

**The Unity and Diversity of Executive Function as shown by Resting-State Functional
Connectivity: An fNIRS Replication and Methods Comparison Study**

A Thesis Presented for the
Master of Arts
Degree
The University of Tennessee, Knoxville

Haylee L. Hudson

May 2025

ABSTRACT

This study is, first and foremost, a direct replication of “The unity and diversity of executive functions and their contributions to complex ‘frontal lobe’ tasks: A latent variable analysis” (Miyake et al., 2000). We recreated the same nine simple tasks, three tasks each for three “components” of executive function, “shifting,” “inhibition,” and “updating.” We also used the same five complex executive (or “frontal lobe”) tasks. First, we used a confirmatory factor analysis to replicate the factor structure of executive functioning identified by Miyake et. al. (2000). Then, we used structural equation modeling to replicate the findings that each complex executive task has one or more component of executive function mapping onto it. We then expanded upon this original study by comparing a *Passive Viewing Paradigm* and *Mind-Wandering Paradigm* for resting-state functional connectivity data collection with functional near-infrared spectroscopy. We compared the functional connectivity profiles between these two resting-state methods, and we used group-level correlation analyses and principal components analysis to better understand the structure of EF from a neurocognitive perspective. Ultimately, we did not replicate the findings from the original EF study but rather endorsed a model where shifting and inhibition are equal. We then used this model to differentially assess how the components map onto the complex EF tasks. Using the neural data, we found that the two resting-state measures are quite different, especially when using them to assess how this data correlates with EF outcomes. Executive function is not as simple as a set of components such as shifting, updating, and inhibition, but rather a complex system in the brain that leads to behavioral outcomes defined as EF skills. From a neurocognitive perspective, EF measures cannot be easily structured into three separable components with an underlying common EF factor, nor can the behavioral structure be easily replicated twenty-five years later.

TABLE OF CONTENTS

Chapter 1: Introduction	1
Chapter 2: Methods.....	6
Participants.....	6
Measures	6
Neural Measures	7
Behavioral Measures.....	8
Shifting.....	8
Inhibition.....	13
Updating.....	15
Complex.....	19
Research Design.....	22
Chapter 3: Procedures	24
Chapter 4: Analytic Approach.....	25
Chapter 5: Results	30
Replication Study.....	30
Preliminary Analyses	30
Confirmatory Factor Analysis.....	34
Structural Equation Model.....	39
Methods Comparison Study.....	49
Executive Function and rsFC Study	54
Chapter 6: Discussion	70
Overall Conclusions.....	70

Limitations	73
Future Directions	74
References.....	75
Vita.....	79

LIST OF TABLES

Table 2.1	6
Table 5.1	31
Table 5.2	32
Table 5.3	33
Table 5.4	33
Table 5.5	35
Table 5.6	40
Table 5.7	41
Table 5.8	42
Table 5.9	43
Table 5.10	44
Table 5.11	45
Table 5.12	47
Table 5.13	48
Table 5.14	49
Table 5.15	50
Table 5.16	55
Table 5.17	59
Table 5.18	63
Table 5.19	66
Table 5.20	67
Table 5.21	68

Table 5.22.....	68
-----------------	----

LIST OF FIGURES

Figure 2.1	8
Figure 2.2	9
Figure 2.3	9
Figure 2.4	11
Figure 2.5	12
Figure 2.6	14
Figure 2.7	14
Figure 2.8	16
Figure 2.9	17
Figure 2.10	19
Figure 2.11	20
Figure 2.12	22
Figure 4.1	26
Figure 4.2	28
Figure 5.1	37
Figure 5.2	38
Figure 5.3	51
Figure 5.4	51
Figure 5.5	53
Figure 5.6	54
Figure 5.7	57
Figure 5.8	58

Figure 5.9	61
Figure 5.10	62

Chapter 1: Introduction

Executive function (EF) can be defined as a set of cognitive skills or processes that are critical for regulating cognitive function and engaging in goal-directed behavior. A seminal study in this field by Miyake et al. (2000) defined EF in terms of its components, specifically shifting, updating, and inhibition, which they deemed the Unity and Diversity model of executive function. The authors used a set of behavioral tasks for each component, shifting (switching between cognitive processes, otherwise called cognitive flexibility), updating (dynamic use of working memory), and inhibition (specifically inhibiting a prepotent response). A confirmatory factor analysis model revealed three correlated components of EF with a common underlying EF factor. The authors then used structural equation modeling (SEM) to parse out the unique contribution of EF components to five commonly used complex “frontal lobe” or “executive” tasks that are used to measure and discuss overall EF in clinical and research applications (Miyake et al., 2000). The rationale for this approach stemmed from the lack of validation in the literature on the relationship between these cognitive processes and how the various existing EF measures tap into these relationships. The authors demonstrated that the three separable components of EF are shown to be assessed differentially by the complex EF tasks. Specifically, they found that the Wisconsin Card Sorting Test assesses shifting, the Tower of Hanoi task assesses inhibition, the Random Number Generation task assesses inhibition and updating differentially, the Operation Span task assesses updating, and the Dual Task was shown to not assess any of these three EF components. Overall, Miyake et al. (2000) demonstrated that EF is organized into separable components sharing a common underlying factor and that these components differentially contribute to complex EF, behaviorally.

The Miyake et al. (2000) model of and definition for EF has been cited repeatedly since its initial publication. Despite the widespread adoption of this model over the last twenty years, no study has attempted to replicate. The first goal of the current investigation was to provide a direct replication and extension of the Miyake et al. (2000) model of EF. Specifically, this study examined whether the unity and diversity model of EF could be replicated behaviorally as these conclusions have been important in the field of EF research. Then, we expanded on these results by examining how this structure of EF is organized functionally in the brain by analyzing resting-state functional connectivity profiles. There were two motivations beyond a lack of direct replication behind this study design. First, defining EF in this way becomes a circular, descriptive definition without a tangible basis (McCraw et. al., 2024). What this means is that, when asked “what is EF?”, this definition only defines it in terms of other cognitive processes (or components). This creates a lack of true definition or examination of EF overall and how these processes come together to create more advanced EF outcomes (such as planning and goal-oriented behavior). By not only assessing the ability to replicate this behavioral/statistical component structure of EF but also attempting to base the structure in neurocognitive terms as well, there is more of an ability to redefine EF itself outside of its component processes. Second, there was a lack of statistical power for the level of significance of all parameters and estimates in the original study. With a lack of statistical power, it is reasonable to question whether this study can be replicated with new data. By conducting this study and expanding on the results from the replication, there may be renewed ability to break from patterns in current EF research in order to develop more understanding of the neurocognitive and developmental processes underlying EF.

Resting-state functional connectivity (rsFC) refers to synchronous activation of spatially remote neural populations specifically measured at rest (Lu, 2010). This can then be used as an intrinsic or a baseline measure of neural structure that can then be related to other measures, for the purposes of this study, behavioral EF measures. However, before discussing the relation between rsFC neural profiles to measures of EF, one must acknowledge the literature gap related to rsFC data collection methods. rsFC data has been used in many studies for many applications, but the methods of collecting this data vary drastically between studies (Bulgarelli, 2019; Eng, 2022; Fernandez, 2022; Hu, 2020; Vanderwal, 2015). Some studies use a video or *Passive Viewing* paradigm where participants are instructed to quietly and calmly watch a set of videos or a movie, and some studies use variations of a *Mind-Wandering* paradigm where participants (with their eyes either open or closed) are instructed to sit quietly and “let their minds wander.” One fMRI study showed better compliance with a video paradigm called INSCAPES with limited differences between the two methods (Vanderwal, 2015), but this still left open many questions about what the consensus should be in adults. This problem is exacerbated when discussing development because “rest” is so undefined and could change so rapidly with age in children (Camacho, 2020).

This study used functional near-infrared spectroscopy (fNIRS) to directly compare a simple *Passive Viewing* and *Mind-Wandering* paradigm in adults, which is the first step to comparing these methods in younger, developing populations. The guiding hypothesis was that we would see better rsFC data using the *Mind-Wandering* paradigm despite better compliance with the *Passive Viewing* paradigm in adults. The data with the *Passive Viewing Paradigm* should show more activity because the participant is receiving more stimulation than they are during the *Mind-Wandering Paradigm*. However, this level of stimulation may be helpful for

participants to remain still, calm, and restful, which may produce less noise (movement and physiological) when compared to the *Mind-Wandering Paradigm*. In adults, these differences may be minimal, but the hypothesis is that the endorsed method will be the *Mind-Wandering Paradigm*, which will then be used in conjunction with the behavioral EF data as follows.

The third goal of this study was to use behavioral EF data and the rsFC data to elucidate more clearly the structure of EF in the brain. Previous literature has shown that there is a relationship between rsFC neural profiles and EF. This study expanded on these results by examining this relationship in the context of the unity and diversity model of EF, specifically. Overall, the literature suggests an association between rsFC and executive function, but the results are varied. Studies have found that expanded rsFC networks (Reineberg, 2015), connectivity between the sensory and default networks (Reineberg, 2018), and increased connectivity mainly within the pre-frontal cortex (Zhao, 2016) are correlated with higher scores and better EF outcomes. Also, a common theme among the literature is that the frontoparietal network is often implicated when comparing rsFC to EF outcomes (Reineberg, 2016). However, studies often find variation in the strength of these associations when examining different components of EF as well as large effects of individual differences among the results (Roye, 2020; Reineberg, 2016).

Considering previous research, the hypothesis guiding this study was that there would be clear correlations between better EF outcomes in all three components and differential connectivity profiles within frontoparietal networks. More specifically, we looked to see if we could elucidate a common neural profile that defines the currently undefined “common factor” between the three EF components as shown by Miyake et al. (2000). Using the rsFC data, we set out to redefine the structure of EF outside of behavior to base this structure in neurocognitive

terms. Using rsFC data to do this can guide regions of interest in further studies in order to better understand what guides EF in the brain. Ultimately, this study provides answers to many questions in the field of EF research, which will guide us to begin understanding these ideas more clearly in developing populations in the future. These questions about EF structure, both behavioral and neural, still need a lot of research. However, these ideas are even less well defined in developing and clinical populations. By developing a better understanding of resting-state methods as well as by developing a clearer definition of EF structure, we will have a better ability to assess these processes between populations and age groups.

Chapter 2: Methods

Participants

Participants were recruited from the student population at the University of Tennessee in Knoxville. A total of sixty-seven participants participated in the study, one participant did not complete the Wisconsin Card Sorting Test and Random Number Generation task, so these two tasks had a sample size of sixty-six. From this sample, a total of sixty-four participants were included in the fNIRS, resting-state functional connectivity analysis. Participants enrolled in psychology courses at the University received research credits as compensation for their participation in the study, and some participants volunteered to participate as well. All participants were between eighteen and thirty years old. Excluded from this study were individuals who were colorblind or left-handed. A breakdown of the study sample demographics is shown in **Table 2.1**.

Measures

This study consisted of two resting-state tasks with functional connectivity neural measures and fourteen tasks with behavioral measures. All computerized tasks were created using MATLAB and PsychToolbox.

Table 2.1: *Demographics breakdown of study sample-count of participants (mean age, years).*

Sex	Ethnicity	Race					N/A	Total
		American Indian/Alaska Native	Asian	Black/ African American	Caucasian	Multiracial		
Female 45(19.6)	Hispanic/Latino	1(18)	0	0	2(20.5)	0	0	3(19.7)
	Non- Hispanic/Latino	0	2(20)	0	37(19.6)	1(18)	2(19)	42(19.6)
Male 21(20.1)	Hispanic/Latino	0	1(20)	0	0	0	0	1(20)
	Non- Hispanic/Latino	0	2(19)	1(19)	13(19.8)	1(18)	1(21)	18(20.2)
N/A 1(20)	N/A	0	0	0	1(21)	0	1(19)	2(20)
	Non- Hispanic/Latino	0	0	0	1(20)	0	0	1(20)
Total		1(18)	5(19.6)	1(19)	54(19.7)	2(18)	4(19.5)	67(19.8)

Neural Measures

All neural activity was measured using functional near-infrared spectroscopy (fNIRS). We used a TechEn CW7 fNIRS system with a probe that has eight sources and sixteen detectors measuring activity from twenty-four channels within the frontal, parietal, and temporal cortices. In fNIRS, near-infrared light is emitted from the light sources, travels through the skull into the cortex, and the detected light is collected. Using Beer Lambert's law, this light is transformed into a measure of both oxygenated and deoxygenated hemoglobin, which is an indirect measure of brain activity (higher oxygenated hemoglobin correlates to more activity in that cortical region). For functional connectivity analysis, the synchronous level of oxygenated hemoglobin between channels provides a measure of connectivity between the two separate neural populations at the same point in time.

Passive Viewing Paradigm: The *Passive Viewing Paradigm* (Kerr-German et. al., 2022) is a resting-state measure that requires participants to watch a series of five videos (stills from each video are shown in **Figure 2.1**) accompanied by classical music with short breaks in between while wearing the fNIRS cap. Participants were told to “remain calm, still, and restful while watching the series of videos.” Each video is approximately one minute in length, providing over four minutes of resting-state neural data.

Mind-Wandering Paradigm: The *Mind Wandering Paradigm* was designed in such a way that it would serve as a direct comparison to the *Passive Viewing Paradigm* measure using fNIRS. In this task, after a fixation, the screen turns green, at which point, the participants were told to “close their eyes and let their minds wander.” The screen returns to fixation at which point the participant was told that they could open their eyes and stop resting. The duration of the rest will increase from thirty seconds to a minute, then to two minutes (for two trials). This task is

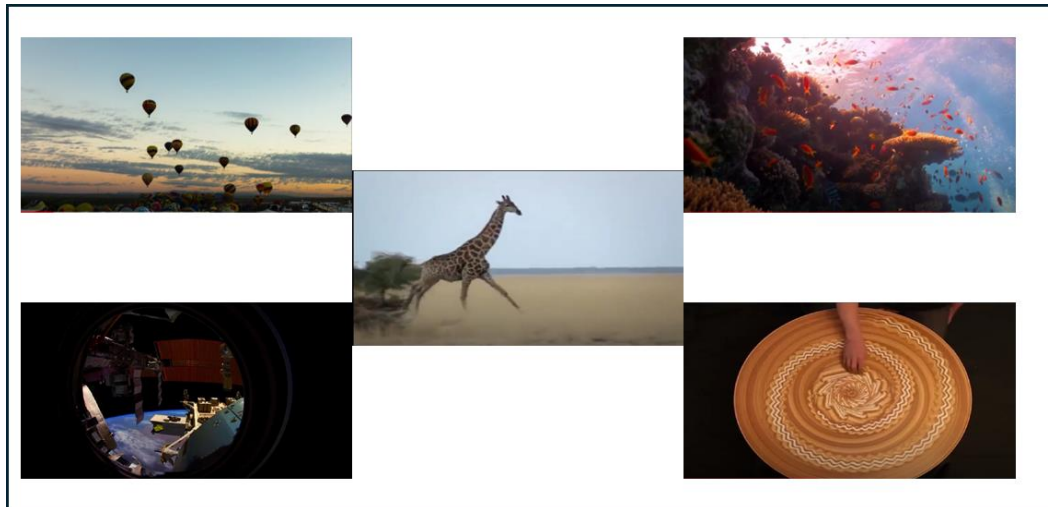


Figure 2.1: *Still Images from the Five Videos in the Passive Viewing Paradigm.*

depicted in **Figure 2.2**. This will give a total of five and a half minutes of resting-state neural data. Looking towards future use of this task with younger populations, this design was meant to create a game with breaks to help with compliance in the task overall. This task was deliberately created to be slightly longer than the *Passive Viewing Paradigm* due to the expectation that compliance would be more difficult leading to more movement and therefore noise in the data.

Behavioral Measures

All the following behavioral tasks were run without collecting fNIRS data. Each task was designed to be as close to the original tasks used by Miyake et. al. (2000), and the additional references were those used by Miyake et. al. (2000) as well. There are four categories of tasks as outlined in the introduction, which are “Shifting,” “Inhibition,” “Updating,” and the complex EF tasks.

Shifting

Plus-Minus Task: The *plus-minus task* (**Figure 2.3**) is a computerized task where the participant is shown a series of thirty, two-digit numbers between ten and ninety-nine. There are three blocks; in the first block, participants add three to each number they see, in the second

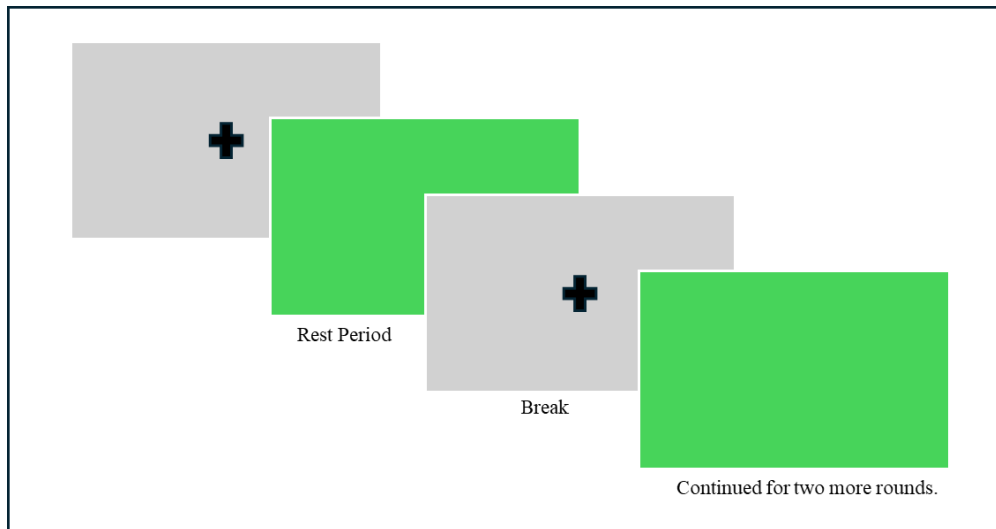


Figure 2.2: Task Design for the Mind-Wandering Paradigm (First Two Rest Periods Shown).

