0.704			0.687

Note: Means for the variable load conditions with an asterisk are significantly different ($p < 0.05$) form their appropriate controls, which appear to their left on the same line of the table.

This study demonstrated that the effect of hysteresis was more immediately significant when the transition between two workload levels was cyclic rather than ramp-up and down. In particular, the cyclic transitioned groups tended to perform better than their continuous 2-back controls immediately after the transition. However, there were no significant sudden changes in workload profiles.

3.1.3.2. Mental workload

In terms of mental workload (ISA rating), the ANOVA yielded as a significant workload profile main effect, $F(5, 103.2) = 4.290$, $p < .001$, $\eta^2 = .08$, but there was no time block effect.

Table 3.10. Mean and standard deviation of ISA rating

	Mean	**SD**
V1	5.76	0.043
V2	3.56	0.033
V3	5.93	0.046
0-back	3.37	0.040
1-back	5.23	0.037
2-back	6.55	0.029

Table 3.11. ANOVA result on ISA

	Effect	**F value**	***p*-value**	**Eta square**
ISA rating	Workload profile	F(5, 103.2) = 4.290	0.0001**	0.08
	Time block	F(7, 378) = 1.081	0.2845	0.01

The 1-back control group was significantly different from the V1 and V3 groups. Over time, the reported mental workload ratings (ISA) of the transitioned groups approached the non-transitioned groups. We concluded that workload levels of individuals who experience a transition would approach those reported by non-transitioned individuals.

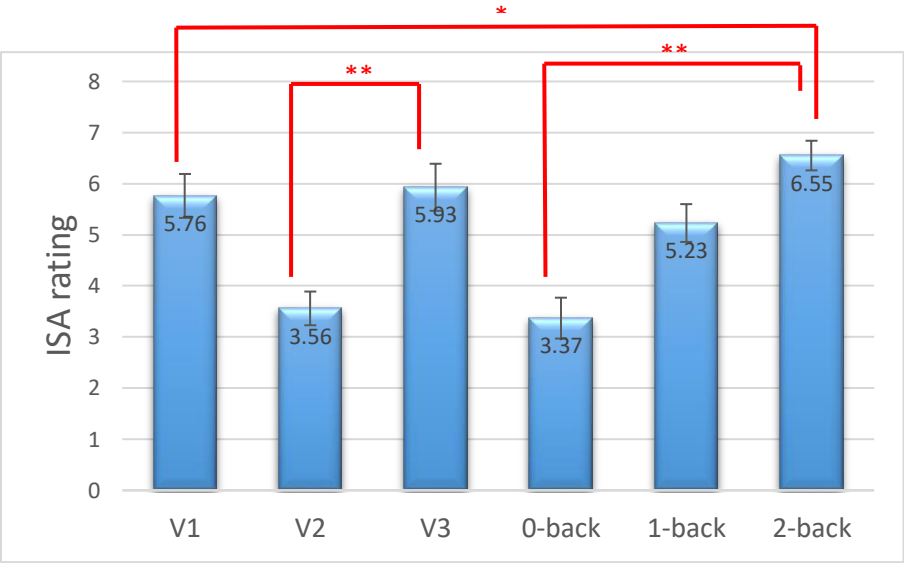


Figure 3.11. Result of Tukey's HSD test on ISA

3.1.3.3. Neurophysiological measures

Effective connectivity analysis

Source localization: On the basis of MVAR modeling (refer to 2.3.1), we intended to calculate the effective connectivity magnitude in both time and frequency domains, thus the Granger causality measures in two domains were introduced in the current study, direct Directed Transfer Function (dDTF) in both time and frequency domain analysis (Kim et al., 2018). In Table 3.12, the cortical regions associated with Brodmann's areas (BAs) were listed. These sources were localized using a single or dual-symmetric equivalent-current dipole model using a four-shell spherical head model co-registered to the subjects' electrode locations by warping the electrode locations to the model head sphere using tools from the EEGLAB DIPFIT plug-in.

Table 3.12. Coordinates of the six independent components with residual variance < 10%

	Talairach coord. (x,y,z)	Location	Lobe	Closest BA	RV (%)
1	16, -96, 6	Cuneus	Occipital	17	8.37
2	64, -37, 4	Middle Temporal Gyrus	Temporal	21	3.33
3	16, -34, 39	Cingulate Gyrus	Motor (Limbic)	5	6.14
4	33, 25, 36	Middle Frontal Gyrus	Frontal	8	2.22
5	48, -57, 5	Temporal Gyrus	Temporal	37	9.19
6	64, -16, 20	Postcentral Gyrus	Parietal	40	1.55

Note: BA: Brodmann Area

We conducted Group average ICA, according to SCCN's group-level analysis pipelines. Having computed the Granger causality measures in both time and frequency domains, we proceeded to construct the effective connectivity graph for each group, respectively. Therefore, we conducted a combination framework of time domain and frequency domain multivariate Granger causality analysis to evaluate the direct causal interactions between time courses.

Time-frequency domain Granger causality analysis: Mean causal information transfer (averaged across all participants) from each IC (column) to all other localized ICs (rows) as measured by the dDTF, is shown in Figure 3.9. Each cell of the matrix shows the time-frequency distribution (1–40 Hz) of information transfer between a respective pair of ICs, (i.e., columns represent the source or FROM and rows represent destination or TO), with the highest information transfer indicated by warm colors.

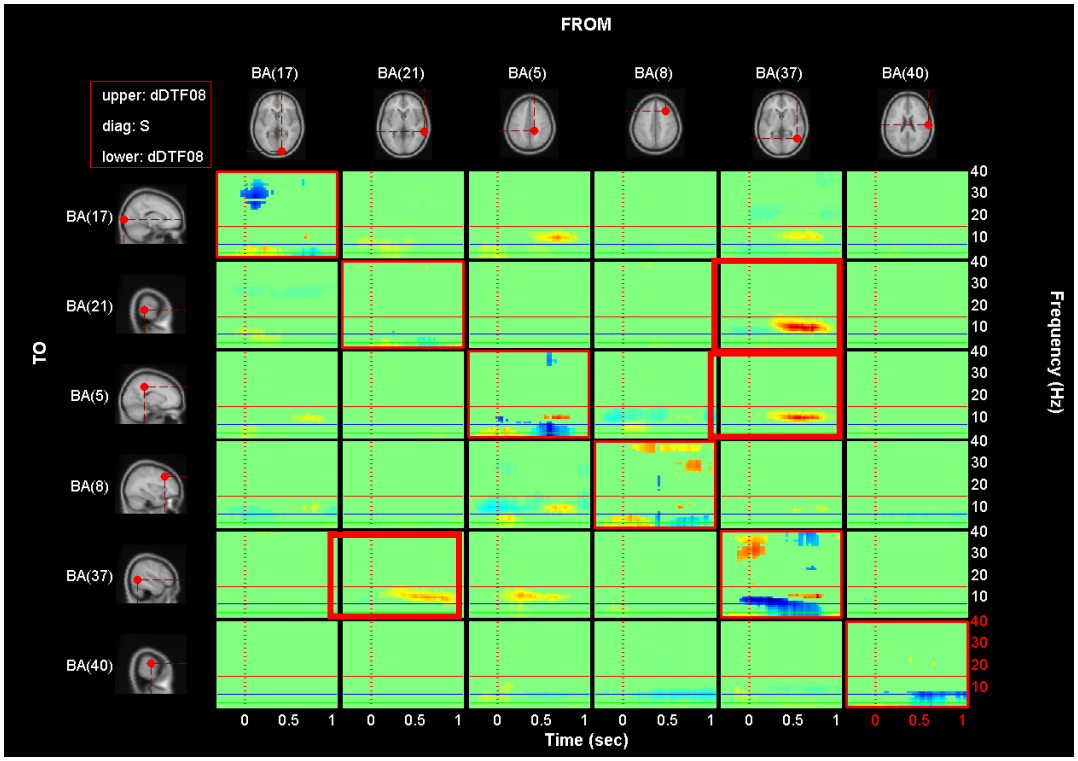


Figure 3.12. Time-Frequency grid

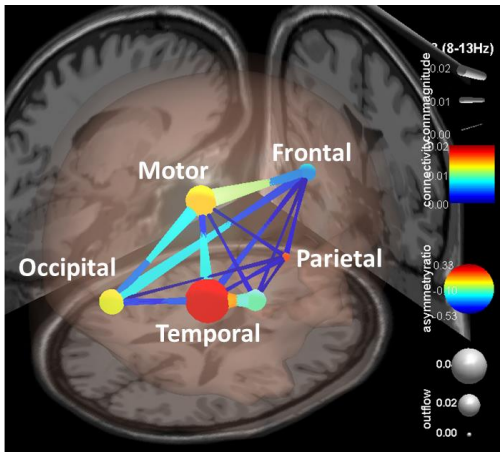
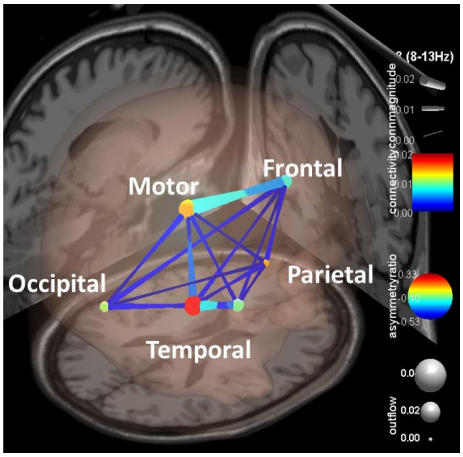
See Figure 3.12. The frequency is on the y-axis and time on the x-axis on each cell of the matrix. The upper and lower triangles of the grid (i.e., above and below the red-bordered diagonal cells, respectively) represent the dDTF (conditional GC) between each pair of sources. Time-Frequency analysis gives information not only of the activation strength of neuronal network structures (i.e., the frequency band power) but also of their timing. The time ranges from -500 ms to 1,000 ms and the dotted vertical lines indicate the n-back stimuli onset and horizontal lines denote green: 3Hz, blue: 7Hz and red: 15 Hz frequencies. On the diagonal, we have plotted the event-related spectral perturbation (ERSP). The most notable feature of the analysis is that the temporal BA(37) and parietal cortex are the strongest drivers of activity between other ICs. The BA(37) ↔ BA(21) and BA(37) ↔ BA(5) are most strongly coupled to each other, but also demonstrate influence on the visual cortex and left and right parietal cortices. The greatest

information flow was seen in the lower frequencies (3–15 Hz), which includes discrete theta and alpha bands with the highest information transfer observed in the alpha band (8–13 Hz).

Evaluating effective connectivity strength: We analyzed information flow between several of these anatomically localized sources of brain activity during trials with correct responses. Each cell image in Table 3.16 shows a transient alpha information flow, and the frames correspond to sessions 2 and 3. It is interesting to note that the alpha rhythm (reflecting synchronized neural activity around the 8-13 Hz range) is associated with memory and cognitive function (Gevins et al., 1997; Tesche & Karhu, 2000). Klimesch et al. (2011) founded the alpha activity in the memory process, and that founding indicated frontal and parietal upper-alpha was involved in the executive function in WM. The images indicated significant different network patterns between transitioned groups and control groups at +500ms from stimulus onset.

See Table 3.13. The color of the node represents the asymmetry ratio of connectivity for that source. The asymmetry ratio indicates whether all connectivity related to that node is inflowing or outflowing. It ranges from -1 to 1. A red color (close to +1) indicates that a node is a causal source, blue (close to -1) means that a node has a role of a causal sink, and green (close to 0) represents a balanced flow. The size of a node represents the amount of information outflow from the source. For the difference between the control and V1 group, the temporal cortex and parietal cortex node are colored red and orange, respectively. This indicates that these two components are the causal source for the whole network. The edge width and color between the temporal cortex and frontal cortex coupling have a sky-blue hue and are thicker than the others. Secondly, similar to behavioral data, under the 2-back condition, there were no significant differences V2 group and V3 group connectivity measures.

Table 3.13. Effective connectivity: V1 (Cyclic profile) group vs. Non-transitioned group

	V1 group	**Non-transitioned group**
Brain network pattern		
From – to	**Temporal – Frontal**	
Outflow	0.198 (0.022)	0.153 (0.013)
Connectivity magnitude	0.079 (0.012)	0.063 (0.008)
Asymmetry ratio	0.796 (0.243)	0.656 (0.127)

Statistical analysis using connectivity measures: Outflow, connectivity magnitude, and causal asymmetry ratio were computed for each IC source. We investigated effective connectivity with a focus on temporo-parietal and temporal-frontal, and their relation to behavioral performance. In the next step, we performed statistical analysis using connectivity measures between frontal and parietal regions were calculated.

In this study, we constructed cortical effective connectivity networks in the EEG IC sources and adopted a graph theoretical framework to analyze the topological variations of the brain network during workload transition compared with the non-transition situation. We revealed that, in the transitioned group (compared with the control task), the alpha-band functional segregation was increased; the network indices of different brain regions in the frontal, temporal and occipital

regions were altered in alpha frequency bands; the reaction time of the subjects was negatively correlated with the average asymmetry ratio and positively correlated with the outflows in the alpha band. Based on this finding, Phase II would let us change the ACT-R model parameters using the topology of the effective brain network and promote understanding of the underlying mechanism of the workload transition effects in ACT-R.

3.2. Phase II: Development of an effective model tuning method in ACT-R

3.2.1. Objectives

EEG (electroencephalograph) provides information about the electrical fluctuations between neurons that characterize brain activity, and measurements of brain activity at resolutions approaching real-time. On the other hand, cognitive architectures such as ACT-R would explain how all the components of the mind work together to generate coherent human cognition. Thus EEG data and ACT-R can provide two aspects to explore the cognitive processes and their neural basis. In phase II, we present a study by combining EEG and ACT-R for investigating workload transition effects and focus on improving the ACT-R model with default parameters based on Phase I outcomes. We focus on behavioral and neural correlates during dual cognitive tasks with or without workload transition. Based on Phase I outcomes, we proposed to systematically update the current ACT-R model to explain workload profile and time effects.

The main goal of Phase II was to validate the proposed method (refer to Fig 3.14) on how to adjust three main ACT-R parameters. Importantly, connectivity information between the main hub nodes and peripheral nodes in the brain (e.g., connectivity magnitude, asymmetry ratio, and causal flow). The study identified three important measures of connectivity between the main hub node and peripheral nodes based on the graph theory – connectivity magnitude, asymmetry ratio

of caudate, and causal flow. We hypothesized the following three new methods to adjust parameters for the ACT-R model based on these three measures. We hypothesized that people, with a strong connection between the cortical hub and frontal/temporal node, better retrieve their memory. We also hypothesized that the time to retrieve memory is proportionate to differences between peak and trough levels of the asymmetry ratio of cortical hub nodes. Lastly, the activation noise is hypothesized to be proportionate to the level of causal flow between the main hub nodes and peripheral nodes, excluding the Parietal-Temporal-Occipital (PTO) network.

3.2.2. Methods

3.2.2.1. ACT-R modeling for n-back task

In designing a dual cognitive ACT-R model, we focused on working memory mechanisms that were limited to the problem state and declarative memory modules. We used the same strategy as Juvina & Taatgen (2007). It perceives the sequence of letters belonging to the n-back task and maintains a problem state (same as a declarative buffer) that contains the previously attended letter and a reference to the n-back letter in declarative memory. Let's assume this model described the 2-back condition. When the model perceived a new letter, the 2-back letter is retrieved from declarative memory and compared against the new letter. The model then gives a response with its hand to press a button for 'target' or 'non-target,' depending on the retrieved letter. If the 2-back letter cannot be retrieved because its activation fell below the retrieval threshold, the model guesses either 'target' or 'non-target' with equal likelihood. Following the response, the problem state is replaced with a new state containing the current letter and a reference to the old problem state. This causes the old state, which contains the 1-back letter, to move into declarative memory (See Figure 3.10). Figure 3.10 shows a timeline of the 2-back model when it perceives a new letter.

