Characterization of the Relationship Between Executive Functioning Scores in Preschoolers and Performance on Various Imitation Tasks

by Gwyneth Francesca McNiff

B.A. in Philosophy, May 2020, Thomas Aquinas College

A Thesis submitted to

The Faculty of
The Columbian College of Arts and Sciences
of The George Washington University
in partial fulfillment of the requirements
for the degree of Master of Science

May 15, 2022

Thesis directed by

Francys Subiaul
Associate Professor of Speech, Hearing & Language Sciences

© Copyright 2022 by Gwyneth McNiff
All rights reserved

Acknowledgments

I would like to thank Dr. Francys Subiaul for his incredible patience, support, and enthusiasm in completing this thesis project. In addition, I would like to thank Adrian Istvan Varallyay of Columbia University for editing and compiling the initial data, and Dr. Chet Sherwood for approving this thesis. Finally, I would like to acknowledge all the faculty and students at the Center for the Advanced Study of Human Paleobiology for their insights and support throughout this degree.

Table of Contents

Acknowledgments.....iii

List of Figures.....v

List of Tables.....vi

Chapter 1: Introduction.....1

Chapter 2: Task Descriptions.....7

Chapter 3: Data Analyses and Results.....10

Chapter 4: Conclusions.....36

Chapter 5: Future Directions……..……38

References.....40

Appendix A.……...……54

Appendix B……55

Appendix C……56

Appendix D……58

Appendix E…….60

Appendix F…….62

List of Figures

Figure 1: Computerized Task. (A) Cognitive and (B) Motor-Spatial Task…….....9

Figure 2: Tower Building Task……...….9

Figure 3: Interaction Plot for Age and FI in the Cognitive Task..…….....19

Figure 4: Interaction Plot for Age and Inhibit in Motor-Spatial Task……..….....25

Figure 5: Interaction Plot for Age and WM in Motor-Spatial Task……...……..…..26

Figure 6: Boxplot of Mean EMI Score by Imitator Type……..……...….....34

Figure 7: Boxplot of Mean GEC Score by Imitator Type……..…….….....34

Figure 8: Boxplot of Mean WM Score by Imitator Type……..…….…...…..35

Figure 9: Boxplot of Mean Plan Score by Imitator Type……...…….....35

List of Tables

Table 1: Description of BRIEF-P measures……..4

Table 2: Number of individuals by task, having data for BRIEF-P, Age, and Task Outcome……..…….…….……10

Table 3: Regression for relationship between Age and GEC……..11

Table 4: Results of the regression for the individual scores in Cognitive Task……....12

Table 5: Results of the regression for the summary scores in Cognitive Task……..12

Table 6: Summarized results of bootstrap analysis for Cognitive Task.....17

Table 7: Regression results for interaction terms in Cognitive Task …..……..17

Table 8: Summarized results of bootstrap analysis for Cognitive Task, interaction terms removed……..……..20

Table 9: Results of regression for individual scores Motor-Spatial Task……..……..21

Table 10: Results of regression for summary scores Motor-Spatial Task……...21

Table 11: Summarized results of bootstrap analysis for Motor-Spatial Task……..23

Table 12: Summarized results of bootstrap analysis for Motor-Spatial Task with Interaction Terms Removed…….……..……...27

Table 13: Regression results for individual scores for Tower Task…….28

Table 14: Regression results for summary scores for Tower Task……..…28

Table 15: Summarized results of bootstrap analysis for the Tower Task……30

Table 16: Summarized results of most contributive predictors to the model in Dominance analyses…….……..32

Chapter 1: Introduction

Imitation is defined as encoding and reproducing an agent's behavior with medium to high fidelity, that is, preserving most to all observed steps leading to a goal (Brugger et al., 2007; Lyons, Young, & Keil, 2007; Marsh, Ropar, & Hamilton, 2014; Whiten & Ham, 1992). Some have argued that imitation is a uniquely human ability (Byrne, 2005), while others believe that non-human primates evince imitation, pointing to evidence for imitation in familiar tasks (de Waal, 2016; Persson et al., 2017; Whiten, 2019). Even others point to a more nuanced picture, where apes display limited imitation learning skills in restricted task domains in comparison to human preschoolers, who possess a range of imitation learning abilities across many different tasks and domains (Subiaul, 2016).

Imitation may account for the level of complexity and range seen in human cultures. According to cumulative cultural evolution, the ability to transfer information with the high-fidelity afforded by imitation, ensures not only that no information is lost across generations but allows the 'ratcheting' of technology by improving on past practices (Boyd and Richerson 1996; Boyd, Richerson, & Heinrich, 2011; Caldwell, Renner, & Atkinson, 2017; Dean, 2014; Lewis & LaLand, 2012; Mesoudi and Thornton, 2018; Tennie, Call, & Tomasello, 2009; Tomasello, 1999; Wasielewski, 2014). In fact, not only are humans primed to imitate, we can even be said to overimitate in contrast to other primates—that is, we copy both causally relevant and irrelevant actions (Horner & Whiten, 2004; Lyons et al., 2007; 2011; Subiual & Schilder, 2014). Tomasello (2019) noted that human children copy every movement of a model, regardless whether the behavior is silly or inefficient, seeming to take on that model's intentions, something that

chimpanzees (*Pan troglodytes*) fail to display (But see Buttelmann et al., 2007 on enculturated apes). This readiness to copy even irrelevant actions has also been seen in human children relative to bonobos (*Pan paniscus*) (Clay & Tennie, 2017).

But what accounts for this propensity compared to other primates? One brain region which has been explored as a possible answer to what accounts for human cognitive abilities is the pre-frontal cortex (PFC) (Herculano-Houzel et al., 2016; Neubert et al., 2014; Schoenemann, 2005; Smaers et al., 2011; Smaers et al., 2017; Semendeferi et al., 2001; van Essen, 2007; but see Barton & Venditti, 2013), which notably has been linked to the development of executive function (EF) (Diamond, 2002; Durston et al., 2006; Moriguchi and Hiraki, 2009; Moriguchi, 2014), as highlighted by multiple case studies of injuries and lesions to this region resulting in reduced executive performance (Baddeley, 1986; Fuster, 2008; Funahashi & Andreau, 2013; Kolb and Whishaw, 1996; Stuss and Benson, 1986).

EF mediates cognitive control, allowing us to respond effectively and appropriately to our environment and plan for and complete goal-directed behavior, especially in novel tasks (Moriguchi, 2014; Rhoads et al., 2018). Although there is some debate about its developmental trajectory, EF is generally agreed to be non-unitary, and can be divided into three core areas: (1) shifting (mental flexibility/task-switching), (2) inhibition (self-regulation/self-control), and (3) updating working memory (WM) with relevant information (Diamond, 2013; Garcia-Madruga et al., 2016; McAlister & Schmitter-Edgecombe, 2016; Moriguchi, 2014).

Many studies have noted correlations between social interaction, social cognition, and EF, suggesting that EF is developed by and plays an important role in social

interactions and vice versa (Ardila et al., 2005; Carlson, 2009; Legare et al., 2018; Marcisziko et al., 2019; Moriguchi, 2014). Some research also points to the potential link between EF deficits in individuals with autism spectrum disorder (ASD) and difficulties learning in social situations (Bradford, Jentzsch, & Gomez, 2015; Hollocks et al., 2014; Li et al., 2018; Rosello et al., 2017). But few studies have explicitly considered how EF and individual differences in EF predict performance on specific social-learning tasks for typical preschoolers (but see Dickerson, Gerhardstein, Zack, & Barr, 2013; Simpson & Riggs, 2011; Subiaul & Schilder, 2014). An important question to consider then is how individual differences in EF are correlated to a range of novel imitation tasks.

This thesis attempts to characterize how imitation learning in three tasks—Cognitive, Motor-Spatial, and Tower Tasks—correlates to individual EF scores in preschool age children (2 to 5 years of age) as reported by the Behavior Rating Inventory of Executive Function-Preschool (BRIEF-P) (Gioia, Espy, & Isquith, 2003; Sherman & Brooks, 2010). The BRIEF-P is a parent completed report that attempts to standardize EF in preschoolers based on their daily behaviors. The BRIEF-P uses five measures of EF (including the customary WM, Inhibit, and Shift), three corresponding indices, and an overall composite score. All nine will be considered in this analysis (see Table 1). Since this thesis characterizes how the individual subcomponents of EF covary with imitative tasks, it may be informative on which parts of EF are more or less important for human imitative abilities early in development. Finally, this thesis addresses how individual differences in EF correspond to performance across all three imitative tasks—namely, it examines whether imitative performance in one task predicts imitative performance in other tasks and whether this variation is predicted by BRIEF-P scores.

Due to the exploratory approach undertaken here, there are no specific predictions about how EF covaries with imitative ability, besides the following (1) older children will have better EFs scores and generally perform better on imitative tasks relative to younger children, and (2) better EF scores indicated by the BRIEF-P (i.e., lower scores) will correlate with better imitative performance.

Table 1: Descriptions of scales and indices used in the BRIEF-P report.

Scale/Index	Description (Gioia, Espy, & Isquith, 2003)
Inhibit	Ability to resist impulses and action.
Shift	Ability to make transitions, tolerate change, switch attention, and problem-solve.
Emotional Control (EC)	Ability to regulate emotional expression
Working Memory (WM)	Ability to hold info in mind to act or generate goals and plans
Plan/Organize	Ability to manage current and future-oriented task demands
Inhibitory Self-Control Index (ISCI) (Composed of EC and Inhibit)	Ability to modulate actions, responses, emotions, and behavior
Flexibility Index (FI) (Composed of EC and Shift)	Ability to move among actions, responses, emotions, and behavior
Emergent Metacognition Index (EMI) (Composed of WM and Plan)	Ability to sustain ideas and activities in WM and problem-solve
Global Executive Composite (GEC) (Composed of Inhibit, Shift, EC, WM, and Plan)	Summary Score

Rationale. During preschool age, imitation skills are undergoing rapid development (Lyons, Young, & Keil, 2007; Speidel et al., 2021; Subiaul & Schilder, 2014; Subiaul et al., 2016), making this age uniquely situated to reveal differences in imitation in relation to EF. In fact, children around and under the age of 2 prefer to use low fidelity means of copying (i.e., emulation: achieving the same goal as an agent, but through idiosyncratic means), while during preschool years there is a marked propensity for high fidelity

copying (i.e., imitation and overimitation) (Bjorklund, 2016; Dickerson, Gerhardstein, Zack, & Barr, 2013; McGuigan, Makinson, & Whiten, 2011; Nielsen, 2006; Nielsen, Moore, & Mohamedally, 2012; Sherwood, Subiaul, & Zawidzki, 2008; Subiaul et al., 2012). Hence, preschoolers are ideal subjects for examining imitation. They are also better suited to the question of how EF relates to imitation as the imitation tasks may not be as novel or engaging to adults, making the relationship between EF and novel imitation not as clear as it is earlier in development. Adults also may be primed to behave in a certain way in an experimental setting, thinking 'this is what they want me to do', whereas children may be motivated more by the task at hand (Whiten et al., 2016).