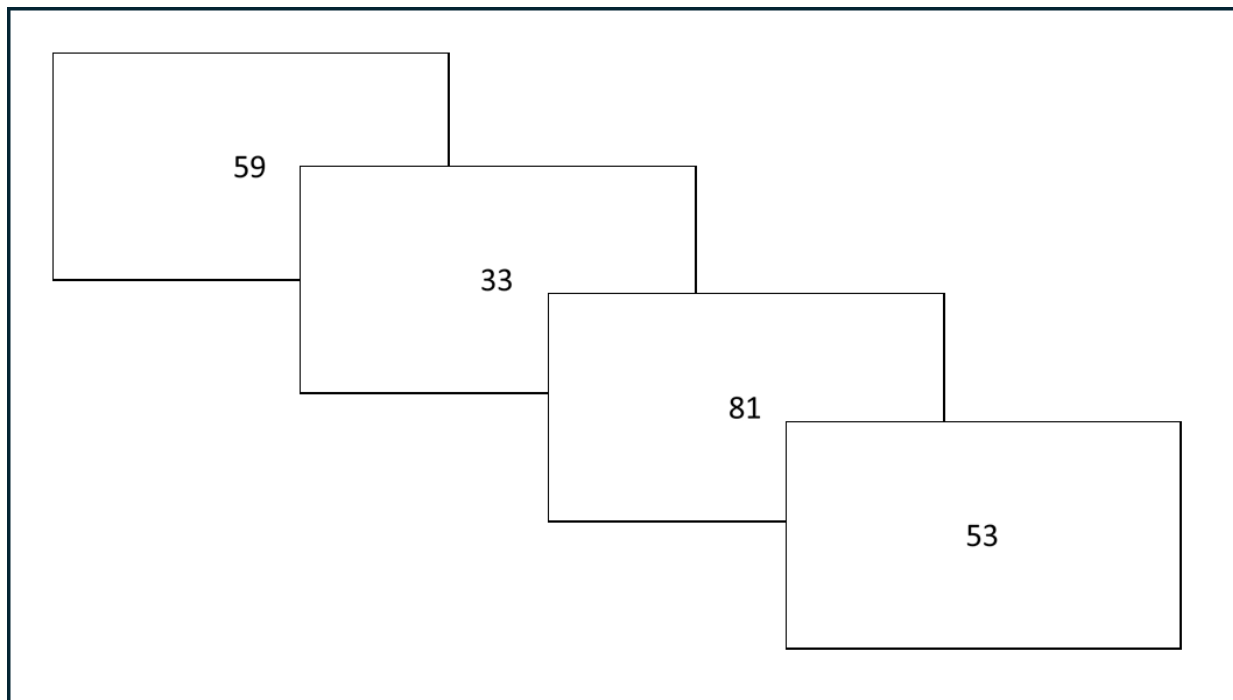


Figure 2.3: Four Trials in a Block for the Plus-Minus Task.

block, they subtract three, and in the third block, they alternate adding and subtracting three for each number. Each time, they verbally state their answer, and the experimenter marks each as correct or incorrect. Accuracy and reaction time were recorded, and the main dependent measure was the difference between the duration to completion of the third block (the shifting block) and the average duration to completion on the first and second blocks giving a measure of shifting cost.

Number-Letter Task: The *number-letter task* (**Figure 2.4**) is a computerized task in which the participants are shown a number-letter pair in one quadrant of the screen, and they are told to indicate whether the number is even or odd if the pair is shown at the top of the screen and whether the letter is a consonant or vowel if presented on the bottom of the screen. The respond by pressing either the “e,” “o,” “c,” or “v” on the keyboard. The first block is thirty-two trials only in the top of the screen with ten practice trials, the second is thirty-two trials only in the bottom of the screen with ten practice trials, and the third rotates clockwise around the screen for one hundred and twenty-eight trials with twelve practice trials. The numbers used are 2, 4, 6, 8, 3, 5, 7, or 9, and the letters used are G, K, M, R, A, E, I, or U. Accuracy and reaction time were recorded, and the main dependent measure was the difference in average reaction time on the switch trials in the third block and the average reaction time on the first two blocks.

Local-Global Task: The *local-global task* (**Figure 2.5**) is a computerized task that consists of the participants being shown a series of Navon figures where a large shape is made of smaller shapes (Navon, 1977). The shapes were either blue or black, and if blue, the participants responded by verbally stating the number of edges on the large shape, and if black, they verbally stated the number of edges on the small shapes, and the experimenter typed the number that the participant says. The shapes are circles (one edge), x’s (two edges), triangles (three edges),

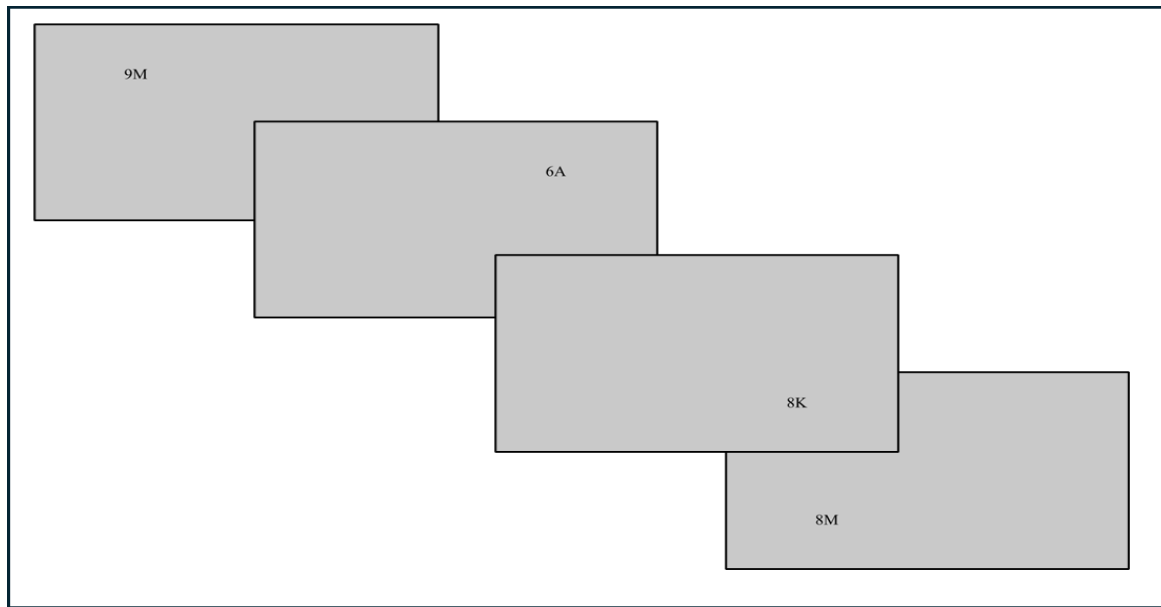


Figure 2.4: Set of Four Trials in the Number-Letter Task.

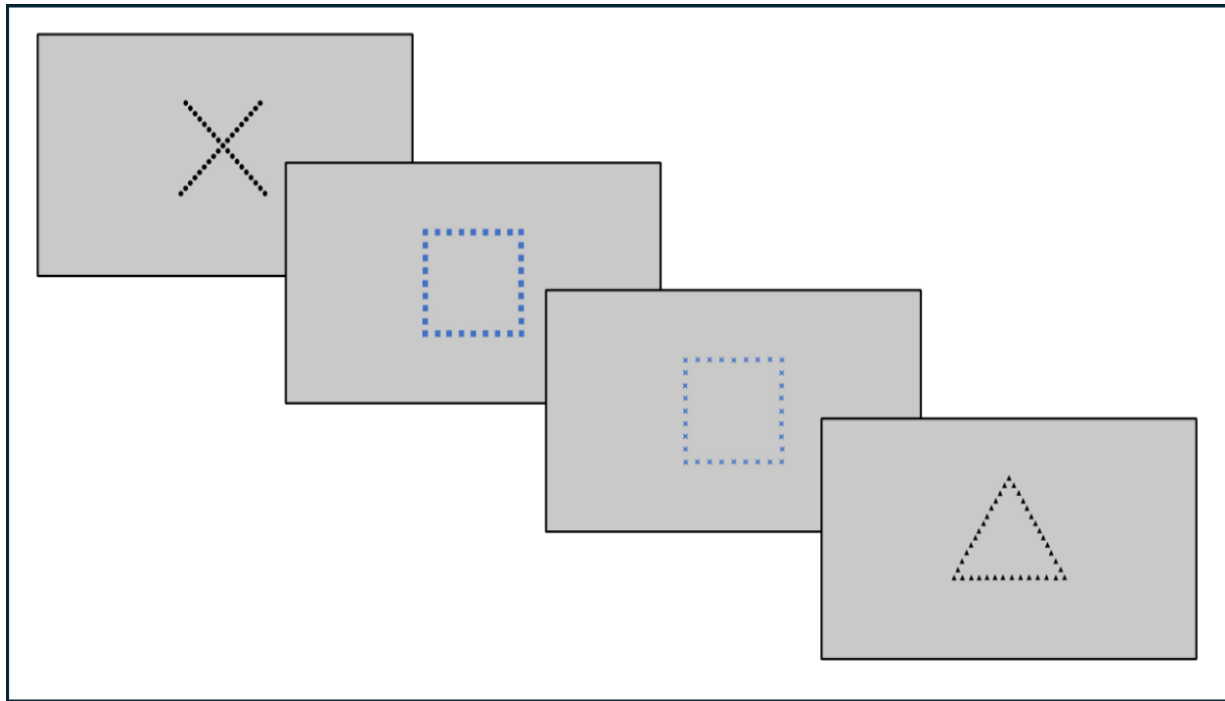


Figure 2.5: *Four Trials (Including Two Switch Trials) of the Local-Global Task.*

or squares (four edges). There is a practice block of thirty-six trials and a target block of ninety-six trials. Half of the trials were switch trials in each block (change of color). Accuracy and reaction time were recorded, and the main dependent measure was the difference between the average reaction time on the switch trials and the non-switch trials.

Inhibition

Antisaccade Task: The *antisaccade task* (**Figure 2.6**) is a computerized task in which participants are shown a fixation for a varied amount of time. They then see a distractor (a black square) on one side of the screen (left or right), which they should inhibit themselves from looking to, followed by a target arrow inside an open square on the opposite side of the screen (both presented 3.4 inches from the center of the screen). The time of the fixation is either 1500 ms, 1750 ms, 2000 ms, 2250 ms, 2500 ms, 2750 ms, 3000 ms, 3250 ms, or 3500 ms, which is varied randomly throughout the task. The distractor is presented for 225 ms, and the target stimulus is presented for 150 ms before being covered by a gray square. The participant will respond by pressing the arrow on the keyboard that corresponds to the direction the target stimulus (the arrow) is pointing (either up, left, or right). There is a block of twenty-two practice trials followed by a block of ninety target trials. Accuracy and reaction time were recorded, and the main dependent measure was the proportion of correctly answered target trials.

Stop-Signal Task: The *stop-signal task* (**Figure 2.7**) is a computerized task where participants see a series of words that are either animals or not animals, and they will respond by pressing either “a” or “o” on the keyboard, respectively. Some trials will have a beep sound as the stimuli is presented, and when this happens, the participant will be told that they should inhibit their response. They first do a practice block of forty-eight trials with no beeps, then thirty-four practice trials with beeps, and then they do a full block of one hundred and

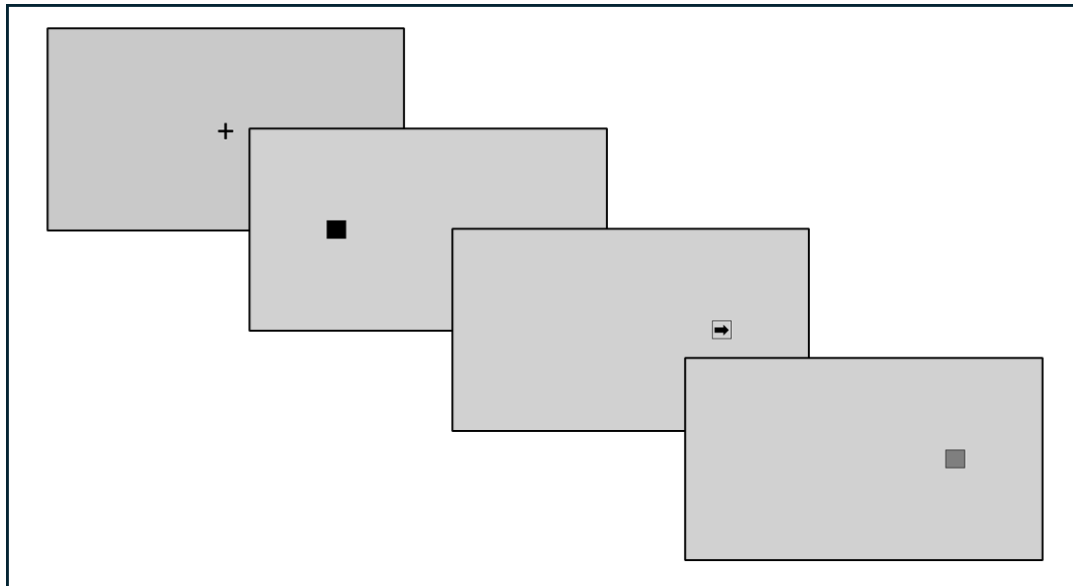


Figure 2.6: *One Trial in the Antisaccade Task.*

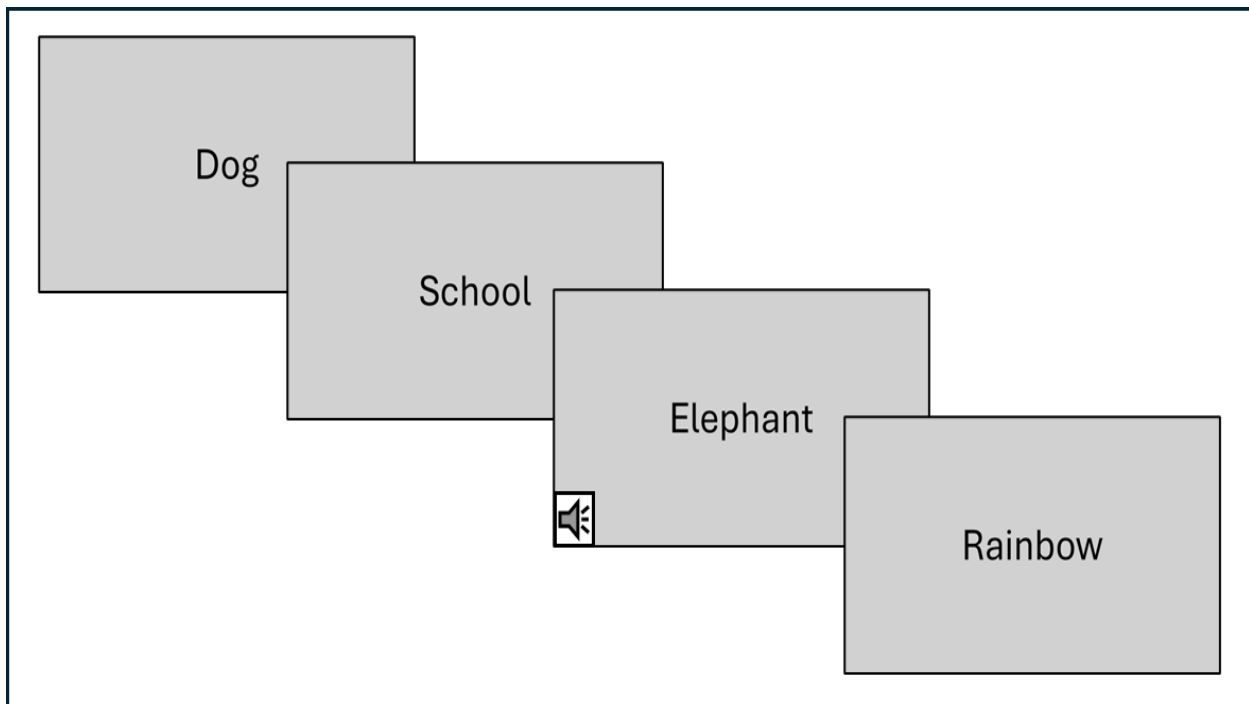


Figure 2.7: *Four Trials of the Stop-Signal Task Showing One Stop Trial.*

ninety-two trials with a beep on forty-eight of these trials. Each stimulus is shown for 1500 ms before switching to the next stimulus regardless of trial type. Accuracy and reaction time were recorded, and the main dependent measure was the proportion of correct “no response” on stop trials (with beeps).

Stroop Task: The *Stroop task* (**Figure 2.8**) is a computerized task where the participant sees multiple types of stimuli. One is a string of asterisks in various colors (red, orange, yellow, green, blue, or purple), the second is a color word written in the same color (congruent trials), and the third is a color word written in any of the other colors (incongruent trials). For each trial, the participant verbally states the color of the stimulus, and the experimenter presses the key corresponding to that color. There are three practice blocks of ten trials each and then one hundred and forty-four target trials (seventy-two asterisk trials, sixty incongruent trials, and twelve congruent trials, all non-blocked). Accuracy and reaction time were recorded, and the main dependent measure was the difference in reaction time between the incongruent and asterisk trials.

Updating

Keep Track Task: The *keep track task* (**Figure 2.9**) is a computerized task where participants see a series of words that belong to six different target categories (relatives, metals, distances, countries, colors, and animals). There are two to three words that belong to each category that are presented in a random order of fifteen. At the beginning of the task, the participant is shown a screen with all words belonging to each target category and given as much time as they want to study the list. They begin with a practice block with three target categories, then three blocks with four target categories, then three blocks with five target categories. On each trial, the word was shown for 1500 ms, and the target categories remained at the bottom

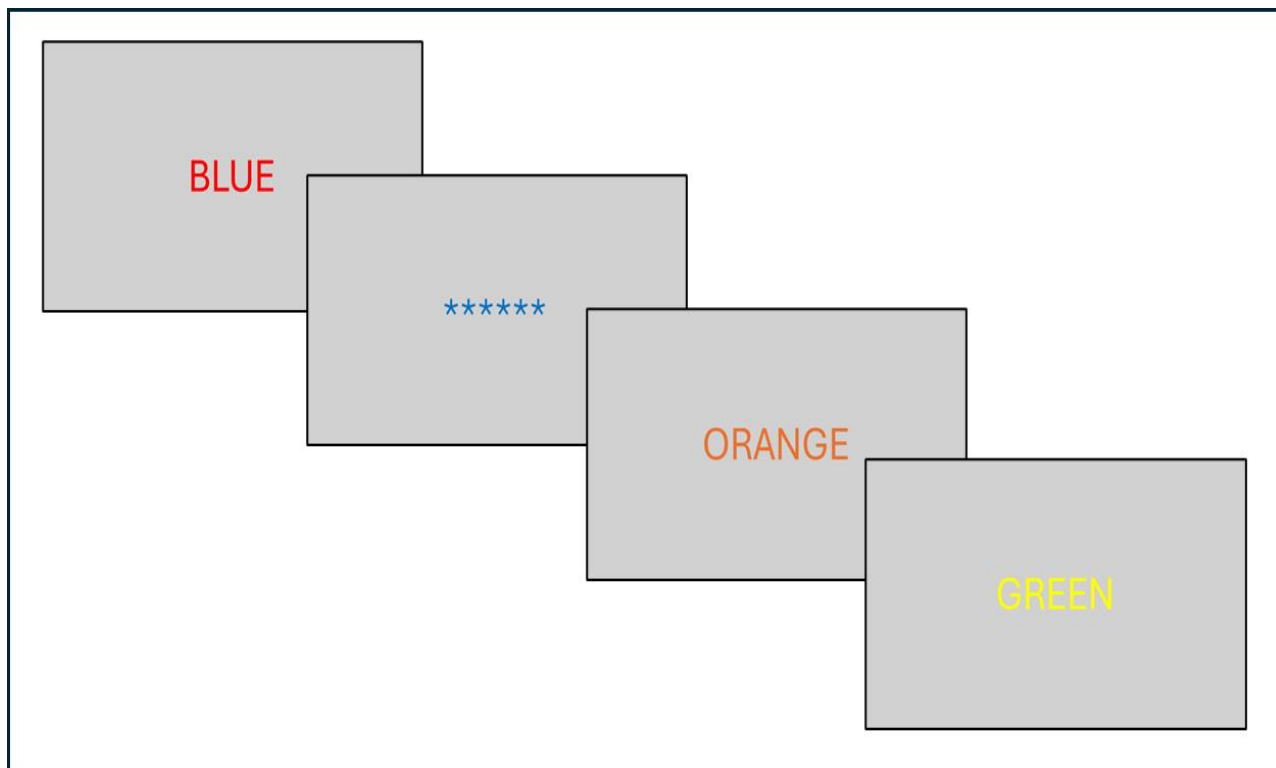


Figure 2.8: *Four Trials (of Each Type) of the Stroop Task.*

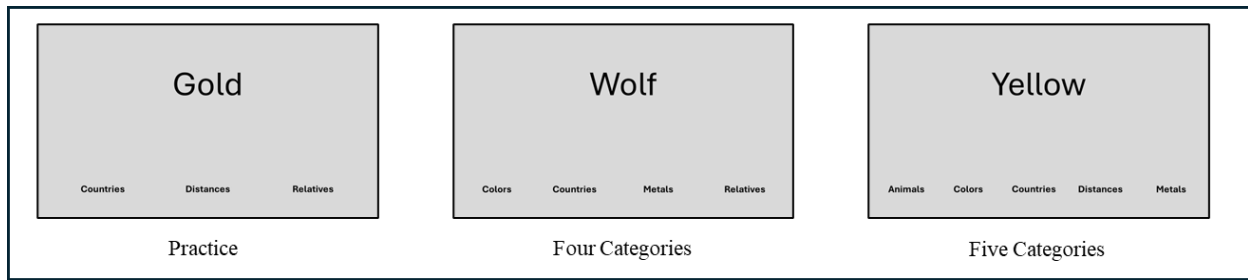


Figure 2.9: Screens from Three Different Blocks in the Keep Track Task.

of the screen the entire time. The participant must remember the last word presented that belongs to each target category, and they verbally recalled the words to the experimenter at the end of each block. There was a total of twenty-seven words to recall, and the proportion of correctly recalled words was the main dependent measure.