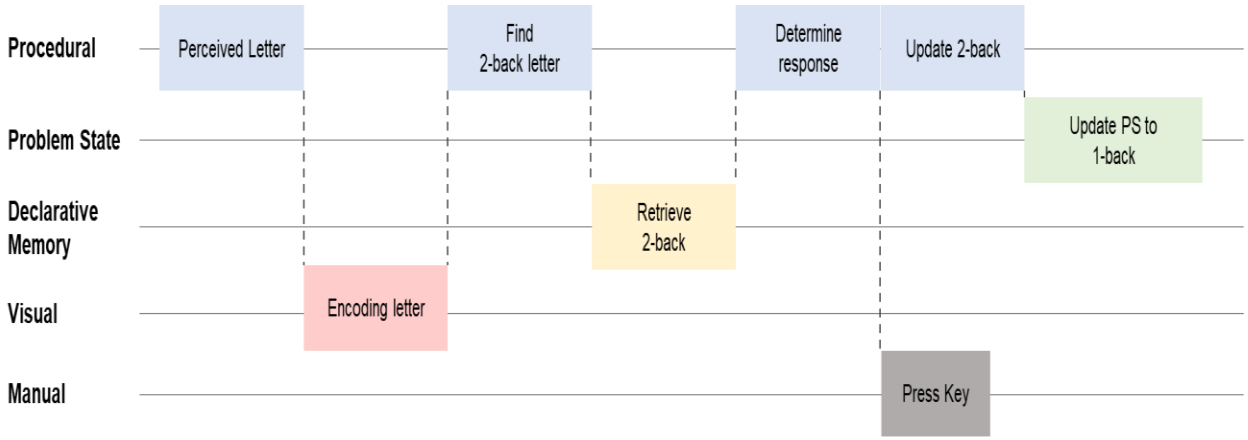


Figure 3.13. Timeline of n-back ACT-R model (revised from Nijboer et al., 2016)

We ran the ACT-R model (Nijboer's n-back model) 250-350 times (corresponding to the number of trials in the dataset) and computed the average activation for different model conditions: correct responses. These average activations, therefore, reflect the probability of a module is active.

3.2.2.2. Procedures

When we compared the *a priori* prediction to the observed data, we hypothesized that the model would not predict the workload transition effect. To accurately capture post-transition performance, model parameters would be adjusted to increase the quantitative model fitness. See Figure 3.14. There are two approaches to improve the ACT-R model.

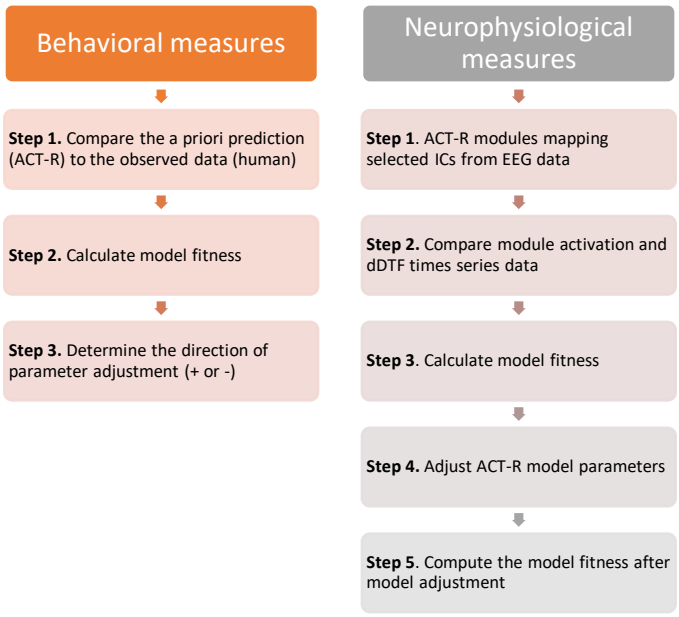


Figure 3.14. Schematic representation of the ACT-R parameter tuning based on behavioral and neurophysiological measures**.**

3.2.2.2.1. Behavioral data-based approach

Step 1. Compare the a priori prediction (ACT-R) to the observed data (human): We compared the simulated data from the ACT-R model with the default setting and Phase I outcomes. Each of the models was run 250-350 times. The data were preprocessed with the n-back ACT-model from Niboer et al. (2016) and then analyzed with Microsoft Excel. The model data and the empirical data were divided into each 2-min time bins, but we only considered sessions 2 and 3. The average proportion of correct responses and the standard deviation per block was computed for the experiment as well as for each of the six workload profiles.

Step 2. Calculate model fitness: We need to calculate model fitness between human and model by calculating R^2 and Root Mean Square Error (RMSE).

Step 3. Determine the direction of parameter adjustment (+/-): Based on the behavioral data-based approach, we could determine the direction of parameter adjustment (+/-). For example,

empirical data shows better/worse performance than the model, and we need to decrease/increase the memory retrieval threshold.

3.2.2.2.2. Neural data-based model fitting

Step 1. ACT-R modules mapping selected ICs from EEG data: See Figure 3.15. In order to compare the EEG data to the ACT-R model, Independent Components (ICs) derived from EEG data were linked to ACT-R buffer activation using dipole fitting. Prezenski & Russwinkel (2016) demonstrated that EEG data could be used to validate ACT-R models. From the results of Phase I, the six selected ICs present below (See Table 3.14).

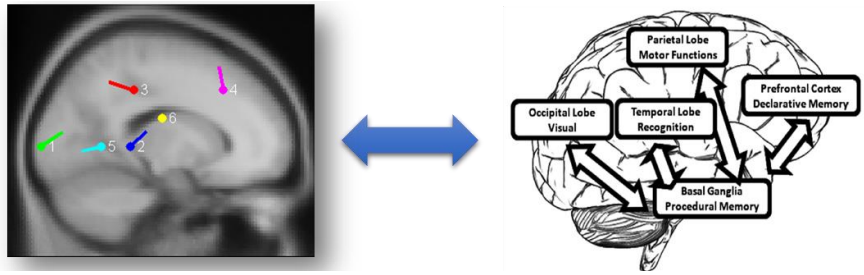


Figure 3.15. ACT-R modules and selected ICs

Table 3.14. Selected ICs and ACT-R module matching

EEG		Brain Region	**ACT-R**	
Talairach Coord. (x,y,z)	IC		Module	Function
16, -96, 6	BA(17)	Occipital	Visual Module	Activated by the displayed n-back stimuli
64, -16, 20	BA(40)	Parietal	Visual Buffer	Place a representation of the letter stimulus in the visual buffer
64, -37, 4/ 48, -57, 5	BA(37 & 21)	Temporal	Declarative Module	Retrieve the instruction about what to do for that target letters within each condition from declarative memory
33, 25, 36	BA(8)	Frontal	Problem state	Placed in the declarative memory buffer

An ACT-R model of the process would, at minimum, predict that the visual module (occipital) would be activated by the displayed n-back stimuli and would place a representation of the letter stimulus in the visual buffer (parietal). Next, the "parietal" representation would be used to retrieve the instruction about what to do for that target letters within each condition from declarative memory (temporal), which in turn would be placed in the problem state (declarative memory buffer) (frontal).

Step 2. Compare module activation and dDTF: If the ICs match specific ACT-R modules, the timing of peaks of buffer activity should match IC-peaks of information flow. In this study, IC-peaks timing of brain connectivity between ICs was observed. The mean causal information flow, as measured by the dDTF, was analyzed between the localized cortical sources at each task difficulty level at the alpha band (8-13Hz). The frequency band was based on Time-Frequency causal flow analysis. In neural data approaches, we selected the best fitting regions corresponding to the problem state and declarative memory, which are part of the fronto-parietal, temporo-parietal network that is frequently associated with working memory, cognitive control, and attentional selection (Cole & Schneider, 2007; Dosenbach et al., 2007). From the PTO network, we got significant graphic measures such as outflow, and causal flow values would let us know how to match scale parameters that reflect the amount of weighting given to the information flow in the retrieval specification. To analyze the effects in the predefined regions, we used source localization according to the outcome of Phase I. The causal flow was then calculated as percent signal change as compared to each trial. We only took the area between -500 ms and +1000ms (0 time means stimuli onset). ACT-R's modules are not constantly in use during the execution of a model but operate for short periods of time (in the order of hundreds of ms). The assumption is that when a module is active, the causal flow increases in the associated brain region.

The output of an ACT-R simulation is a time course for when and how long the activations of each module involved in the task. If D(t) is a 0–1 demand function that indicates whether a module is active at time t, the causality at time t can be calculated by convolving D(t). We report only significant results concerning those ICs of interest for the correlated modules.

Step 3. Calculate model fitness: We need to calculate model fitness between human and model by calculating R^2 and Root Mean Square Error (RMSE).

Step 4. Adjust ACT-R model parameters: After model fitting, we need to decide which parameters would be adjusted. In this study, a new set of ACT-R parameters was assigned for memory retrieval function based on literature. Anderson et al. (1998) demonstrated that there are systematic variations between the retrieval threshold and latency-scaling factor across studies. Based on Wong et al. (2012)'s meta-analysis, the three most frequently modified were retrieval threshold, module activation noise, and latency-scaling factor.

- **Retrieval threshold:** the minimal amount of activation required to retrieve a chunk.
- **Activation noise**: the amount of random activation added or subtracted to individual chunk activations.
- **Latency-scaling factor:** it scales the retrieval times of chunks based on their activation.

As mentioned earlier, the current study focuses on memory retrieval and their buffers' causality flows. Tsai et al. (2019) found that the fronto-parietal network is indeed crucial for successful memory retrieval. We utilized these findings to construct a value for adjustment of the ACT-R declarative module. If an object or annotation chunk falls below this threshold, the chunk cannot be retrieved anymore. ACT-Rs chunk activation is calculated as a function of the chunk's creation time and the number (and points in time) of references of the chunk and the spreading

activation of related chunks. Therefore, those chunks will probably fall below the retrieval threshold that is retrieved the less and are far away from the focus.

In this study, we constructed cortical effective connectivity networks in the EEG IC sources and adopted a graph theoretical framework to analyze the topological variations of the brain network during workload transition compared with the non-transition situation. Based on the results of Phase I, below parameter mapping formulas were examined in this section. Models were run in each variation group's condition with these updated parameters:

(a) Connectivity magnitude (Temporal – frontal): Retrieval threshold

$$\tau \propto \frac{\sum dDTF\ from\ Temporal\ to\ frontal}{Total\ dDTF\ in\ network}$$

(b) Asymmetry ratio of Temporal:

$$F \propto Time_{Max\frac{Inflow_{temporal}-Outflow_{temporal}}{Inflow_{temporal}+Outflow_{temporal}}} - Time_{Min\frac{Inflow_{temporal}-Outflow_{temporal}}{Inflow_{temporal}+Outflow_{temporal}}}$$

(c) Causal flow (temporal -> parietal): Activation noise

$$\varepsilon \propto \sum_{i=1}^{10} WMC_i(Outflow_{temporal} - Inflow_{temporal})$$

Step 5. Compute the model fitness after model adjustment: To validate the proposed parameter mapping method in this study, model fitness between human and updated model was calculated using R^2 and Root Mean Square Error (RMSE).

3.2.3. Results

3.2.3.1. Behavioral measures

When we compare a pre-model prediction to the human performance data across the non-transitioned and transitioned group, we found that the prediction has a good qualitative fit to the data, taking into account only working memory load within the non-transitioned group. However, as we expected, the pattern that was not predicted was that the performance of some of the transitioned group (See Table 3.15).

Table 3.15. Cognitive dual-task behavioral fit for all relevant models. (Pre-model vs. data)

Workload profile	RT		Accuracy	
	R^2	RMSE	R^2	RMSE
0-back	0.886	0.15	0.893	0.12
1-back	0.878	0.13	0.828	0.08
2-back	0.852	0.20	0.893	0.17
V1	0.693	0.36	0.782	0.24
V2	0.785	0.14	0.838	0.17
V3	0.795	0.25	0.842	0.13

Note. RMSE: Root Mean Square Error

Pre-model made fewer mistakes as compared to V2 and V3 groups but more mistakes than the V1 group. The Pre-model performed slightly worse than the V1 participants did. In order to estimate what changes were required to the Pre-model 1 to have better fit the transitioned group data, we need to make an adjustment on parameters that can result in Post-model 1. To accurately capture Session 2 and 3's V1 performances, we could consider the possibility of hysteresis effects on them based on Phase I outcomes. Therefore, a new production rule has to update a lower retrieval threshold, which determines the minimal amount of activation required to retrieve a chunk.

The latency-scaling factor also should be altered because it scales the retrieval times of chunks based on their activation.

3.2.3.2. Neurophysiological measures

In Table 3.16, the EEG predictions of the model are presented. The numbers represent the total causal flows of each transitioned group of the problem state module and declarative memory. It should be noted that the scale of the model results is not meaningful: the ICs' causality time series data is a relative measure, and thus, we can only compare the activation patterns. Partial correlations of the EEG data and the connectivity values derived from the ACT-R model revealed a negative association between temporal and parietal nodes for the high load (2-back) condition ($p = 0.003$).

Table 3.16. Fit measures of the neural data between Model and data from Phase I

Workload profile	Problem State (Frontal)		Declarative module (Temporal)	
	R^2	RMSE	R^2	RMSE
0-back	0.734	0.34	0.723	0.42
1-back	0.783	0.23	0.764	0.27
2-back	0.803	0.33	0.792	0.32
V1	0.720	0.38	0.689	0.44
V2	0.773	0.27	0.782	0.38
V3	0.789	0.30	0.829	0.22

Note. RMSE: Root Mean Square Error

The problem state of simulation results shows that the V1 group had the lowest activity for all conditions involving 2-back, instead of the highest. The prediction for declarative memory represented the data marginally better ever for V2 and V3 groups. The low amount of predicted activity during gradually increase workload transition occurred did not fit with the temporal node information flow. In contrast to the model, the EEG data indicated that the problem state was used

less in the transitioned group than control groups, but the declarative module was used more for the V1 group than the corresponding control. The data indicate that participants use declarative memory more after experiencing a gradual workload transition than our model predicted. Taken together, this indicates that participants used a different memory strategy to keep track of the previous letter stimuli than the model predicted.

3.2.3.3. Adjustment of parameters

We tested the hypothesis that the number of causal flows of a temporal node could predict the degree of information flow for memory retrieval processing, as indexed by the sum of outgoing interactions in the alpha-band. The indices were judged on the basis of their relative strength in exhibiting expected causality flows within temporal-parietal-frontal pathways. In Table 3.17, the proposed parameters were listed.

Table 3.17. Proposed a new set of parameters

Parameter	Description	Default Value	Adjusted Value
F	Retrieval latency factor	0.1	0.13
$\boldsymbol{\tau}$	Retrieval duration threshold	0.1	0.037
$\boldsymbol{\varepsilon}$	Moment-to moment noise	0.1	0.035

To validate the proposed set of parameters in this study, model fitness between humans and model was calculated using R^2 and Root Mean Square Error (RMSE). See Table 3.18.