Finally, children are ideal models when considering the question from an evolutionary perspective. Chimpanzees and other primates who are used as models for human cognitive evolution are frequently surpassed by human children on imitation tasks (Herrmann et al., 2007; Subiaul, 2016; Tennie, Call, & Tomasello, 2006; but see Russell et al., 2011 on encultured apes). While we frequently use chimpanzees as models for human evolution, we diverged from the chimpanzee-human last common ancestor 5 to 7 million years ago (Hobolth et al., 2007, 2011; Patterson et al., 2006; Sarich & Wilson, 1967; Scally et al., 2012) and have taken our own paths of evolution due to different social and ecological pressures (Reader & Laland, 2002; Renner, Patterson, & Subiaul, 2020). As our closest extant relatives, other primates can help us understand what makes humans distinct; however, they are unlikely to tell us how uniquely human cognition evolved. Although not subscribing to Haeckel's "ontogeny recapitulates phylogeny" (Haeckel, 1866; Olsson, Levit, & Hoßfeld, 2017) observing the cognitive developmental processes that children undergo is a useful proxy for how components of uniquely human

cognition may have evolved. Hence, examining imitation and EF in preschoolers may provide insight into the ancestral state of human cognition. While examining imitation and EF in other primate species is beyond the scope of this thesis, the results of analyzing imitation in children may also be relevant for future comparative cognitive research.

Chapter 2: Task Descriptions

The three tasks examined in this thesis are the Cognitive Task, Motor-Spatial Task, and Tower Task. The Cognitive and Motor-Spatial Tasks are both 2-dimensional, while the Tower Task is 3-dimensional. In the typical procedure for the Cognitive Task, images of three objects appear on a screen simultaneously—these objects are distinctive enough to not cause confusion (Fig. 1A). Participants are asked to select images in a specific order based on 'cognitive' (i.e., ordinal) representation rather than 'spatial' representation. To ensure that participants are not memorizing the spatial arrangement, the location of the pictures are randomized after each trial, correct and incorrect. In contrast, to avoid encoding item identity, in the procedure for the Motor-Spatial Task (Fig. 1B), three identical images appear and the spatial arrangement of the images remains fixed. The content of the images, however, changes from trial to trial (Renner et al., 2018; Subiaul et al., 2012).

The Cognitive Task requires participants to copy abstract ordinal rules—for example, imitating the sequence microphone-scenery-chicken (Fig. 1A) (Terrace et al., 2003; Subiaul et al., 2004; 2012). In contrast, Motor-Spatial Tasks require learning motor-spatial-based rules—for example, down-right-up (Fig. 1B) (Subiaul et al., 2012). These tasks experimentally reveal differing results in children of different ages suggesting that the distinction is an important one—for instance, 3-year-olds successfully imitate novel cognitive sequences and have difficulty imitating novel motor-spatial sequences, while 4-year-olds succeed at both (Subiaul et al., 2012; Subiaul, 2019).

The Tower Task differs from the Cognitive and Motor-Spatial Tasks in that it tests an individual's propensity to build an optimal tower after observing models build components of that tower; however, it is still believed to require imitation skills (Subiaul & Stanton, 2020). The tower is built from Dado cubes of various sizes and colors and Dado squares, both of which have slits in the side allowing them to be stacked. The cubes are used to build the base of the tower, and the squares to build the apex of the tower (Fig. 2). For summative imitation, the condition of the Tower Task examined in this thesis, one model builds the base of the tower by rotating and stacking cubes on top of each other, and another model builds the apex by conjoining the two flat squares (Fig. 2A). The children do not see the base and apex combined before being asked to build their own tower using the cubes and squares; and the models disassemble the tower pieces at the end of each demonstration. The typical procedure is to have children observe three live demonstrations of about 30 seconds each (Fig. 2A) and then ask them to build their own tower (while not explicitly asking them to imitate) (Fig. 2B) to see if they incorporate the elements learned from each demonstrator in their final product (Subiaul and Stanton, 2020).

Measures. For the Cognitive and Motor-Spatial tasks, the task outcomes are reported as 0s (failure) and 1s (success), from first trial accuracy. For the Tower Task, the Target Tower measure was used, where 0 indicated the optimal tower was not built and 1 indicates that it was. This measure is of summative imitation (i.e., putting together the two elements of the tower learned by imitation from different demonstrators), which is considered to be comparable to imitation (Subiaul & Stanton, 2020).

Figure 1. Computerized Task. (A) Cognitive and (B) Motor-Spatial Task.

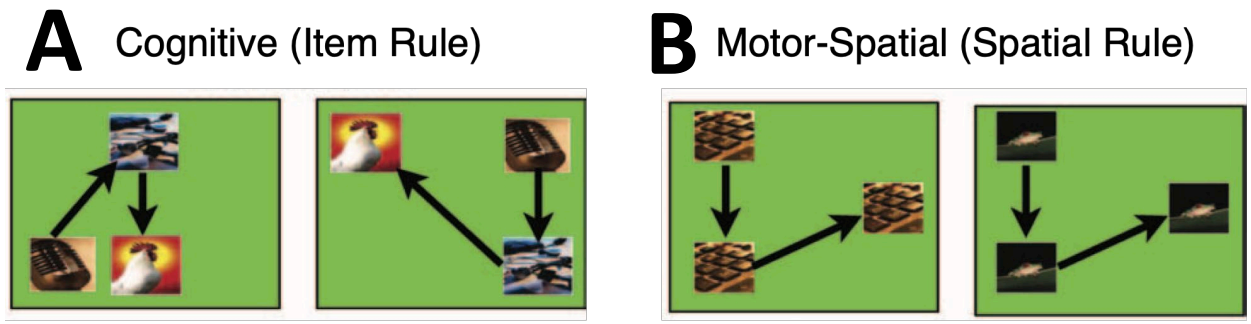
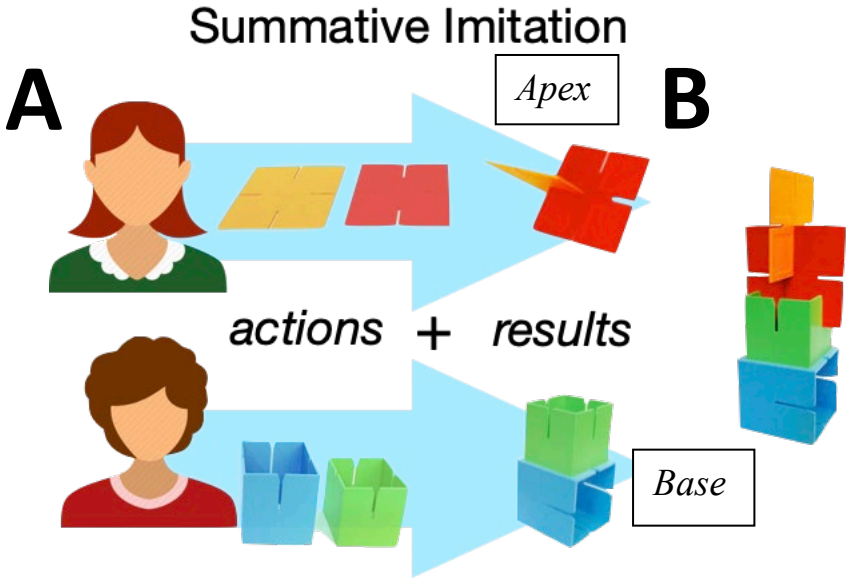


Figure 2. Tower Building Task. (A) Demonstration: Two different models demonstrated how to build different tower elements (Base and Apex); (B) Target Response.



Chapter 3: Data Analyses and Results

Data. The sample size is 596 children ranging in ages from 24 to 72 months with parents from largely high socioeconomic backgrounds. The data were collected by the Subiaul Social Cognition Lab in separate studies. The method of analysis is by binomial logistic regressions, dominance analysis, and bootstrapping through R programming version 4.0.2. Significance level is $p \leq 0.05$. All BRIEF-P reports are RAW scores.

Table 2: Number of individuals by task, having data for BRIEF-P, Age, and Task Outcome.

Data with Age and BRIEF-P	No Task	Cog Task	MS Task	Tower Task	MS.Cog Tasks	Cog.Tower Tasks	MS. Tower Tasks
# of Individuals	596	147	157	167	140	16	16

Preliminary Results. Preliminary data analysis reveals that there is no significant effect of age in years on the composite EF scores (see Table 3). And contrary to prediction, scatterplots of GEC and Age in both years and months reveal no strong linear trends of greater age corresponding with lower executive functioning scores (i.e., better EF as reported by the BRIEF-P); the Pearson's correlation coefficient was negative and close to zero for age and GEC in all three tasks. There was a significant effect of age on both the Cognitive and the Motor-Spatial Tasks (p-values < 0.001) consistent with prior work (c.f., Subiaul & Schilder, 2014; Subiaul et al., 2015; Subiaul & Stanton, 2020); however, there was no significant effect of age on performance of the Tower Task ($p=0.38$). Finally, gender (male=1, female=0) had no significant effect on imitation tasks or BRIEF-P scores, and was removed from the rest of the analysis. The data were examined to meet the criteria of binomial logistic regressions (see Appendix A).

Table 3: Regression for relationship between Age and GEC. ***No significant interaction.***

	AgeYears		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	4.62	4.22 – 5.02	**<0.001**
GEC	-0.00	-0.01 – 0.00	0.129
Observations	597		
R^2 / R^2 adjusted	0.004 / 0.002		

A. Cognitive Task Analyses

Binomial Logistic Regression. A total of 147 individuals from the dataset had values for the BRIEF-P and cognitive imitation and were used in this analysis. In the initial generalized linear model, Age, as was shown in the initial regressions, had a significant effect on the Cognitive Task outcome for the individual scores; and Age, EMI, and FI were significant for the summary scores (see Tables 4-5).

Table 4. Results of the regression for the individual scores in Cognitive Task. ***Only Age is significant.***

	Cognitive.Imit		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.23	0.01 – 5.30	0.354
AgeYears	2.20	1.28 – 3.95	**0.006**
Plan	0.93	0.73 – 1.16	0.515
WM	0.89	0.76 – 1.05	0.184
Inhibit	0.94	0.82 – 1.07	0.354
Shift	1.13	0.95 – 1.36	0.162
Emotion	1.11	0.91 – 1.37	0.293
Observations	102		
R^2 Tjur	0.194		

Table 5. Results of the regression for the summary scores in Cognitive Task. ***Age, EMI, and FI were all significant.***

	Cognitive.Imit		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.24	0.01 – 5.24	0.355
AgeYears	2.20	1.28 – 3.94	**0.005**
EMI	0.90	0.82 – 0.99	**0.028**
ISCI	0.95	0.85 – 1.06	0.339
FI	1.15	1.02 – 1.31	**0.033**
Observations	102		
R^2 Tjur	0.192		

An Analysis of Deviance Table was used to further interpret the results of the initial model. A greater difference between the null deviance (i.e., deviance of the model with only the intercept) and the residual deviance indicates that the addition of a given predictor makes a better model (i.e., adding this variable explains the amount of variation better than when it is not included) (Azen & Traxel, 2009). Correspondingly, a small deviance value means that adding the variable does not contribute much to the model (Azen & Traxel, 2009).