Tone Monitoring Task: The *tone monitoring task* is a computerized task where participants see a fixation while hearing a series of beeps at different pitches (880 Hz, 440 Hz, and 220 Hz). Each beep is presented for 500 ms, and they are presented 2500 ms apart. They first completed a guided practice block of fourteen trials, then a practice block of twenty-five trials, and then four target blocks of twenty-five trials.

Their goal was to respond by pressing a key each time the fourth of each pitch sounds. In the twenty-five trials, there are eight of each pitch and one added (of any of the pitches) randomly mixed (for the first, guided practice block, there are four of each pitch with one additional mixed in). If the participant is incorrect, they will see a red screen, and the count will restart (which they are made aware of at the beginning of the task). Overall, there were twenty-four possible correct responses, and, with accuracy and reaction time being recorded, the main dependent measure was the proportion of correct responses.

Letter Memory Task: The *letter memory task* (**Figure 2.10**) is a computerized task where participants see a series of consonants (Morris & Jones, 1990) in series of either five, seven, nine, or eleven letters. The participants must remember and verbally recall the last four (or fewer if four had not yet been presented) letters that have been presented after each letter, and at the end of each set of letters, they recalled the final four letters in that set to the experimenter. Each letter was presented for 2000 ms. They completed a practice set of five and then another of seven letters, and then they completed twelve target sets of randomly varying lengths. There were

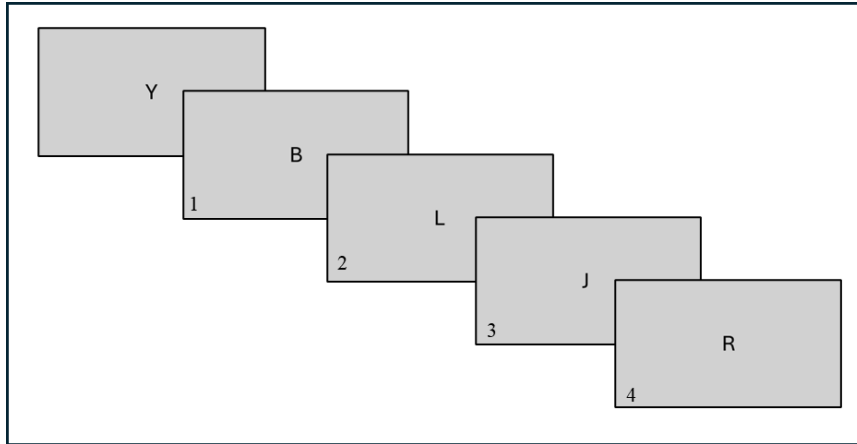


Figure 2.10: *One (Five-Letter) Series in the Letter Memory (Response is Marked).*

forty-eight possible correct letters to recall, and the main dependent measure was the proportion of correct letters recalled overall.

Complex

Wisconsin Card Sorting Test: The *Wisconsin Card Sorting Test* (**Figure 2.11**) is a computerized task in which participants are shown a series of target “cards” with four reference “cards” shown above them. Each “card” has a certain number (one, two, three, or four) of shapes (circle, cross, square, or star) of a certain color (red, blue, yellow, or green). Each target will match by one dimension to three of the reference cards, and one reference card will not match at all. However, only one type of match (number, shape, or color) is correct for each target card. The participant must indicate which reference card is the match by pressing one, two, three, or four on the keyboard corresponding to that card. They briefly see a screen that tells them whether they are correct or incorrect, and they will use that feedback to figure out the sorting rule. After eight correct responses, the rule will change, but they will not be told when the rule changes (again using the feedback to figure out the new rule). If a response was not given in three seconds, they were told they timed out, and this counted as an incorrect trial. After a practice block of thirty trials, they continued until they correctly sorted fifteen sets. The main dependent

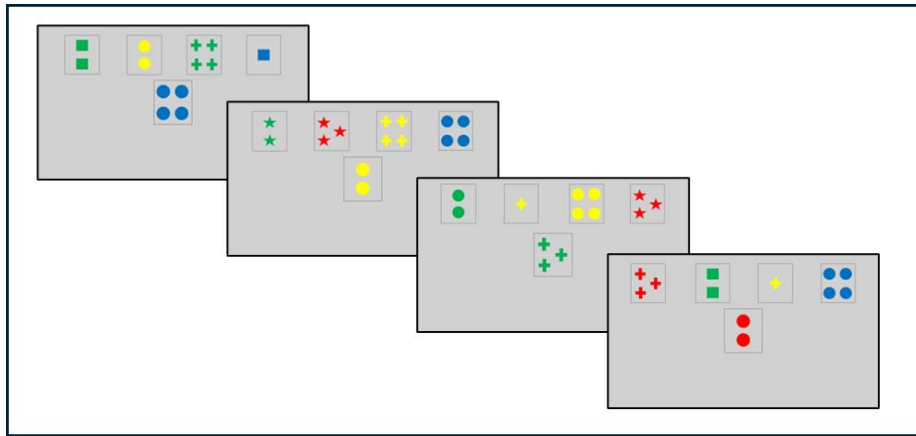


Figure 2.11: *Four Trials from the Wisconsin Card Sorting Test.*

measure was the number of perseverative errors that they made (how many times they sort by the previous rule after the rule changes).

Tower of Hanoi Task: The *Tower of Hanoi* task was not computerized; it was done with a physical version of the task. This task has three pegs with disks of ascending sizes. The goal is to take a given configuration and create a goal configuration in as few moves as possible. A move cannot consist of a larger disk being placed on a smaller disk, and only one disk can be moved at one time. The participants saw an image of the goal configuration, which they were allowed to study for as long as they desired. Then, they were given the physical task set in a different starting configuration. They will start with a practice two-disk problem, then they will do two practice four-disk problems that require at least eleven moves to complete. Then, they will complete two target problems with four disks that require at least fifteen moves to complete. The configurations were created by Humes et. al. (1997). The main dependent measure was the number of moves taken to complete the two target configurations.

Random Number Generation Task: The *random number generation task* is a computerized task where participants see a fixation on the screen and hear a beep every 800 ms. Each time they hear a beep; they type a number between one and nine with the instructions that

they are to make the string of numbers as random as possible. They first do a practice block of ten trials and then a target block of one hundred and sixty-two trials. Analysis of these results uses a program designed by Towse and Neil (1998) called RgCalc and principal components analysis with the calculated randomness indices. There were two components stated by this original study that relate to measures of randomness. One is “prepotent associates,” which is a measure of the stereotypical sequences (such as counting) that may be used. The other is “equality of response usage,” or the frequency of the same numbers/sequences that are used. These were the main dependent measures for this task. This is expanded upon in the results section.

Operation Span Task: The *operation span task* (**Figure 2.12**) is a computerized task where participants are shown a series of simple math equations followed by a word. The math equations consist of a simple multiplication or division problem followed by the addition or subtraction of a one-digit number (Turner et. al., 1989). These equations were either true or false, and the participant must read the equation out loud and state verbally whether it is true or false, then the experimenter indicated their answer using the keyboard. They then read the word presented to them directly after the equation. These words are the most frequently used four to six letter common nouns (Francis et. al., 1982). The equation is shown until a response is given or until eight seconds pass, and the word is presented immediately afterward for 750 ms. The goal of the task is for the participants to remember all the words presented and recall them verbally to the experimenter at the end of each block. There is first, three practice blocks with two pairs, then four target blocks each of two, three, four, and five pairs for a total of fifty-six words presented. The main dependent measure is the proportion of correctly recalled words.

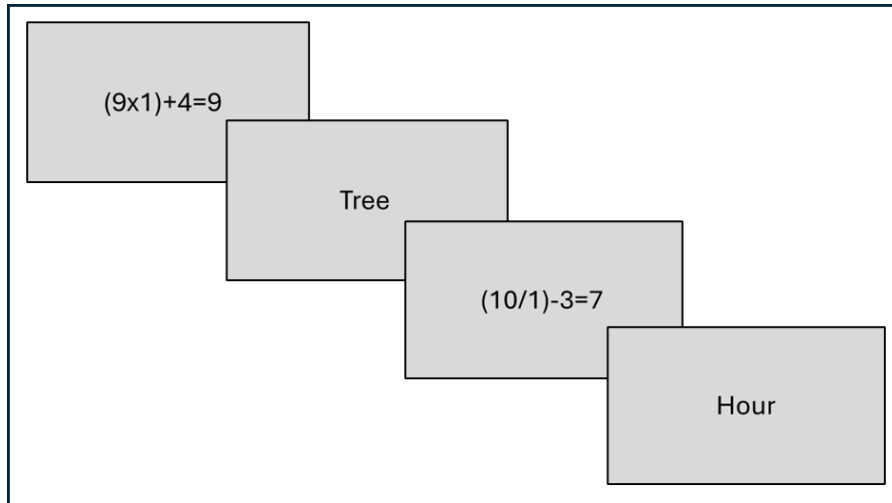


Figure 2.12: One Set of Two Trials in the Operation Span Task.

Dual Task: The *dual task* is a partially computerized task and partially physical task. In this task, participants had three three-minute blocks. In the first three minutes, participants are told to complete as many mazes as possible (Ekstrom et. al., 1976). Then, they completed a verbal word generation task in the second three minutes. They were given a letter every twenty seconds, and in those twenty seconds they were told to verbally say as many words starting with that letter as possible (while avoiding proper nouns and function words). The letters are balanced with approximate numbers of words in the English language that start with said letter (Solso et. al., 1982). In the final three minutes, the participants do both tasks at the same time. The main dependent measure was calculated with the equation below.

$$\left(\frac{nMaze_{single} - nMaze_{dual}}{nMaze_{single}} \right) + \left(\frac{nWord_{single} - nWord_{dual}}{nWord_{single}} \right) \Bigg/ 2$$

Research Design

The basic structure of this study was a latent variable analysis replicating the study done by Miyake et. al. (2000). Each group of “simple” tasks is a latent variable (shifting, inhibition, and updating), and each “simple” task is a manifest variable. The first goal of the study was to

use a confirmatory factor analysis to replicate the factor structure proposed by Miyake et. al. (2000), which showed that the three components were separable while sharing a common EF factor. Then, structural equation modeling was used to demonstrate in what way each complex EF task assessed the three EF components.

There were multiple analyses completed for the fNIRS, resting-state functional connectivity portion of this study, which is explained in further sections of this report. The data was processed, and whole-brain functional connectivity as well as full functional connectivity matrices were created at the individual level. Then, the two methods were compared at the group level using a paired t-test as well as visually with whole-brain connectivity and functional connectivity matrices. Then, this data was analyzed for correlations at the group level between all of the behavioral tasks used in this study. These correlation analyses were done with both tasks as well to further understand resting-state task differences as well as elucidate resting-state neural underpinnings of the EF tasks themselves.

Chapter 3: Procedures

The study took place over two sessions that both lasted an hour and a half and took place on separate days within two weeks of each other. In the first session, the participants signed an informed consent form followed by filling out a brief demographic questionnaire (asking sex, ethnicity, race, handedness, and vision correction). The participants then completed both resting-state tasks while wearing the fNIRS cap. The order for the two resting-state tasks was balanced between all participants to eliminate the possibility of an order effect, changing the functional connectivity analyses between the two tasks. The fNIRS cap was then removed, and the participant completed the session one task battery as determined by Miyake et. al. (2000). These were (in this order) the *antisaccade task*, *number-letter task*, *keep track task*, *stop-signal task*, *local-global task*, *Stroop task*, and *letter memory task*. Session two consisted of only behavioral tasks in the following order: the *plus-minus task*, *tone monitoring task*, *operation span task*, *random number generation task*, *Tower of Hanoi task*, *dual task*, and *Wisconsin Card Sorting Test*. For all tasks, the trial order was pre-randomized, but this was fixed for each participant to eliminate as many order effects as possible (Miyake et. al., 2000).

Chapter 4: Analytic Approach

We used CFA and SEM to replicate the findings from Miyake et al. 2000. The procedure for analyzing the data we collected was directly replicated from the original study. This includes the chosen dependent measures for each task as outlined above, the methods for analyzing outliers in the data, and the models being assessed as well.

Using CFA, we sought to replicate the finding that a three-factor model of executive functioning with an underlying common factor is the model that best represents the data. Correlations among latent factors will be estimated. We compared this model to three two-factor models (meaning two of the three components are the same construct), an independent three-factor model (meaning the components do not measure any common construct), and a one-factor model (meaning the three components are unitary). We used the same or equivalent fit indices as used by Miyake et. al. (2000) to see which model fits the data the best. These indices are the χ^2 statistic (small values indicate a better fit), Akaike's Information Criterion (AIC) (lower, including negative values indicate a better fit), the standardized root mean-squared residual (SRMR) (values less than 0.08 indicate a close fit), Bentler's Comparative Fit Index (CFI) (values that exceed 0.90-0.95 indicate a good fit), and the Tucker-Lewis Index (TLI) (values that exceed 0.90-0.95 indicate a good fit). Specifically, we used the TLI index rather than the IFI index, which is the only difference in comparison to the original study. To compare the relative fit between models, the χ^2 difference test was used, comparing each to the full three-factor model. The diagram of the full three-factor model is shown in **Figure 4.1**, below.

Next, using SEM, we examined which of the three components of EF are tapped by each of the five complex executive tasks. Each complex task has been previously theorized to assess a certain component or combination of components of EF. These theories from previous literature

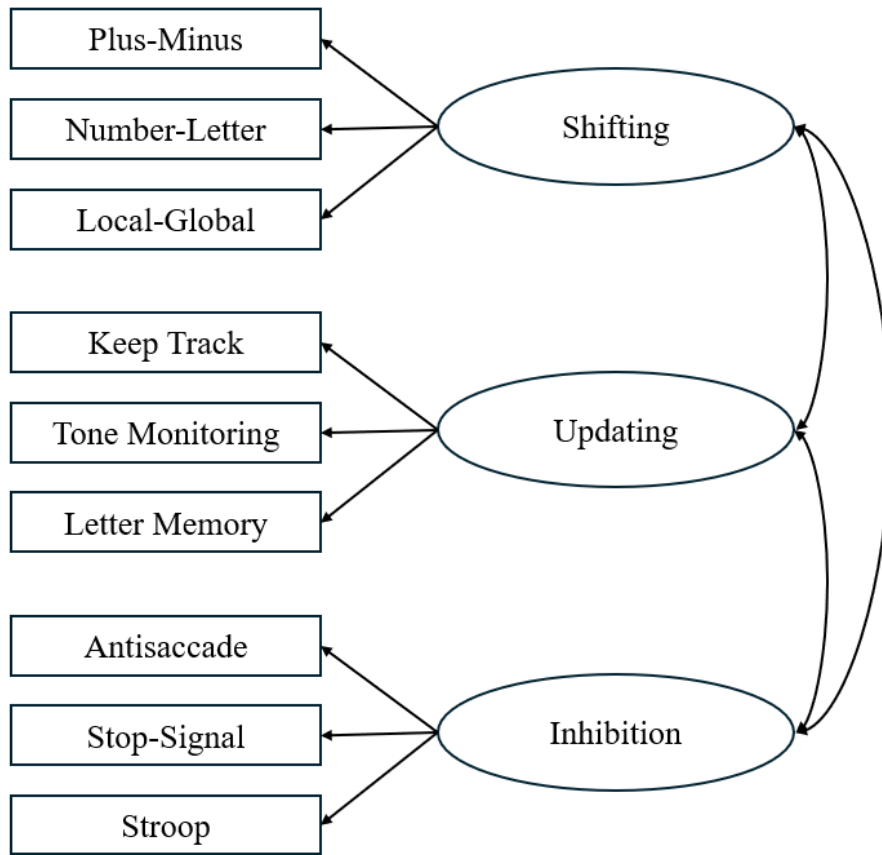


Figure 4.1: CFA Diagram Showing Full Three-Factor Model

informed the models being tested, which are the same models tested by Miyake et. al. (2000). For each complex task (in addition to the specific theorized models, which are outlined below) we compared these to a full three-path model (meaning all three components are used to perform the task) and a no-path model (meaning none of these three EF components are used to perform the task-that they were constrained to zero in the model). For the Wisconsin Card Sorting Test, we tested a one-path model from shifting, a one-path model from inhibition, and a two-path model from both shifting and inhibition. For the Tower of Hanoi task, we tested a one-path model from inhibition. For the random number generation task, we tested a one-path model from inhibition and a one-path model from updating. However, for this task, we had two separate components, one being the “prepotent associates” and the other being the “equality of response usage,” each model was used for these components, respectively. For the operation span task, we tested a one-path model from shifting, a one-path model from updating, and a two-path model from shifting and updating. Finally, for the dual task, we tested a one-path model from shifting. The three-path model to a complex EF task is shown in **Figure 4.2**, below.