Table 3.18. Fit measures between behavioral data and updated ACT-R model.

n-back Performance		Before tuning		After tuning	
		R^2	RMSE	R^2	RMSE
RT	**V1**	0.693	0.36	0.936	0.10
	V2	0.785	0.14	0.882	0.04
	V3	0.795	0.25	0.885	0.12
Accuracy	**V1**	0.782	0.24	0.913	0.18
	V2	0.838	0.17	0.872	0.11
	V3	0.842	0.13	0.869	0.02

Note. RMSE: Root Mean Square Error

Table 3.19. Fit measures between neural data and updated ACT-R model.

Module Activation		Before tuning		After tuning	
		R^2	RMSE	R^2	RMSE
Problem State	**V1**	0.720	0.38	0.854	0.12
	V2	0.773	0.27	0.897	0.09
	V3	0.789	0.30	0.889	0.13
Declarative Memory	**V1**	0.689	0.44	0.879	0.20
	V2	0.782	0.38	0.890	0.13
	V3	0.829	0.22	0.892	0.06

Note. RMSE: Root Mean Square Error

Consequently, we modified the parameters based on effective connectivity, which was able to fit both behavioral and EEG data (See Table 3.19). Thus, the EEG data and effective connectivity proved to be a crucial tool to evaluate the plausibility of ACT-R models.

3.3. Discussions of Study 1

In summary, during Phase 1, the cyclic variation group (V1) showed a better performance than the control group during post workload transition periods, which confirmed our hypothesis. However, abrupt changes in workload (V2 and V3) could have a negative impact on performance, which was consistent with other studies. Interestingly, only the V1 group showed significant differences in neutral data compared to the control group. The V1 group, who experienced gradual changes in task difficulty, exhibited continued activation of the temporal cortex, which functioned as a causal

source within the brain network, throughout the 2-back condition (high task demand). Multiple studies have shown that the temporal cortex is related to working memory and cognition. Specifically, recent studies highlighted that the temporal cortex plays a crucial role in retrieving the memory. (Ranganath et al., 2004; Dolcos et al., 2005; Tsai et al., 2019).

Based on the results from Phase I, Phase II adjusted the parameters of the ACT-R model to capture the workload transition effects, which was not detected by the existing ACT-R model with the default settings. Previous studies tried to find optimal parameters through trial and error. Unlike the previous studies, this study utilized systematic methods, which are based on effective connective measures from neural data, to adjust the parameters and evaluate model fitness. As a result, the model fitness improved, with both behavioral data and neural data fitting, showing higher model fitness than the criteria (> 0.8). The major finding of Study 1 is the development of a new method to adjust the parameters of the ACT-R model by connecting the ACT-R module to the cortical node which functions as a main hub within the brain network. We could conclude that the new method is more streamlined and efficient than adjusting parameters through trial and error. Study 2 focused on 1) how workload transitions effects materialize and 2) how the new method performs under more complex tasks.

4. STUDY 2: ACT-R modeling of workload transition in AF-MATB tasks

Similar to study 1, study 2 has two primary goals: 1) to confirm the hysteresis effect and 2) to validate the proposed method from the study 1 on how to adjust major parameters in ACT-R. Study 2 involves a more ecologically valid interface instead of a simple cognitive task. Figure 4.1 below outlines the two phases of study 2, with each phase to achieve the aforementioned goals. By incorporating neural correlates and ACT-R production rule processing, the current study suggested a novel way to adjust three significant parameters of ACT-R - retrieval threshold, latency-scaling factor, and activation noise. These parameters are essential building blocks of a substantial ACT-R architecture and are associated with human memory store and retrieval processing mechanisms. This study validated that the simulated data with adjusted parameters exhibit a better fit with human data than default parameters.

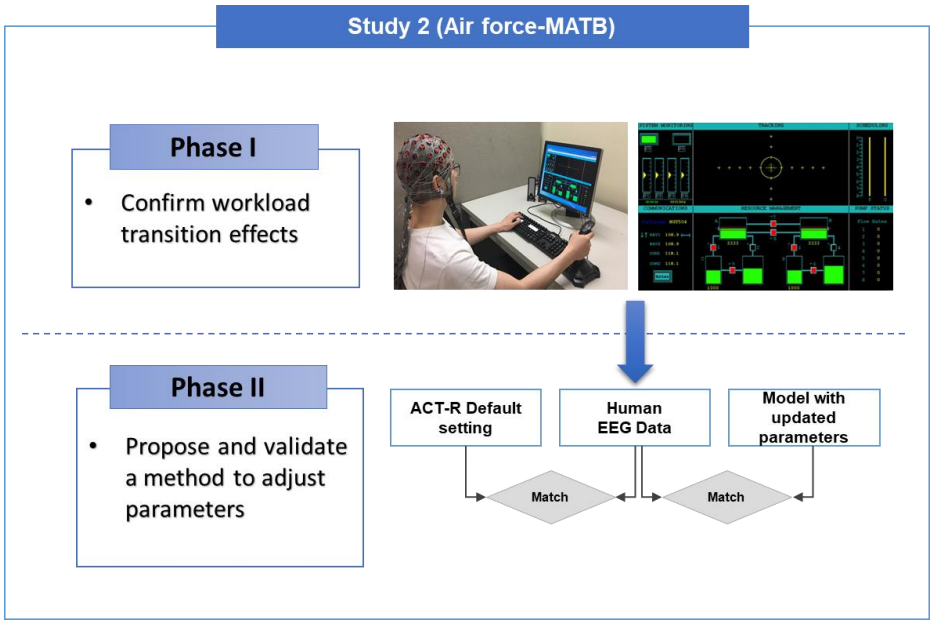


Figure 4.1. Framework of Study 2 with two phases

4.1. Phase I: Workload transition effects in multitasking

4.1.1. Objectives and hypotheses

The main goal of Phase I in Study 2 is to systematically investigate how a hysteresis effect can impact human performance and cognition in a multitasking environment. In Phase I tested the effect of task demand transitions at multiple levels of analysis including behavioral performance, subjective rating, and brain effective connectivity by using Air Force Multi-Attribute Task Battery (AF-MATB) task which is an ecologically validated platform for multitasking research. In phase I, three sets of hypotheses were tested for each transition profile (Refer to Figure 3.2).

Performance hypotheses.

H4.1: The post-transition performance of a V1 group (cyclic variation) should be superior to that of a non-transitioned group.

H4.1 (a): Participants' averaged correct response ratio for System monitoring would be higher under the cyclic workload profiles than under the non-transitioned conditions.

H4.1 (b): Participants' averaged RMSD of Tracking would be lower under the cyclic workload profiles than under the non-transitioned conditions.

H4.1 (c): Participants' averaged reaction time of Communication would be shorter under the cyclic workload profiles than under the non-transitioned conditions.

H4.2: The post-transition performance of a V2 group (Ramp-up) should be inferior to that of a non-transitioned group.

H4.2 (a): Participants' averaged correct response ratio for System monitoring would be lower under the ramp-up workload profiles than under the non-transitioned conditions.

H4.2 (b): Participants' averaged RMSD of Tracking would be higher under the ramp-up workload profiles than under the non-transitioned conditions.

H4.2 (c): Participants' averaged reaction time of Communication would be longer under the ramp-up workload profiles than under the non-transitioned conditions.

H4.3: The post-transition performance of a V3 group (Ramp-down) should be inferior to that of a non-transitioned group.

H4.3 (a): Participants' averaged correct response ratio for System monitoring would be lower under the ramp-down workload profiles than under the non-transitioned conditions.

H4.3 (b): Participants' averaged RMSD of Tracking would be higher under the ramp-down workload profiles than under the non-transitioned conditions.

H4.3 (c): Participants' averaged reaction time of Communication would be longer under the ramp-down workload profiles than under the non-transitioned conditions.

Mental workload hypotheses.

H4.4: The post-transition mental workload of a V1 group (cyclic variation) should be lower than a non-transitioned group.

H4.4 (a): Participant's averaged NASA-TLX composite score would be lower under the cyclic workload profiles than under the non-transitioned conditions.

H4.5: The post-transition mental workload of a V2 group (Ramp-up) should be higher than a non-transitioned group.

H4.5 (a): Participant's averaged NASA-TLX composite score under the ramp-up workload profiles than under the non-transitioned conditions.

H4.6: The post-transition mental workload of a V3 group (Ramp-down) should be higher than a non-transitioned group.

H4.6 (a): Participant's averaged NASA-TLX composite score would be higher under the ramp-up workload profiles than under the non-transitioned conditions.

Neurophysiological hypotheses.

H4.7: The post-transition information flow of the cortical hub node of a V1 group (cyclic variation) should be stronger than a non-transitioned group.

H4.7 (a): Participants' averaged connectivity magnitude of the hub node would be higher under the cyclic workload profiles than under the non-transitioned conditions.

H4.7 (b): Participants' averaged outflow of the hub node would be higher under the cyclic workload profiles than under the non-transitioned conditions.

H4.7 (c): Participant's averaged asymmetry ratio of the hub node would be higher under the cyclic workload profiles than under the non-transitioned conditions.

H4.8: The post-transition information flow of the cortical hub node of a V2 group (ramp-up) should be weaker than a non-transitioned group.

H4.8 (a): Participants' averaged connectivity magnitude of the hub node would be lower under the ramp-up workload profiles than under the non-transitioned conditions.

H4.8 (b): Participants' averaged outflow of the hub node would be lower under the ramp-up workload profiles than under the non-transitioned conditions.

H4.8 (c): Participants ' averaged asymmetry ratio of the hub node would be lower under the ramp-up workload profiles than under the non-transitioned conditions.

H4.9: The post-transition information flow of the cortical hub node of a V3 group (ramp-down) should be weaker than a non-transitioned group.

H4.9 (a): Participants' averaged connectivity magnitude of the hub node would be lower under the ramp-down workload profiles than under the non-transitioned conditions.

H4.9 (b): Participants' averaged outflow of the hub node would be lower under the ramp-down workload profiles than under the non-transitioned conditions.

H4.9 (c): Participant's averaged asymmetry ratio of the hub node would be lower under the ramp-down workload profiles than under the non-transitioned conditions.

4.1.2. Methods

4.1.2.1. Participants

The same participants who completed all sessions in Study 1 were invited to the experiment in Study 2. After two participants were recruited, data analysis was performed to adjust the sample size. JMP was used to calculate the sample size. The statistical power of the model was set up at a conventional power level, 0.8 (Cohen, 1992; Thomas & Juanes, 1996).

4.1.2.2. Experimental task

AF-MATB: Participants were trained on the Air Force Multi-Attribute Task Battery (AF-MATB, Miller, 2010) (see Figure 4.2(b)), a computer-based multitasking environment designed to evaluate operator performance and mental workload while human operators perform a benchmark set of tasks similar to activities that aircraft crew members perform in flight (i.e., system monitoring, resource management, communications, and tracking tasks) (Comstock & Arnegard, 1992).

The MATB was developed by the National Aeronautics and Space Administration (NASA) to evaluate human performance in a multitasking environment. The AF-MATB requires the human operator to simultaneously monitor and respond to four independent tasks on one computer screen. AF-MATB is a research paradigm where participants perform a tracking task while concurrently monitoring warning lights and dials, responding to computer-generated auditory requests to adjust radio frequencies, and managing simulated fuel flow rates using various key presses. The current version of AF-MATB has serial and digital port-triggering functions, allowing the task to interface with neurophysiological acquisition systems. This change facilitated time-syncing between AF-MATB and acquired neurophysiological data and aided in the integration of state-based information into acquired data.

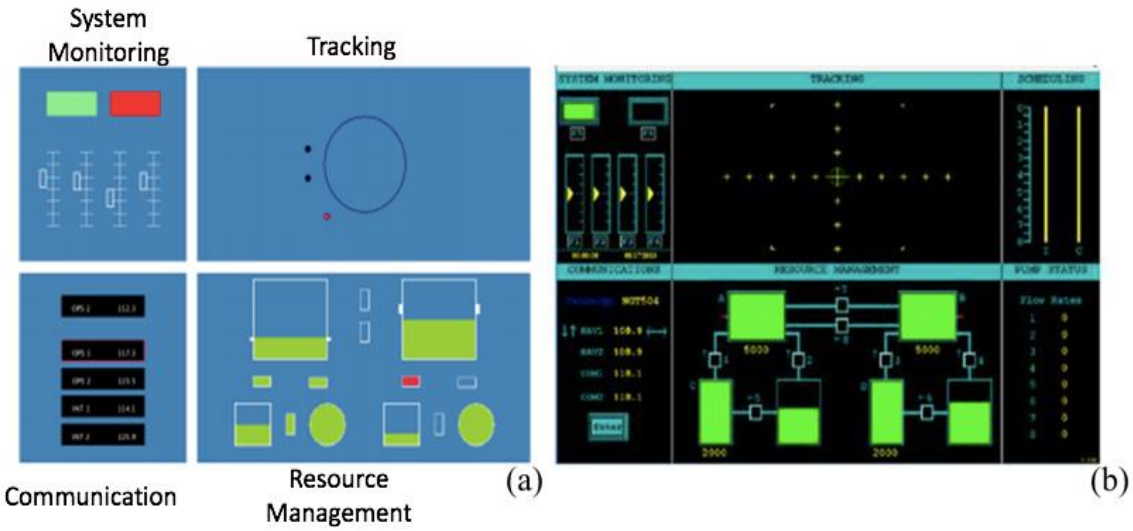


Figure 4.2. Experimental task interface (a) mMATB (b) AF-MATB

The tasks consist of Systems Monitoring, Communication, Targeting, and Resource Management. For this study, each of the four tasks was equally weighted; therefore no task had greater importance than another task.

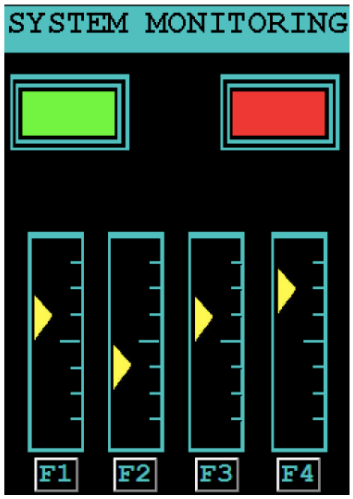


Figure 4.3. System Monitoring task

The System Monitoring task is located in the top left corner of the MATB window and consists of two subtasks: lights and dials. See Figure 4.3. The two rectangles at the top represent the lights. The participant is asked to keep the left light in the on status “displaying green” and the right light in the off status “displaying black.” If the lights switched from these initial conditions, selecting F5 or F6 keys reset the lights. Beneath the lights are four vertical columns which represent dials. Throughout the task, yellow markers within the dials continuously oscillated one location above and below the center of the dial. Occasionally, the yellow marker shifted toward the top or the bottom of the dial and began oscillating around a new location. When this event occurred, participants were to select the corresponding F1–F4 keys to reset the dials.

Figure 4.4. Communication task

The Communication task is located in the bottom left corner of the MATB window. See Figure 4.4 The objective of the communications task is to alter the channel and frequency in response to an auditory cue. An audible message instructed the participants to modify a specific communication channel to a given frequency. The participants navigate to the appropriate channel and set the frequency by selecting the up, down, left, and right arrow keys. Participants only have to attend to communications when their "callsigns" were addressed directly, as extraneous auditory information also occurred for different callsigns.