In the Analysis of Deviance Table, adding Age reduced the Null deviance from 141.25 to 132.83, which is a significant amount, its Deviance Residual was the largest at 8.417, and its p-value was significant (0.004). Adding Plan reduced the Null deviance a significant amount as well from 132.83 to 126.01, the Deviance Residual was the second

largest at 6.822, with a p-value of 0.009. These results suggest that Age and Plan both explain the variation in the data better than the null model, or the other variables. (The only other variable that appeared as contributive besides these was Shift which had a larger Deviance Residual (3.500), and a p-value equal to 0.061; however, these results were not significant.) The McFadden index (a pseudo R^2 for logistic regressions) for this model was 0.152. A McFadden index between 0.2 and 0.4 is an excellent fit, so the index suggests for the overall model, the regression is not an excellent fit, but it is not poor (Domencich & McFadden, 1975).

The same steps were repeated for the summary scores, which resulted in EMI having the largest Deviance Residual of all measures (almost as large as Age: 8.417 vs 8.022) and a significant p-value of 0.005; FI had a Deviance Residual of 4.858 and p-value of 0.028. These results parallel the results of the regression and Analysis of Deviance Table for the individual measures, since EMI is composed of WM and Plan and FI is composed of Shift and EC. The McFadden index was 0.151, similar to the individual scores, suggesting that the regression had a moderate fit.

Finally, for GEC and Age, the generalized linear model had a p-value of 0.005 for Age and p-value of 0.023 for GEC, indicating that they both were significant. In the Analysis of Deviance table, Age and GEC again both had significant p-values of 0.004 and 0.017 respectively. The Deviance Residual for Age was 8.417 and for GEC 5.651. These results suggest that the Cognitive Task outcomes were explained by both Age and the overall EF score, with age contributing to the model more. The McFadden index for this analysis, however, was 0.010, indicating a poor fit.

Summary. From the results of the Analysis of Deviance, Age and Plan were the best predictors of task outcome. For the summary results, Age was slightly better than EMI (composed of WM and Plan), but both were better than the null model. FI (composed of Shift and EC) was also contributive to the model.

Dominance Analysis. To explore the relative importance of different predictors in the regression analysis, a dominance analysis was performed (Azen & Traxel, 2009). Dominance analysis defines importance by comparing pairs of predictors (from a selected model) across all subset models (subsets of the predictors) to make the determination of relative importance or dominance (Azen & Traxel, 2009). In ordinary least squares regression, dominance analysis defines the additional contribution of any given predictor to a given subset model as the change in R^2 when the predictor is added to the model; in the case of a logistic regression, the pseudo R^2 McFadden Index is used instead.

Dominance analysis considers one predictor to completely dominate another, if its additional contribution to every possible subset model is greater than that of the other predictor. If complete dominance analysis is not established, the average of the additional contributions can be considered, which is the conditional dominance. Finally, the general dominance is the overall average of a predictor compared to the overall average of another predictor. Even though in a given dominance analysis one predictor may be established as dominant over another, however, this does not mean that better predictors do not exist, just that they are not present in the model (Azen & Traxel, 2009).

This dominance analysis was applied to the Cognitive Task. The results of an initial complete dominance matrix showed that Age dominated every measure except WM, where dominance could not be established between the pair. WM also dominated

Plan, but no other dominances could be established between the rest of the variables. For the summary scores, Age dominated FI and ISCI, and EMI dominated ISCI and FI. These results were concordant with the initial regression and Analysis of Deviance table, where Age and EMI were most contributive to the model. For the conditional dominance analysis, Age dominated every other variable and WM dominated Inhibit, Emotion, and Plan. For the summary scores, Age dominated all variables, and EMI dominated ISCI and FI. In the general dominance for the individuals scores, again Age overcame all other variables, with WM having the next highest effect. For summary scores, Age again was ranked the highest, with EMI being the next highest. In all dominance conditions for the composite score, Age dominated GEC.

Looking at the Average Contribution by Predictor (a way to represent general dominance) revealed the following ranking for the individual components: Age (0.06)> WM (0.031)>Inhibit (0.021)>Plan (0.02)>Shift (0.012)>Emotion (0.008). And, for the summary scores: Age (0.06)>EMI (0.049)>ISCI (0.022)>FI (0.021); Age (0.06)> GEC (0.039). As these results show, when the components of EF are considered individually, Age is contributive by a greater amount to the model than when the components are considered in their summary forms. This suggests that there is some explanatory power that emerges when the individual measures are grouped.

Summary. For the dominance analysis, looking at the individual measures, Age and WM contribute the most to the model; and looking at the summary, Age and EMI. Despite Plan being suggested from the regression and Analysis of Deviance to be a good predictor, no dominance could be established of it over any other predictor in the dominance analysis and it contributed less than WM or Inhibit in general dominance. On

the other hand, there was inconclusive dominance established between Plan and the other variables, meaning that Plan was not consistently dominated. In agreement with the previous analysis, however, EMI, which is composed of WM and Plan, contributed the most to the model. Overall, then Age and EMI were the most predictive of task outcome.

Bootstrapping. To establish the robustness of these results, a bootstrap analysis was performed. The bias reported for general dominance analysis was low for both individual and summary scores, and standard error (SE) was small (between 0.17-0.37), which suggests robustness (where robustness is defined as a bias near or equal to 0, SE near or equal to 0, and reproducibility ranging from 0.7 to 1) (Azen & Traxel, 2009). The most robust results from the bootstrap analysis were Age over Shift and Emotion, which had 0.695 and 0.735 reproducibility respectively. Age being dominant over Plan and WM, however, had poor reproducibility. This result suggests that Age cannot reliably be established as dominant over WM and Plan, like it can for the other individual measures. Since WM and Plan both appeared as significant to the model in different analyses, these results are not surprising.

For the summary scores, EMI over ISCI had a reproducibility of 0.766, which is a fairly robust outcome, and confirms that EMI contributes more to the model than ISCI. While its dominance over FI was not particularly robust this makes sense in light of FI's significant results in the initial regression. Age was established as fairly robustly dominant over FI, however, not over ISCI, which is a somewhat surprising result. Overall, these outcomes suggest that WM and Age, and the combination of WM and Plan (EMI) contributed the most to the model (see Table 6; extended results in Appendix B).

Table 6: Summarized results of bootstrap analysis for Cognitive Task. ***1 indicates dominance, 0 indicates being dominated, and 0.5 indicates no dominance was established.***

I	k	SE	Dominance i>k	REP
Age	Plan	0.267	1	0.455
Age	Inhibit	0.286	1	0.451
Age	Shift	0.245	1	0.695
Age	Emotion	0.225	1	0.735
Plan	WM	0.382	0	0.466
Age	ISCI	0.265	1	0.506
Age	FI	0.313	1	0.622
EMI	ISCI	0.313	1	0.766
EMI	FI	0.265	1	0.508
Age	GEC	0.483	1	0.632

Cognitive Age Interaction. Next, to ensure there were no interactions, a regression was performed with interaction terms. An interaction occurs if the relation between one predictor (e.g., Plan) and the response variable (in this case, Cognitive Imitation) depends on the value of another independent variable (e.g., Age), meaning that their effect on the response cannot be considered independently (Fisher, 1926). Age was chosen as the interaction term since previous research has suggested that executive function develops with age (Subiaul & Schilder, 2014; Zmyj et al., 2010).

The results of a generalized linear model with an age interaction included, showed no significant p-values for the individual scores, so there were no interactions occurring. For the summary scores, however, the generalized linear model revealed FI as significant (p-value of 0.030), and the interaction term of FI and Age having a p-value of 0.051 (see Table 7). Looking at the Analysis of Deviance Table, the interaction term of Age and FI was significant (p-value=0.036) and the Deviance Residual was 4.392. These results indicated that there is likely some interaction between Age and FI. For GEC and Age,

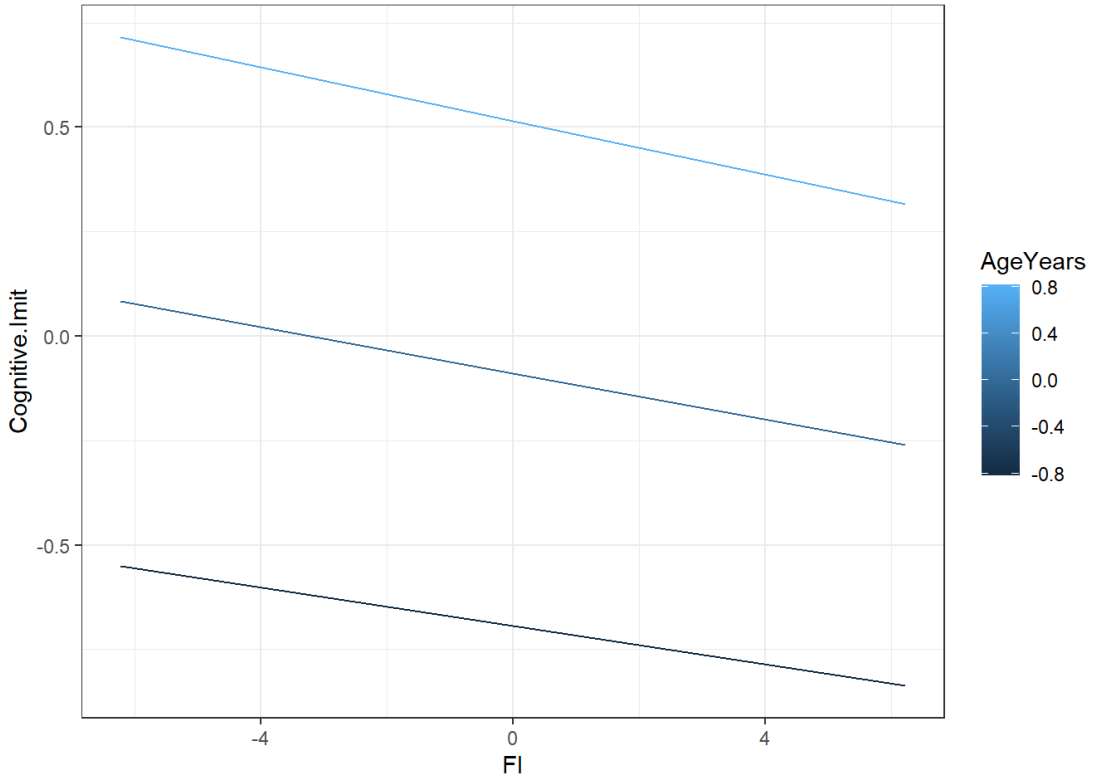
there was no interaction. To visualize this interaction, an interaction plot for a binary outcome variable was created (Heyas & Matthes, 2009; Heyas, 2017), where parallel lines indicate there is no interaction and converging or intersecting lines indicates an interaction (Fig. 3; see Appendices C & F). Although this interaction is slight, FI and Age were removed from the following analysis to see if the results changed in any way.

Table 7: Regression Table for Interaction Terms in the Cognitive Task for summary scores. ***Only FI was significant, and the Interaction term for Age and FI was near significance.***

	Cognitive.Imit		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.03	0.00 – 2258.68	0.502
AgeYears	3.58	0.24 – 51.84	0.341
EMI	0.83	0.45 – 1.53	0.556
ISCI	0.57	0.26 – 1.11	0.127
FI	2.74	1.20 – 7.48	**0.029**
AgeYears * EMI	1.01	0.88 – 1.16	0.843
AgeYears * ISCI	1.13	0.96 – 1.36	0.161
AgeYears * FI	0.82	0.66 – 0.99	0.051
Observations	102		
R^2 Tjur	0.231		

Summary. For the individual scores, there were no significant interactions, however, for the summary scores FI and Age interacted; hence, FI and Age will be removed from the following analysis.