As described above, we also used functional connectivity analyses for the resting-state component of this study. For these analyses, we used NIRS-KIT, a MATLAB toolbox (Hou et al., 2021) in conjunction with Homer 3 (Huppert, et. al., 2009) and SPM12. For pre-processing, the standard NIRS-KIT process of removing motion, physiological, and other artifacts was used. This pre-processing method specifically entailed the TDDR method for motion correction and the Infinite Impulse Response filtering method (set to 0.01 to 0.08 Hz as these low frequencies are important for spontaneous neural activity found at rest) (Hou, et. al., 2021). Then, we calculated whole-brain functional connectivity profiles and full functional connectivity matrices at the individual level across the full sample for both the Passive Viewing Paradigm and Mind-

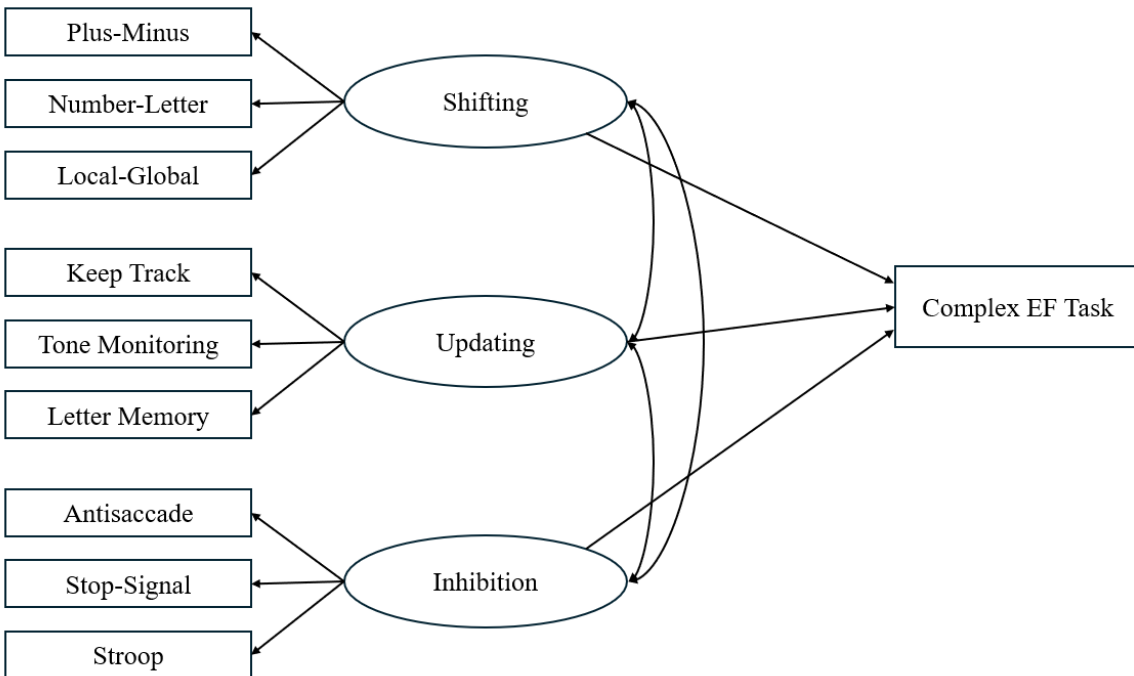


Figure 4.2: SEM Model Showing Three-Path Model with a Complex EF Task

Wandering Paradigm. The whole brain functional connectivity profiles were calculated by calculating the total connectivity at each channel to each other channel, and the functional connectivity profiles are calculated without reducing each channel to one value. Therefore, for the whole brain analyses, there is one value for each of the twenty-four channels, and for the functional connectivity matrices, there is a value for each connection (i.e. channel one to channel two is a single value). For the methods comparison portion of this study, paired t-tests using both the whole brain and functional connectivity matrices were calculated at the group level between the two resting-state tasks. We also directly compared average connectivity values using whole brain and functional connectivity matrix profiles. This group average was calculated by taking the mean connectivity at every channel (or channel connection) across all participants.

After comparing the functional connectivity profiles from both tasks, we used the data from both tasks to analyze the neural structure of EF as demonstrated using neurocognitive measures. First, group-level correlations between the whole brain functional connectivity profiles were calculated for each behavioral task (all nine simple tasks and all five complex tasks including both components of the Random Number Generation task). This allowed us to visually compare how resting-state functional connectivity data is correlated with each behavioral task separately. After visually assessing these relationships, this was further examined by looking specifically at the channels with significant correlations to each of the behavioral tasks. Finally, to develop a possible statistical structure of EF from the resting-state functional connectivity data, principal components analysis was done using the correlations calculated in NIRS-KIT. With this, we were able to understand not only how the neural data correlated to each task but also how these correlations cluster together to create components of EF. The results and further details about all of these analyses can be found in the results section of this study.

Chapter 5: Results

Replication Study

Preliminary Analyses

The first steps in the analysis required transforming and trimming the data for the nine tasks used in the confirmatory factor analysis. This study follows the procedures for trimming and transforming the data as done in the original study (Miyake et. al., 2000) as adapted to fit the needs of the current study's data. To trim the reaction time data for the Number Letter, Local-Global, and Stroop tasks, we used a two-step process. Only target trials with a reaction time greater than 200 ms were included. The first step was to replace any trial with a reaction time outside of the bounds stated originally (Miyake et. al., 2000) across all participants with these values. The bounds were 300 to 3500 ms for the Number Letter task, 500 to 3500 for the Local-Global task, and 400 to 2000 ms for the Stroop task. The second step was a within-subjects analysis where any reaction time that was greater than three standard deviations from the individual subject's mean was replaced with that value. For the plus-minus task, any duration that was more than three standard deviations from the overall, between-subjects mean duration was replaced with that value for each condition separately. An average of 4.1% of observations were affected by these trimming procedures across these four tasks. For proportion of correct responses, we used arcsine transformations as done in the original study, this was done for the Keep Track, Letter Memory, Tone Monitoring, Antisaccade, and Stop-Signal tasks. We visually assessed within each construct for any outliers as well as calculating leverage, but no participants were excluded as outliers. The descriptive statistics and normality values for these nine tasks can be seen in **Table 5.1**. In this table, for the proportion correct measures, the descriptive statistics before arcsine transformation are included in brackets.

Table 5.1: Descriptive statistics for all simple EF tasks after transformations and RT trimming.

Task	n	Mean (SD)	Range	Skewness	Kurtosis
Plus-Minus	67	6.29 s (10.33)	-13.08 to 35.01	0.93	3.72
Number-Letter	67	688.54 ms (305.67)	178.94 to 1531.87	0.76	3.24
Local-Global	67	143.61 ms (114.31)	-111.71 to 448.45	0.15	3.06
Keep Track	67	0.93 (0.17)	0.44 to 1.38	-0.53	4.39
		[0.63 (0.15)]	[0.19 to 0.96]		
Tone Monitoring	67	1.14 (0.30)	0.36 to 1.57	-0.05	2.32
		[0.77 (0.21)]	[0.12 to 1.00]		
Letter Memory	67	0.98 (0.21)	0.36 to 1.57	-0.01	3.50
		[0.68 (0.18)]	[0.12 to 1.00]		
Antisaccade	67	1.40 (0.14)	1.01 to 1.57	-0.51	2.96
		[0.95 (0.06)]	0.72 to 1.00		
Stop-Signal	67	1.42 (0.16)	0.96 to 1.57	-0.95	3.43
		[0.95 (0.07)]	[0.67 to 1.00]		
Stroop	67	132.54 ms (76.74)	32.09 to 374.75	0.87	3.39

After trimming and transforming the data of the nine simple tasks, all values were adjusted so that higher values indicated better performance, and standardized (z-score) values were used for the remainder of the confirmatory factor analysis and structural equation modeling.

For the complex EF tasks, any value greater than three standard deviations from the mean was replaced with the three standard deviations from the mean value for that task, and the descriptive statistics for these tasks can be seen in **Table 5.2**. An average of 1.8% of the observations were affected by this trimming procedure. After calculating descriptive statistics, all measures were again adjusted so that higher values indicated better performance, and standardized (z-score) values were used for the structural equation modeling with these tasks.

The preliminary analyses for the Random Number Generation task were more in depth and followed the steps done in the original study as outlined here as well (Miyake et. al., 2000). After analyzing each participant's raw data using RgCalc, a set of fifteen randomness indices were used for this analysis.

Table 5.2: Descriptive statistics for the complex EF tasks after outlier trimming.

Task	n	Mean (SD)	Range	Skewness	Kurtosis
WCST	66	39 errors (11)	23 to 75	1.31	5.33
TOH	67	42 moves (10)	30 to 71	0.99	3.65
RNG C1	66	0 (1)	-6.45 to 0.95	-4.42	27.81
RNG C2	66	0 (1)	-2.49 to 5.38	2.27	14.32
RNG C3	66	0 (1)	-2.26 to 2.09	-0.19	2.70
Operation Span	67	47 words (7)	27 to 56	-0.48	2.23
Dual Task	67	0.07 (0.15)	-0.39 to 0.39	-0.73	3.87

These indices are TPI, combined adjacency (A), Runs, Evan’s RNG, R, Coupon, Mean RG, Mode RG (smallest value if multiple were available), phi2, phi3, phi4, phi5, phi6, phi7, and RNG2. More information can be found about these randomness indices and the selection of these fifteen specifically in the two original studies (Miyake et. al., 2000; Towse & Niel, 1998).

Percentage values were used as decimals, and all indices were adjusted so that higher values indicated better performance. Principal components analysis with Promax rotation was then performed with these indices. As done in the study being replicated and based on a scree plot, we extracted three components using this data. Our component one was similar to the “equality of response usage” component, and our component two was similar to the “prepotent associates” component. Our third component was also in line with the third component found in the original study (Miyake et. al., 2000), which was not used for further analyses. The loadings for all three components can be seen in **Table 5.3**, and the descriptive statistics after trimming procedures and standardizing the component values can be seen in **Table 5.2** with the other complex tasks.

The bivariate correlations among all tasks can be seen in **Table 5.4**. It would be expected that all correlations are low and positive with significance being within tasks of the same component of EF (Miyake et. al., 2000). This, however, is not the case in our data, which is difficult to interpret considering the original study. This does not inherently affect the theories

Table 5.3: Component loadings after PCA and Promax rotation for the Random Number Generation task.

Measure	Component 1	Component 2	Component 3
TPI	0.371	0.669	-0.131
A	0.292	0.723	0.000
Runs	-0.253	1.033	-0.070
Evan's RNG	0.694	0.341	0.049
R	1.048	-0.094	0.080
Coupon	1.050	-0.185	0.025
Mean RG	1.040	-0.122	0.072
Mode RG	0.069	0.103	-0.786
phi2	-0.526	0.199	0.273
phi3	-0.701	-0.035	0.271
phi4	0.130	-0.014	0.806
phi5	0.010	-0.106	0.710
phi6	0.038	0.201	0.679
phi7	-0.169	0.537	0.309
RNG2	0.846	0.205	0.148

Table 5.4: Correlation Coefficients for all Tasks and both RNG task components, * indicates significance ($p < 0.05$).

	PM	NL	LG	KT	TM	LM	AS	SS	ST	WC	TH	R1	R2	OS
PM	-													
NL	.204	-												
LG	.150	.072	-											
KT	-.088	.133	-.211	-										
TM	.032	.134	.103	.371*	-									
LM	.141	-.130	-.129	.209	.402*	-								
AS	.040	.107	.011	.223	.174	-.009	-							
SS	.275*	.087	.235	.095	.107	.125	.110	-						
ST	.150	.198	.072	.187	.287*	.235	.096	.362*	-					
WC	-.137	.045	-.193	.254*	.490*	.400*	.068	.152	.199	-				
TH	.011	-.004	-.105	.415*	.136	.140	-.002	.013	.074	.166	-			
R1	.263*	.073	.123	.110	.316*	.201	.106	.326*	.143	.190	.282*	-		
R2	-.186	.063	-.116	.016	-.162	-.141	-.002	-.326*	-.111	.012	-.258*	-.867*	-	
OS	-.043	.211	-.115	.425*	.515*	.469*	.223	.086	.375*	.488*	.278*	-.002	.110	-
DT	.232	.021	.063	.037	-.098	.010	-.096	.091	.059	-.110	.131	-.091	.019	.101

going into the CFA and SEM analyses, but it is interesting. This may be in line with the idea that there is little reliability between EF tasks, so they do not necessarily correlate in the patterns that are expected. From the significance of some of the correlations, one may wonder if there should be different paths assessed in the SEM analyses with the complex tasks-for example, maybe the WCST task taps into the updating construct rather than the shifting construct as it is significantly correlated with all three updating tasks. The bivariate correlations themselves are not as informative as the further analyses but may lead to new questions about this structure of EF that could be analyzed deeper using CFA or SEM as well.

Confirmatory Factor Analysis

In the confirmatory factor analysis, six models were estimated. The full-three factor model, which is theorized based on the original study to be the best fitting model indicates that the three components of EF are separable but share a common factor (Miyake et. al., 2000). As seen in **Table 5.5**, this model has the lowest chi-square value (which is also non-significant), the lowest SRMR, and high CFI/TLI values, which indicates an overall good fit to the data. Note here that all the AIC values are very high, but it is the relative magnitude of AIC that indicates model fit rather than the absolute values. In this case, the AIC value for the full three-factor model is the second lowest, indicating a good fit to the data as well. To fully understand which model is the best, all other models were compared directly to this full three-factor model using chi-square difference tests.

The one-factor model assumes that all three latent constructs are the same-that the components are completely unitary. This was done by setting all latent correlations to 1.0 in the model.

Table 5.5: Fit indices for all models in confirmatory factor analysis, selected model of best fit is bold, * indicates significance ($p < 0.05$).

Model	df	χ^2	AIC	SRMR	CFI	TLI
Full Three-Factor	24	27.057	1721.094	0.069	0.926	0.889
One-Factor	27	40.304*	1728.342	0.090	0.677	0.569
Shifting=Inhibition	25	27.984	1720.021	0.071	0.928	0.896
Shifting=Updating	25	31.190	1723.227	0.078	0.850	0.784
Inhibition=Updating	25	32.112	1724.149	0.076	0.827	0.751
Independent Factors	27	40.197*	1728.324	0.111	0.680	0.573

As seen in **Table 5.5**, this model had a significant and higher chi-square value, higher AIC, higher SRMR, and lower CFI/TLI than the full three-factor model, which indicates worse fit. The chi-square difference test ($\chi^2_{(3)} = 13.247, p = 0.004$) is significant, which means that this one-factor model fits the data significantly worse than the full three-factor model. Next, three different two-factor models were analyzed; each would represent that those two constructs are unitary while the third is separate. These were done by setting one latent correlation to 1.0 while allowing the others to correlate freely. The model in which shifting and inhibition are unitary has a higher chi-square value and SRMR value, but it also has a lower AIC and higher CFI/TLI values. This makes it harder to determine by the fit indices whether it is a worse or better fit than the full three-factor model. The chi-square difference test ($\chi^2_{(1)} = 0.414, p = 0.520$) is not significant, which will be addressed momentarily. The second two-factor model represents shifting and updating being unitary, and it has worse fit indices than the full three-factor model. Also, the chi-square difference test ($\chi^2_{(1)} = 4.133, p = 0.042$) is significant, so this model is statistically worse than the full three-factor model. The final two-factor model represents updating and inhibition being unitary and again has worse overall fit than the full three-factor model based on all fit indices, and the chi-square difference test ($\chi^2_{(1)} = 5.055, p = 0.025$) is again significant indicating worse fit than the full three-factor model. Finally, the last model is an

independent three-factor model, and this would mean that the three latent constructs are not correlated and are completely distinct. This was done by setting the three latent correlations to 0.0 in the model. Overall, this model has worse fit indices and a significant chi-square value, and the chi-square difference test ($\chi^2_{(3)} = 13.140, p = 0.004$) is significant, indicating statistically worse fit than the full three-factor model.

After statistically comparing all six of these models from the confirmatory factory analysis, it is clear that the full three-factor model is a better fit to the data than all except for the two-factor model where shifting is equal to inhibition. Because there are fewer freely estimated parameters in this two-factor model, and the chi-square difference test was not significant, this model cannot be ignored. In this study, these components of EF may be completely unitary rather than distinct constructs. Based on this analysis; to complete the replication of the original study, the full three-factor model will be used as the base for structural equation modeling with the complex EF tasks. However, we will also perform the structural equation modeling portion of the study using the two-factor model where shifting and inhibition are the same construct. A visual depiction of both selected models (with the parameter estimates and error estimates included) are shown in **Figure 5.1** (full three-factor model) and **Figure 5.2** (two-factor model where shifting equals inhibition). Note that, possibly due to sample size being low, the parameter estimates are not consistent within the constructs despite the two models being similar overall. Also, the original estimated correlation between the shifting and inhibition constructs is only 0.690, which is not equivalent to one. Theoretically, this makes it interesting that the two-factor (shifting equals inhibition model) where the correlation between the two is set to one is not statistically worse than the full three-factor model.

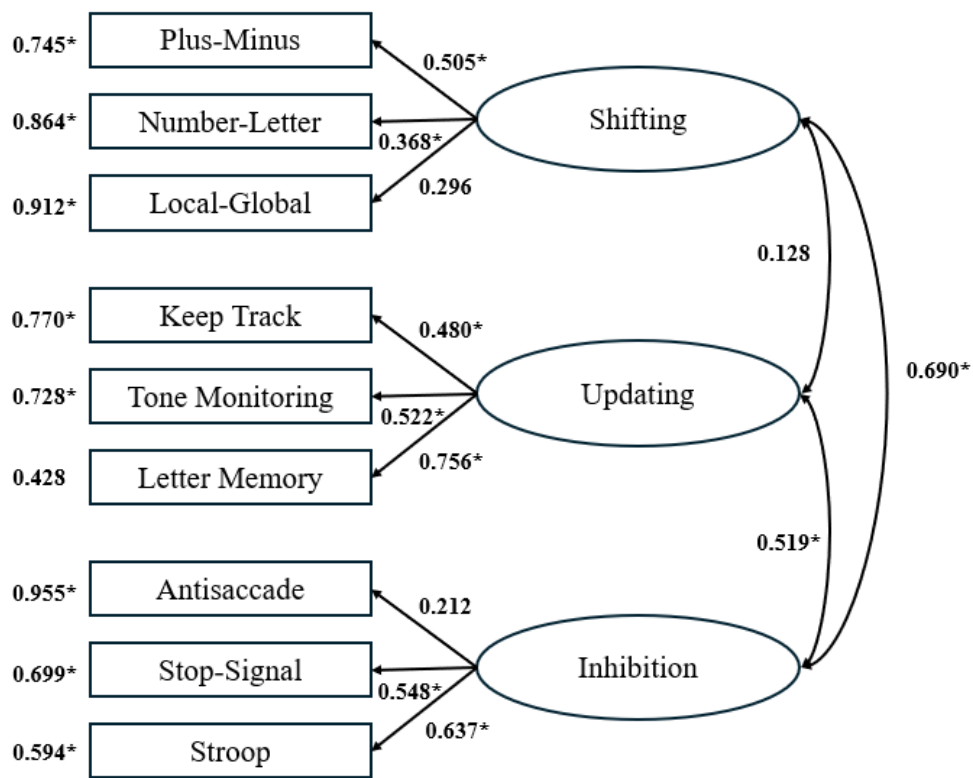


Figure 5.1: Diagram of full three-factor model, * indicates significance ($p < 0.05$)

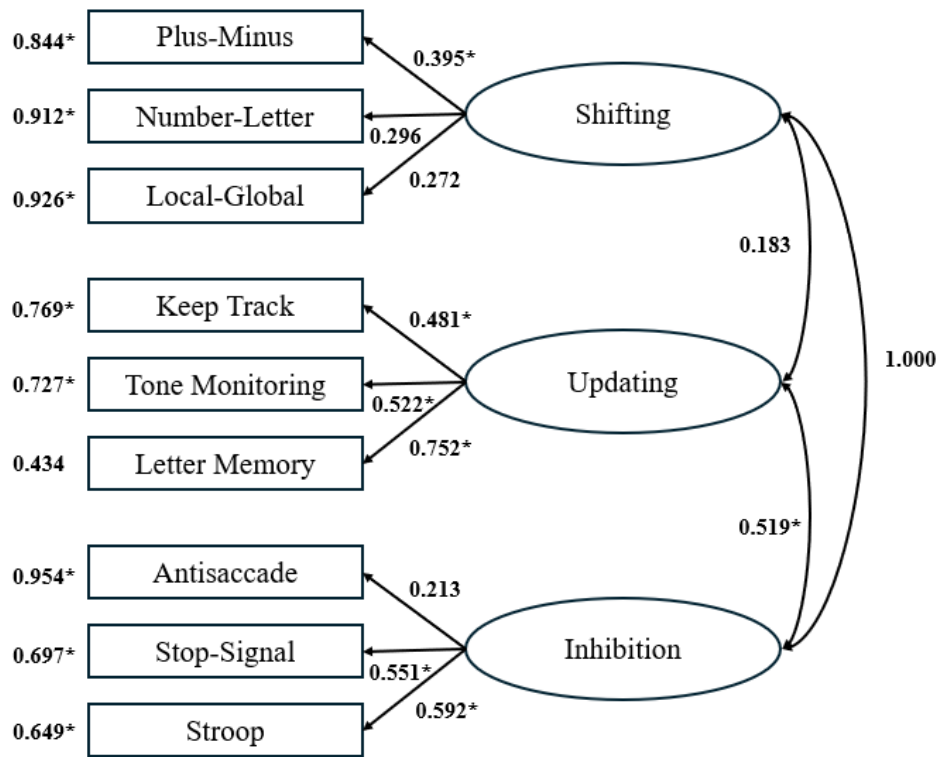


Figure 5.2: Diagram of two-factor model where shifting and inhibition are the same construct, * indicates significance ($p < 0.05$).