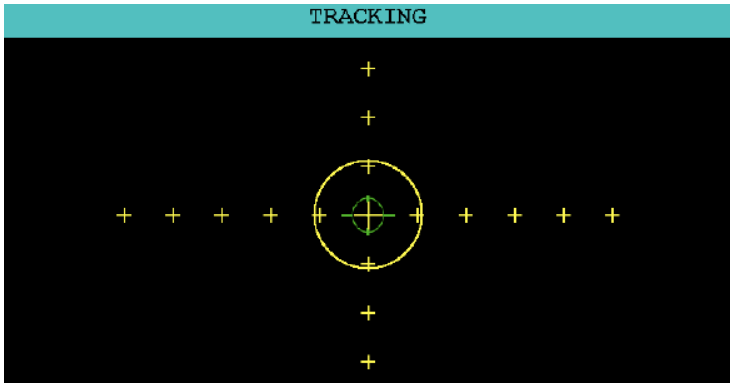


Figure 4.5. Tracking task

The Tracking task is located in the top right corner of the MATB window. See Figure 4.5. Throughout the task, the green cursor drifts around the window. The objective is to maintain the green cursor within a larger yellow circle found in the center of the window by using a joystick.

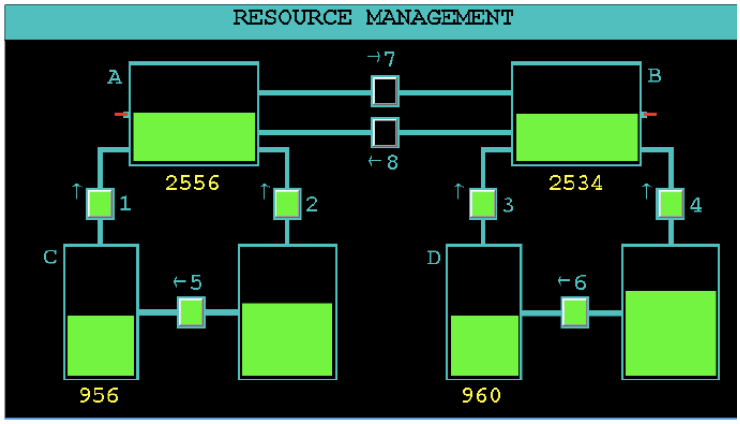


Figure 4.6. Resource management task

The Resource Management task is located in the bottom right corner of the MATB window. See Figure 4.6. The objective of the Resource Management task is to maintain a fluid level within a specific range in two primary tanks. This was accomplished by turning "on" and "off" four reservoir tanks by using number keys. It is important to note that the fluid level is continuously flowing, and therefore must be monitored continuously. Flow is sometimes hindered, as flow between tanks can be disrupted without input by the participant, forcing the participant to adjust flow between the tanks. The performance in each subtask is individually scored.

4.1.2.3. Experimental design and independent variables

The independent variable is the difficulty level and is manipulated by the event rate in each of the subtasks. Table 4.1 was used as a reference in our research and is contained in Bowers' study

(Bowers, 2013). For the tracking component, difficulty was manipulated by adjusting speed and number of directional changes.

Table 4.1. Event rates in each level of task difficulty

	Sequence Length		Communications		Tracking	System Monitoring		Fuel Management	
	Seconds	Minutes	Target	Distractor	Difficulty	Lights	Gauges	Failures	Shut-offs
Easy	120	2	2	2	Easy	6	3	3	3
Medium	120	2	3	2	Medium	7	4	4	4
Hard	120	2	4	2	Hard	8	4	5	5

The current study identifies the distinct three difficulty levels of AF-MATB (Easy, Medium, and Hard). Refer to Table 4.2. Each block consisted of one difficulty level. There were six experimental conditions in all, with 10 subjects allocated randomly to each group. The experimental design follows Study 1.

The workload was manipulated by three difficulty levels (Easy, Medium, and Hard) that represent varying levels of demand on human operators during multitasking. The variations can be categorized into two groups, experimental and control. First, the experimental group was created with three types of difficulty level variations during a session: 1) **Cyclic** variation (V1): a cyclic variation of the load across successive trial blocks (e.g., E $\rightarrow$ M $\rightarrow$ H $\rightarrow$ H $\rightarrow$ M $\rightarrow$ E), 2) **Ramp-up** variation (V2): a significant increase in the levels of workload from the lowest (Easy) to the highest level (Hard) on the fourth trial block during each session, and 3) **Ramp-down** variation (V3): a significant decrease in the levels of workload from the highest (Hard) to the lowest level (Easy) on the fourth trial block during each session, which is the reversed V2 pattern. Three control groups performed the entire session at a single load level (i.e., levels E for 'Easy', M for 'Medium', H for 'Hard').

Table.4.2. The sequence of the difficulty levels by six variation groups during sessions

		Session 1				Session 2				Session 3			
Block		1	2	3	4	5	6	7	8	9	10	11	12
Experimental Variation	V1	E	M	M	H	H	M	M	E	E	M	M	H
	V2	E	E	E	H	E	E	E	H	E	E	E	H
	V3	H	H	H	E	H	H	H	E	H	H	H	E
Control Variation	C1	E	E	E	E	H	H	H	H	M	M	M	M
	C2	M	M	M	M	E	E	E	E	H	H	H	H
	C3	H	H	H	H	M	M	M	M	E	E	E	E

4.1.2.4. Dependent variables

This study measured several data at behavioral, neurophysiological and neural connectivity levels to investigate workload transition effects, respectively: (1) task performance, (2) mental workload, and (3) brain activity.

(1) **Task performance**

Within the AF-MATB system, there are several performance measures. (Bower et al., 2014).

- **System monitoring task**: We recorded an average correct reaction time (in seconds), and mean percent correct responses. Correct responses occurred when participants pressed the appropriate button to respond to an event in the System Monitoring task.
- **Tracking task**: Root Mean square (RMS) was measured. RMS is the distance between the moving cursor and the center point of the Tracking task.

- **Resource management task**: In this task, the mean deviation from the visible target value (2500 units of fuel) was calculated to assess performance.
- **Communication task**: Performance for the Communications task was assessed in terms of percent correct responses (hits). In this task, correct responses occurred when the participants follow the audio directions and switch the radio to the correct channel and frequency. The mean reaction time of participants' correct responses measured in seconds.

(2) Mental workload

The mental workload was measured by NASA-TLX that includes five components: mental demand (MD), physical demand (PD), temporal demand (TD), and the individual's perceived level of performance (PE), effort (EF), and frustration (FR) (Hart and Staveland, 1988). The NASA TLX was administered at the end of each session, and participants were asked to complete it with respect to the workload they experienced during the session. TLX composite scores were calculated by computing the average of the scores for the six sub-scales for each session (Bower, et al., 2014).

(3) Brain activity

Independent Component Analysis (ICA) was performed to remove artifacts, and all retained ICs were localized using DIPFIT. The Source Information Flow Toolbox (SIFT) for EEGLAB was used to evaluate effective connectivity, the causal flow of information between brain sources (Delorme et al., 2011). For this analysis, please refer to section 3.1.2.4 and 3.1.2.5. After the connectivity evaluation, brain networks' measures were calculated, such as inflow/outflow, causal flow, and asymmetry ratio.

4.1.2.5. Procedure

At the beginning of the study, participants performed a practice session under each static-load condition before the actual AF-MATB experiment begins. The practice session was about less than 20 minutes long, and it was repeated until participants reached an accuracy of at least 65 percent correct responses. We used NASA TLX to assess mental workload during each session. The total experiment, including EEG preparation, task training, task run, and breaks took about 1.5 hours. Figure 4.7 presents an experimental set-up that housed an EEG acquisition system and a computer system that displays an AF-MATB.

Figure 4.7. Experimental set-up in Study 2

4.1.3. Results and discussions

4.1.3.1. Behavioral measures

Performance on each of the four subtasks was analyzed separately. The repeated-measure ANOVA test was conducted to evaluate the main effect of IVs, such as the workload profile and time block effects, the same as Study 1. Nevertheless, the continuous Resource Management task failed to yield evidence of any transition effect.

There was no missing data, and for a whole set of analyses, performance for each block was calculated. The two parametric assumptions of analysis of variance were assessed in JMP® (SAS Institute Inc., Cary, NC) before conducting ANOVA. The data were checked for skewness, kurtosis, normality, and homogeneity of variance prior to analysis.

System monitoring task: The System Monitoring task data were based on responses to malfunctions with the gauges and lights and were analyzed in terms of percent correct responses and reaction time, which were measured as a function of total responses. Correct responses occurred when participants pressed the appropriate button to respond to an event in the System Monitoring task. The mean correct response ratio (CRR) was calculated and plotted in Figure 4.8.

We found the main effects of workload profile, $F(5, 103.2) = 5.72$, $p = .0002$, $\eta^2 = .07$. This result is generally consistent with the pattern observed previously in Study 1. The paired t-tests were used to determine which conditions had significantly different means from the others. The result showed that the hard control group was significantly different from V1 and V2 groups. Interestingly, the V2 group was significantly different from the easy control group.

Table 4.3. Mean and standard deviation of CRRs

	Mean	SD
V1	0.654	0.023
V2	0.704	0.012
V3	0.588	0.022
0-back	0.735	0.017
1-back	0.667	0.011
2-back	0.609	0.020

Table 4.4. Results of the ANOVA on CRR

	Effect	F value	*p* levels	Eta square
CRR In SM	Workload profile	F(5, 103.2) = 5.72	0.0002	0.07
	Time block	F(7, 378) = 7.88	0.0001**	0.10

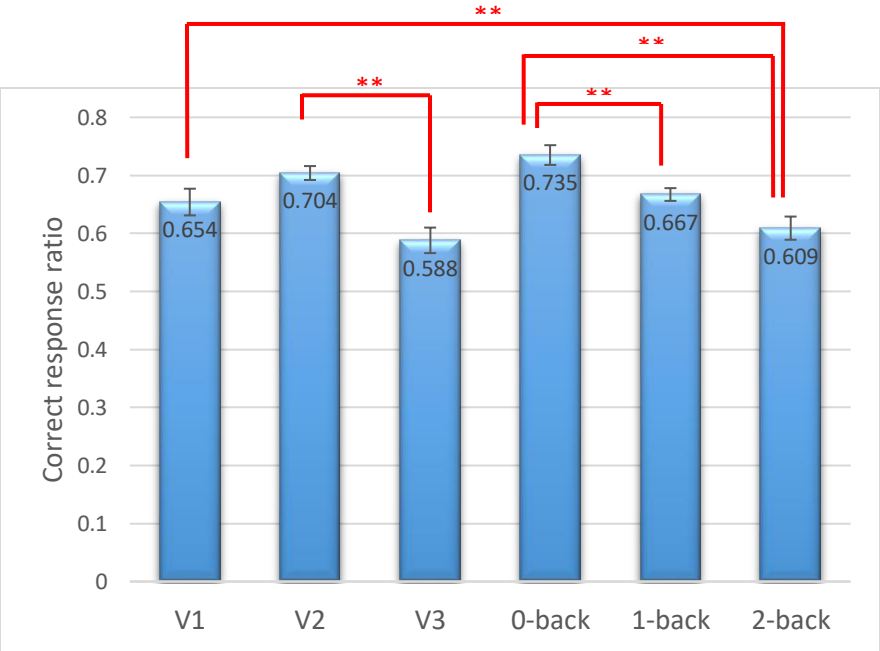


Figure 4.8. Result of Tukey's HSD test on CRR

The main effect of time block was also found, $F(7, 378) = 7.88$, $p < 0.001$, $\eta^2 = .10$. For further analysis, a simple t-test was performed to test the significance of the differences between

the control groups and variation groups with the same n-back level (Table 4.5). The result of the paired t-test showed that the average CRR of the V1 (0.649; p = 0.003) was significantly higher than the constant hard control group (0.603). Therefore, this result supports the H4.1 (a) that the cyclic workload profile affects performance. Under easy condition, the mean CRR of V2 group (0.699, p < 0.001) were significantly lower than non-transitioned group (0.742). This result supports the position H 4.2 (a), which indicates that the ramp-up workload profile negatively affects performance.

Table 4.5. Mean correct response ratio: Control vs. Transitioned group on CRR

		Mean correct response ratio			
Workload transition		Control	V1	V2	V3
System Monitoring	Easy	0.732	0.719		
		0.742		0.699*	
		0.727			0.703
	Medium	0.652	0.628		
	Hard	0.603	0.649*		
		0.598		0.601	
		0.583			0.592

Note: Means for the variable load conditions with an asterisk are significantly different (**; p <0.001, *; p<0.05) form their appropriate controls, which appear to their left on the same line of the table.

Tracking task: The success in the task was measured by the root mean square distance (RMSD), which is the mean deviation of the cursor from the center. The ANOVA revealed the effects of workload profiles on participants' RMSD in the tracking task. The main effects of workload profile, $F(5, 103.2) = 7.23$, $p < .001$, $\eta^2 = .22$.

Table 4.6. Mean and standard deviation of RMSDs

	Mean	**SD**
V1	77.54	3.23
V2	67.77	5.23
V3	73.64	4.43
0-back	60.54	6.32
1-back	86.04	2.23
2-back	88.75	4.30

Tukey's HSD test showed that the hard control group was significantly different from V1 and V2 groups. See Table 4.9.

The main effects time block, $F(7, 378) = 5.38$, $p < .001$, $\eta^2 = .10$. For further analysis, the paried t-test was performed to test the significance of the differences between the control groups and variation groups with the same n-back level. The result of the paired t-test showed that the average RMSD of the V1 (73.56, $p < 0.001$) was significantly lower than the constant hard control group (87.07). Therefore, this result supports the H4.1 (a) that the cyclic workload profile affects performance.

Table 4.7. Results of ANOVA on RMSD

	Effect	**F value**	***p* levels**	**Eta square**
RMSD In Tracking	Workload profile	F(5, 103.2) = 7.23	0.0001**	0.22
	Time block	F(7, 378) = 5.38	0.0001**	0.10

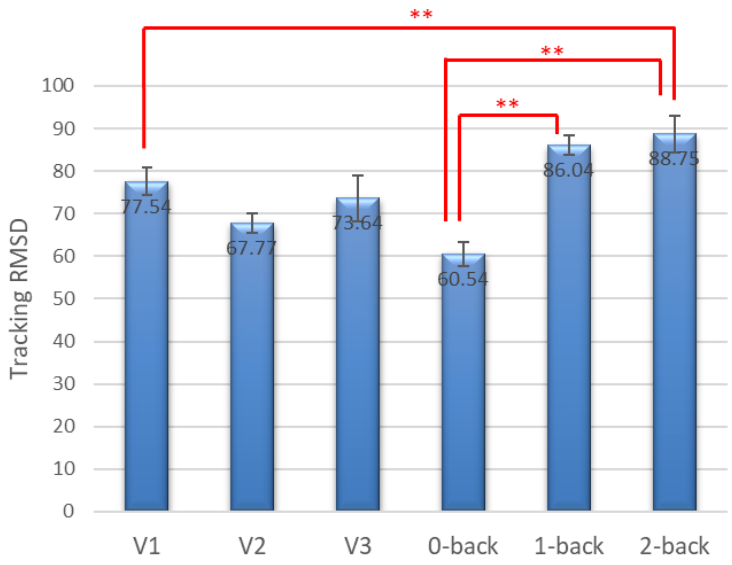


Figure 4.9. Result of Tukey's HSD test on RMSD

Table 4.8. Mean RMSD: Control vs. Transitioned group on RMSD

		RMSD			
Workload transition		Control	V1	V2	V3
Tracking	Easy	61.70	59.03		
		62.10		63.08	
		60.94			62.32
	Medium	85.68	83.62		
	Hard	87.07	73.56**		
		85.31		83.20	
		92.85			88.32

This study examined the effects of transitions in task demand on AF-MATB performance. Task demand was manipulated through variations in the background event rate of given tasks. The post-transition performance of cyclic workload profile in (Cyclic; V1 group) showed better performance (higher CRR, smaller RMSD, and shorter RT) than continuous hard control groups. Interestingly, in the tracking task, the V2 group (ramp-up) showed inferior performance than the continuous hard group.