Figure 3: Interaction Plot of Age and FI in the Cognitive Task. ***The lines are nearly parallel, which suggests that there is not much interaction. The shades of blue represent age, and the lines themselves are showing the interaction between FI and Age (see Appendices for more details on plot generation).***



Cognitive Analysis with Interaction Terms Removed. The analysis was repeated removing Age and FI. The results of the regression and Analysis of Deviance, suggest that Plan contributes the most to the model in agreement with the previous section's results. For the summary scores, again in agreement, EMI contributed the most to the model, unsurprisingly, as it is the composite of Plan and WM. In the dominance analysis, WM again rather than Plan appeared as most contributive; and EMI for the summary scores. Overall, the results of the bootstrap (see Table 8) confirmed that EMI is the most robust of the BRIEF-P measures in this analysis.

Table 8: Summarized results of bootstrap analysis for Cognitive Task. ***1 indicates dominance, 0 indicates being dominated, and 0.5 indicates no dominance was established. WM dominating Plan and Inhibit had poor reproducibility; and looking at their general dominance rank, Plan and Inhibit are very close in contribution. EMI being dominant over ISCI was a highly robust result. In general, however, these results were more robust than when Age and FI were included.***

i	K	SE	Dominance i>k	REP
Plan	WM	0.408	0	0.560
Plan	Shift	0.191	0.5	0.844
Plan	Emotion	0.198	0.5	0.822
Plan	Inhibit	0.386	0.5	0.404
WM	Inhibit	0.402	1	0.499
WM	Shift	0.238	0.5	0.698
WM	Emotion	0.243	0.5	0.666
Inhibit	Shift	0.206	0.5	0.810
Inhibit	Emotion	0.229	0.5	0.738
Shift	Emotion	0.239	0.5	0.771
EMI	ISCI	0.314	1	0.889

B. Motor-Spatial Task Analyses

Binomial Logistic Regression. A total of 157 individuals had data for both the Motor-Spatial Imitation task and the BRIEF-P reports. Age was shown in the initial regressions to have a significant effect on the outcomes of Motor-Spatial Imitation tasks, and in the generalized linear model was the only measure of significance (p-value < 0.001) (Tables 9-10).

Table 9: Results of regression for individual scores in Motor-Spatial Task. *Only Age appeared as significant.*

	MS.Imit		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.00	0.00 – 0.09	**0.002**
AgeYears	3.17	1.82 – 5.88	**<0.001**
Plan	1.19	0.95 – 1.50	0.133
WM	1.00	0.84 – 1.18	0.987
Inhibit	0.96	0.84 – 1.09	0.504
Shift	1.04	0.89 – 1.23	0.597
Emotion	0.90	0.73 – 1.09	0.275
Observations	109		
R^2 Tjur	0.192		

Table 10: Results of regression for summary scores in Motor-Spatial Task. *Only Age appeared as significant.*

	MS.Imit		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.00	0.00 – 0.10	**0.002**
AgeYears	3.16	1.83 – 5.82	**<0.001**
EMI	1.08	1.00 – 1.17	0.071
ISCI	0.93	0.83 – 1.04	0.199
FI	1.01	0.91 – 1.11	0.907
Observations	109		
R^2 Tjur	0.177		

The results of the Analysis of Deviance Table established Age as the most significant predictor of outcomes again. No other variables appeared as significant in this analysis. In fact, the contrast in Deviance Residual for Age (17.249) and the Deviance Residual for Plan (2.058), which was the highest of any variables, showed how much more Age was affecting the outcome. The McFadden index was 0.150, suggesting a moderate fit.

Again, for the summary scores only Age was statistically significant in both the generalized linear model and the Analysis of Deviance table. In the initial generalized linear model, EMI had a p-value of 0.071 but in the Analysis of Deviance table its p-value was 0.272 and its Deviance Residual while of the highest, was still very low. The McFadden index was 0.141, which again is a moderate fit. For the composite score in the generalized linear model, in agreement with these results, GEC had no significant effect while Age was significant (p-value < 0.001). For the Analysis of Deviance Table, Age

again had a deviance residual of 17.249 and a significant p-value <0.001. The McFadden index was 0.120, which is a comparable fit to the rest of the analysis.

Summary. The results of the regression and Analysis of Deviance, suggest that Age was most predictive of task outcome.

Dominance Analysis. The results from the complete, conditional, and general dominance analyses indicated that Age overcame every other measure, consistent with the regression results. Plan overcame WM, Inhibit, Emotion, and Shift. Dominance could not be established between the rest of the variables in complete and conditional. For the summary results, Age was dominant over every other variable, while EMI dominated ISCI and FI. Finally, Age was dominant over GEC.

Looking at the Average Contribution by predictor showed the following ranking for the individual components: Age (0.12)>Plan (0.015)>Emotion (0.008)>WM (0.003)>Inhibit (0.002)>Shift (0.001); for the summary scores: Age (0.12)>EMI (0.014)> ISCI (0.006)>FI (0.001). Finally, Age (0.118)> GEC (0.002).

Summary. For the dominance analysis of the individual scores, consistent with the previous section, Age and Plan increased fit the most. For the summary scores, Age followed by EMI increased fit the most. While EMI and Plan did appear highest compared to the other variables, it is clear how much more Age contributed to the model.

Bootstrapping. To examine the robustness of these results, a bootstrap analysis was performed. The SE across all individual results ranged from 0.07-0.3, which is within acceptable range. For the reproducibility results in complete dominance, Plan did not show strong reproducibility of dominance over any other component (see Table 11). Age,

however, dominated all the individual variables with 0.901 reproducibility or higher (see Table 11). Results were similar in conditional and general dominance analysis.

For the summary scores, the SE ranged from 0.05-0.34, with EMI over ISCI having the largest SE at 0.343. Age again was the most robust, dominating all the other variables with 0.936 reproducibility or higher. EMI over ISCI and FI had a reproducibility of 0.501 and 0.514 respectively. Results were similar for conditional and general dominance analyses. Finally, with a SE of 0.045, Age was dominant over GEC with 0.998 reproducibility. The bias reported for all was small, suggesting that there is not much difference between the general dominance outcome and the estimated bootstrap values.

Summary. The bootstraps results suggest that the only robust outcome of this analysis was Age dominating the other measures; the reproducibility for both Plan and EMI dominating other measures were poor.

Table 11. Summarized results of bootstrap analysis for Motor-Spatial Task. ***1 indicates dominance, 0 indicates being dominated, and 0.5 indicates no dominance was established.***

i	k	SE	Dominance i>k	REP
Age	Plan	0.165	1	0.901
Age	WM	0.112	1	0.947
Age	Inhibit	0.084	1	0.974
Age	Shift	0.073	1	0.978
Age	Emotion	0.097	1	0.967
Plan	WM	0.303	1	0.534
Plan	Inhibit	0.285	1	0.430
Plan	Shift	0.280	1	0.401
Plan	Emotion	0.262	1	0.265
Age	EMI	0.134	1	0.936
Age	ISCI	0.084	1	0.974
Age	FI	0.052	1	0.989
EMI	ISCI	0.343	1	0.501
EMI	FI	0.315	1	0.514

Motor-Spatial Age Interaction. Since Age had such a large statistical effect compared to the other variables, an interaction analysis between Age and the other variables was performed. The generalized linear model showed that WM and Inhibit were significant (0.017 and 0.011 resp.) and the interaction terms of Age and WM and Age and Inhibit showed statistically significant interactions, p=0.016 and p=0.013 respectively. The Analysis of Deviance Table confirmed that the interaction of Age and Inhibit had the largest Deviance Residual compared to other variables and interactions, and was statistically significant (p=0.008). For the summary scores, there were no statistically significant results (extended results in Appendix D).

Summary. The results of this analysis suggest that WM, Inhibit, and Age should be removed as they interact (see Fig. 4-5).

Figure 4: Interaction Plot for Age and Inhibit in Motor-Spatial Task. ***The lines are converging, suggesting that interaction is occurring.***

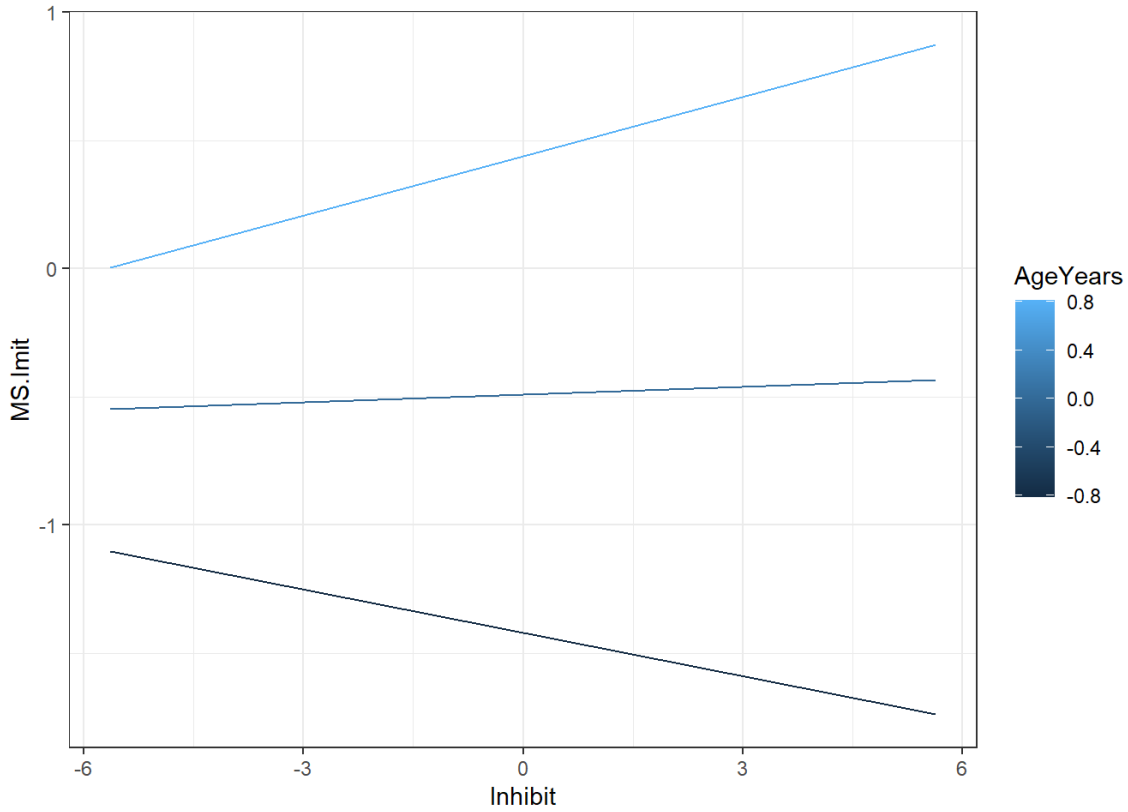
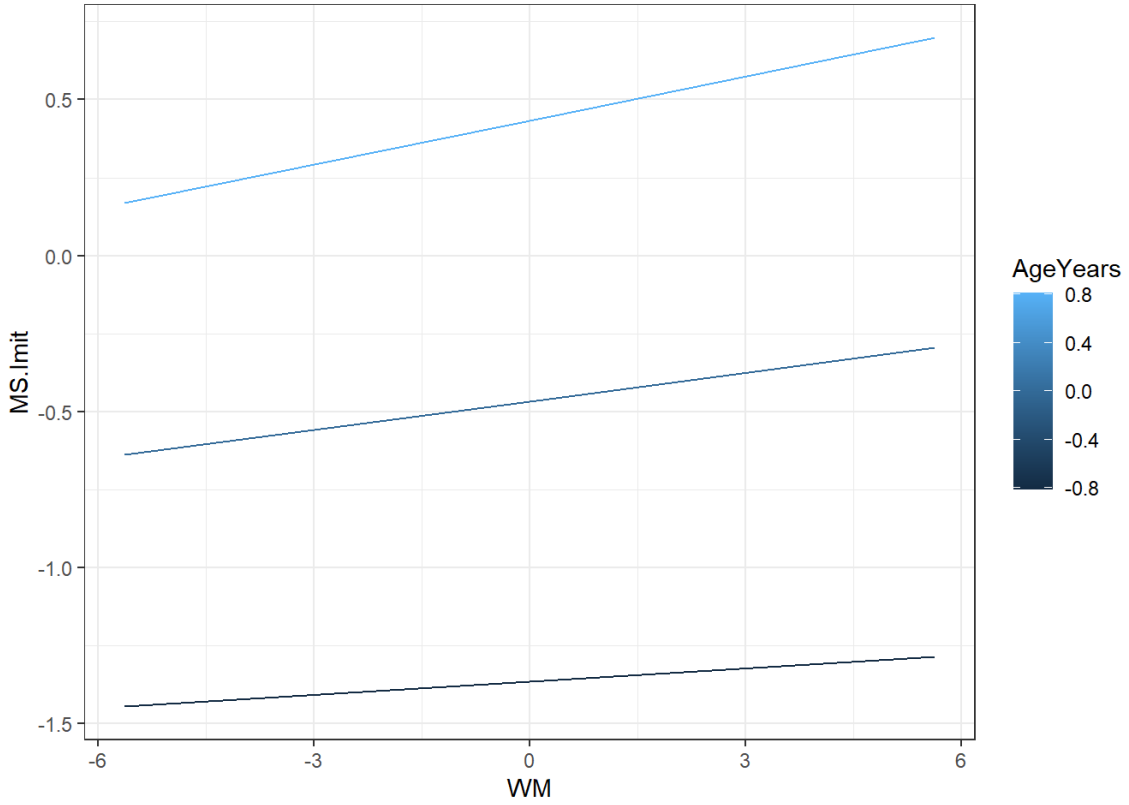