Structural Equation Modeling

For all five complex EF tasks (including both Random Number Generation task components), the structural equation models used were the same as those in the original study. These original decisions were based on theories for how performance/outcomes on these complex EF tasks would relate to the three EF components being assessed by the simple tasks-shifting, updating, and inhibition (Miyake et. al., 2000). For each task a full three-path and a no paths model was also estimated. These models were estimated with the base structure of the full three-factor confirmatory factor analysis model for replication of the original study. However, based on the model comparison results from the confirmatory factor analysis, all complex tasks and models were estimated using the two-factor (where shifting is the same as inhibition) as the base model as well. To do this, all models being estimated using structural equation modeling were performed in the same way, but in all cases, the latent correlation between shifting and inhibition was set to 1.0.

Wisconsin Card Sorting Test

In the analysis of the WCST, both shifting and inhibition are theorized to contribute to performance on this task, so a two-path (shifting and inhibition) model and a one-path model from each were analyzed (Miyake et. al., 2000). In the original study, the one-path model from shifting was the best fit to the data. However, here, in **Table 5.6** this model has one of the worst fits, and the coefficient of this path is not significant. Chi-square difference tests were again calculated to directly compare these models to one another. The two-path model from shifting and inhibition ($\chi^2_{(1)} = 0.061, p = 0.805$) is not statistically a worse fit than the full three-path model. The remaining models were then compared to this two-path model. The one-path from shifting model ($\chi^2_{(1)} = 27.454, p < 0.001$) was statistically worse than the two-path model.

Table 5.6: Fit indices for all path models for WCST, selected model of best fit is bold, * indicates significance ($p < 0.05$).

Model	df	χ^2	SRMR	CFI	Coefficients		
					Shifting	Updating	Inhibition
Full Three-Path	30	31.959	0.071	0.964	1.133	0.351	1.015
Two-Path (Shifting, Inhibition)	31	32.020	0.071	0.981	-2.010	-	1.958
One-Path (Shifting)	32	59.474*	0.116	0.499	-0.155	-	-
One-Path (Inhibition)	32	44.516	0.087	0.772	-	-	0.712*
No Paths	33	60.308*	0.115	0.502	-	-	-

The one-path from inhibition model ($\chi^2_{(1)} = 12.496, p = 0.0004$), while being a better fit than the shifting model, was also statistically worse than the two-path model. Finally, the no paths model ($\chi^2_{(2)} = 28.288, p < 0.001$) was also statistically a worse fit than the two-path model. From these statistical tests, the two-path model from shifting and inhibition was endorsed as the best fitting model to our data.

These models were then estimated using the two-factor model that was endorsed through confirmatory factor analysis as the base model, and these results are shown in **Table 5.7**. The same chi-square difference test method was used to determine the model of best fit. Using this base structure, there are only two latent constructs, so the path analyses are slightly different. First, the one-path model from the shifting/inhibition construct was compared to the full two-path model from this and the updating construct. The chi-square difference test shows ($\chi^2_{(1)} = 8.303, p = 0.004$) that the one-path model was a statistically worse fit to the data. The no paths model ($\chi^2_{(2)} = 22.435, p < 0.001$) was also statistically worse than the full two-path model. Therefore, the endorsed model when shifting and inhibition are the same construct is the full two-path model.

Table 5.7: Fit indices for all path models for WCST with the two-factor (shifting equals inhibition) CFA model as the base, * indicates significance ($p < 0.05$).

Model	df	χ^2	SRMR	CFI	Coefficients	
					Shifting/Inhibition	Updating
Full Two-Path	33	39.961	0.083	0.873	-0.050	0.683*
One-Path (Shifting/ Inhibition)	34	48.246	0.088	0.740	0.712	-
No Paths	35	62.396*	0.119	0.500	-	-

This is still a different result from the original study, and it differs from the first models (with the full three-factor model as the base). This may indicate that in this study, the shifting and inhibition constructs differ from the constructs in the original study. Or it could mean that the WCST task for this study is using the updating construct as well as the combined shifting and inhibition construct.

The two endorsed path models were also compared directly using a chi-square difference test. This test ($\chi^2_{(2)} = 7.941, p = 0.019$) was significant, which means that the endorsed model using the two-factor (shifting=inhibition) structure was a worse fit to the data than the two-path model using the full three-factor base model. This provides evidence for endorsement of the full three-factor model for this study. The negative path coefficients for both endorsed models to give pause, however, because this implies a negative or inverse relationship between that component and the complex task being assessed, which is not theoretically sound.

Tower of Hanoi

For the Tower of Hanoi task, there were many theories discussed in the original study for how the three EF components were contributing to task performance, but the overall premise was that it requires inhibition to strategize the fewest number of moves to make on the task (Miyake et. al., 2000). Therefore, the only model assessed was a one-path model from inhibition, and the model fit results are shown in **Table 5.8**. This was the endorsed model in the original study.

Table 5.8: Fit indices for all path models for TOH, selected model of best fit is bold * indicates significance ($p < 0.05$).

Model	df	χ^2	SRMR	CFI	Coefficients		
					Shifting	Updating	Inhibition
Full Three-Path	30	35.248	0.072	0.887	0.520	0.876	-0.769
One-Path (Inhibition)	32	40.453	0.082	0.818	-	-	0.147
No Paths	33	41.260	0.086	0.822	-	-	-

Based on chi-square difference tests, this one-path model ($\chi^2_{(2)} = 5.205, p = 0.074$) was not statistically worse than the full three-path model. The no paths model can then be compared to this one-path model, and this comparison ($\chi^2_{(1)} = 0.807, p = 0.369$) demonstrated that the no paths model was also not statistically worse than the one-path model. Based on this, the no paths model must be endorsed as the model of best fit for this task.

These three models were then estimated using the two-factor model as the base, which is shown in **Table 5.9**. The one-path model from shifting/inhibition ($\chi^2_{(1)} = 4.314, p = 0.038$) was statistically worse than the full two-path model. The no paths model was then compared to the full two-path from model ($\chi^2_{(2)} = 4.697, p = 0.096$), and it was not a statistically worse fit. Therefore, despite changing the base model to the two-factor model where shifting equals inhibition, the no paths model was still endorsed as the model of best fit for this task. The base model used for this task did not change the outcome, and for both, the outcome in this study was different than the original result. This could be further evidence that there is something different regarding the inhibition construct in this study, or it could indicate that for this study, the TOH task is not mainly tapping into inhibition but possibly a different cognitive process.

Table 5.9: Fit indices for all path models for TOH with the two-factor (shifting equals inhibition) CFA model as the base, * indicates significance ($p < 0.05$).

Model	df	χ^2	SRMR	CFI	Coefficients	
					Shifting/Inhibition	Updating
Full Two-Path	33	38.661	0.078	0.878	0.164	0.470*
One-Path	34	42.975	0.088	0.807	0.102	-
(Shifting/Inhibition)						
No Paths	35	43.358	0.091	0.820	-	-

The two endorsed models were then compared directly with a chi-square difference test. This test ($\chi^2_{(2)} = 2.098, p = 0.350$) is not significant, which means the no paths model for the two-factor design is not a statistically worse fit than the no paths model with the full three-factor design. Therefore, for this task, we further endorse the two-factor (shifting=inhibition) design.

Random Number Generation

As described previously, the Random Number Generation task is broken into components that differentially represent one of the EF components. Specifically, from the principal components analysis, two components equating those found in the original study were extracted. Component one in this study is the “equality of response usage” components, which is theorized to represent a measure of updating (Miyake et. al., 2000), so this is the path that was tested for component one. The original study endorsed this one-path from updating model as the best fit model. The data for this model, which is shown in **Table 5.10**, while having a significant coefficient estimate, shows worse overall fit than the three-path model. The chi-square difference between these two models ($\chi^2_{(2)} = 6.923, p = 0.031$) also showed that the one-path model from updating was statistically worse than the full three path model. Then, the no paths model can also be compared to the full three-path model ($\chi^2_{(3)} = 15.265, p = 0.002$), and this was also statistically worse than the full three-path model. Therefore, the full three-path model was endorsed as the model that is the best fit in this study.

Table 5.10: Fit indices for all path models for RNG, selected model of best fit is bold * indicates significance ($p < 0.05$).

Model	df	χ^2	SRMR	CFI	Coefficients		
					Shifting	Updating	Inhibition
C1: Full Three-Path	30	29.215	0.069	1.000	0.273	0.305	0.195
C1: One-Path (Updating)	32	36.138	0.079	0.894	-	0.428*	-
C1: No Paths	33	44.480	0.100	0.706	-	-	-
C2: Full Three-Path	30	30.758	0.072	0.977	0.054	0.092	-0.358
C2: One-Path (Inhibition)	32	30.979	0.072	1.000	-	-	-0.438*
C2: No Paths	33	38.205	0.087	0.841	-	-	-

The models for component two of the Random Number Generation task, are also shown in **Table 5.10**. This component represents the “prepotent associates” component, which is a measure of inhibition. The original study only tested a one-path from inhibition model, which was endorsed as the model of best fit (Miyake et. al., 2000). A chi-square difference test ($\chi^2_{(2)} = 0.221, p = 0.895$) demonstrated that this one-path model was not statistically worse than the full-three path model. The no paths model was then compared to this one-path model ($\chi^2_{(1)} = 7.226, p = 0.007$), and the no paths model was statistically worse than the one-path model. Based on this, the one-path from inhibition model was endorsed as the model of best fit for this component, but with a negative path coefficient, this is showing an opposite result from the original study.

The same analyses were then performed using the two-factor model as the base, and this is shown in **Table 5.11**. For component one, the one-path from updating model ($\chi^2_{(1)} = 6.761, p = 0.009$) was statistically worse than the full two-path model as was the no paths model ($\chi^2_{(2)} = 15.239, p = 0.001$). Therefore, the full two-path model was endorsed as the model of best fit for this “equality of response usage” component of this task. This differs from

Table 5.11: Fit indices for all path models for RNG with the two-factor (shifting equals inhibition) CFA model as the base, * indicates significance ($p < 0.05$).

Model	df	χ^2	SRMR	CFI	Coefficients	
					Shifting/Inhibition	Updating
C1: Full Two-Path	33	31.328	0.074	1.000	0.429*	0.264
C1: One-Path (Updating)	34	38.089	0.083	0.895	-	0.432*
C1: No Paths	35	46.567	0.105	0.703	-	-
C2: Full Two-Path	33	32.035	0.076	1.000	-0.395*	-0.102
C2: One-Path (Shifting/Inhibition)	34	32.504	0.076	1.000	0.433*	-
C2: No Paths	35	40.293	0.092	0.838	-	-

the original study, and because it is associated with the updating construct rather than shifting or inhibition, does not follow with being caused specifically by the shifting or inhibition constructs as the other differences seem to be. This could be due to small differences in the component structure of this task between the two studies, or it could be due to differences in how this component is being affected by all three constructs.

The data for component two of this task using the two-factor base model is also provided in **Table 5.11**. The one-path from shifting/inhibition model ($\chi^2_{(1)} = 0.469, p = 0.493$) was not statistically worse than the full two-path model. The no paths model was then compared to the one-path from shifting/inhibition model ($\chi^2_{(1)} = 7.789, p = 0.005$) and was statistically a worse fit to the data. Despite changing the base model, for component two of this task, the one-path from shifting/inhibition model was endorsed, which is a similar result as the original study as the coefficient is positive and significant. This may indicate that component two for this study is like that of the original study, but this result was achieved when inhibition and shifting are the same construct. Because of this, shifting may also play a role in this “prepotent associates” component of this task.

For both components for the Random Number Generation task, a chi-square difference test can be used to directly compare the two endorsed models. For component one, this test ($\chi^2_{(3)} = 2.113, p = 0.550$) is not significant. Therefore, the full three-two model with the two-factor design is not statistically worse than the full three-path model using the full three-factor design. For component two, ($\chi^2_{(2)} = 1.525, p = 0.466$), the same is true. The one-path model from shifting/inhibition is no worse using the two-factor design as the one-path from inhibition alone using the full three-factor design.

Operation Span

The Operation Span task was originally theorized that both shifting and updating would play a part in performance on this task, so a two-path model from shifting and updating, and a one-path model from each was assessed. The original study endorsed the one-path model from updating as the model of best fit (Miyake et. al., 2000). The model fit data for this task is shown in **Table 5.12**. The two-path model from shifting and updating, when compared to the full three-path model ($\chi^2_{(1)} = 0.126, p = 0.723$) was not statistically worse, so the remaining models were compared to this model. The one-path model from shifting ($\chi^2_{(1)} = 4.345, p = 0.037$) was statistically a worse fit to the data than the two-path model. The one-path model from updating ($\chi^2_{(1)} = 0.014, p = 0.906$) was not statistically worse than the two-path model, so this was used to test the no paths model fit. The no paths model ($\chi^2_{(1)} = 36.122, p < 0.001$) was statistically worse than the one-path model from updating. Therefore, the one-path model from updating is the model of best fit in this study.

Table 5.12: Fit indices for all path models for Operation Span, selected model of best fit is bold
* indicates significance ($p < 0.05$).

Model	df	χ^2	SRMR	CFI	Coefficients		
					Shifting	Updating	Inhibition
Full Three-Path	30	36.467	0.075	0.917	0.160	0.715*	-0.174
Two-Path (Shifting, Updating)	31	36.593	0.075	0.928	-0.022	0.815*	-
One-Path (Shifting)	32	40.938	0.092	0.885	7.340	-	-
One-Path (Updating)	32	36.607	0.075	0.941	-	0.811*	-
No Paths	33	72.729*	0.135	0.490	-	-	-

The same model analyses were performed using the two-factor model as the base, and this is shown in **Table 5.13**. The one-path from shifting/inhibition model, when compared to the full two-path model ($\chi^2_{(1)} = 8.316, p = 0.004$) was a statistically worse fit. The one-path from updating model ($\chi^2_{(1)} = 0.064, p = 0.800$) was not statistically worse than the full two-path model, so this was used for the comparison to the no paths model. The no paths model ($\chi^2_{(1)} = 34.670, p < 0.001$) was statistically worse than the one-path from updating model. This entails that the one-path from updating model is again endorsed as the model of best fit for this task when using the two-factor base structure, which aligns with the original study and the analysis with the three-factor base model in this study.

The two endorsed one-path from updating models were then compared directly using a chi-square difference test. This test ($\chi^2_{(2)} = 3.551, p = 0.169$) being non-significant indicates that the model using the two-factor base design is not statistically worse than the model using the full three-factor base design. Therefore, the two-factor base design is again endorsed for the operation span task despite endorsing the same one-path from updating model as the original study. It could be that the updating construct is more stable in this study than the others.

Table 5.13: Fit indices for all path models for Operation Span with the two-factor (shifting equals inhibition) CFA model as the base, * indicates significance ($p < 0.05$).

Model	df	χ^2	SRMR	CFI	Coefficients	
					Shifting/Inhibition	Updating
Full Two-Path	33	40.094	0.081	0.909	0.047	0.775*
One-Path	34	48.410	0.091	0.815	0.852*	-
(Shifting/Inhibition)						
One-Path	34	40.158	0.081	0.921	-	0.807*
(Updating)						
No Paths	35	74.828*	0.138	0.489	-	-

Dual Task

For the dual task, it was originally hypothesized that shifting would be the primary component assessed by this task, so only the one-path model from shifting was estimated. The original study endorsed the no paths model despite the original theories (Miyake et. al., 2000). The model fit data is shown in **Table 5.14**. Using chi-square difference tests, the one-path model from shifting ($\chi^2_{(2)} = 0.673, p = 0.714$) was not statistically worse than the full three-path model. The no paths model was then compared to the one-path model ($\chi^2_{(1)} = 2.59, p = 0.107$) and was also not a statistically worse fit for the data. This no paths model was endorsed as the best model in this study as well, so the same result as the original study was achieved for this task when using the full three-factor model as the base model. While this is the same result as the original study, it is again not theoretically sound that the shifting component would have an inverse relationship with Dual Task performance (as shown by the negative path coefficient), which may be further evidence that the shifting construct in this replication study is unstable or different than that of the original study.

Table 5.14: Fit indices for all path models for Dual Task, selected model of best fit is bold* indicates significance ($p < 0.05$).

Model	df	χ^2	SRMR	CFI	Coefficients		
					Shifting	Updating	Inhibition
Full Three-Path	30	30.564	0.068	0.986	-0.419	0.058	0.138
One-Path (Shifting)	32	31.237	0.069	1.000	0.292	-	-
No Paths	33	33.827	0.073	0.979	-	-	-

This analysis was repeated using the two-factor model as the base model, and the results are shown in **Table 5.15**. The one-path from shifting/inhibition model ($\chi^2_{(1)} = 1.279, p = 0.258$) was not statistically worse than the full two-path model. The no paths model was then again compared to this one-path from shifting/inhibition model ($\chi^2_{(1)} = 0.751, p = 0.386$) and was not statistically worse. Therefore, the no paths model was endorsed again despite changing the base model to the two-factor model.

Both no paths models for the Dual Task were compared directly using a chi-square difference test. This test ($\chi^2_{(2)} = 2.089, p = 0.352$) is non-significant, so the model using the two-factor base design was endorsed again. Overall, other than the WCST results, the two-factor where shifting equals inhibition structure was endorsed over the full three-factor design, which contradicts the original study. Also, the only complex EF tasks that gave the same result as theorized in this study were the “prepotent associates” random number generation component (only when using the two-factor design), the operation span task, and the dual task.

Methods Comparison Study

The next part of this study was to directly compare the fNIRS data collected from the two resting-state functional connectivity measures, the *Passive Viewing Paradigm* and the *Mind-Wandering Paradigm*. The first step, as outlined previously, was to clean the data of motion, physiological, and other noise artifacts using the pre-processing method.

Table 5.15: Fit indices for all path models for Dual Task with the two-factor (shifting equals inhibition) CFA model as the base, * indicates significance ($p < 0.05$).