4.1.3.2. Mental workload

The NASA TLX was administered at the end of each session, and participants were asked to complete it with respect to the workload they experienced during that session. TLX composite scores were calculated by computing the average of the scores for the six sub-scales for each session (Christ et al., 1993; Hendy, et al., 1993; Nygren, 1991). Then, the composite scores were collapsed across session 2 and 3 within each workload transition condition to create average scores.

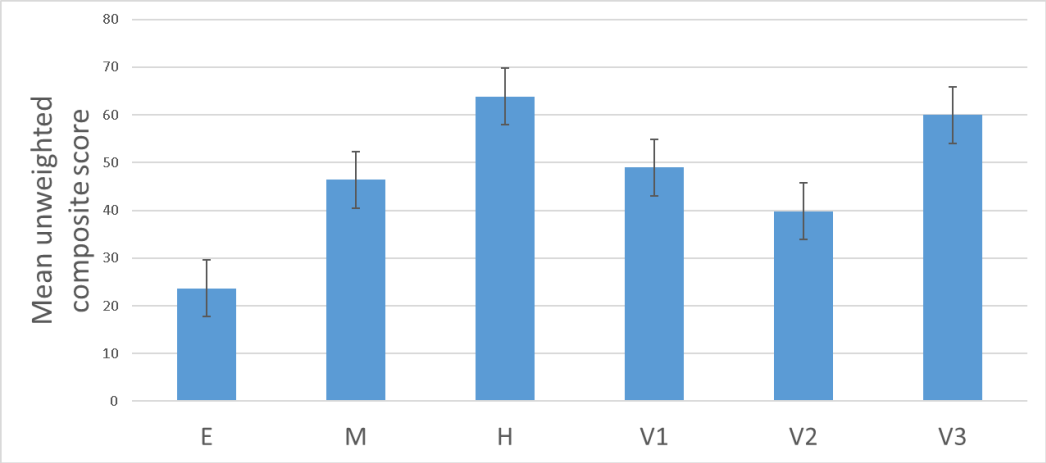


Figure 4.10. Mean unweighted composite NASA TLX scores

Figure 4.10 illustrates the mean, unweighted TLX composite scores of each group. The result of ANOVA showed the significant main effect of workload profile, $F(5, 103.2) = 8.27$, $p < .01$. The mean composite scores for the V1 groups reported a smaller mental workload than the continuous hard group. The score of V2 groups was significantly different from the control easy group. These results serve to caution that workload measurements obtained through the NASA-TLX at the end of an experimental session containing variations in task demand do not simply reflect an averaging of the participants' demand experiences.

4.1.3.3. Neurophysiological measures

Effective connectivity analysis

Source localization: We examined the spatial distribution of time-frequency information flow dynamics in the source domain using a combination of causal flow metrics and dipole fitting source localization. This revealed that causal source and sink hubs emerged during the multitasking. In Table 4.9, six selected ICs are presented. We focused on the parietal-temporal-occipital (PTO) association area, which is located in the cerebral cortex of the human brain. As its name implies, the PTO includes portions of the parietal, temporal and occipital lobes. This

association area—one of three in the cortex—is responsible for the assembly of auditory, visual, and somatosensory system information. Meaning is assigned to stimuli in the PTO, which outputs to numerous other areas of the brain, notably the limbic and prefrontal association areas, which are involved in memory and language recognition.

Table 4.9. Coordinates of the six independent components with residual variance < 10%

	Talairach coord. (x,y,z)	Location	Lobe	Closest BA	RV (%)
1	16, -1, -63	Caudate	Temporal	48	3.72
2	-48, -45, 28	Inferior Parietal	Parietal	39	5.74
3	49, 10, -21	Temporal Gyrus	Temporal	38	9.41
4	-12, -34, 0	Fusiform Gyrus	Temporal	36	7.37
5	36, -74, -2	Lateral Occipital Gyrus	Occipital	19	7.77
6	-7, -6, 34	Cingulate Gyrus	Ventral Anterior Cingulate	24	8.49

As we mentioned, AF-MATB is one of the representative multitasking ecological interfaces, and it contains audio, visual, and motor task. Like Study 1, temporal, parietal, and occipital ICs were found. In addition to this, due to the property of task complexity, we found ventral anterior cingulate cortex (vACC), BA (24) and Caudate, BA (48) which is one of the basal ganglia parts. Previous studies demonstrated that BA (24) is associated with visuospatial attention and the limbic system. Besides, several fMRI studies found BA (24)'s functions, such as response to executive functions (e.g., deductive reasoning, inductive reasoning, and mental timekeeping) and different types of working memory. Functional studies have disclosed the unexpected complexity of BA (38) functions. Because of its location in the brain, it is understandable that BA

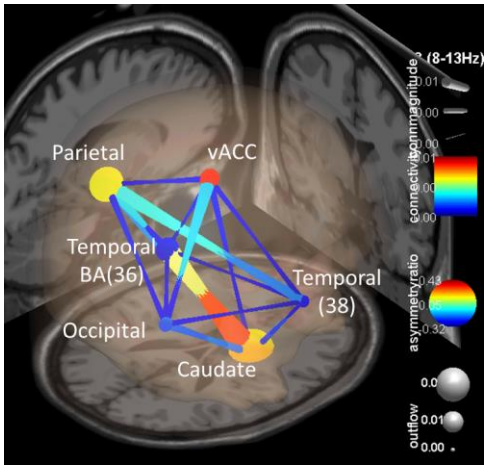
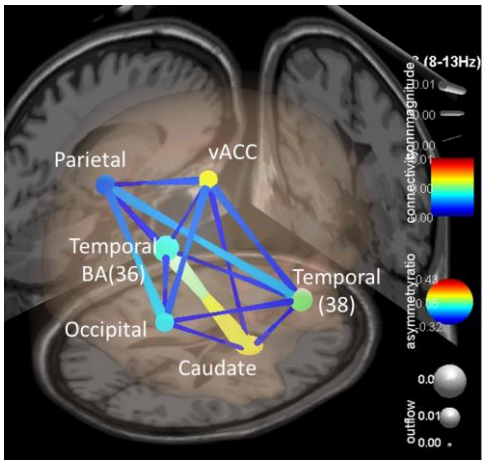
(38) participates in language processes, emotion, executive functions, and memory. Diverse studies support BA (38) contribution to multimodal memory retrieval.

Additionally, it seems to contribute to some complex auditory processing. During AF-MATB, participants were asked to attend audio stimuli continuously. Finally, BA (48) area is located in the hippocampus regions. According to the brain functional imaging studies, the hippocampus participates in a wide variety of memory processes, including working memory, episodic memory, and memory retrieval. In ACT-R, the caudate area would match the procedural module, and its primary function is a kind of comparator that determines whether the current information is new (and should be stored in memory), or it is old (and no storage is necessary). Unlike study 1, BA (48) and temporal regions were causal sources, and it would affect ACT-R model parameters adjustment.

Evaluating effective connectivity strength: We analyzed differences in information flow between the non-transitioned (control) group and the transitioned group. Table 4.10 shows two sessions of a causal transient alpha information flow and, each cell indicated significant different network patterns between transitioned groups and control groups at right after the workload transition has occurred. From the result of Phase I in study 2, only High difficulty condition shows significant differences between control and transitioned group so that we report brain network patterns under high difficulty level.

For the difference between control and V1 group, the IC 1, 2, and 6 nodes are colored orange, yellow, and red, respectively. Caudate BA (48) node shows red node color across the groups and conditions, which indicates that the component is the causal source for the network. The edge width and color between the temporal cortex and caudate coupling have a yellow-green hue and are thicker than the others.

Table 4.10. Effective connectivity: Control vs. Transitioned group

	V1	**Non-transitioned**
Brainnetwork pattern		
From – to	**Caudate – Temporal**	
Outflow	0.310 (0.017)	0.251 (0.012)
Connectivity magnitude	0.092 (0.013)	0.071 (0.010)
Asymmetry ratio	0.825 (0.192)	0.701 (0.118)

Statistical analysis using connectivity measures: Outflow, connectivity magnitude, and causal asymmetry ratio were computed for each IC source. We conducted statistical analysis using connectivity measures between caudate and temporal regions were calculated. We focused on the PTO association identical to study 1. In addition, the caudate regions played the role of a cortical hub of the brain network. Across all groups, the connectivity between the caudate and temporal cortex was significant, and in the next phase, these measures would match the procedural and declarative modules activations.

4.2. Phase II: Evaluating the feasibility of the proposed method in multitask

4.2.1. Objectives

The main goal of Phase II was to validate the methods on how to adjust three main ACT-R parameters, developed and updated in Study 1. The Phase II of Study 2 is different from that in Phase II of Study 1 in three ways: Additional considerations of findings on 1) workload transition effects measured in an ecologically more valid environment (e.g., a basic dual cognitive task vs. AF-MATB), 2) five ACT-R modules (the motor, declarative, imaginal, goal, and procedural) compared to Study 1 with only the declarative and imaginal modules considered, and 3) validated the method based on connectivity information between the main hub nodes and peripheral nodes in the brain (e.g., connectivity magnitude, asymmetry ratio, and causal flow).

4.2.2. ACT-R modeling of AF-MATB tasks

Our models were built in the ACT-R cognitive architecture. A version of ACT-R runs in Java, on a server that can communicate with a client browser. This arrangement is made possible by SIMCog-JS (Halverson et al., 2015). SIMCog-JS (Simplified Interfacing for Modeling Cognition – JavaScript) allows models to interact with browser-based software while requiring little modification to the task code. The modeler specifies how elements in the interface are translated into ACT-R chunks; the software allows keyboard and mouse interaction with JavaScript code, and it allows sending ACT-R commands from the external software. Our framework is based on SIMCog-JS (See Figure 4.11).

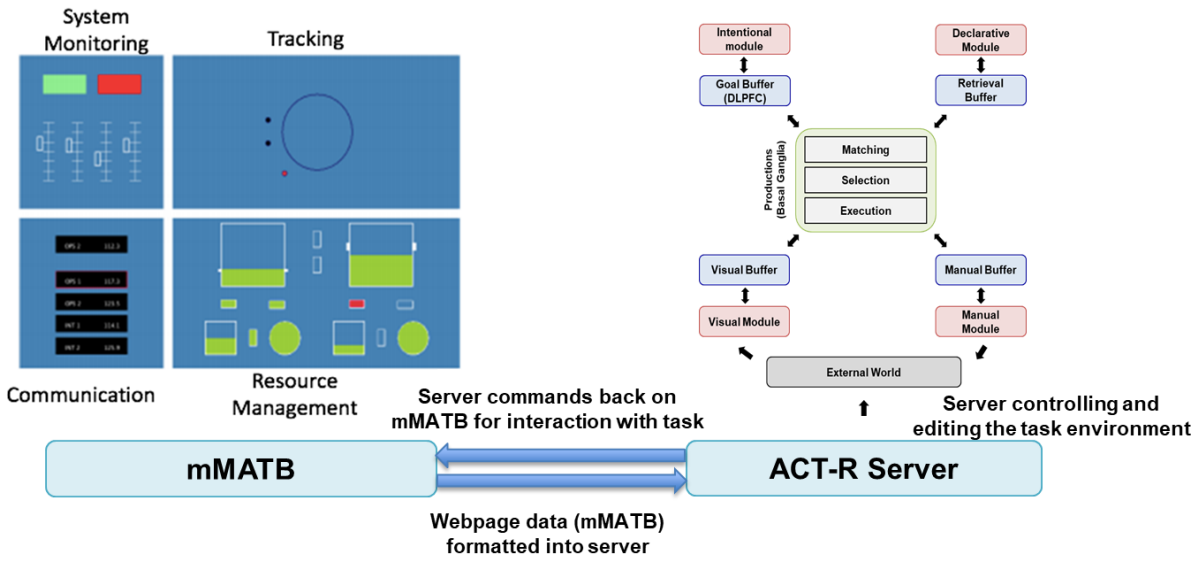


Figure 4.11. SIMCog-js Framework: mMATB (revised from Halverson et al., 2015 and adapted from Kim et al., 2018)

In the mMATB-JS, all task parameters were matched to the AF-MATB setting. Task demands were manipulated by increasing the event rate in each of the subtasks. SIMCog-js allows for the logging of detailed model data. To turn on data logging, edit the SIMCog.js file and set the *recordActrLogFile* (found in the Debugging and Data Recording Settings, under screen object declarations) to true. The system generated data on the model's task execution. Every time a change is detected, the following is recorded: Time since the start of the task, Declarative memory contents, Events and ACT-R buffer contents (Halverson et al., 2015). The output of model simulations included behavioral data like that produced in human observations, such as response time and error rate. Over the course of an mMATB session, the time series data for joystick movement on the screen and various button presses were recorded, and the utility of multitasking behavior when comparing human-to-model data was examined.

To fit the model to the data, we first created a model using the set of parameter values that were updated in the ACT-R architecture in Study 1. In Phase II, we need to manually adjust the

additional parameters until we obtain a model fit that was close to the data. We excluded outliers that are more than three standard deviations of the average performance data (Reaction time, RMSD, etc.) of each subtask.

We assumed that memory retrieval is optimal if the expected retrieval time is minimized for a particular error rate (Brown et al., 2009). Optimality in this sense means continuously computing the posterior probability for each option until one of these exceeds a criterion value (Dragalin et al., 1999) and this value is to be determined the magnitude of hysteresis effects. These assumptions would be in line with Study 1's outcomes. Brain connectivity dynamic measures are also involved in the retrieval threshold value and have been shown to compute activation values for a certain decision criterion. Due to the complexity of the task, the mMATB model has more dependent accumulating processes than the dual-task ACT-R model in Study 1. We changed the spreading activation from each task influences not only the workload transition but also the way the accumulating processes affect each subtask. In order to assess what changes were required to the model to fit the data better; we made several changes to capture post-transition performance accurately. In Study 2, a new set parameter was changed within sensory modules such as visual and motor modules in the ACT-R model due to AF-MATB properties. We follow Study 1's procedures in this section.

4.2.3. Results

4.2.3.1. Behavioral measures

One aim of this study was to predict behavioral performance and neural data changes of the participants. Thus, the proportion of correct responses of the participants was compared to the proportion of correct responses of each group and condition. To validate the proposed parameter mapping method in this study, model fitness between human and model were calculated using R^2 and Root Mean Square Error (RMSE). Each of the models was run 250-350 times. The data were preprocessed with custom Java SimCog files and then analyzed with Microsoft Excel. The model data and the empirical data were divided into each 2-min time bins, but we only considered sessions 2 and 3. The average proportion of correct responses and the standard deviation per block was computed for the experiment as well as for each of the six workload transition types.

Table 4.11. AF-MATB task behavioral fit with the default setting of mMATB

Subtask	**System monitoring**		**Tracking**		**Communication**	
	CRR		RMSD		RT	
Workload profile	R^2	RMSE	R^2	RMSE	R^2	RMSE
E	0.863	0.12	0.828	0.13	0.853	0.03
M	0.883	0.18	0.889	0.15	0.877	0.10
H	0.892	0.23	0.899	0.17	0.892	0.16
V1	0.786	0.40	0.726	0.37	0.728	0.33
V2	0.798	0.38	0.830	0.29	0.789	0.25
V3	0.736	0.23	0.827	0.22	0.799	0.36

Note. CRR: Correct response ratio; RMSD: Root mean square distance; RT: Response time

As we expected, the pattern that was not predicted was the performance of the transitioned group. Like Study 1, the pre-model 2 simulated data show slightly worse performance as compared to V2 and V3 groups, but for V1, the model data worse than human performance. To accurately capture V1 performances in sessions 2 and 3, we should consider the possibility of hysteresis

effects on them based on Phase I outcomes. Therefore, a new production rule would update as a lower retrieval threshold and the latency-scaling factor.