Figure 5: Interaction Plot for Age and WM in Motor-Spatial Task. ***The lines are nearly parallel, however, they each have different slopes, indicating some interaction.***



Motor-Spatial Analysis with Interaction Terms Removed. The variables that had the greatest effects in the initial regression and Analysis of Deviance were Plan and Emotion, and correspondingly EMI and ISCI, but none of these results were significant. The results of the dominance analysis showed that Plan and EMI contributed most to the models in agreement with the analysis where interaction terms were included. The only somewhat robust result was Plan being dominant over Shift (see Table 12).

Table 12: Summarized results of bootstrap analysis for Motor-Spatial Task with Interaction Terms Removed. ***1 indicates dominance, 0 indicates being dominated, and 0.5 indicates no dominance was established. While EMI was dominant over ISCI and FI, these results were not very robust. The results overall, however, were more robust than when interaction terms were included, but nowhere near as robust as Age was in the initial model.***

i	k	SE	Dominance i>k	REP
Plan	Shift	0.386	1	0.649
Plan	Emotion	0.412	1	0.496
Shift	Emotion	0.395	0	0.483
EMI	ISCI	0.390	1	0.554
EMI	FI	0.374	1	0.555
ISCI	FI	0.384	0.5	0.398

C. Tower Task Data Analyses

Binomial Logistic Regression. A total of 167 individuals had data for both the Tower Task and the BRIEF-P scores. Age was not significant in the generalized linear model as reported before, and the Analysis of Deviance table also shows it to have an insignificant p-value and a small Deviance Residual. All the other measures were also nonsignificant in the initial regression (see Tables 13-14).

Table 13: Regression results for individual scores for Tower Task. ***No significant results.***

	MPTower		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.14	0.01 – 3.47	0.233
AgeYears	1.17	0.72 – 1.91	0.528
WM	1.06	0.94 – 1.20	0.339
Plan	1.02	0.83 – 1.25	0.827
Inhibit	1.05	0.96 – 1.16	0.303
Shift	0.92	0.79 – 1.07	0.302
Emotion	0.92	0.78 – 1.06	0.267
Observations	116		
R^2 Tjur	0.050		

Table 14: Regression results for summary scores for Tower Task. ***No significant results.***

	MPTower		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.11	0.00 – 2.64	0.177
AgeYears	1.19	0.74 – 1.94	0.474
EMI	1.06	0.98 – 1.15	0.132
ISCI	1.03	0.95 – 1.12	0.500
FI	0.91	0.82 – 1.01	0.082
Observations	116		
R^2 Tjur	0.046		

In the Analysis of Deviance Table, Shift had the largest Deviance Residual (2.855) but this result was not significant (p = 0.091). The McFadden index moreover was very low at about 0.041, suggesting that the data did not fit this regression well. For the initial regression for the summary scores, no variables were significant. In the Analysis of Deviance table, FI had a p-value of 0.070 and its Deviance Residual was the largest at 3.305 and the second largest was EMI with a Deviance Residual of 1.702 and a p-value of 0.192. Again, the McFadden index was 0.038, suggesting that the regression model is not a great fit for the observed data. Finally, for GEC and Age the generalized linear model and Analysis of Deviance was not significant and the fit was poor.

Summary. None of the results of the regression were significant. For the Analysis of Deviance, Shift had the highest Deviance Residual but no significant p-value.

Correspondingly, FI (composed of Shift and Emotional Control), had the highest Deviance Residual, but again this was not significant.

Dominance Analysis. For the complete dominance analysis, both WM and Inhibit dominated Plan, but the rest of the relationships were indeterminate. For the summary scores, EMI overcame ISCI and Age. In the conditional dominance analysis, both WM and Inhibit overcame Plan and Age. For the summary scores, EMI overcame ISCI and Age. In the general dominance analysis, WM overcame every other variable and Shift overcame Emotion, Inhibit, Plan and Age. Age was dominated by every variable. For the summary scores, EMI overcame the rest of the variables, and again Age was dominated by all others. Finally, in all dominance conditions, for the composite score, GEC dominated Age.

Looking at the Average Contribution by predictor showed the following ranking for the individual components: Shift (0.009)=WM (0.009)>Emotion (0.008)=Inhibit (0.008)>Plan (0.004)>Age (0.003); and for the summary scores, EMI (0.015)>FI (0.014)> ISCI (0.005)> Age (0.003). Finally, GEC (0.004)> Age (0.003).

Summary. In the dominance analysis, consistent with the results of the regression Shift and FI were most contributive to the model. WM and EMI were also, however, equally contributive, which is consistent with the results for the previous tasks.

Bootstrapping. The results of the bootstrapping analysis showed that bias was low and SE varied from about 0.3 to 0.4. In the complete, conditional, and general dominance analyses, WM dominated Plan with a low reproducibility of 0.377. The rest of the measures were all incomplete dominance. For the summary scores, the bias again was

low and the SE ranged from 0.3 to 0.4. In the complete, conditional, and general dominance analyses, EMI dominated both Age and ISCI with a reproducibility of 0.572, which is not robust, but not poor. Finally, Age was dominated by GEC with a SE of 0.5, low bias, and reproducibility of 0.555.

Summary. None of the results for the individual scores were robust. EMI dominating Age, however, had a reproducibility of 0.572 as did EMI dominating ISCI, which while not robust, is moderate.

Table 15: Summarized results of bootstrap analysis for Tower Task. ***1 indicates dominance, 0 indicates being dominated, and 0.5 indicates no dominance was established.***

i	k	SE	Dominance i>k	REP
WM	Plan	0.352	1	0.377
WM	Inhibit	0.369	0.5	0.455
WM	Shift	0.294	0.5	0.654
WM	Emotion	0.294	0.5	0.779
Plan	Inhibit	0.326	0	0.296
Plan	Shift	0.244	0.5	0.736
Plan	Emotion	0.219	0.5	0.798
Age	EMI	0.349	0	0.572
EMI	ISCI	0.383	1	0.572
Age	GEC	0.497	0	0.555

Tower Task Age Interaction. For the generalized linear model with Age as an interaction term, no results were significant. Again, in the Analysis of Deviance, no results were significant, although the interaction term of Age and Plan had a p-value of 0.097 and a Deviance Residual of 2.753, just below Inhibit which had the highest Deviance Residual at 2.855. Since no interactions were significant, the Tower Task re-analysis was not reported (but see Appendix E).

D. Summary

Overall, in line with this thesis' predictions, age was a significant predictor of task performance (except in the Tower Task), dominating the other variables. In the Cognitive Task, Age and WM had the highest predictive power. The summary scores showed again Age as most predictive, and then EMI; Age also overcame GEC. When Age was removed, the ranking for the variables were similar. In the Motor-Spatial Tasks, Plan and EMI contributed the most to the model, once interaction terms were removed, but these results were not robust. Age in the initial Motor-Spatial analysis, however, was not only significant but was dominant in almost 100% of models, and the fit was much better when Age was included, suggesting that Age in the Motor-Spatial task had the best explanatory power; however, both EMI and Plan were the next highest in this analysis.

The main difference between the results of the Cognitive and Motor-Spatial Tasks was Age's contribution. Even though Age was also highlighted as the most contributive to the model in the Cognitive Task, it was much more strongly affecting the model in the Motor-Spatial Task. The influence of WM and EMI in the Cognitive Task was more robust than Plan and EMI in the Motor-Spatial Task, however, it was interesting that for both these tasks similar parts of EF were implicated. The third highest predictor, however, for the Motor-Spatial was Emotion and for the Cognitive it was Inhibit.

In contrast to both the Cognitive and Motor-Spatial Tasks, Age was the least predictive variable in the Tower Task, with Shift and WM being comparably predictive. Although Shift had appeared in the Cognitive Task regression as potentially promising, it did not end up ranking high in the dominance analysis. For the summary scores, EMI and FI were the most contributive, consistent with the individual scores; finally, the Tower

Task was the only task where GEC overcame Age (c.f., Table 16). EMI being the most contributive for the BRIEF-P summary scores was a shared result across all three tasks. Interestingly, both Inhibit and Emotion which had appeared separately in the Motor-Spatial and Cognitive Tasks were tied for second most contributive in the Tower Task dominance analysis. None of these results in the Tower Task, however, were robust.

Table 16: Summarized results of most contributive predictors to the model in dominance analyses.

Task	Highest Individual Scores	Highest Summary Scores	Age vs. GEC
Cognitive Task	Age>WM WM>Inhibit	Age>EMI EMI>ISCI	Age
Motor-Spatial Task	Age>Plan Plan>Emotion	Age>EMI EMI>ISCI	Age
Tower Task	Shift=WM>Inhibit=Emotion	EMI>FI	GEC

E. Between Tasks Analysis

Logistic Regressions Between Tasks. To determine if performance of a given individual on one task predicted performance of that individual on other tasks, binomial logistic regressions were performed. For a total of 140 children, Motor-Spatial Tasks and Cognitive Tasks were not significantly predictive of each other ($p \approx 0.275$) in a generalized linear model, and had a Pearson's correlation of around 0.09, indicating little to no correlation. There was also not enough overlapping data on children who performed Motor-Spatial, Cognitive, and Tower Tasks to be able to analyze whether they were predictive of each other. The p-values for Motor-Spatial and Cognitive being predictive of Tower performance were unsurprisingly close to 1, given the scant overlap.