Model	df	χ^2	SRMR	CFI	Coefficients	
					Shifting/Inhibition	Updating
Full Two-Path	33	33.895	0.075	0.977	0.245	0.184
One-Path	34	35.174	0.077	0.970	0.141	-
(Shifting/Inhibition)						
No Paths	35	35.925	0.078	0.976	-	-

This and all further analyses were completed using the NIRS-KIT toolbox for resting-state functional connectivity analyses (Hou et. al., 2021). After this was done, whole-brain functional connectivity analyses as well as functional connectivity matrices were calculated at the individual level, which were then used for all of the following group-level analyses.

Using the whole-brain functional connectivity analyses, a paired t-test comparing the two tasks as well as an average functional connectivity measure for all channels was calculated. This group-level average functional connectivity measure is done by calculating a mean of all connectivity between one channel and all other channels. This was done over the entire time-course for the task, and the mean was then also calculated across all subjects' individual-level data. These whole-brain functional connectivity group averages are shown in **Figure 5.3** (for the Passive Viewing Paradigm) and **Figure 5.4** (for the Mind-Wandering Paradigm).

From the whole-brain group average analyses, the results indicate a deviation from the guiding hypothesis for this portion of the study. Across the brain, there is, on average, more functional connectivity (higher spontaneous activity levels between spatially remote neural populations) during the Mind-Wandering Paradigm than the Passive Viewing Paradigm. Because the adults in this study did well with complying with both tasks and the same pre-processing steps removed noise from both tasks, this analysis does indeed contradict the guiding hypothesis behind this study.

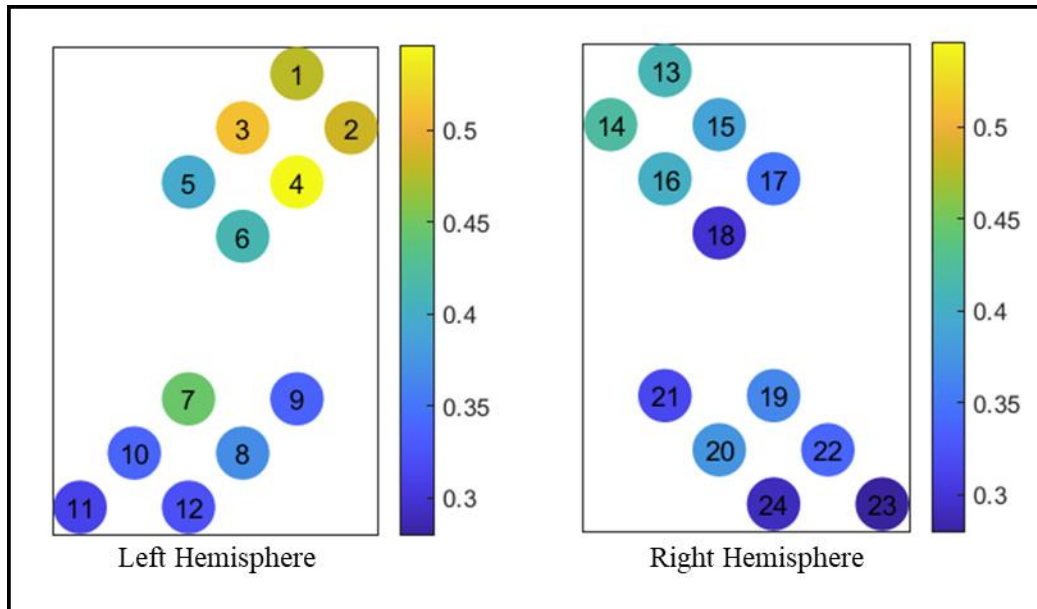


Figure 5.3: Whole-Brain functional connectivity group average for the Passive Viewing Paradigm.

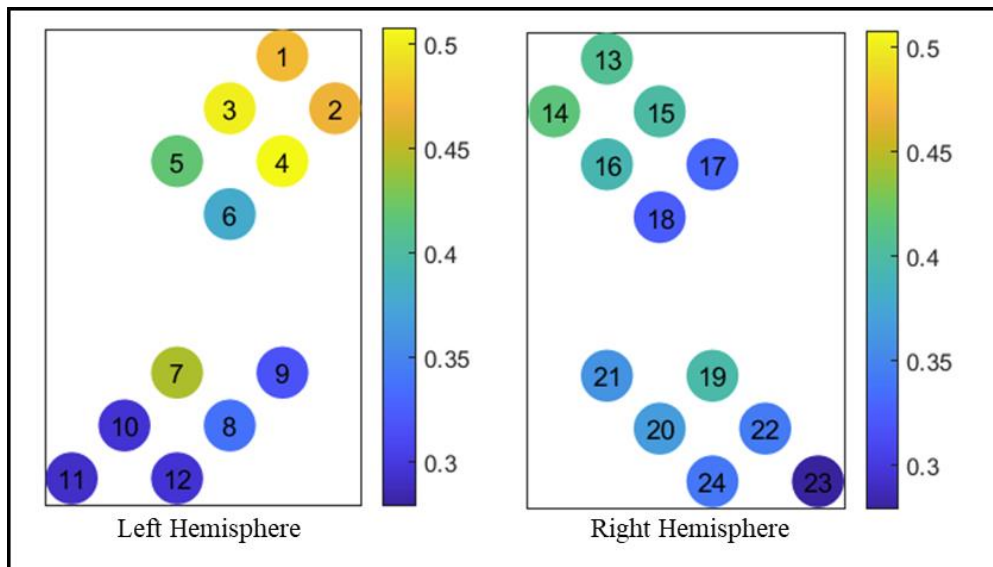


Figure 5.4: Whole-Brain functional connectivity group average for the Mind-Wandering Paradigm.

However, a group-level paired t-test was also performed on the whole-brain functional connectivity analyses between the two measures. For the whole brain oxygenated hemoglobin connectivity levels, there were no channels with a statistically significant ($p < 0.05$) difference. The visual differences in the group-level average are still indicative that there was more spontaneous neural activity during the Mind-Wandering Paradigm, but further questions were answered using the functional connectivity matrix approach.

The full functional connectivity matrix calculates values based on each channel's connectivity with each other channel. A group average, like the one done with the whole brain functional connectivity values, was calculated using the individual functional connectivity matrices. The full group average functional connectivity matrix for each separate task is shown in **Figure 5.5** and **Figure 5.6** for the Passive Viewing Paradigm and the Mind-Wandering Paradigm, respectively.

It is difficult to understand the differences between the two tasks' average functional connectivity by channel in the two matrices, visually. However, to better understand what differences were present between the two tasks, a paired t-test was done at the group level to find channels where there were significant differences in the connectivity value. This t-test indicated seven statistically significant channel connectivity differences between the two tasks. In **Figure 5.5** and **Figure 5.6**, these significant differences are outlined in red. Note that there are two squares representing each value (i.e. channel two to channel five is shown twice), so there are fourteen squares representing significant differences between the two tasks. Visually, it is still difficult to interpret these differences, so **Table 5.16** shows the group average for each task and the p-value for the statistical significance from the paired t-test for each significant connectivity value difference.

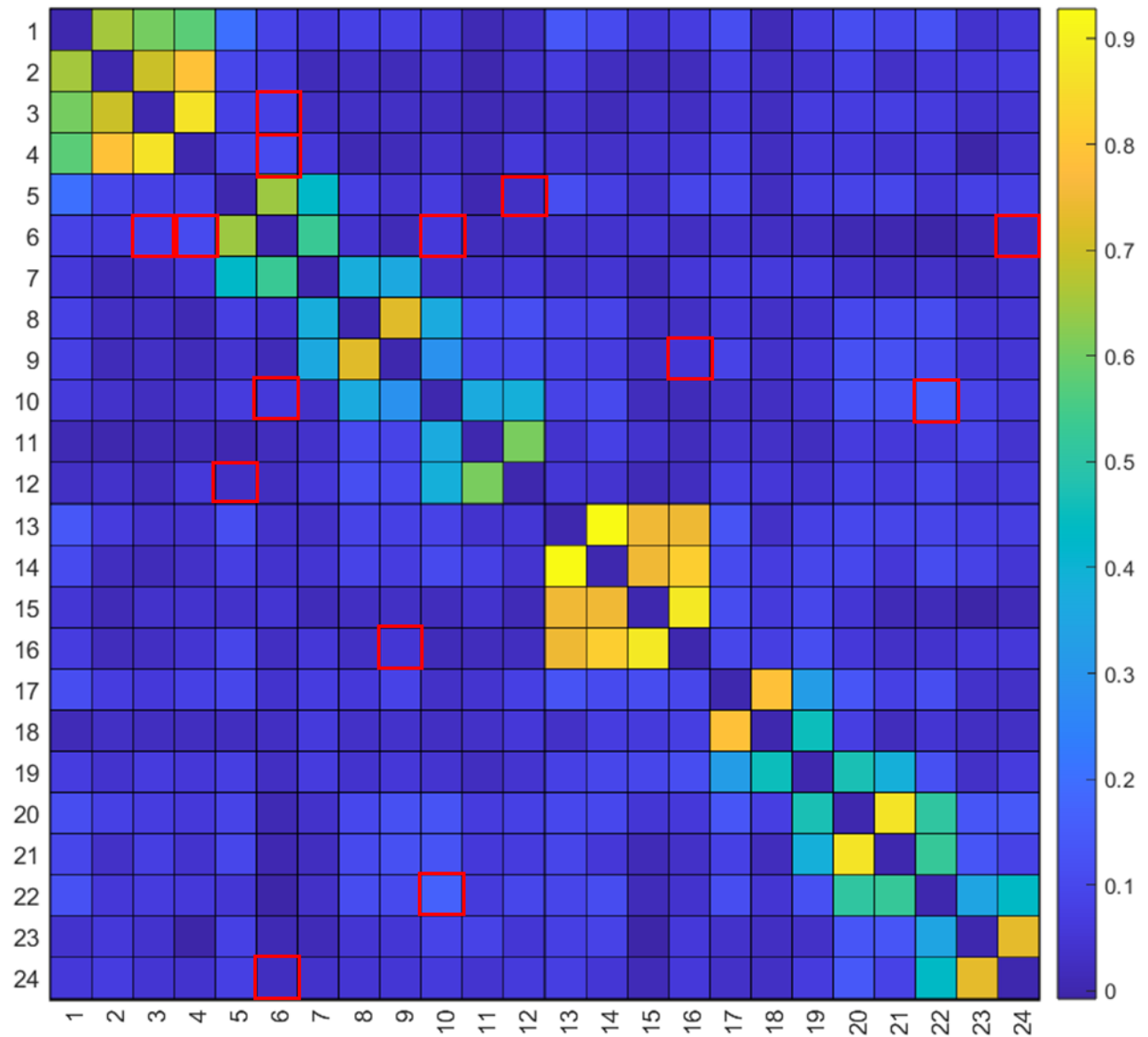


Figure 5.5: Functional connectivity matrix for group average on the Passive Viewing Paradigm (significance is denoted with a red outline, $p < 0.05$).

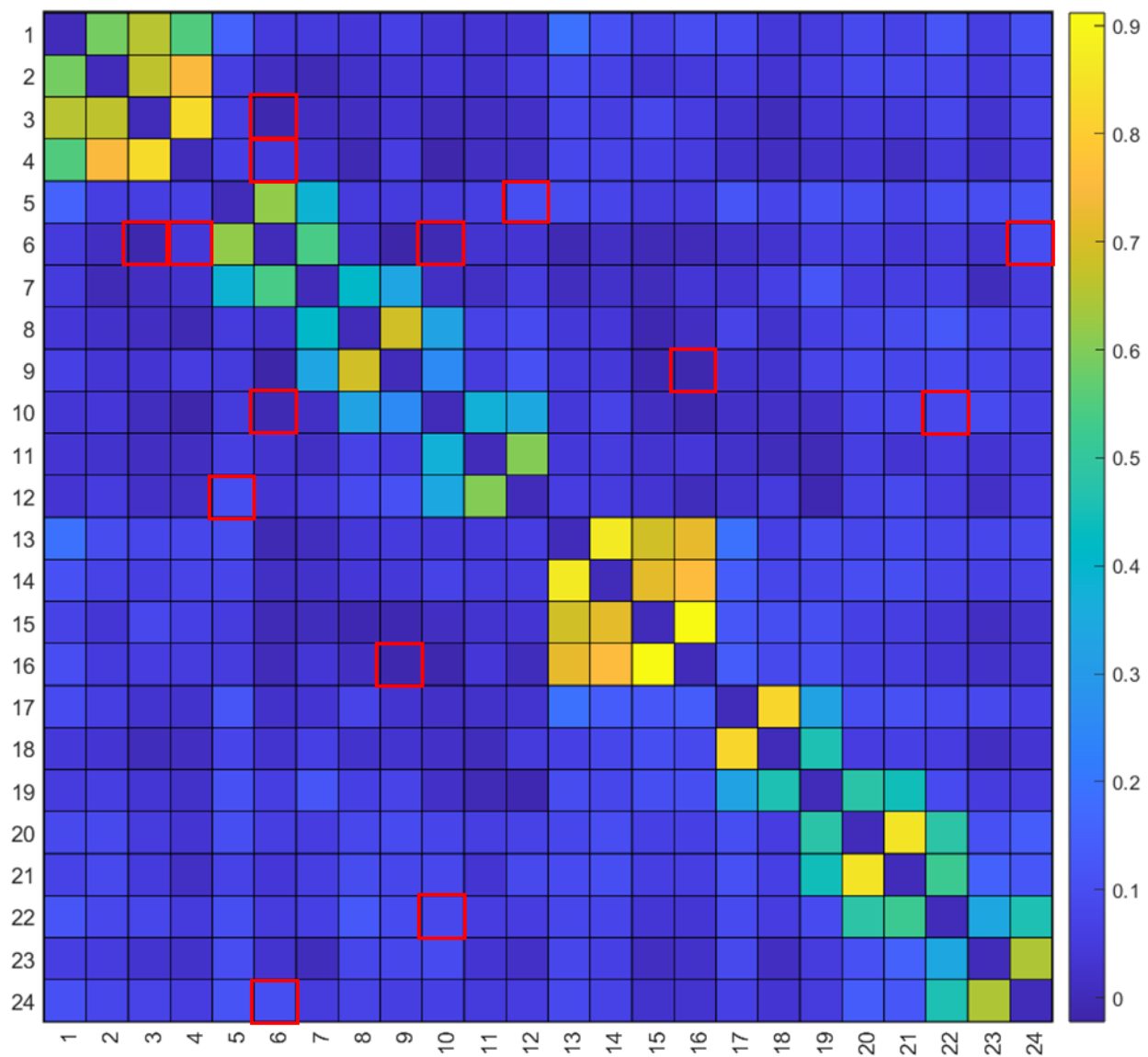


Figure 5.6: Functional connectivity matrix for group average on the Passive Viewing Paradigm (significance is denoted with a red outline, $p < 0.05$).

Table 5.16: Statistically significant functional connectivity group averages between resting-state tasks.

Channels	Regions	Average (PVP)	Average (MWP)	p-value (<0.05)
3 $\leftrightarrow$ 6	IFC $\leftrightarrow$ ITC	0.0825	-0.0133	0.0181
4 $\leftrightarrow$ 6	IFC $\leftrightarrow$ ITC	0.1090	0.0460	0.0386
5 $\leftrightarrow$ 12	ITC $\leftrightarrow$ IPC	0.0320	0.1031	0.0435
6 $\leftrightarrow$ 10	ITC $\leftrightarrow$ IPC	0.0583	-0.0070	0.0310
6 $\leftrightarrow$ 24	ITC $\leftrightarrow$ rPC	0.0237	0.1032	0.0113
9 $\leftrightarrow$ 16	IPC $\leftrightarrow$ rFC	0.0513	-0.0124	0.0283
10 $\leftrightarrow$ 22	IPC $\leftrightarrow$ rPC	0.1672	0.0871	0.0108

From **Table 5.16**, one can see that of the seven significant differences between the average functional connectivity by channel for the two tasks, only two of them are higher for the Mind-Wandering Paradigm. This means that most of the significant differences between the two tasks indicate higher functional connectivity values for the Passive Viewing Paradigm. This does agree with our guiding hypothesis. Visually, the average connectivity for each channel may seem greater in the Mind-Wandering Paradigm than the Passive Viewing Paradigm, but statistically, this difference is not significant. It was expected that the differences between the two resting-state tasks would be minimal especially in adult participants, and this is demonstrated by the analysis done to compare the two tasks. Within these minute differences, however, the Passive Viewing Paradigm does demonstrate higher levels of functional connectivity than the Mind-Wandering Paradigm. While it is still difficult to determine a “best method” for collecting this data in adult participants, this data allows for future analyses to focus on what may cause these differences between the two types of tasks. One step in this direction is to use both sets of data in the rsFC and EF study rather than selecting one over the other. By continuing the methods comparison while analyzing correlations to behavioral data, a better understanding of the difference between the tasks can be discovered.

Executive Function and rsFC Study

To analyze the structure of executive function in neurocognitive terms, the first steps were the same as for the methods comparison study. First, all data for both tasks was processed, and whole brain functional connectivity profiles were created at the individual level. Then, these whole brain functional connectivity profiles were analyzed at the group-level by examining correlations to all fifteen behavioral EF tasks. This included all nine simple tasks and all five complex EF tasks including both components for the Random Number Generator task. This analysis was performed for both resting-state measures. For the Passive Viewing Paradigm, all of the whole brain functional connectivity correlational profiles with all behavioral tasks are shown in **Figure 5.7** (simple EF tasks) and **Figure 5.8** (complex EF tasks).

In these figures, higher values indicate greater magnitude correlation coefficients. Because all of the task dependent measures were set so that higher values represented better performance, the higher values in this correlation analysis indicate a positive correlation or direct relationship, which can be interpreted as demonstrating that stronger connectivity correlates with better task performance. **Figure 5.7** shows the nine simple EF tasks organized by latent construct, or EF component, from the first part of this study-shifting, updating, and inhibition (by column). Visually, it is difficult to see trends or patterns based on the EF component or task, and there is a lot of variation in where in the brain stronger resting-state connectivity can be interpreted as indicating better task performance. Some tasks such as the Antisaccade task, Number-Letter task, and Tone Monitoring task do seem to indicate that stronger connectivity in the frontal cortex indicates better EF task performance as predicted. However, this cannot be said for all tasks, and some even show the opposite.