4.2.3.2. Neurophysiological measures

From the results of Phase I, the six selected ICs present below in Table 4.12.

Table 4.12. Selected ICs and ACT-R module matching

EEG			**ACT-R**	
Talairach Coord. (x,y,z)	IC	Brain Region	Module	Function
16, -1, -63	BA(48)	Caudate	Procedural module	Activated by the production rules
-48, -45, 28	BA(39)	Parietal	Visual buffer	Place a representation of the stimulus in the visual buffer
49, 10, -21/ -12, -34, 0	BA(38/36)	Temporal	Declarative module	Retrieve the instruction about what to do within each condition from declarative memory
-7, -6, 34	BA(24)	Limbic	Motor module	Perceive and do the action (press the button or control joystick)
36, -74, -2	BA(19)	Occipital	Visual module	Perceive visual stimuli

Table 4.13. Fit measures of the neural data between Model and data from Phase I

Workload profile	**Procedural (Caudate)**		**Declarative (Temporal)**		**Visual (Occipital)**		**Motor (Limbic)**	
	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
E	0.883	0.120	0.896	0.110	0.936	0.032	0.875	0.034
M	0.893	0.078	0.887	0.083	0.863	0.036	0.889	0.036
H	0.897	0.089	0.869	0.076	0.913	0.053	0.867	0.463
V1	0.653	0.157	0.642	0.265	0.783	0.132	0.687	0.165
V2	0.789	0.163	0.792	0.360	0.732	0.143	0.795	0.138
V3	0.882	0.223	0.775	0.331	0.897	0.203	0.668	0.154

In Table 4.13, the EEG predictions of the model are presented. The problem state of simulation results shows that the V1 group had the lowest activity for all conditions involving 2-

back, instead of the highest. The prediction for declarative memory represented the data marginally better ever for V2 and V3 groups. The low amount of predicted activity during gradually increase workload transition occurred did not fit with the temporal node information flow. In contrast to the model, the EEG data indicated that the problem state was used less in the transitioned group than control groups, but the declarative module was used more for the V1 group than the corresponding control. The data indicate that participants use declarative memory more after experiencing a gradual workload transition than our model predicted. Taken together, this indicates that participants used a different memory strategy to keep track of the previous letter stimuli than the model predicted.

4.2.3.3. Adjustment of parameters

The main hypothesis of this study is that post-workload transition behavior can be modeled as an adaptive response to effective connectivity in certain areas of the brain. This is represented by a reduction in the source between the Declarative module and other modules, corresponding to connectivity between the temporal cortex and other areas (ACC, Motor, Parietal, and Fusiform). This reduction in source activations only affects the Declarative module, so connectivity between the Procedural module (Caudate) and other modules is maintained from Simcog mMATB. Based on our neurological outcomes, we found a significant reduction in connectivity for the V1 group compared to control from caudate to temporal regions, corresponding to reduce connectivity from the procedural module to the Declarative module of ACT-R. Koshino et al. (2007) found reduced connectivity from frontal regions, including the vACC to the fusiform gyrus, corresponding to reduced connectivity from the Declarative module to the Visual module. In particular, Kim et al. (2018) found a pattern of significant effective connectivity in AF-MATB multitasking from the

subcortical caudate area of the brain to areas across temporal, parietal, and occipital lobes. Models were run in each variation group's condition under these candidate definitions of a memory retrieval index:

(a) Connectivity magnitude (Caudate – temporal): Retrieval threshold

$$\tau \propto \frac{\sum dDTF\ from\ Caudate\ to\ temporal}{Total\ dDTF\ in\ network}$$

(b) Asymmetry ratio of caudate:

$$F \propto Time_{Max\frac{Inflow_{caudate}-Outflow_{caudate}}{Inflow_{caudate}+Outflow_{caudate}}} - Time_{Min\frac{Inflow_{caudate}-Outflow_{caudate}}{Inflow_{caudate}+Outflow_{caudate}}}$$

(c) Causal flow (caudate -> parietal): Activation noise

$$\varepsilon \propto Outflow_{caudate} - Inflow_{caudate}$$

Table 4.14. Proposed a new set of parameters

Parameter	Description	Pre-Model 1	Post-Model 1
F	Retrieval latency factor	0.1	0.23
$\boldsymbol{\tau}$	Retrieval duration threshold	0.1	0.026
$\boldsymbol{\varepsilon}$	Moment-to moment noise	0.1	0.047

We tested the hypothesis that the number of causal flows of a temporal node could predict the degree of information flow for memory retrieval processing, as indexed by the sum of outgoing interactions in the alpha-band. The indices were judged based on their relative strength in exhibiting expected causality flows within caudate-temporal pathways. In Table 4.14, the proposed parameters were listed.

To validate the proposed set of parameters in this study, model fitness between human and model was calculated using R^2 and Root Mean Square Error (RMSE) (See Table 4.15).

Table 4.15. Fit measures between behavioral data and updated model

Performance		Before tuning		After tuning	
		R^2	RMSE	R^2	RMSE
System Monitoring CRR	**V1**	0.786	0.40	0.953	0.013
	V2	0.798	0.38	0.926	0.032
	V3	0.736	0.23	0.913	0.043
Tracking RMSD	**V1**	0.726	0.37	0.905	0.065
	V2	0.830	0.29	0.894	0.053
	V3	0.827	0.22	0.913	0.043
Communication RT	**V1**	0.728	0.33	0.913	0.036
	V2	0.789	0.25	0.896	0.034
	V3	0.799	0.36	0.904	0.048

Note. CRT: Correct response ration; RMSD: Root Mean Square Distance; RT: Reaction Time

Model 2-post shows that a very close match to the participant data is possible. The procedural module activation fits very well with caudate area information flows.

Table 4.16. Fit measures between neural data and updated model

Module Activation		Before tuning		After tuning	
		R^2	RMSE	R^2	RMSE
Procedural Module	**V1**	0.653	0.157	0.932	0.036
	V2	0.789	0.163	0.953	0.034
	V3	0.882	0.223	0.943	0.046
Declarative Module	**V1**	0.642	0.265	0.916	0.051
	V2	0.792	0.36	0.923	0.023
	V3	0.775	0.331	0.933	0.031
Visual Module	**V1**	0.783	0.132	0.945	0.016
	V2	0.732	0.143	0.913	0.025
	V3	0.897	0.203	0.927	0.032
Motor Module	**V1**	0.687	0.165	0.931	0.047
	V2	0.795	0.138	0.926	0.051
	V3	0.668	0.154	0.933	0.063

4.3. Discussions of STUDY 2

Subsequent to the STUDY1, we wanted to validate the updated model by applying it to different tasks. In study 2, we utilized a broader area of ACT-R architecture to understand a more realistic, multitasking environment. By understanding neural correlates and ACT-R production rule processing, in the current study suggested a novel way to scale the significant three parameters of ACT-R. These parameters are essential building blocks of a substantial ACT-R architecture, and they mainly involved human memory store and retrieval processing mechanisms. Based on the literature and analysis of our neurological data set, we could figure out the way to calculate how much we need to adjust parameters. In the current study, we utilized various information flow measures to rescale the following three parameters; retrieval threshold, latency-scaling factor, and activation noise. We validated the simulated data with updated parameters have a better fit with human data than default parameters.

5. General discussion

5.1. Workload transition effects

The workload transition effect is broadly defined as the influence of previous states on current states. Performance hysteresis is the effect of workload transition types on the subject's performance and is usually considered applicable when a system responds differently to identical inputs depending on the direction the system is driven (Farrell, 1999; Helton et al., 2004). Specifically, while sudden changes in workload tended to deteriorate performance, performance improves with increasing workload, reaches its peak at the highest difficulty, and subsequently decreases (Cox-Fuenzalida et al., 2006). Alternatively, the gradual transitions result in performance improvement in the post-transition period (Moroney et al., 1995).

5.1.1. Task type and workload profile

There were two reasons behind the inconsistent results of the previous studies – inconsistency in workload profiles and the types of tasks participants were asked to perform. In order to overcome these issues, this study consists of two different experiments. Participants performed a simple task in the first experiment and more complex workload tasks in the second one to increase ecological validity. It is becoming increasingly apparent that a simple explanation addressing responses to a change in workload level fails to address the complexities involved in the more dynamic effect of workload history on performance. The importance of this research is evident for many real-world jobs, particularly those involving safety sensitive occupations. Adjustments in task difficulties through two different types of tasks were utilized to better understand the effect of workload history on performance. In addition, further exploration of factors potentially responsible for post-training effects needs to be further investigated.

Next, this study examined the effects of workload transition while addressing several designs and methodological limitations that challenged the internal and external validity of previous studies (Cox- Fuenzalida, 2007). Three types of workload profiles were used in this study (cyclic variation and ramp-down and ramp-up), which allows us to collect much more time points to investigate workload transition effects than the simplified version. With the three types of workload profiles, we can investigate multiple aspects of workload transition effects; (1) transition direction (increase or decrease), (2) degree of transition (moderate change (L-M-H) or extreme change (L-H)), and (3) time-oriented- distinction of change (gradual change or abrupt change).

5.1.2. Memory retrieval and temporal cortex

By analyzing neural data, we found the reason why the V1 group outperformed the control group. We observed significant differences in the strength of connectivity between temporal-frontal and caudate-temporal regions. The V1 group showed continued activation of the temporal regions throughout the experiment. This finding is consistent with the previous studies that the temporal regions govern memory retrieval functions (Gelinas, 2019). The better performance of the V1 group implies that the temporal regions also play an important role in performing multitasking. We observed that the temporal cortex has a strong information flow that might involve the efficiency of memory retrieval (Salvucci et al., 2004). In particular, under the multitasking environment, temporal regions could contribute to the memory retrieval of task-relevant information that occurs more quickly and with less chance of error (misretrieval or failure to retrieve).

5.2. Proposed method on how to adjust the ACT-R model

The previous ACT-R models are based on behavioral results, which provide limited insight into processes within the brain. As ACT-R modules have neural correlates (Anderson, 2007; Borst et al., 2015), neuroimaging data can be used to evaluate the cognitive resources utilized during a particular task. However, our study confirmed the existence of the hysteresis effect and found that the previous models with default parameters could not explain the hysteresis effect. Our study proposed a new method to adjust parameters that incorporates the hysteresis effect and results in better model fitness. The new method is based on the results of EEG effective connectivity. Specifically, our study is focused on memory retrieval mechanisms, which are associated with PTO networks.

Study 1 found that the temporal node functions as the main hub for simple cognitive tasks, whereas Study 2 identified the caudate node as a critical hub for complex multi-tasking. These findings suggest that the level of the task's complexity determines which node will be primarily used when one experiences changes in workload. The study identified three important measures of connectivity between the main hub node and peripheral nodes based on the graph theory – connectivity magnitude, asymmetry ratio of caudate, and causal flow. We proposed the following three new methods to adjust parameters for the ACT-R model based on these three measures. We postulated that the connectivity measures are correlated to ACT-R modules activation (i.e., higher connectivity magnitude reflects increased activation of relevant modules). We found that the proposed methods below were meaningful to adjust parameters in ACT-R. Based on these findings, we could conclude that there is a high level of synchronization between ACT-R modules and EEG data based on effective connectivity.

6. Conclusions and future research

This chapter presents a summary of research findings (Section 6.1), summarizes implications for ACT-R and Neuroergonomics research (Section 6.2), and identifies avenues for future research (Section 6.3).

6.1. Summary of research findings

Past research has shown inconsistent results on workload transition or Hysteresis effects at the behavioral, perceived workload, or neurophysiological level, resulting in three different hypotheses: Enhancement, Deterioration, or No change. In addition, little attention has been paid to a quantitative modeling approach, driven by the neurophysiological activity of human operators, to modeling and simulating workload transition effects. Hence, this research explored effective connectivity-based ACT-R modeling of workload transition in multitasking. This study has met its goals by addressing all the research goals and questions specified in 1.2 and 1.3.

This research addressed the first research question, "What is the nature of workload transition effects?", by examining the behavioral, perceived workload, and neurophysiological measures in both a basic cognitive task (in Phase I of Study 1) and a more ecologically valid task (AF-MATB) (in Phase I of Study 2). The findings of the present study confirmed workload transition effects at the three measurement levels in both task settings.

The second research question, "What is the nature of the effects of workload transition variation types at the behavioral, perceived workload, and neurophysiological levels?" supported the "Enhancement" hypothesis. The cyclic transition type (V1) was shown to significantly affect task performance, perceived mental workload, and brain connectivity, as compared to other transition types. We investigated the difference in connectivity flow between the V1 and control

group and identified differentiating brain regions that coincide with the ACT-R module associated with memory retrieval.

Last but not least, answers to the third research question "What is a systematic way that can improve the fitness of the ACT-R model that simulated human performance during workload transitions?" outlined a parameter tuning framework using both behavioral measures and effectivity connectivity of human operators. This research found that the proposed method for tuning ACT-R parameters can be used to improve the ACT-R model's fit accounting for Hysteresis. Through Studies 1 and 2, we validated that the ACT-R model with the revised parameters resulted in better model fitness than the model with default parameters.

6.2. Contributions and implications of this research

6.2.1. The nature of workload transition effects

The findings of this research have many implications for practitioners. There have been multiple studies on the workload transition effects based on behavioral data, which produced inconsistent results. Our study supported the enhancement hypothesis that cyclic workload profiles result in better performance than non-transition profiles, specifically under difficult workload conditions. These findings would benefit those who build training programs targeted for multitasking.

Workload history (more specifically, a workload shift) has significant implications for many work environments. These implications are particularly salient in safety sensitive occupations where individuals are confronted with varying levels of workload demand. These people often undergo regular trainings in order to better cope with emergency situations. The results of the experiment showed that the participants who went through the gradual changes

exhibited longer activation of the temporal area and performed better than other groups. This result suggests that the design of training programs, which incorporates gradual workload changes, would potentially better prepare trainees for abrupt workload changes in real situations.

Future research is needed to observe how the participants who went through trainings with different workload profiles react in real situations such as driving, medical operation, air traffic controller and power plant operator. Workload history (more specifically, a workload shift) has significant implications for many work environments. These implications are particularly salient in safety sensitive occupations where individuals are confronted with varying levels of workload demand. These people often undergo regular trainings in order to better cope with emergency situations. The results of the experiment showed that the participants who went through the gradual changes exhibited longer activation of the temporal area and performed better than other groups. This result suggests that the design of training programs, which incorporates gradual workload changes, would potentially better prepare trainees for abrupt workload changes in real situations.

Future research is needed to observe how the participants who went through trainings with different workload profiles react in real situations such as driving, medical operation, air traffic controller and power plant operator.

In the neuroergonomics field, this study may contribute to better understand the neural mechanism behind multitasking. Through neural analysis, the study found that people with better performance exhibited the strong activation of the temporal regions which are involved in retrieving task-relevant information. We also found that cyclic workload profiles might better induce continued activation of the temporal regions than other workload profiles. Therefore, we believe that cyclic workload profiles would be implemented to develop training programs to

augment multi-tasking performance. Further study is needed to verify this finding by utilizing brain stimulation in the temporal regions.