Better Imitators Predicted by BRIEF-P Scores? High imitators were defined as children who were successful in at least two out of three tasks, and low imitators were defined as those who were only successful in one task. Each individual was assigned the label High or Low Imitator, where High corresponded to a score of 2 or 3 indicating they were successful in at least 2 tasks, and Low when they were either totally unsuccessful or only completed one task successfully. For analysis, individuals were simply categorized as High or Low. The data again were sparse for addressing this question (n=16), and unsurprisingly the p-values were close to 1. However, looking at boxplots for GEC, EMI, WM, and Plan (Fig. 6-9) show some potentially promising results if more data were available: the mean (represented by points) for each BRIEF-P measure is slightly lower for high imitators compared to low imitators, suggesting that a better EF score could be predictive of better imitation across different tasks.

It should be noted, however, that the mean is particularly susceptible to outliers and the median (represented by horizontal lines) shows no clear trends. Overall, more data are needed with performance across all three tasks. It would also be interesting to divide the imitators not only into high and low, but into categories representing which tasks they succeeded in (e.g., group MC as motor-spatial and cognitive success, or MCT as motor-spatial, cognitive, and tower etc.) to see what patterns emerge. Although a neural network was considered for extending the data points, there was not enough consistency in results or input for prediction through neural network.

Figure 6: Boxplot of Mean EMI Score by Imitator Type. ***Mean score (represented by the point) of EMI for Low Imitators is slightly higher than for High Imitators. The horizontal line represents the median, and the vertical lines represent the range (minimum to maximum).***

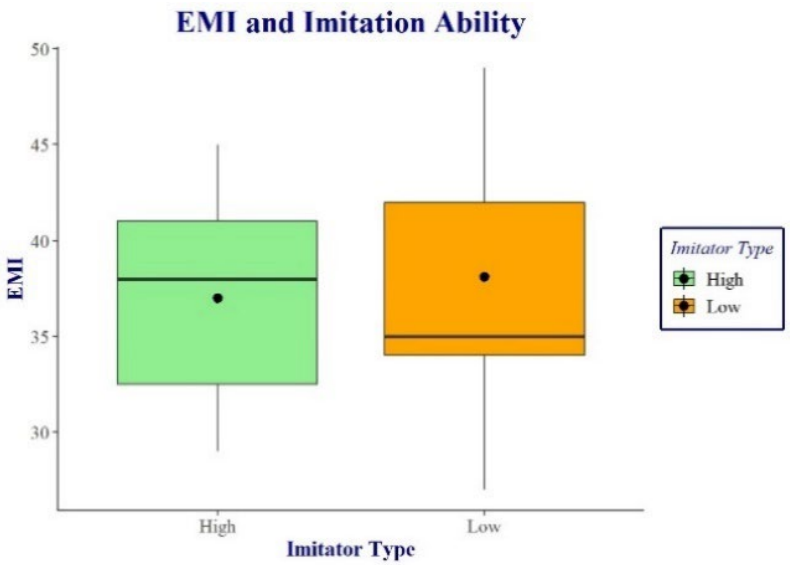


Figure 7: Boxplot of Mean GEC Score by Imitator Type. ***Mean score (represented by the point) of GEC for Low Imitators is slightly higher than for High Imitators. The horizontal line represents the median, and the vertical lines represent the range (minimum to maximum).***

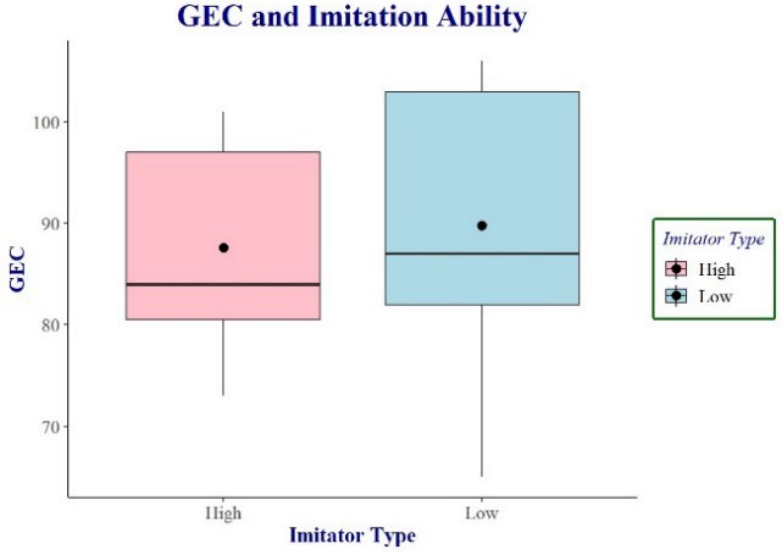


Figure 8: Boxplot of Mean WM Score by Imitator Type. ***Mean score (represented by the point) of WM is slightly higher for Low Imitators than for High Imitators. The horizontal line represents the median, and the vertical lines represent the range (minimum to maximum).***

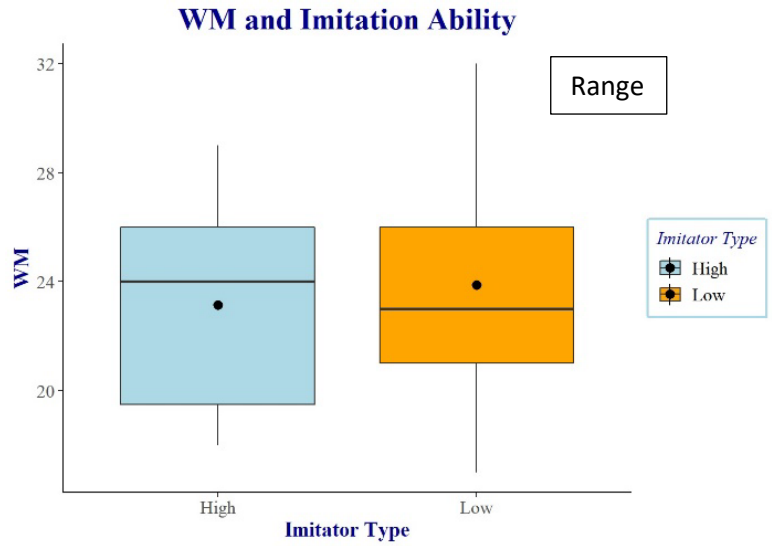
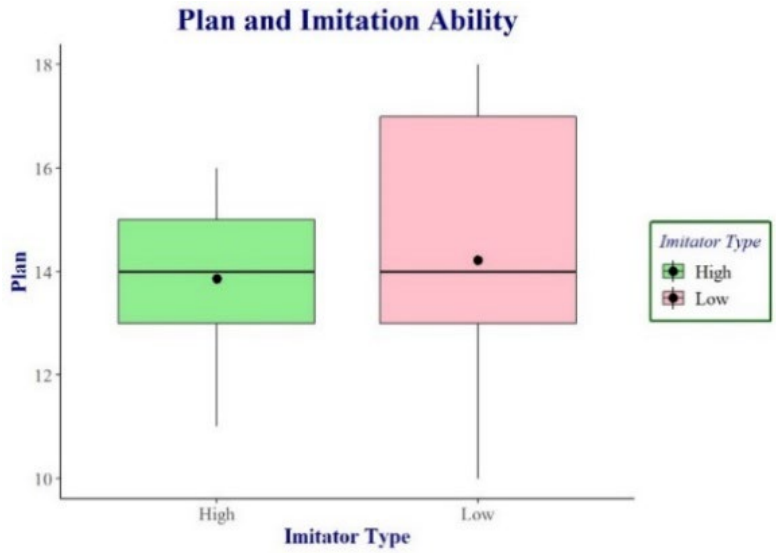


Figure 9: Boxplot of Mean Plan Score by Imitator Type. ***Mean score of Plan is slightly higher for Low Imitators than for High Imitators.***



Chapter 4: Conclusions

Age was the most predictive of task outcome, except in the Tower Task. For the Cognitive Task, WM and Inhibit had the highest predictive power; and for the summary scores, EMI was the most predictive, which is expected as it is composed of WM and Plan. Plan, in the initial regression and Analysis of Deviance, had appeared as significant for the Cognitive Task; and its ranking in the dominance analysis was just behind Inhibit (0.02 for Plan versus 0.021 for Inhibit)—although it was not dominate in the dominance analysis, it was not consistently dominated. For the Motor-Spatial Task, Plan contributed the most to the model and again EMI; however, in contrast to the Cognitive Task, when interaction terms were preserved, the reproducibility of Age was nearly 100%, while the other variables had around 50% reproducibility.

Even though the Cognitive Task agreed with the Motor-Spatial Task that Age was most significant, the reproducibility of age over other variables was not much better than the reproducibility of Plan, Shift, and WM over other variables; in fact, EMI over ISCI had a reproducibility of almost 80%, making EMI dominating other EF measures a more robust result in the cognitive analysis than Age dominating other measures. The general rankings for the components of EF remained consistent when Age was removed, however, for both the Cognitive and Motor-Spatial tasks, the results of the dominance analysis were more robust.

In contrast to both the Cognitive and Motor-Spatial Tasks, Age was the least predictive variable in the Tower Task, with WM and Shift being the most predictive and Emotion and Inhibit tying for second. For the summary score, EMI appeared as the most predictive, as it did for the other tasks. These results, however, were not robust. The

Tower Task, however, did differ from the other tasks in that the regression fit was much poorer.

That the Tower Task was less straightforward than the other two tasks perhaps reflects its increased demands: involving not only the social learning aspect, but the motor manipulation of 3D objects, and consideration of optimal construction. Further, as noted in the introduction, the measure for the Tower Task was of summative imitation. Even so, the results are promising that EF is related to social learning ability, and particularly that planning, updating working memory, and inhibiting are vital to task success. As noted above, interestingly emotional control was also tied to inhibit in the Tower Task, and was the next most predictive after Age and Plan in the Motor-Spatial Task.

From the results of the first two imitation tasks, the conclusion that age alone is predicting task outcome would be more convincing than the conclusion that EF is playing a significant role. However, the results from the Tower Task where the same EF measures appeared on top while Age was ranked the lowest, suggests that EF is particularly (perhaps, uniquely) important for imitation success in that task. This conclusion is further corroborated by reports of deficits in EF affecting social learning in autistic and ADHD children (Bradford, Jentzsch, & Gomez, 2015; Hollocks et al., 2014; Li et al., 2018; Rosello et al., 2017; but see Beck, 2016).

Chapter 5: Future Directions

The measurements of executive functioning were based on parent reports of behavior. While they have ways to control for error, they are still indirect. In addition to BRIEF-P reports, it could be enlightening then to record brain activity through EEG/ERP to more precisely pinpoint which subcomponents of EF are being employed (see review Downes, Bathelt, & Haan, 2017). For instance, some studies have suggested that the dorsolateral prefrontal cortex is connected to WM and planning and the orbitofrontal cortex is related to inhibition and emotional control (Szczepanski & Knight, 2014; Yuan & Raz, 2014).