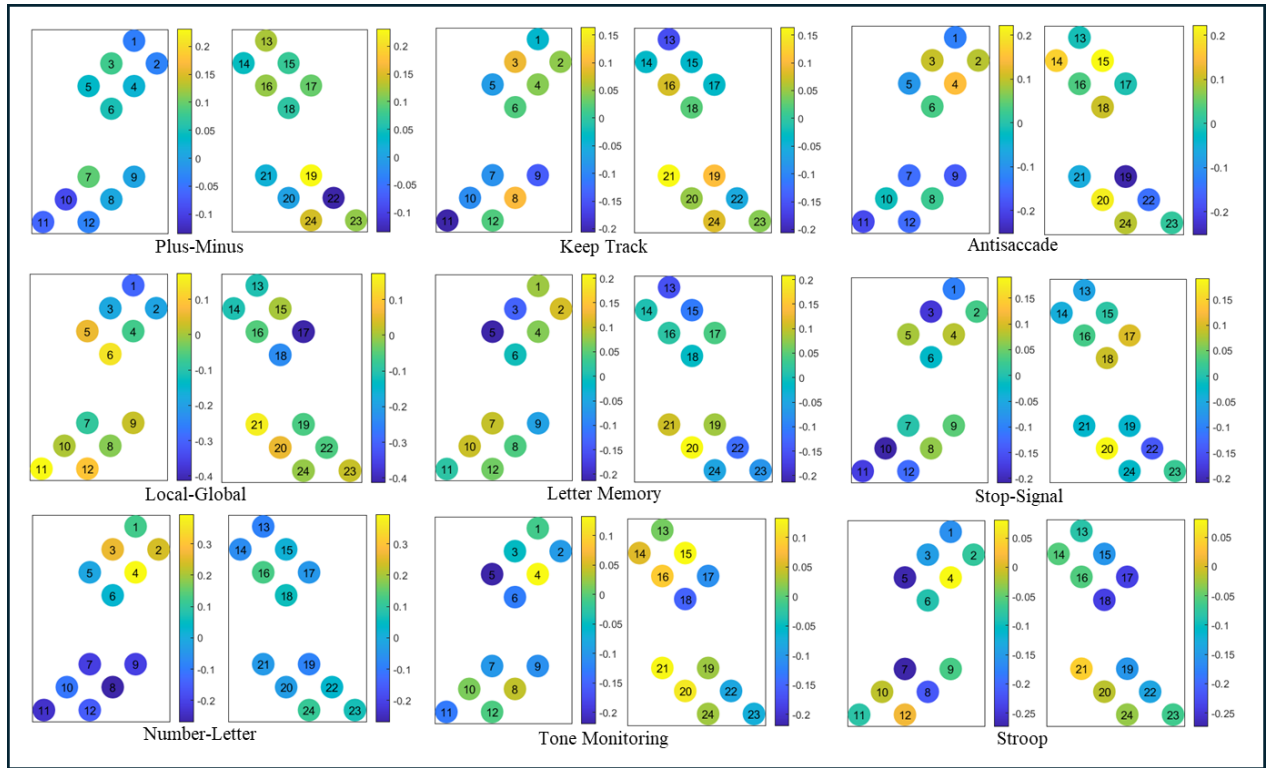


Figure 5.7: Whole-brain functional connectivity correlation profiles for the Passive Viewing Paradigm and all nine simple EF tasks.

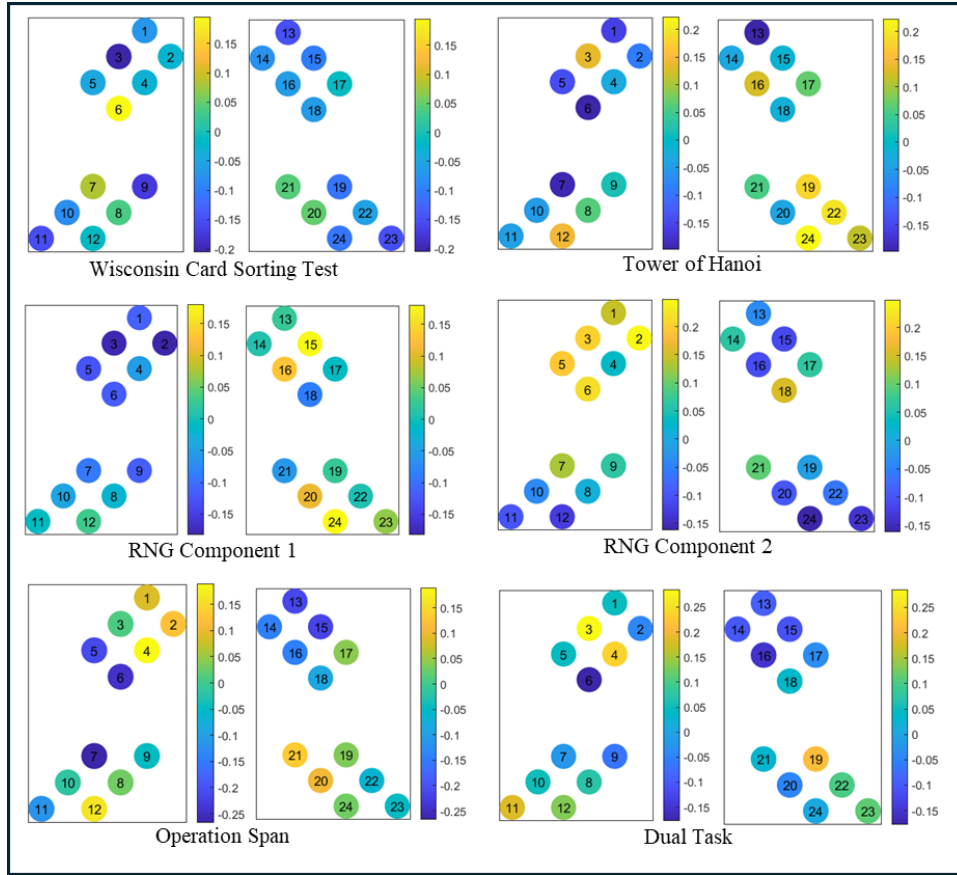


Figure 5.8: Whole-brain functional connectivity correlation profiles for the Passive Viewing Paradigm and all complex EF tasks.

The same variation and nuance can be seen for the complex EF tasks in **Figure 5.8**. However, only the WCST and TOH tasks seem to go against the premise that stronger resting-state functional connectivity indicates better EF task performance. To attempt to achieve a better understanding of this nuance, **Table 5.17** shows the significant correlations found through this analysis.

From the significant correlations in **Table 5.17**, more can be understood about the relationship between resting-state functional connectivity during the Passive Viewing Paradigm and the EF tasks in this study. Not all of the tasks had significant correlations with the resting-state functional connectivity data, and an interesting finding is that the majority of the significant correlations are negative. For the Local-Global task, decreased functional connectivity in the left frontal cortex and right temporal cortex are associated with better performance. For the Number-Letter task, increased connectivity in the left frontal cortex indicates better task performance as does decreased connectivity in the left parietal cortex.

Table 5.17: *Significant correlations by channel for the Passive Viewing Paradigm and behavioral EF tasks.*

Task	Channel	Region	Correlation	p-value (<0.05)
Local-Global	1	lFC	-0.294	0.0185
Local-Global	17	rTC	-0.412	<0.0001
Number-Letter	3	lFC	0.264	0.0348
Number-Letter	4	lFC	0.395	0.0013
Number-Letter	8	lPC	-0.268	0.0325
Antisaccade	19	rTC	-0.251	0.0452
Stroop	5	lTC	-0.270	0.0312
Stroop	7	lTC	-0.272	0.0293
RNG C2	2	lFC	0.249	0.0472
Operation Span	7	lTC	-0.266	0.0339
Dual Task	3	lFC	0.285	0.0223

Better Antisaccade task performance is associated with decreased connectivity in the right temporal cortex, and better Stroop task performance is associated with decreased connectivity in the left temporal cortex. Better performance on the “prepotent associates” component of the Random Number Generator task is associated with stronger connectivity in the left frontal cortex. Better Operation Span performance is associated with less connectivity in the left temporal cortex. Finally, better dual task performance is associated with stronger left frontal cortex connectivity. The guiding hypothesis was that better EF performance would be associated generally with stronger connectivity in the frontal cortex or parietal cortex. The Number-Letter, Random Number Generation, and Dual Task agree with this principle. However, most tasks are shown to correlate more with weaker or decreased connectivity in the temporal cortex for the Passive Viewing Paradigm. One reason for this may be that the resting-state task itself requires the participant to inhibit themselves and could therefore be deemed an inhibitory or inhibition task (Camacho et. al., 2020). If this is the case, it would make sense that participants with decreased connectivity could be “resting” better, which would mean they are inhibiting themselves more and therefore this would associate with EF performance. This is an interesting line of further inquiry.

Next, the same analyses were performed using the Mind-Wandering Paradigm to further understand differences between these resting-state fNIRS methods as well as the relationships between resting-state functional connectivity and EF outcomes. As with the Passive Viewing Paradigm, a visual representation of the correlations between the whole brain functional connectivity profiles and each EF task are shown in **Figure 5.9** (simple EF tasks) and **Figure 5.10** (complex EF tasks).

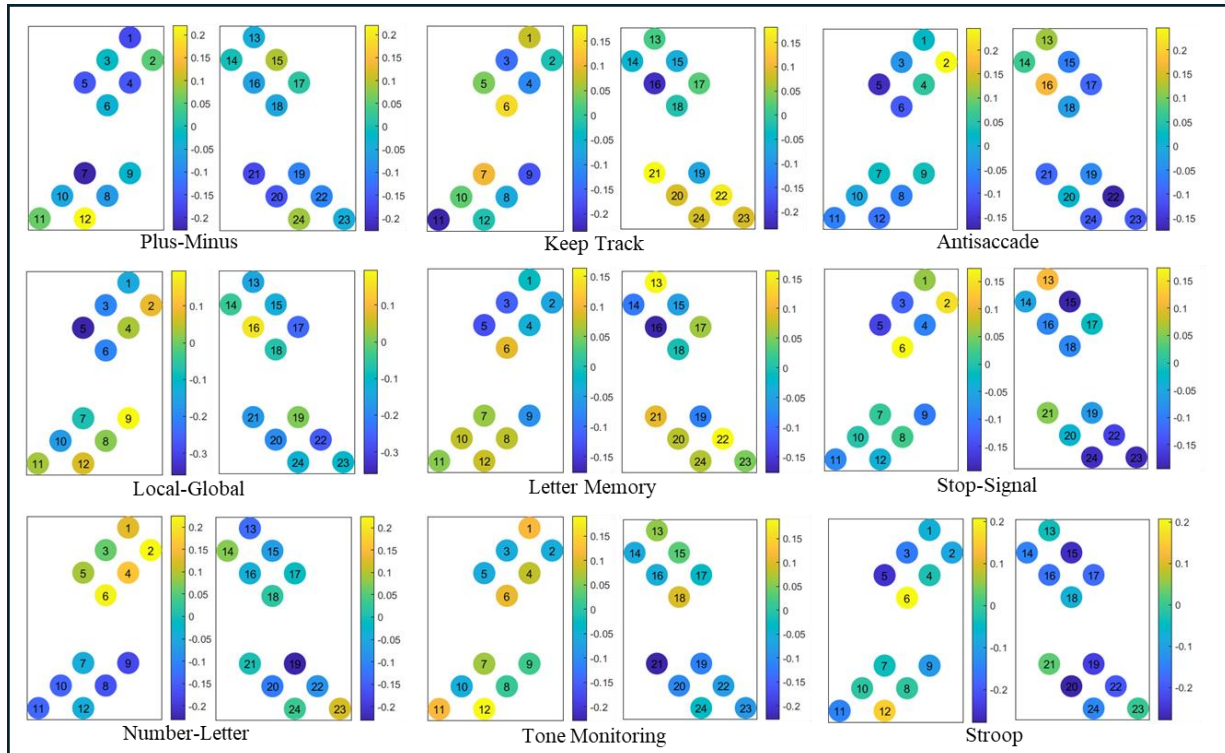


Figure 5.9: Whole-brain functional connectivity correlation profiles for the Mind-Wandering Paradigm and all nine simple EF tasks.

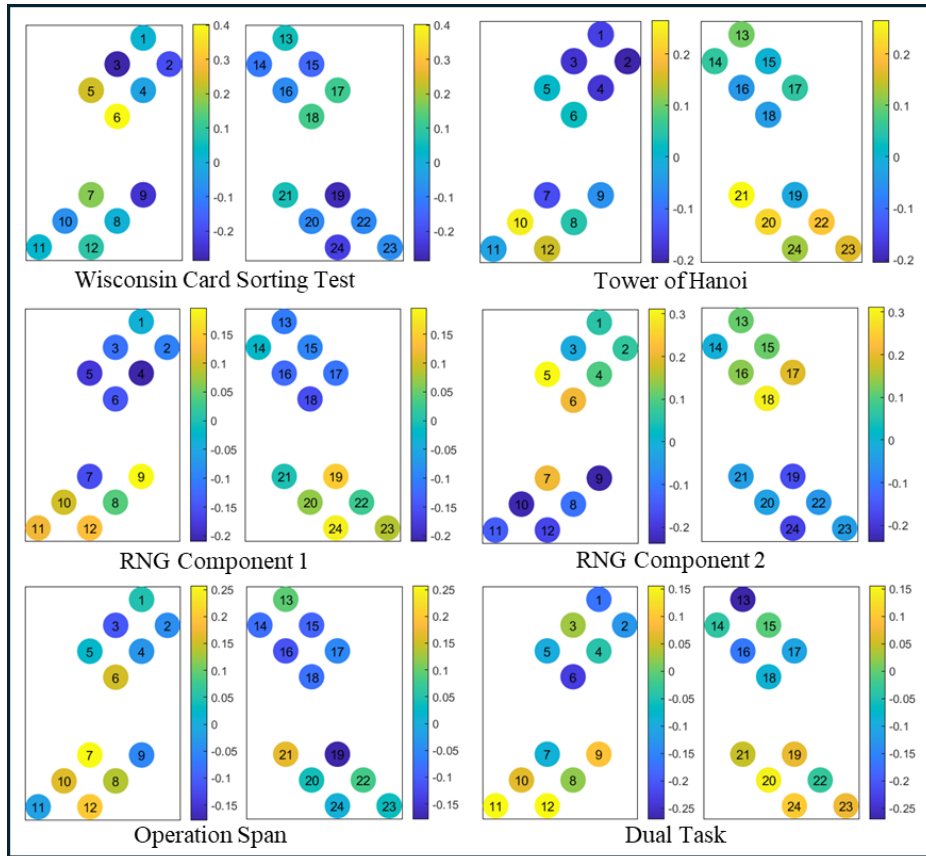


Figure 5.10: Whole-brain functional connectivity correlation profiles for the Mind-Wandering Paradigm and all complex EF tasks.

As seen in **Figure 5.9** and **Figure 5.10**, there are a lot of nuances in how the whole brain functional connectivity profiles for the Mind-Wandering Task correlate with the simple and complex EF measures. From the figures depicting the correlations with the simple EF tasks, there is not much difference between the Mind-Wandering and Passive Viewing Paradigms. One difference, however, is that the memory tasks (Keep Track, Letter Memory, and Tone Monitoring) do seem to have some stronger correlations with the data from the Mind-Wandering Paradigm. This is similarly true for both Random Number Generation components as well as the Operation Span task and Dual Task when looking at the complex EF tasks. The nuance and more specific differences can be broken down and better understood by analyzing **Table 5.18**, which shows the significant correlations between the Mind-Wandering Paradigm resting-state connectivity and the EF tasks.

The significant correlations shown in **Table 5.18** again do not include all the tasks, and a lot of them are negative. The tasks with significant correlations differ from those shown to

Table 5.18: *Significant correlations by channel for the Mind-Wandering Paradigm and behavioral EF tasks.*

Task	Channel	Region	Correlation	p-value (<0.05)
Local-Global	5	ITC	-0.359	0.0036
Local-Global	22	rPC	-0.262	0.0367
Antisaccade	2	lFC	0.246	0.0499
Stroop	5	ITC	-0.252	0.0499
Stroop	15	rFC	-0.259	0.0391
Stroop	20	rPC	-0.277	0.0266
WCST	3	lFC	-0.285	0.0223
WCST	6	ITC	0.403	<0.0001
WCST	19	rTC	-0.272	0.0299
TOH	10	lPC	0.250	0.0463
TOH	21	rPC	0.265	0.0341
RNG C2	5	ITC	0.312	0.0120
RNG C2	18	rTC	0.297	0.0171
Operation Span	7	ITC	0.257	0.0403
Dual Task	13	rFC	-0.269	0.0314

correlate significantly with the Passive Viewing Paradigm as well. Better performance on the Local-Global task is associated with decreased connectivity in the left temporal cortex and right parietal cortex. Better performance on the antisaccade task is associated with stronger connectivity in the left frontal cortex. Better performance on the Stroop task is correlated with weaker connectivity in the left temporal cortex, right frontal cortex, and right parietal cortex. Some important differences here compared to the Passive Viewing Paradigm are the antisaccade task correlation, which here is closer to what would be expected from previous literature. The results from the Local-Global and Stroop Tasks are like that seen in **Table 5.17**, but the right frontal cortex is implicated here for the Stroop task. It is also interesting that the Number-Letter task has no significant correlations while there were three significantly correlated channels with the Passive Viewing Paradigm data.

There are also interesting differences when looking at the complex EF tasks, which have more significant associations with the Mind-Wandering Paradigm. Better performance on the WCST is correlated with decreased connectivity in the left frontal cortex and right temporal cortex, which is not what would be expected as well as increased connectivity in the left temporal cortex. Better Tower of Hanoi performance is associated with stronger connectivity in the left and right parietal cortices. While not shown in the results from the Passive Viewing Paradigm, this bilateral correlation with the parietal cortex does make sense based on what this EF task is. The “prepotent associates” component of the Random Number Generation task is associated with stronger connectivity in both the left and right temporal cortices. The Operation Span task performance is associated with stronger connectivity in the left temporal cortex, which is the opposite of what was found in the Passive Viewing Paradigm correlations. Finally, better performance on the Dual Task is correlated with decreased connectivity in the right frontal

cortex, which is also opposite what was found for the left frontal cortex in the Passive Viewing Paradigm. Overall, there are more differences than similarities between the correlational results from the two different resting-state measures at the group level. This would indicate that the two tasks differ in how they assess rest in the brain because rest should correlate equally to these measures if it is truly being measured in the same way. Regarding developing an understanding of how EF is structured in the brain, this data demonstrates that EF tasks have different associated neural networks. Different regions of the brain may be associated more with a task itself (either positively or negatively) rather than in a certain way with a “component” of EF or EF as one overarching cognitive process. From previous literature, it is also clear that there are many individual differences in how resting-state neural data is correlated with EF outcomes (Roye, et. al., 2020; Reineberg et. al., 2016). While there is evidence supporting associations with connectivity in the frontoparietal network from this data (Reineberg, et. al., 2016), it is possible that more could be understood by examining this data at the individual, region-specific level rather than at the group, whole brain level.

The guiding principle behind expanding the replication study to include resting-state functional connectivity analyses was to see if there was a statistical structure of EF that could be based in neurocognitive measures rather than behavioral measures alone. For this, principal components analysis was performed on all the channel correlations for each task (for both the Passive Viewing Paradigm and Mind-Wandering Paradigm). For the Passive Viewing Paradigm, the results from this analysis are shown in **Table 5.19**.

As described above, this analysis was done using the whole brain channel correlations at the group level for the Passive Viewing Paradigm and behavioral EF tasks. The goal was to see if certain components would be extracted made up of specific sets of the EF tasks. There was no

Table 5.19: Pattern matrix (loadings) for the components extracted using PCA on the correlations between the Passive Viewing Paradigm and all EF tasks.