6.2.2. Effective connectivity-based ACT-R parameter tuning method

A combination of experimental and computational assessment is often the best approach. Advances in simulator technologies are creating experiences with a greater sense of presence such that simulators become increasingly accurate estimates of operational performance. Nevertheless, task performance measures are key for postulating predictive models based on other operator-state factors that can be evaluated in constructive simulations with many replications; virtual simulations are typically restricted to a few replications and so can consider commensurately fewer conditions.

Multiple studies have covered information-processing procedures under multitasking environments. However, most of these studies are focused on high-level concepts and lack processing details, which require a computational model. Only a few computational models factoring workload transitions have been developed and can be applied to the limited number of tasks (Cohen, 1998; McClelland, 2000; Kriete & Noelle, 2005).

This study proposes a new method based on the ACT-R cognitive architecture that incorporates workload transition effects and manages information processing (Anderson et al., 2004). This study focuses mainly on the declarative module for memory retrieval. Previous studies manually adjusted parameters through iterations, while our study proposes a systematical way to adjust parameters by reflecting the connectivity between the ACT-R model and brain regions. We believe this study benefits other ACT-R modelers by proposing a way to calibrate general key parameters, which is diagnostic to the type of underlying tasks.

Lastly, the findings of this research could function as a tutorial for practitioners who want to build a neurocognitive model using ACT-R. While there are many tutorials for fMRI and ACT-R, only a few are available for EEG, specifically effective connectivity combined with ACT-R.

6.3. Research limitations and future work

6.3.1. Limitations of the study

Although this study contributes to the body of research in the domain of human-systems engineering, specifically, the research on the nature of workload transition effects in multitasking, ACT-R modeling, and Neuroergonomics, it suffers from limitations that may affect interpretation and generalization of the findings, described as follows.

Compared to Study 1, we expected more significant outcomes from the dependent variables from Study 2 given more subtasks and multi-modality stimuli. Despite the many performance measures provided by the AF-MATB task, Study 2 resulted in fewer significant outcomes than expected. Specifically, Study 2 featured a communication task using audio cues, which is not included in Study 1. However, we could not find any significant workload transition effects from the communication task. Bower (2013) pointed out that the communications task was not ideal for examining workload transitions because of the task property. There are time constraints on the communications task. It takes time for the instructions in each communication to be spoken, and then participants are given 15 seconds to comply. We believe that this 15-second lag was too long to identify the impacts on the workload profile

We tried to measure the mental workload through self-reporting. However, self-reporting seemed to interfere with continuous workload profiles. Therefore, future research will need workload measures that minimize this type of interference.

Lastly, this study used WMC to solely categorize groups. However, the WMC might impact workload transition effects. Future research will further investigate differences in performance and neural data between the high WMC and the low WMC in terms of workload transition effects.

6.3.2. Directions for future research

Various recommendations for future research can be made, which are related to cognitive modeling using ACT-R, real-time BCI systems, and neuroergonomics areas.

As our proposed tuning measure is the first systemic approach to adjust the ACT-R parameters, to our best knowledge, it will require further validation by applying to different settings and experiments. We intend to integrate our method with fMRI as EEG connectivity patterns have lower spatial resolution and need fMRI data analysis to identify activated areas more accurately. Thus, combining EEG, fMRI, and ACT-R will provide a powerful approach to investigate the spatiotemporal features, neural structures, and neurobiological processes related to the activated areas.

Mullen et al. (2015) have demonstrated that elements of effective connectivity can be applied to a real-time BCI setting and decode brain states using the spectro-temporal features. With the development of online recursive ICA (Hsu et al., 2014) that allows for robust online artifact rejection and source identification, this proposed approach can be utilized for real-time BCI applications. Combing ACT-R and EEG data would allow us to find the best features to explain workload transition effects and multitasking behavior in real-time analysis.

Neuroergonomics aims to integrate cognitive neuroscience methods with human factors (Parasuraman et al., 2012). One of the important ergonomic issues is to understand how human

operators cope with unexpected workload transitions (Huey & Wickens, 1993). By combining EEG with effective connectivity analysis, we demonstrated how researchers might identify and analyze the understudied effects of workload transition. This work presents several potential research areas for neuroergonomics.

By identifying the effects of workload transitions, researchers can create interventions that specifically target these effects. Potential interventions might include cognitive training to improve specific cognitive abilities (McKendrick et al., 2014), adaptive aiding (Wilson & Russell, 2007), or non-invasive brain stimulation such as tDCS to enhance the processing capacity of specific brain regions or networks (Clark & Parasuraman, 2014). In order to develop an algorithm that detects changes in workload, EEG data that reflect the hysteresis effect can be one of the important features of the algorithm to improve the strength of the classifier. By delving deeper into potential physiological hysteresis effects, EEG based-closed loop system would be an appropriate continuation of this work.

REFERENCES

Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: An activation-based model. *Cognitive Science*, *26*(1), 39–83.

Anderson, J. R. (1983). *Cognitive science series. The architecture of cognition.* Hillsdale, NJ, US.

Anderson, J. R. (2009). *How can the human mind occur in the physical universe?* Oxford University Press.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An Integrated Theory of the Mind, *111*(4), 1036–1060. https://doi.org/10.1037/0033-295X.111.4.1036

Anderson, J. R., Carter, C. S., Fincham, J. M., Qin, Y., Ravizza, S. M., & Rosenberg-lee, M. (2008). Using fMRI to Test Models of Complex Cognition. *Cog*, *32*(8), 1323–1348. https://doi.org/10.1080/03640210802451588

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, *63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Baker, C. H. (1963). Signal duration in a vigilance task. *Science*, 136, 46-47.

Bothell, D. (2010). Modeling Space Fortress: CMU Effort [PowerPoint slides].

Bowers, M. A. (2013). August 2013.

Bowers, M. A., Christensen, J. C., & Eggemeier, F. T. (2014). The effects of workload transitions in a multitasking environment. *Proceedings of the Human Factors and Ergonomics Society*, *2014–Janua*, 220–224. https://doi.org/10.1177/1541931214581046

Brillinger, D. R. (2001). *Time series: data analysis and theory*. Siam.

Brown, S. D., Steyvers, M., & Wagenmakers, E. -J. (2009). Observing evidence accumulation during multi-alternative decisions. *Journal of Mathematical Psychology*, *53*(6), 453–462.

Bullmore, E. T., & Sporns, O. (2009). Complex brain networks : graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*(3), 186. https://doi.org/10.1038/nrn2575

Butts, C. T. (2009). Revisiting the Foundations of Network Analysis. *Science*, *325*(5939), 414–417.

Byrne, M. D., & Anderson, J. R. (2001). Serial Modules in Parallel : The Psychological Refractory Period and Perfect Time-Sharing, 1–78.

Cassenti, D. N., Kerick, S. E., & McDowell, K. (2011). Observing and modeling cognitive events through event-related potentials and ACT-R. *Cognitive Systems Research*, *12*(1), 56–65. https://doi.org/10.1016/j.cogsys.2010.01.002

Chen, X., Bailly, G., Brumby, D. P., Oulasvirta, A., & Howes, A. (2015). The emergence of interactive behavior: A model of rational menu search. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems* (pp. 4217–4226). ACM. https://doi.org/10.1145/2702123.2702483

Clark, V. P., & Parasuraman, R. (2014). Neuroenhancement: Enhancing brain and mind in health and in disease. *Neuroimage*, *85*(889–894).

Cohen, J. (1988). Statistical power analysis for the behavioral sciences Lawrence Earlbaum Associates. Hillsdale, NJ, 20–26.

Cohen, J. (1992). Statistical Power Analysis, *1*(3), 98–101.

Comstock, J. R., & Arnegard, R. J. (1992). Task Battery Description Monitoring Running the Task Battery, (January).

Cooper, R. P. (2007). The Role of Falsification in the Development of Cognitive Architectures : Insights from a Lakatosian Analysis. *Cognitive Science*, *31*(3), 509–533.
Cox-Fuenzalida, L.-E. (2007). Effect of Workload History on Task Performance. *Human Factors*, *49*(2), 277–291.
Delorme, A., Mullen, T., Kothe, C., Akalin Acar, Z., Bigdely-Shamlo, N., Vankov, A., & Makeig, S. (2011). EEGLAB, SIFT, NFT, BCILAB, and ERICA: New tools for advanced EEG processing. *Computational Intelligence and Neuroscience*, *2011*. https://doi.org/10.1155/2011/130714
Deprez, S., Vandenbulcke, M., Peeters, R., Emsell, L., Amant, F., & Sunaert, S. (2013). The functional neuroanatomy of multitasking: Combining dual tasking with a short term memory task. *Neuropsychologia*, *51*(11), 2251–2260. https://doi.org/10.1016/j.neuropsychologia.2013.07.024
Dragalin, V. P., Tartakovsky, A. G., & Veeravalli, V. V. (1999). Multihypothesis sequential probability ratio tests. I. Asymptotic optimality. *IEEE Transactions on Information Theory*, *45*(7), 2448–2461.
Endsley, M. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, *37*, 32–64.
Engle, R. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23.
Faubel, C., & Schöner, G. (2008). Learning to recognize objects on the fly: a neurally based dynamic field approach. *Neural networks*, *21*(4), 562-576.
Forstmann, B. U., Wagenmakers, E., Eichele, T., Brown, S., & John, T. (2011). Reciprocal Relations Between Cognitive Neuroscience and Cognitive Models : Opposites Attract ? *Trends Cogn Sci*, *15*(6), 272–279. https://doi.org/10.1016/j.tics.2011.04.002.Reciprocal
Friston, K. J. (1994). Functional and Effective Connectivity in Neuroimaging : A Synthesis. *Human Brain Mapping*, *2*(1–2), 56–78.
Friston, K., Moran, R., & Seth, A. K. (2013). Analysing connectivity with Granger causality and dynamic causal modelling. *Current Opinion in Neurobiology*. https://doi.org/10.1016/j.conb.2012.11.010
Gartenberg, D., Veksler, B., Gunzelmann, G., & Trafton, J. G. (2014). An ACT-R process model of the signal duration phenomenon of vigilance. In *Proceedings of 58th annual meeting of the Human Factors and Ergonomics Society*.
Gelinas, J. (2019). Ripples for memory retrieval in humans. *Science*, *363*(6430), 927-928.
Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, *7*(4), 374–385. https://doi.org/10.1093/cercor/7.4.374
Geweke, J. (1982). Measurement of Linear Dependence and Feedback Between Multiple Time Series. *Journal of the American Statistical Association*, *77*(378), 304–313.
Goldinger, S. D., Kleider, H. M., Azuma, T., & Beike, D. R. (2003). "BLAMING THE VICTIM" UNDER MEMORY LOAD. *Psychological Science*, *14*(1), 81–85.
Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, *27*(4), 591–635.
Granger. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods A. *Econometrica*, *37*(3), 424–438.
Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology. General*, *121*, 480–

506.
Gray, W. D., Schoelles, M. J., & Sims, C. R. (2005). Adapting to the task environment : Explorations in expected value Action editor : Christian Schunn, *6*, 27–40. https://doi.org/10.1016/j.cogsys.2004.09.004
Halverson, T., Reynolds, B., & Blaha, L. M. (2015). SIMCog-JS: simplified interfacing for modeling cognition–JavaScript. In *The International Conference on Cognitive Modeling* (pp. 39–44).
Hancock, P. A., & Parasuraman, R. (1992). Human factors and safety in the design of Intelligent Vehicle-Highway Systems (IVHS). *Journal of Safety Research*, *23*(4), 181–198.
Hancock, P. A., & Verwey, W. B. (1997). Fatigue, workload and adaptive driver systems. *Accident Analysis and Prevention*, *29*(4), 495–506.
Hancock, P. A., Williams, G., Manning, C. M., Hancock, P. A., & Williams, G. (2009). Influence of Task Demand Characteristics on Workload and Performance Influence of Task Demand Characteristics on Workload and Performance, *8414*. https://doi.org/10.1207/s15327108ijap0501
Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, *52*(C), 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9
Helton, W.S., Shaw, T.H., Warm, J.S., Matthews, G., Dember, W.N., & Hancock, P. . (2004). Demand transitions in vigilance: Effects on performance efficiency and stress. In *Human Performance, situation awareness and automation: Current research and trends* (pp. 258–262). Erlbaum: Mahwah, NJ.
Horwitz, Barry, and A. R. B. (2004). Brain network interactions in auditory, visual and linguistic processing. *Brain and Language*, *89*(2), 377–384.
Hsu, S. H., Mullen, T., Jung, T. P., & Cauwenberghs, G. (2014, August). Online recursive independent component analysis for real-time source separation of high-density EEG. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 3845-3848). IEEE.
Jaccard, J., Becker, M.A., & Wood, G. (1984). Pairwise multiple comparison procedures: a review. *Psychological Bulletion*, *96*(3), 589–596.
Jonathan P.Gluckman, Joel S. Warm, william N. Dember, roger R. rosa. (1993). Demand Transition and sustained attention. *The Journal of General Psychology*, *120*(3), 323–337.
Kamiński, M., Ding, M., Truccolo, W. A., & Bressler, S. L. (2001). Evaluating causal relations in neural systems : Granger causality , directed transfer function and statistical assessment of significance. *Biological Cybernetics*, *85*(2), 145–157.
Kamiński, M., Ding, M., Truccolo, W. A., & Bressler, S. L. (2001). Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biological Cybernetics*, *85*(2), 145–157. https://doi.org/10.1007/s004220000235
Kane, M. J., & Engle, R. W. (2003). Working-Memory Capacity and the Control of Attention : The Contributions of Goal Neglect , Response Competition , and Task Set to Stroop Interference. *Journal of Experimental Psychology: General*, *132*(1), 47–70. https://doi.org/10.1037/0096-3445.132.1.47
Kangasrääsiö, A., Athukorala, K., Howes, A., Corander, J., Kaski, S., & Oulasvirta, A. (2017). Inferring cognitive models from data using approximate Bayesian computation. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 1295–

1306). New York: ACM.
Ketola, M., Jiang, L. P., & Stocco, A. (2019). Comparing Alternative Computational Models of the Stroop Task Using Effective Connectivity Analysis of fMRI Data. *bioRxiv*, 647271.
Kieras, D. E., Meyer, D. E., Ballas, J. A., & Lauber, E. J. (2000). Modern computational perspectives on executive mental processes and cognitive control: Where to from here. *Control of Cognitive Processes: Attention and Performance, XVIII*, 681–712.
Kim, N., Kim, W., Yun, M. H., & Nam, C. S. (2018). Behavioral and Neural Correlates of Hysteresis Effects during Multitasking, 2010–2012. https://doi.org/10.1177/1541931218621003
Kim, N., Wittenberg, E., & Nam, C. S. (2017). Working Memory Capacity , Memory Load , and Cognitive Control Network, 46–48. https://doi.org/10.1177/1541931213601506
Konig, C. J., Buhner, M., Murling, G., & König, C. J. (2005). Working Memory , Fluid Intelligence , and Attention Are Predictors of Multitasking Performance , but Polychronicity and Extraversion Are Not Working Memory , Fluid Intelligence , and Attention Are Predictors of Multitasking Performance , but Polychronicit. *Human Performance*, *18*(3), 243–266. https://doi.org/10.1207/s15327043hup1803
Korzeniewska, A., Mańczak, M., Kamiński, M., Blinowska, K. J., & Kasicki, S. (2003). Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method. *Journal of Neuroscience Methods*, *125*(1–2), 195–207. https://doi.org/10.1016/S0165-0270(03)00052-9
Krulewitz, J. E., Warm, J. S., & Wohl, T. H. (1975). Effects of shifts in the rate of repetitive stimulation on sustained attention. *Perception & Psychophysics*, *18*(4), 245–249.
Kus, R., Kaminski, M., & Blinowska, K. J. (2004). Determination of EEG activity propagation: pair-wise versus multichannel estimate. *IEEE Transactions on Biomedical Engineering*, *51*(9), 1501–1510.
Laird, J. E. (2012). *The SOAR Cognitive Architecture*. MIT press.
Langley, P., Laird, J. E., & Rogers, S. (2009). Cognitive architectures : Research issues and challenges. *COGNITIVE SYSTEMS RESEARCH*, *10*(2), 141–160. https://doi.org/10.1016/j.cogsys.2006.07.004
Lebiere, Christian, and J. R. A. (1993). A connectionist implementation of the ACT-R production system. In *the fifteenth annual conference of the Cognitive Science Society* (pp. 635–640).
Lebiere, C., Anderson, J. R., Bothell, D., Lebiere, C., & Anderson, J. R. (2001). Multi-Tasking and Cognitive Workload in an ACT- R Model of a Simplified Air Traffic Control Task.
Leendert van Maanen, Hedderik van Rijn, N. T. (2012). RACE / A : An Architectural Account of the Interactions Between Learning , Task Control , and Retrieval Dynamics. *Cognitive Science*, *36*, 62–101. https://doi.org/10.1111/j.1551-6709.2011.01213.x
Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*(3), 375–419.
Liao, W., Mantini, D., Zhang, Z., Pan, Z., Ding, J., Gong, Q., … Chen, H. (2010). Evaluating the effective connectivity of resting state networks using conditional Granger causality. *Biological Cybernetics*, *102*(1), 57–69. https://doi.org/10.1007/s00422-009-0350-5
Lin., Chuang, C. H., Kerick, S., Mullen, T., Jung, T. P., Ko, L. W., .& McDowell, K. (2016). Mind-wandering tends to occur under low perceptual demands during driving. *Scientific reports*, 6, 21353.)
Logie, R. H. (2011). The Functional Organization and Capacity Limits of Working Memory.