Moreover, future studies should also consider how EF is correlated to emulation, particularly how it relates to whether an individual overimitates, imitates, or emulates given task complexity as previous research has noted that imitating or emulating may be related to WM demands and cognitive load (i.e., anything that increases task complexity and constrains what is encoded and recalled) (Speidel et al., 2021; Subiaul & Schilder, 2014). According to the model developed by Speidel et al. (2021), conditions of high cognitive load limit encoding and retrieval and increase exploratory goal-directed responses, meaning lower fidelity and higher rates of emulation. On the other hand, when there are conditions of low cognitive load, encoding and retrieval are high, meaning higher fidelity and lower rates of emulation. According to this model as well, older individuals in general are able to handle larger cognitive loads than younger individuals.

Besides, as noted above, emulation is the preferred mode of non-human primates, so noting what influences imitation and emulation and how it relates to EF could be informative for comparative psychology. In addition, it is important to consider EFs

predictive power in asocial learning as compared to social learning conditions across various types of comparable tasks to determine whether its relationship is the same in both or if it is more significant in one.

Although EF is involved in the performance of complex and goal-oriented tasks outside of social situations, there is a real role it is playing in particularly social interactions that should not be ignored, given both the results of this analysis and previous literature exploring EF's role in the development of social cognition. The potential inextricability of EF and social experience is highlighted by Doebel (2020) in her approach to understanding EF. In contrast to the component-process view of EF (i.e., considering how associated components like WM develop), she suggests that the development of EF is better understood as the emergence of skills in using control for specific goals, where goals activate and are shaped by knowledge, beliefs, norms, values, and preferences. A key direction of future research is then, as she points out, to understand how mental content influences the way children engage control, which in turn may help explain individual, group, and developmental differences in EF (Doebel, 2020). Incorporating more participants from diverse cultures and familial backgrounds then is a critical area of future research, to help us better understand what characterizes the relationship between EF and social learning.

References

Aiken, L. S., & West, S. G. (1991). Multiple regression: Testing and interpreting interactions. Thousand Oaks, CA: Sage Publications Inc.

Allison, P. D. (2014). Measures of fit for logistic regression. SAS Global Forum, Washington, DC.

Ardila, A., Rosselli, M., Matute, E., & Guajardo, S. (2005). The Influence of the Parents' Educational Level on the Development of Executive Functions. *Developmental Neuropsychology*, *28*(1), 539-560. doi: 10.1207/s15326942dn2801_5

Azen, R., & Traxel, N. (2009). Using Dominance Analysis to Determine Predictor Importance in Logistic Regression. *Journal of Educational and Behavioral Statistics*, *34*(3), 319–347. https://doi.org/10.3102/1076998609332754

Baddeley, A. (1986). *Working memory.* Clarendon Press/Oxford University Press.

Barton, R., & Venditti, C. (2013). Human frontal lobes are not relatively large. *Proceedings Of The National Academy Of Sciences*, *110*(22), 9001-9006. doi: 10.1073/pnas.1215723110

Beck, S., Williams, C., Cutting, N., Apperly, I., & Chappell, J. (2016). Individual differences in children's innovative problem-solving are not predicted by divergent thinking or executive functions. *Philosophical Transactions Of The Royal Society B: Biological Sciences*, *371*(1690), 20150190. doi: 10.1098/rstb.2015.0190

Bewick, K. C., Raymond, M. J., Malia, K. B., & Bennett, T. L. (1995). Metacognition as the ultimate executive: Techniques and tasks to facilitate executive functions. *NeuroRehabilitation*, *5*(4), 367-375.

Bjorklund, D. F. (2016). Incorporating Development Into Evolutionary Psychology: Evolved Probabilistic Cognitive Mechanisms. Evolutionary Psychology. https://doi.org/10.1177/1474704916670166

Bjorklund, D. F., & Pellegrini, A. D. (2000). Child development and evolutionary psychology. *Child development*, *71*(6), 1687-1708.

Boyd, R., & Richerson, P. J. (1996). Why culture is common, but cultural evolution is rare. In W. G. Runciman, J. M. Smith, & R. I. M. Dunbar (Eds.), *Evolution of social behaviour patterns in primates and man* (pp. 77–93). Oxford University Press.

Boyd, R., Richerson, P., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings Of The National Academy Of Sciences*, *108*(supplement_2), 10918-10925. doi: 10.1073/pnas.1100290108

Bradford, E., Jentzsch, I., & Gomez, J. (2015). From self to social cognition: Theory of Mind mechanisms and their relation to Executive Functioning. *Cognition*, *138*, 21-34. doi: 10.1016/j.cognition.2015.02.001

Brass, M., Ruby, P., & Spengler, S. (2009). Inhibition of imitative behaviour and social cognition. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *364*(1528), 2359–2367. https://doi.org/10.1098/rstb.2009.0066

Brugger, A., Lariviere, L. A., Mumme, D. L., & Bushnell, E. W. (2007). *Doing the Right Thing: Infants' Selection of Actions to Imitate From Observed Event Sequences. Child Development, 78(3), 806–824.* doi:10.1111/j.1467-8624.2007.01034.x

Bryce, D., Whitebread, D., & Szűcs, D. (2015). The relationships among executive functions, metacognitive skills and educational achievement in 5 and 7 year-old children. *Metacognition and Learning, 10*(2), 181–198. https://doi.org/10.1007/s11409-014-9120-4

Buss, A., & Lowery, K. (2020). Inhibitory Control and Executive Function. *Encyclopedia Of Infant And Early Childhood Development*, 183-193. doi: 10.1016/b978-0-12-809324-5.23669-9

Buttelmann, D., Carpenter, M., Call, J., & Tomasello, M. (2007). Enculturated chimpanzees imitate rationally. *Developmental science*, *10*(4), F31-F38.

Byrne, R. (2005). Social Cognition: Imitation, Imitation, Imitation. *Current Biology*, *15*(13), R498-R500. doi: 10.1016/j.cub.2005.06.031

Caldwell, C., Renner, E., & Atkinson, M. (2017). Human Teaching and Cumulative Cultural Evolution. *Review Of Philosophy And Psychology*, *9*(4), 751-770. doi: 10.1007/s13164-017-0346-3

Carlson, S. (2009). Social origins of executive function development. *New Directions For Child And Adolescent Development*, *2009*(123), 87-98. doi: 10.1002/cd.237

Clay, Z., & Tennie, C. (2017). Is Overimitation a Uniquely Human Phenomenon? Insights From Human Children as Compared to Bonobos. *Child Development*, *89*(5), 1535-1544. doi: 10.1111/cdev.12857

de Waal, Frans. (2016). *Are We Smart Enough To Know How Smart Animals Are?* New York, NY: W. W. Norton & Company, Inc.

Dean, L. G., Vale, G. L., Laland, K. N., Flynn, E., & Kendal, R. L. (2014). Human cumulative culture: a comparative perspective. *Biological reviews*, *89*(2), 284-301.

Diamond, A. (2002). "Normal development of prefrontal cortex from birth to young adulthood: cognitive functions, anatomy, and biochemistry," in *Principles of Frontal Lobe Function,* eds D. T. Stuss and R. T. Knight (New York: Oxford University Press), 466–503.

Diamond, A. (2013). Executive Functions. *Annual Review Of Psychology*, *64*(1), 135-168. doi: 10.1146/annurev-psych-113011-143750

Dickerson, K., Gerhardstein, P., Zack, E., & Barr, R. (2012). Age-related changes in learning across early childhood: A new imitation task. *Developmental Psychobiology*, n/a-n/a. doi: 10.1002/dev.21068

Doebel, S. (2020). Rethinking Executive Function and Its Development. *Perspectives On Psychological Science*, *15*(4), 942-956. doi: 10.1177/1745691620904771

Domencich, T. A., & McFadden, D. (1975). *Urban travel demand-a behavioral analysis* (No. Monograph).

Downes, M., Bathelt, J. and De Haan, M. (2017), Event-related potential measures of executive functioning from preschool to adolescence. *Dev Med Child Neurol*, 59: 581-590. https://doi.org/10.1111/dmcn.13395

Durston, S., Davidson, M. C., Tottenham, N., Galvan, A., Spicer, J., Fossella, J. A., et al. (2006). A shift from diffuse to focal cortical activity with development. *Dev. Sci.* 9, 1–8. doi: 10.1111/j.1467-7687.2005.00454.x

Fisher, R.A. (1926). The Arrangement of Field Experiments. Journal of the Ministry of Agriculture of Great Britain, 33, 503–513.

Frederick L. Coolidge and Thomas Wynn (2001). Executive Functions of the Frontal Lobes and the Evolutionary Ascendancy of Homo Sapiens. *Cambridge Archaeological Journal, 11, pp 255-260*. doi:10.1017/S0959774301000142

Friedman, N., & Robbins, T. (2021). The role of prefrontal cortex in cognitive control and executive function. *Neuropsychopharmacology*. doi: 10.1038/s41386-021-01132-0

Funahashi, S., & Andreau, J. (2013). Prefrontal cortex and neural mechanisms of executive function. *Journal Of Physiology-Paris*, *107*(6), 471-482. doi: 10.1016/j.jphysparis.2013.05.001

Fuster, J.M., 2008. The Prefrontal Cortex. Fourth Edition, Elsevier, London, UK

Gabrys, R., Tabri, N., Anisman, H., & Matheson, K. (2018). Cognitive Control and Flexibility in the Context of Stress and Depressive Symptoms: The Cognitive Control and Flexibility Questionnaire. *Frontiers In Psychology*, *9*. doi: 10.3389/fpsyg.2018.02219

García-Madruga, J., Gómez-Veiga, I., & Vila, J. (2016). Executive Functions and the Improvement of Thinking Abilities: The Intervention in Reading Comprehension. *Frontiers In Psychology*, *7*. doi: 10.3389/fpsyg.2016.00058

Gioia, G. A., Espy, K. A., & Isquith, P. K. (2003). *The Behavior Rating Inventory of Executive Function-Preschool version (BRIEF-P)*. Odessa, FL: Psychological Assessment Resources.

Goswami, U. (Ed.). (2002). *Blackwell handbook of childhood cognitive development.* Blackwell Publishing. https://doi.org/10.1002/9780470996652

Haeckel E (1866) Generelle morphologie der organismen, 2 Bde. Georg Reimer, Berlin.

Hardy, M. A. (1993). Regression with dummy variables (Vol. 93). Sage.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.

Hayes, A. F., & Matthes, J. (2009). Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. Behavior Research Methods, 41(3), 924-936.