Task	Component				
	1	2	3	4	5
Plus-Minus	0.188	-0.078	0.035	-0.809	-0.017
Local-Global	0.502	0.149	-0.155	0.390	0.289
Number-Letter	-0.256	-0.194	0.198	0.038	-0.868
Keep Track	0.071	0.458	0.518	-0.378	-0.289
Letter Memory	0.025	0.840	0.226	0.143	0.089
Tone Monitoring	0.624	0.253	0.103	0.079	-0.412
Antisaccade	0.147	0.007	-0.327	-0.073	-0.860
Stop-Signal	-0.003	0.454	-0.225	-0.563	-0.169
Stroop	0.482	0.135	0.138	0.603	-0.400
WCST	-0.165	0.836	-0.276	-0.040	0.126
TOH	0.441	-0.146	0.715	-0.277	0.082
RNG C1	0.929	-0.093	-0.099	-0.260	-0.038
RNG C2	-0.929	0.178	-0.088	-0.018	-0.140
Operation Span	-0.027	0.395	0.756	0.194	-0.194
Dual Task	-0.165	-0.192	0.820	0.136	0.138

specific number of factors extracted, it was done based on the eigenvalues and scree plot. For this analysis, specifically, five components were extracted as shown in **Table 5.19**. This was performed with a direct oblimin rotation to allow the components to correlate freely. The tasks were selected as part of a component if the loading was greater than 0.3, or if none were greater than 0.3, the component with the largest loading was selected (and denoted with an asterisk).

This is shown in **Table 5.20**.

In relation to the original component structure (Miyake et. al., 2000), these components are not easily interpretable. First, despite extracting five components, the fifth has no significant loading, so this can be interpreted as a four-component structure.

Many of the tasks had significant loadings onto multiple components, and four had very low or negative loadings for all the components. Also, they are not split in any organized way based on task type (i.e. shifting, updating, and inhibition) or a clear way based on whether they

Table 5.20: Proposed statistical structure of EF using the Passive Viewing Paradigm (* denotes tasks with no loading >0.3).

	Component				
	1	2	3	4	5
Task	Plus-Minus*	Keep Track	Number-Letter*	Local-Global	
	Local-Global	Letter Memory	Keep Track	Stroop	
	Tone Monitoring	Stop-Signal	TOH		
	Antisaccade*	WCST	Operation Span		
	Stroop	RNG C2*	Dual Task		
	TOH	Operation Span			
	RNG C1				

are simple or complex tasks. However, what this does seem to support, which was indicated by the analysis of the correlations themselves as well, is that each task, while all under the umbrella of EF, associates differentially in the brain. In other words, EF is not as simple as a set of separate but correlated components (Miyake et. al., 2000) but rather a dynamic system with many processes associated with it and as properties of it (McCraw et. al., 2024). As one final methods comparison analysis, this PCA was repeated with the correlations from the Mind-Wandering Paradigm, which is shown in **Table 5.21**.

Using the same PCA method as with the Passive Viewing Paradigm, there were four components extracted using the Mind-Wandering Paradigm rsFC data. The structure created by selecting tasks with a coefficient greater than 0.3 (or the largest loading if necessary) is shown in **Table 5.22**.

As seen with the Passive Viewing Paradigm, there is no clearly interpretable statistical structure based on the components or types of tasks. For the Mind-Wandering Paradigm, however, there is not an unused component, and fewer tasks do not load on any component highly. This structure is also very different than the structure created with the Passive Viewing Paradigm overall. Not only does this further support the idea that the component view of EF is

Table 5.21: Pattern matrix (loadings) for the components extracted using PCA on the correlations between the Mind-Wandering Paradigm and all EF tasks.

Task	Component			
	1	2	3	4
Plus-Minus	-0.266	-0.070	0.710	0.123
Local-Global	-0.371	-0.124	0.379	-0.691
Number-Letter	0.670	0.009	0.102	0.025
Keep Track	0.142	0.608	-0.431	0.338
Letter Memory	-0.251	0.817	-0.019	0.189
Tone Monitoring	0.290	0.206	0.840	-0.015
Antisaccade	0.058	-0.092	-0.124	-0.921
Stop-Signal	0.173	0.686	-0.150	-0.636
Stroop	0.055	0.786	0.392	-0.217
WCST	0.603	0.548	0.188	0.262
TOH	-0.506	0.450	-0.269	0.400
RNG C1	-0.939	0.135	0.082	0.032
RNG C2	0.965	-0.129	-0.067	0.147
Operation Span	-0.062	0.906	0.001	0.046
Dual Task	-0.786	-0.115	0.146	0.263

Table 5.22: Proposed statistical structure of EF using the Mind-Wandering Paradigm (* denotes tasks with no loading >0.3).

	Component			
	1	2	3	4
Task	Number-Letter	Keep Track	Plus-Minus	Keep Track
	Antisaccade*	Letter Memory	Local-Global	TOH
	WCST	Stop-Signal	Tone Monitoring	Dual Task*
	RNG C2	Stroop	Stroop	
		WCST		
		TOH		
		RNG C1		
		Operation Span		

not well supported when developing a structure from neurocognitive data rather than behavior alone, but it also further supports the idea that the two resting-state tasks used in this study are, in fact, different. The hypothesis that there is a specific structure of EF that can be found may be true, but from this data, it seems that the definition of EF in this structured, component-focused method does not hold up. While resting-state neural data does differentially correlate with EF task outcomes (based on regions in the frontal, parietal, and temporal cortices), this study did not yield one simple answer to this question.

Chapter 6: Discussion

Overall Conclusions

To summarize, this study consisted of a replication study, methods comparison study, and an examination of how resting-state neural data correlates with executive function performance. The goal of this study was to develop a better understanding of the component view of EF and resting-state functional connectivity using fNIRS.

We were unable to replicate Miyake's component view of EF using confirmatory factor analysis and structural equation modeling and the same battery of tasks used in the original study. In the original study, they endorsed a full three-factor view of EF, which indicated that the components (shifting, updating, and inhibition) were separable but correlated to each other due to an underlying EF factor (Miyake et. al., 2000). However, using confirmatory factor analysis, we endorsed a two-factor model where shifting and inhibition were the same construct (not separable). This discrepancy could have many causes. One could be that because there is generally limited reliability in EF tasks (Miyake et. al., 2000), while we replicated the tasks to be the same as in the original study, slight differences could have changed the statistical structure we deduced. However, the inability to replicate this study could also support the idea that the component view does not fully define executive function. Executive function may not be as simple as a set of component processes that are distinct and correlated behaviorally.

The original study then performed structural equation modeling with five complex EF tasks (including two components from the Random Number Generation task). Unlike the confirmatory factor analysis portion of this study, we were able to partly replicate this portion. We first used the full three-factor model as the base model in order to complete the replication, but because we endorsed the two-factor model in the CFA, we had to assess the SEM models

using this as the base model as well. Ultimately, this set of analyses led to stronger support for the two-factor model because each endorsed SEM model was directly compared between the two base models, and for each of them, other than the WCST, the model using the two-factor model was the endorsed best fitting model. Despite this, we only replicated the endorsed model for the “prepotent associates” component of the Random Number Generation task, the Operation Span task, and the Dual Task when compared to the original study. For the other complex tasks, we endorsed different paths to the EF components as shown in the results section.

After the replication study, we compared the two resting-state methods used-the Passive Viewing Paradigm and the Mind-Wandering Paradigm. From this analysis, we deduced that the two methods of collecting resting-state functional connectivity data using fNIRS are different. From direct comparison, only some channels were statistically different between the tasks, which is in line with the guiding hypothesis that the task differences would be minimal in adults. However, from part of this analysis, it was indicated that there was more activity during the Mind-Wandering Paradigm than in the Passive Viewing Paradigm, which was not expected. We continued the methods comparison portion of the study in conjunction with the analysis of the EF measures to better understand the task differences. Ultimately, there were very different EF profiles between the two resting-state tasks as well. This demonstrates that using a variety of tasks and defining the data as resting-state functional connectivity is causing discrepancies in research using these methods. From the analyses done in this study, further examination of task differences needs to be performed, and there are still many questions as to why these tasks were so different in comparison to the EF measures. A guiding theory after completing these analyses is that because watching videos is stimulating to the brain, the Mind-Wandering Paradigm may be the best way to collect resting-state data. Without stimulation, the brain is more likely to be at

true rest than it is while being stimulated with visual and auditory stimulation. However, a competing theory is that the Mind-Wandering Paradigm may have more significant correlations to EF tasks because it, itself, is more similar to a measure of inhibition than “rest.” There are theories that resting-state tasks measure inhibition (Camacho et. al., 2020), so it could be that without the stimulation given during the Passive Viewing Paradigm, participants are having to use inhibitory control mechanisms during the Mind-Wandering Paradigm. With more analyses, it could be that this task is more similar to other inhibition EF tasks than rest.

The final portion of the study was to use the neurocognitive measures from the resting-state functional connectivity profiles to reframe the component view of executive function. We used whole brain, group level analyses for this section of the study. Ultimately there were a lot of nuances in the results for how the resting-state functional connectivity data correlated with each of the simple and complex EF tasks. From this nuance it was clear that associations from the frontal, parietal, and temporal cortices at rest are indicative of differential outcomes on EF tasks. However, this was not organized based on a component or set of components such as shifting, updating, and inhibition. Moreover, after running PCA on the correlations from the whole-brain functional connectivity analyses, the component structure extracted was not organized based on the components originally defined. What is clearer is that the component view of EF is too simple to capture the nuance within the dynamic process that is executive function and other cognitive processes defined as component processes of executive function. Newer ideas about how EF is organized and how EF develops in this brain such as dynamic systems (McCraw et. al., 2024) can capture this nuance, and we should not limit EF to being studied on the basis of behavioral measures organized in this component processes manner.

Limitations

The biggest limitation of this study is the sample size. The original study had a sample of 137 participants, which is more than double the sample size of the current study (Miyake et. al., 2000). Despite being larger than our current sample size, based on power analyses done prior to beginning this replication study, the sample size necessary to have enough power to detect statistical significance on all parameters in the CFA and SEM models is much larger. Therefore, some of the inability to replicate the original study, and some of the difficulties in interpreting the results of the fNIRS portions of this study could stem from having such a small sample of data. Further research will continue to grow this sample and replicate the analyses completed now to develop a better understanding of the results found here. Beyond having a larger sample for replication purposes, the data from the behavioral measures (after trimming and transforming the data) had high kurtosis values. While there were no major outliers in the data, these kurtosis levels could have had large effects on the results from the CFA and SEM analyses. This could be lowered by increasing the sample size of the study as well, but different trimming/transformation procedures or more strict outlier analyses could also help lower these values. If the kurtosis values were not minimized by increasing the sample size or by adapting the trimming procedures, the CFA and SEM analyses could be done using a different estimation method than the Maximum Likelihood method that was used in the original study. However, for the purposes of the replication study, the estimation method was not changed for this study.

A second limitation not having to do with sample size or statistics is that despite previous literature, analysis of task differences and EF correlations with the resting-state functional connectivity data was not performed at the individual level. It has been demonstrated that the nuance within how resting-state functional connectivity data correlates to EF outcomes has a lot

to do with differences between individuals (Reineberg et. al., 2016). All analyses performed for this study were done at the group level. By analyzing the differences and correlations at the individual level, a clearer “best method” or EF structure could be deduced. This is a line of further inquiry with the data in this current study.

Future Directions

One of the primary motivations behind this study was to have a starting point for which to understand these ideas in developing populations. Not only is there a lot to learn about executive function development, but the idea of rest is not well understood in developing populations either (Camacho et. al., 2020). By having the data from this study, we can expand this study to directly compare these results to results from analyzing this data in children. The main future direction will be to repeat this methods comparison study in younger children and directly compare that data with data from adults as well. However, it is first necessary to continue to understand the nuance in adult resting-state functional connectivity. Part of this would be to expand this work to look at the individual differences between the two resting-state tasks as discussed above. Beyond the methods comparison study, this opens new avenues of studying the associations between resting-state functional connectivity data and EF outcomes. We can use the data from this study to inform specific regions of interest in tandem with specific EF task data as well. Not only this, but we can also expand the fNIRS data to include task-based data using regions of interest from this first study. By further researching these questions, we can develop a clearer understanding of what EF is and how it can be studied in the brain. Once we do that, the path is even clearer to understand these ideas in developing populations as well.

References

- Bulgarelli, C., Blasi, A., C.J.M. de Klerk, C., Richards, J.E., Hamilton, A., Southgate, V., (2019). Fronto-temporoparietal connectivity and self-awareness in 18-month-olds: A resting state fNIRS study. *Developmental Cognitive Neuroscience*, 38. <https://doi.org/10.1016/j.dcn.2019.100676>.
- Camacho, C.M., Quinones-Camacho, L.E., Perlman, S.B., (2020). Does the child brain rest?: An examination and interpretation of resting cognition in developmental cognitive neuroscience. *NeuroImage*, 212. <https://doi.org/10.1016/j.neuroimage.2020.116688>.
- Ekstrom, R. B., French, J. W., Harman, H. H., Derman, D., (1976). *Kit of Factor-Referenced Cognitive Tests*. Educational Testing Service.
- Eng, C.M., Pocsai, M., Fulton, V.E., Moron, S.P., Thiessen, E.D., Fisher, A.V., (2022). Longitudinal investigation of executive function development and task-based, teacher reports, and fNIRS multimethodology in 4- to 5-year-old children. *Developmental Science*. <https://doi.org/10.1111/desc.13328>.
- Fernandez, Z., Scheel, N., Baker, J. H., Zhu, D. C., (2022). Functional connectivity of cortical resting-state networks is differentially affected by rest conditions. *Brain Research*. <https://doi.org/10.1016/j.brainres.2022.148081>
- Francis, W. N., Kucera, H., Mackie, A. W., (1982). *Frequency Analysis of English Usage: Lexicon and Grammar*. Houghton Mifflin Company.
- Hou, X., Zhang, Z., Zhao, C., Duan, L., Gong, Y., Li, Z., Zhu, C., (2021). NIRS-KIT: toolbox for both resting-state and task fNIRS data analysis. *Neurophotonics* 8. <https://doi.org/10.1117/1.NPh.8.1.010802>

- Hu, Z., Liu, G., Dong, Q., Niu, H., (2020). Applications of Resting-State fNIRS in the Developing Brain: A Review From the Connectome Perspective. *Frontiers in Neuroscience*, 14. <https://doi.org/10.3389/fnins.2020.00476>.
- Humes, G. E., Welsh, M. C., Retzlaff, P., Cookson, N., (1997). Towers of Hanoi and London: Reliability and Validity of Two Executive Function Tasks. *Assessment*, 4(3), 249-257. doi: 10.1177/107319119700400305.
- Huppert, T. J., Diamond, S. G., Franceschini, M.A., Boas, D. A., (2009). HomER: A review of time-series analysis methods for near-infrared spectroscopy of the brain. *Appl. Opt.* 48. <https://doi.org/10.1016/j.neuroimage.2007.09.031>
- Kerr-German, A., White, S.F., Santosa, H., Buss, A. T., Doucet, G. E., (2022). Assessing the relationship between maternal risk for attention deficit hyperactivity disorder and functional connectivity in their biological toddlers. *European Psychiatry*, 65(1), e66, 1-9. <https://doi.org/10.1192/j.eurpsy.2022.2325>
- Lu, C.M., Zhang, Y.J., Biswal, B.B., Zang, Y.F., Peng, D.L., Zhu, C.Z., (2010). Use of fnirs to assess resting state functional connectivity. *J. Neurosci. Methods* 186 (2), 242–249. doi:10.1016/j.jneumeth.2009.11.010.
- McCraw, A., Sullivan, L., Lowery, K., Eddings, R., Heim, H. R., & Buss, A. T. (2024). Dynamic Field Theory of Executive Function: Identifying Early Neurocognitive Markers. *Monographs of the Society for Research in Child Development*, 89(3), 7-109. <https://doi.org/10.1111/mono.12478>
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe"

- tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49-100.
<https://doi.org/10.1006/cogp.1999.0734>
- Morris, N., and Jones, D. M., (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, 81, 111-121.
<https://doi.org/10.1111/j.2044-8295.1990.tb02349.x>
- Navon, D., (1977). Forest Before Trees: The Precedence of Global Features in Visual Perception. *Cognitive Psychology*, 9, 353-383. [https://doi.org/10.1016/0010-0285\(77\)90012-3](https://doi.org/10.1016/0010-0285(77)90012-3)
- Reineberg, A.E., Andrews-Hanna, J.R., Depue, B.E., Friedman, N.P., Banich, M.T., (2015). Resting-state networks predict individual differences in common and specific aspects of executive function. *NeuroImage* 104, 69-78.
[dx.doi.org/10.1016/j.neuroimage.2014.09.045](https://doi.org/10.1016/j.neuroimage.2014.09.045)
- Reineberg, A.E., Banich, M.T., (2016). Functional Connectivity at Rest is Sensitive to Individual Differences in Executive Function: A Network Analysis. *Human Brain Mapping* 37, 2959-2975. doi: 10.1002/hbm.23219
- Reineberg, A.E., Gustavson, D.E., Benca, C., Banich, M.T., Friedman, N.P., (2018). The Relationship Between Resting State Network Connectivity and Individual Differences in Executive Functions. *Front. Psychol.* 9 (1600). doi: 10.3389/fpsyg.2018.01600
- Roye, S., Castagna, J.P., Calamia, M., De Vito, A.N., Lee, T., Greening, S.G., (2020). Relationships between multiple dimensions of executive functioning and resting-state networks in adults. *Neuropsychologia* 141.
doi.org/10.1016/j.neuropsychologia.2020.107418

- Solso, R. L., Juel, C., Rubin, D. C., (1982). The Frequency and Versatility of Initial and Terminal Letters in English Words. *Journal of Verbal Learning and Verbal Behavior*, 21, 220-235.
[https://doi.org/10.1016/S0022-5371\(82\)90581-3](https://doi.org/10.1016/S0022-5371(82)90581-3)
- Towse, J. N., & Neil, D. (1998). Analyzing human random generation behavior: A review of methods used and a computer program for describing performance. *Behavior Research Methods, Instruments & Computers*, 30(4), 583-591. <https://doi.org/10.3758/BF03209475>
- 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](https://doi.org/10.1016/0749-596X(89)90040-5)
- Vanderwal, T., Kelly, C., Eilbott, J., Mayes, L.C., Castellanos, F.X., (2015). Inscapes: A movie paradigm to improve compliance in functional magnetic resonance imaging. *NeuroImage*, 122(15), 222-232. <https://doi.org/10.1016/j.neuroimage.2015.07.069>
- Zhao, J., Liu, J., Jiang, X., Zhou, G., Chen, G., Ding, X.P., Fu, G., Lee, K., (2016). Linking Resting-State Networks in the Prefrontal Cortex to Executive Function: A Functional Near Infrared Spectroscopy Study. *Front. Neurosci.* 10 (452). doi: 10.3389/fnins.2016.00452

Vita

Originally from Lee's Summit, Missouri, Haylee Hudson has been interested in science and research throughout her entire life. After graduating as valedictorian from Lee's Summit High School, she then completed her Bachelor of Science degree in Neuroscience and Chemistry at Creighton University in Omaha, Nebraska where her career in research began. Haylee then continued working as a research assistant studying executive function, attention, and ADHD at Boys Town National Research Hospital for a year before beginning her graduate studies at the University of Tennessee, Knoxville. There, she began pursuing her Master of Arts in experimental psychology, specifically the cognitive and developmental sciences program with a focus on the neural and behavioral components of executive function. Haylee is moving into a career in data science where she hopes to continue developing skills in data analysis, programming, and research methods. Haylee is incredibly grateful for her experiences in research and time at the University of Tennessee, Knoxville as well as her support from her partner, dog, friends, and family who have helped her get to this point in her career.