Current Directions in Psychological Science, *20*(4), 240–245.
Luximon, A.,& Goonetilleke, R. S. (2001). Improvement of the Subjective Workload Analysis Technique. *Ergonomics*, *44*, 229–243.
Margaret A. Bowers, James C. Christensen, and F. T. E. (2016). The Effects of Workload Transitions in a Multitasking Environment. *Conference Proceedings HFES*, *298*(704).
Matessa, M. (2008). An ACT-R Representation of Information Processing in Autism ACT-R. *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society*, 2168–2173.
Matthews, G., & Desmond, P. A. (2002). Task-induced fatigue states and simulated driving performance. *The Quarterly Journal of Experimental Psychology Section A*, *55*(2), 695–689.
Matthews, M. L. (1986). The Influence of Visual Workload History on Visual Performance, *28*(6), 623–632.
McClelland, J. L. (2000). The basis of hyperspecificity in autism: A preliminary suggestion based on properties of neural nets. *Journal of Autism and Developmental Disorders, 30*(5):497-502.
McKendrick, R., & Parasuraman, R. (2014). Using functional Near Infrared Spectroscopy (fNIRS) to Evaluate the Neurocognitive Effects of Transient Events: Design Matrix Mixed Effects Analysis. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 235–239). CA: Los Angeles, CA: SAGE Publications.
Miller, W. D. (2010). The U.S. Air Force-Developed Adaptation of the Multi-Attribute Task Battery for the Assessment of Human Operator Workload and Strategic Behavior. Retrieved from http://www.dtic.mil/docs/citations/ADA537547
Morgan, J. F. (2008). Hysteresis effects in driving. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, *69*(6–B), 3885. Retrieved from http://gateway.library.qut.edu.au/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2008-99240-132&site=ehost-live&scope=site
Morgan, J. F., & Hancock, P. A. (2011). The effect of prior task loading on mental workload: An example of hysteresis in driving. *Human Factors*, *53*(1), 75–86. https://doi.org/10.1177/0018720810393505
Moroney, B.W., Warm, J.S., & Dember, W. N. (1995). Effects of demand transitions on vigilance performance and perceived workload. In *the Human Factors and Ergonomics Society 39th annual meeting*.
Mullen, T. (2010). Source Information Flow Toolbox (SIFT) Theoretical Handbook and User Manual. *Swartz Center for Computational Neuroscience*, 1–69.
Mullen, T. R., Kothe, C. A., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., ... & Cauwenberghs, G. (2015). Real-time neuroimaging and cognitive monitoring using wearable dry EEG. *IEEE Transactions on Biomedical Engineering*, *62*(11), 2553-2567.
Nijboer, M., Borst, J., Rijn, H. Van, & Taatgen, N. (2016). Contrasting single and multi-component working-memory systems in dual tasking. *COGNITIVE PSYCHOLOGY*, *86*, 1–26. https://doi.org/10.1016/j.cogpsych.2016.01.003
Oaksford, M. (2002). How does it fit? *Trends in Cognitive Sciences, 6*(10), 412-413.
Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4
Parasuraman, R., Sheridan, T. B., & Wickens, C. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs.

Journal of Cognitive Engineering and Decision Making, *2*(2), 140–160.
Pitt, M. A., Myung, I. J., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review, 109*(3), 472–491
Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59.
Redick, T. S., & Engle, R. W. (2006). Working memory capacity and attention network test performance. *Applied Cognitive Psychology*, *20*(5), 713–721. https://doi.org/10.1002/acp.1224
Reid, G.B. & Nygren, T. E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. In *Human mental workload* (pp. 185–214). Amsterdam: Elsevier.
Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review, 6*(2), 135–147. Elsevier Ltd.
Rosenberg-Lee, Miriam, Marsha C. Lovett, and J. R. A. (2009). Neural correlates of arithmetic calculation strategies. *Cognitive, Affective, & Behavioral Neuroscience, 9*(3), 270–285. https://doi.org/10.3758/CABN.9.3.270
Salvucci, D. D. (2005). A Multitasking General Executive for Compound Continuous Tasks. *Cognitive Science*, *29*(3), 457–492.
Salvucci, D. D. (2006). Modeling Driver Behavior in a Cognitive Architecture. *Human Factors*, *48*(2), 362–380.
Salvucci, D. D., Kushleyeva, Y., & Lee, F. J. (2004). Toward an ACT-R General Executive for Human Multitasking. *ICCM*, 267–272.
Salvucci, D. D., & Taatgen, N. A. (2008). Threaded Cognition : An Integrated Theory of Concurrent Multitasking. *Psychological Review*, *115*(1), 101–130. https://doi.org/10.1037/0033-295X.115.1.101
Samuelson, L. K., Smith, L. B., Perry, L. K., & Spencer, J. P. (2011). Grounding word learning in space. *PLoS ONE*, *6*(12), 1–13.
Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., & Wolpaw, J. R. (2004). BCI2000: A general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering*, *51*(6), 1034–1043. https://doi.org/10.1109/TBME.2004.827072
Scharinger, C., Soutschek, A., Schubert, T., & Gerjets, P. (2015). When flanker meets the n-back: What EEG and pupil dilation data reveal about the interplay between the two central-executive working memory functions inhibition and updating. *Psychophysiology*, *52*(10), 1293–1304. https://doi.org/10.1111/psyp.12500
Schweizer, T. A., Kan, K., Hung, Y., Tam, F., Naglie, G., & Graham, S. J. (2013). Brain activity during driving with distraction: an immersive fMRI study. *Frontiers in Human Neuroscience*, *7*(February), 1–11. https://doi.org/10.3389/fnhum.2013.00053
Seth, A. K. (2005). Causal connectivity of evolved neural networks during behavior. *Network*, *16*, 35–54.
Sharbrough, F., Chatrian, G., Lesser, R., Luders, H., Nuwer, M., & Picton, T. (1991). American Electroencephalographic Society Guidelines for Standard Electrode Postition Nomenclature. *Clinical Neurophysiology*, (8), 200–202.
Smart, P. R., Scutt, T., Sycara, K., & Shadbolt, N. R. (2016). Integrating ACT-R Cognitive Models with the Unity Game Engine, 74–86.
Sridharan D, Levitin DJ, M. V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proc Natl Acad Sci USA*, *105*, 12569–12574.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *Amerian Scientist*, *57*(4), 421–457.
Stevens MC, Pearlson GD, C. VD. (2009). Changes in the interaction of resting-state neural networks from adolescence to adulthood. *Human Brain Mapping*, *30*, 2356–2366.
Sun, R. (2006). The CLARION cognitive architecture: Extending cognitive modeling to social simulation. In *Cognition and multi-agent interaction* (pp. 79–99).
Sun, R. (2007). Cognitive Social Simulation Incorporating Cognitive Architectures. *IEEE Intelligent Systems*, *22*(5).
Taatgen, N. (2005). Modeling Parallelization and Flexibility Improvements in Skill Acquisition : From Dual Tasks to Complex Dynamic Skills. *Cognitive Science*, *29*, 421–455.
Taatgen, N. A., & Lee, F. J. (2003). Production Compilation : A Simple Mechanism to Model Complex Skill Acquisition, *45*(1), 61–76.
Taatgen, N., & Anderson, R. (2010). The Past , Present , and Future of Cognitive Architectures, *2*, 693–704. https://doi.org/10.1111/j.1756-8765.2009.01063.x
Taatgen, N., Lebiere, C., & Anderson, J. (2006). Modeling paradigms in ACT-R. In *Cognition and multi-agent interaction: From cognitive modeling to social simulation* (pp. 29–52).
Tattersall, A. J., Foord, P. S., Tattersall, J., & Foord, S. (2007). An experimental evaluation of instantaneous self- assessment as a measure of workload, *139*(May). https://doi.org/10.1080/00140139608964495
Thomas, L. E. N., & Juanes, F. (1996). The importance of statistical power analysis : an example from Animal Behaviour, 856–859.
Thornton, D. C. (1985). An Investigation of the "Von Restorff" Phenomenon in Post-Test Workload Ratings. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *29*(8), 760–764.
Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154. https://doi.org/10.1016/0749-596X(89)90040-5
Ungar, N. R. (2005). *DEMAND TRANSITION, TRACKING ACCURACY, AND STRESS: RESOURCE-DEPLETION AND -ALLOCATION MODELS*. University of Cincinnati.
Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*(1), 104–132.
Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498–505. https://doi.org/10.3758/BF03192720
Usher, Marius, and J. L. M. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological Review*, *108*(3), 550.
Vugt, M. K. Van. (2014). NeuroImage Cognitive architectures as a tool for investigating the role of oscillatory power and coherence in cognition. *NeuroImage*, *85*, 685–693. https://doi.org/10.1016/j.neuroimage.2013.09.076
Vugt, M. K. Van, & Groningen, A. G. (2012). Relating ACT-R buffer activation to EEG activity during an attentional blink task. In *Nele Rußwinkel| Uwe Drewitz| Hedderik van Rijn (eds.)* (p. 218).
Watson, J. M., Bunting, M. F., Poole, B. J., & Conway, A. R. A. (2005). Individual differences in susceptibility to false memory in the paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 76–85. https://doi.org/10.1037/0278-7393.31.1.76
Wickens, C. D., Lee, J., & Becker, S. G.(1998). *An introduction to human factors engineering*

Second Edition.
Wiener. (1958). *Nonlinear problems in random theory*. MIT press.
Wilson, G. F., & Russell, C. A. (2007). Performance enhancement in an uninhabited air vehicle task using psychophysiologically determined adaptive aiding. *Human Factors*, *49*(6), 1005–1018.
Witkowski, M., Tomczak, M., Karpowicz, K., Solnik, S., & Przybyla, A. (2019). Effects of fencing training on motor performance and asymmetry vary with handedness. *Journal of motor behavior*, 1-8.
Wong, T. J., Cokely, E. T., & Schooler, L. J. (2010, August). An online database of ACT-R parameters: Towards a transparent community-based approach to model development. In *Proceedings of the 10th international conference on cognitive modeling* (pp. 282-286). Philadelphia, PA: Drexel University.
Zhou, Dongli, Wesley K. Thompson, and G. S. (2009). MATLAB Toolbox for functional connectivity. *Neuroimage*, *47*(4), 1590–1607. https://doi.org/10.1016/j.neuroimage.2009.05.089.MATLAB

APPENDICES

Appendix A: Demographic questionnaire

Participant ID: _____

Gender: ____Male ____Female

Age: _____

Handedness: ____ Right ____ Left

Native/Primary language: ____ English ____ Other

1. Are you in overall good health?

2. Do you have any previous experience with cognitive dual task and MATB or similar systems?

Appendix B: Handedness survey

Please indicate with a one (1) your preference in using your left or right hand in the following tasks. Where the preference is so strong you would never use the other hand, unless absolutely forced to, put a two (2).

If you are indifferent, put a one in each column (1 | 1).

Some of the activities require both hands. In these cases, the part of the task or object for which hand preference is wanted is indicated in parentheses.

Task / Object	Left Hand	Right Hand
1. Writing		
2. Drawing		
3. Throwing		
4. Scissors		
5. Toothbrush		
6. Knife (without fork)		
7. Spoon		
8. Broom (upper hand)		
9. Striking a Match (match)		
10. Opening a Box (lid)		
Total checks:	LH =	RH =
Cumulative Total	CT = LH + RH =	
Difference	D = RH – LH =	
Result	R = (D / CT) × 100 =	
Interpretation: (Left Handed: R < -40) (Ambidextrous: -40 ≤ R ≤ +40) (Right Handed: R > +40)		

Appendix C: Instruction of automated OSPAN task

The automated OSPAN task includes items (letters) to remember and a distracting activity in the form of math problem solving (Unsworth et al., 2005). There are three practice sessions (letter span, math problem, both of them combined) and one experimental session.

1. Practice Session I

Letters appear one at a time on the screen and you are required to remember the letters in the presented order. At recall, you will see a 4 X 3 matrix of letters (F, H, J, K, L, N, P, Q, R, S, T, and Y). After recall, the program provides feedback about the number of letters correctly recalled in the current set.

2. Practice Session II

This session involves simple mathematical problem solving. You need to solve several True/False math problems (e.g., (1*2) + 1= 3 True or False). During the task, the computer will calculate the mean time required to solve the equations for you.

3. Practice Session III

During this session, you need to perform both the letter recall and math portions together. First, a letter will be presented. After a few seconds the letter disappears and a math problem is presented. The math problem is then replaced with a "True/False" prompt. Once the you respond "True" or "False", another letter appears on the screen and the sequence is repeated.

4. OSPAN Test Trial

After completion of practice sessions, the program will progress to the real trials similar to the final session of practice set. The experimental trial consists of 75 letters and 75 math problems, broken down into smaller segments with a maximum segment length of 7 sets. There are 75 trials in the experimental session and score range of this test is 0 – 75.

At the end of the task, the program will report five scores: OSPAN score, total number correct, math errors, speed errors, and accuracy errors. The OSPAN score, uses the traditional absolute scoring method: the sum of all perfectly recalled sets. The second score, “total number correct,” is the total number of letters recalled in the correct position. Three types of errors are reported. “Math errors” represent the total number of task errors, which is then broken down into “speed errors” and “accuracy errors”. ”Speed errors” quantify the amount of times participants run out of time in attempting to solve a given math operation. “Accuracy errors” represent the number of incorrectly solved math operations.