Herrmann, E., Call, J., Hernàndez-Lloreda, M., Hare, B., & Tomasello, M. (2007). Humans Have Evolved Specialized Skills of Social Cognition: The Cultural Intelligence Hypothesis. *Science*, *317*(5843), 1360-1366. doi: 10.1126/science.1146282

Heyes, C. (2016). *Who Knows? Metacognitive Social Learning Strategies. Trends in Cognitive Sciences, 20(3), 204–213.* doi:10.1016/j.tics.2015.12.007

Hobolth, A., Christensen, O., Mailund, T., & Schierup, M. (2007). Genomic Relationships and Speciation Times of Human, Chimpanzee, and Gorilla Inferred from a Coalescent Hidden Markov Model. *Plos Genetics*, *3*(2), e7. doi: 10.1371/journal.pgen.0030007

Hobolth, A., Dutheil, J., Hawks, J., Schierup, M., & Mailund, T. (2011). Incomplete lineage sorting patterns among human, chimpanzee, and orangutan suggest recent orangutan speciation and widespread selection. *Genome Research*, *21*(3), 349-356. doi: 10.1101/gr.114751.110

Hollocks, M. J., Jones, C. R., Pickles, A., Baird, G., Happé, F., Charman, T., & Simonoff, E. (2014). The association between social cognition and executive functioning and symptoms of anxiety and depression in adolescents with autism spectrum disorders. *Autism Research*, *7*(2), 216-228.

Horner, V., & Whiten, A. (2004). Causal knowledge and imitation/emulation switching in chimpanzees (Pan troglodytes) and children (Homo sapiens). *Animal Cognition*, *8*(3), 164-181. doi: 10.1007/s10071-004-0239-6

Jaccard, J. (2001). Interaction effects in logistic regression (Vol. 135). Thousand Oakes, CA: Sage.

Johnson, P.O., & Neyman J. (1936). Tests of certain linear hypothesis and their application to some educational problems. Statistical Research Memoirs, 1, 57–93.

Kaufman, S. (1995). Difficulties with Animal Models: Chimpanzee Research. *Alternatives to Laboratory Animals*, *23*(5), 592–597. https://doi.org/10.1177/026119299502300508

Kolb, B. and Whishaw, I.Q. (1996) Fundamentals of human neuropsychology. 4th Edition, W.H. Freeman and Company, New York.

Legare, C., Dale, M., Kim, S., & Deák, G. (2018). Cultural variation in cognitive flexibility reveals diversity in the development of executive functions. *Scientific Reports*, *8*(1). doi: 10.1038/s41598-018-34756-2

Lewis, H., & Laland, K. (2012). Transmission fidelity is the key to the build-up of cumulative culture. *Philosophical Transactions Of The Royal Society B: Biological Sciences*, *367*(1599), 2171-2180. doi: 10.1098/rstb.2012.0119

Lewis, H., & Laland, K. (2012). Transmission fidelity is the key to the build-up of cumulative culture. *Philosophical Transactions Of The Royal Society B: Biological Sciences*, *367*(1599), 2171-2180. doi: 10.1098/rstb.2012.0119

Li, B., Atyabi, A., Kim, M., Barney, E., Ahn, A. Y., Luo, Y., ... & Shic, F. (2018, April). Social influences on executive functioning in autism: design of a mobile gaming platform. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-13).

Lyons, D., Damrosch, D., Lin, J., Macris, D., & Keil, F. (2011). The scope and limits of overimitation in the transmission of artefact culture. *Philosophical Transactions Of The Royal Society B: Biological Sciences*, *366*(1567), 1158-1167. doi: 10.1098/rstb.2010.0335

Lyons, D.E., Young, A.G., & Keil, F.C. (2007). The hidden structure of overimitation. Proceedings of the National Academy of Sciences, USA, 104 (50), 19751–19756. doi:10. 1073/pnas.0704452104

Marciszko, C., Forssman, L., Kenward, B., Lindskog, M., Fransson, M., & Gredebäck, G. (2020). The social foundation of executive function. *Developmental Science*, *23*(3). doi: 10.1111/desc.12924

Marsh, L. E., Ropar, D., & Hamilton, A. F. (2014). The social modulation of imitation fidelity in school-age children. *PloS one*, *9*(1), e86127. https://doi.org/10.1371/journal.pone.0086127

McAlister, C., & Schmitter-Edgecombe, M. (2016). Everyday functioning and cognitive correlates in healthy older adults with subjective cognitive concerns. *The Clinical neuropsychologist*, *30*(7), 1087–1103. https://doi.org/10.1080/13854046.2016.1190404

McGuigan, N., Makinson, J., & Whiten, A. (2011). From over-imitation to super-copying: Adults imitate causally irrelevant aspects of tool use with higher fidelity than young children. *British Journal Of Psychology*, *102*(1), 1-18. doi: 10.1348/000712610x493115

Mesoudi, A., & Thornton, A. (2018). What is cumulative cultural evolution?. *Proceedings Of The Royal Society B: Biological Sciences*, *285*(1880), 20180712. doi: 10.1098/rspb.2018.0712

Moriguchi, Y. (2014). The early development of executive function and its relation to social interaction: a brief review. *Frontiers In Psychology*, *5*. doi: 10.3389/fpsyg.2014.00388

Moriguchi, Y., and Hiraki, K. (2009). Neural origin of cognitive shifting in young children. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6017–6021. doi: 10.1073/pnas.0809747106

Moriguchi, Y., Kanda, T., Ishiguro, H., & Itakura, S. (2010). Children perseverate to a human's actions but not to a robot's actions. *Developmental Science*, *13*(1), 62-68. doi: 10.1111/j.1467-7687.2009.00860.x

Neubert, F., Mars, R., Thomas, A., Sallet, J., & Rushworth, M. (2014). Comparison of Human Ventral Frontal Cortex Areas for Cognitive Control and Language with Areas in Monkey Frontal Cortex. *Neuron*, *81*(3), 700-713. doi: 10.1016/j.neuron.2013.11.012

Nielsen, M. (2006). Copying actions and copying outcomes: Social learning through the second year. *Developmental Psychology*, *42*(3), 555-565. doi: 10.1037/0012-1649.42.3.555

Nielsen, M., Moore, C., & Mohamedally, J. (2012). Young children overimitate in third-party contexts. *Journal of Experimental Child Psychology*, *112*(1), 73-83.

Nishimura, Y., Ikeda, Y., & Higuchi, S. (2018). The relationship between inhibition of automatic imitation and personal cognitive styles. *Journal Of Physiological Anthropology*, *37*(1). doi: 10.1186/s40101-018-0184-8

Olsson, L., Levit, G., & Hoßfeld, U. (2017). The "Biogenetic Law" in zoology: from Ernst Haeckel's formulation to current approaches. *Theory In Biosciences*, *136*(1-2), 19-29. doi: 10.1007/s12064-017-0243-4

Patterson, N., Richter, D., Gnerre, S., Lander, E., & Reich, D. (2006). Genetic evidence for complex speciation of humans and chimpanzees. *Nature*, *441*(7097), 1103-1108. doi: 10.1038/nature04789

Persson, T., Sauciuc, G., & Madsen, E. (2017). Spontaneous cross-species imitation in interactions between chimpanzees and zoo visitors. *Primates*, *59*(1), 19-29. doi: 10.1007/s10329-017-0624-9

Reader, S., & Laland, K. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings Of The National Academy Of Sciences*, *99*(7), 4436-4441. doi: 10.1073/pnas.062041299

Renner E, Patterson EM, Subiaul, F (2020). Specialization in the vicarious learning of novel arbitrary sequences in humans but not orangutans. Philosophical

Transactions of the Royal Society, B. 375: 20190442. http://dx.doi.org/10.1098/rstb.2019.0442

Renner, E., White, JP., Hamilton, AFdC., Subiaul, F. (2018) Neural responses when learning spatial and object sequencing tasks via imitation. *PLoS ONE*, 13(8): e0201619. doi: 10.1371/journal.pone.0201619

Rhoads, C. L., Miller, P. H., & Jaeger, G. O. (2018). Put your hands up! Gesturing improves preschoolers' executive function. *Journal of Experimental Child Psychology, 173,* 41–58. https://doi.org/10.1016/j.jecp.2018.03.010

Roebers, C. M. (2017). *Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. Developmental Review, 45, 31–51.* doi:10.1016/j.dr.2017.04.001

Roselló, B., Berenguer, C., Navío, P., Baixauli, I., & Miranda, A. (2017). Executive functioning, social cognition, pragmatics, and social interaction in attention deficit hyperactivity disorder and autism spectrum disorder. *Current Developmental Disorders Reports*, *4*(3), 72-77.

Russell, J., Lyn, H., Schaeffer, J., & Hopkins, W. (2011). The role of socio-communicative rearing environments in the development of social and physical cognition in apes. *Developmental Science*, *14*(6), 1459-1470. doi: 10.1111/j.1467-7687.2011.01090.x

Santiesteban, I., White, S., Cook, J., Gilbert, S., Heyes, C., & Bird, G. (2012). Training social cognition: From imitation to Theory of Mind. *Cognition*, *122*(2), 228-235. doi: 10.1016/j.cognition.2011.11.004

Sarich, V., & Wilson, A. (1967). Immunological Time Scale for Hominid Evolution. *Science*, *158*(3805), 1200-1203. doi: 10.1126/science.158.3805.1200

Schoenemann, P. T., Sheehan, M. J., and Glotzer, L. D. (2005). Prefrontal white matter volume is disproportionately larger in humans than in other primates. *Nat. Neurosci.* 8, 242–252.

Semendeferi, K., Armostrong, E., Schleicher, A., Zilles, K., and Van Hoesen, G. W. (2001). Prefrontal cortex in humans and apes: a comparative study of area 10. *Am. J. Phys. Anthropol*. 114, 224.

Sherman, E.M.S, & Brooks, B.L. (2010) Behavior Rating Inventory of Executive Function – Preschool Version (BRIEF-P): Test Review and Clinical Guidelines for Use. *Child Neuropsychology*, 16:5, 503-519. doi: 10.1080/09297041003679344

Sherwood, C., Subiaul, F., & Zawidzki, T. (2008). A natural history of the human mind: tracing evolutionary changes in brain and cognition. *Journal Of Anatomy*, *212*(4), 426-454. doi: 10.1111/j.1469-7580.2008.00868.x

Shultz, S., & Dunbar, R. (2010). Species differences in executive function correlate with hippocampus volume and neocortex ratio across nonhuman primates. *Journal of comparative psychology (Washington, D.C.: 1983)*, *124*(3), 252–260. https://doi.org/10.1037/a0018894

Simpson, A., & Riggs, K. (2011). Under what conditions do children have difficulty in inhibiting imitation? Evidence for the importance of planning specific responses. *Journal Of Experimental Child Psychology*, *109*(4), 512-524. doi: 10.1016/j.jecp.2011.02.015

Smaers, J., Gómez-Robles, A., Parks, A., & Sherwood, C. (2017). Exceptional Evolutionary Expansion of Prefrontal Cortex in Great Apes and Humans. *Current Biology*, *27*(5), 714-720. doi: 10.1016/j.cub.2017.01.020

Smaers, J., Steele, J., Case, C., Cowper, A., Amunts, K., & Zilles, K. (2011). Primate Prefrontal Cortex Evolution: Human Brains Are the Extreme of a Lateralized Ape Trend. *Brain, Behavior And Evolution*, *77*(2), 67-78. doi: 10.1159/000323671

Speidel, R., Zimmermann, L., Green, L., Brito, N. H., Subiaul, F., & Barr, R. (2021). Optimizing imitation: Examining cognitive factors leading to imitation, overimitation, and goal emulation in preschoolers. *Journal of experimental child psychology*, *203*, 105036.

Stuss DT, Benson DF (1986) *The frontal lobes*. New York: Raven Press.

Subiaul (2012). Cognitive Imitation: Insights into the Development and Evolution of Social Learning. In The Oxford Handbook of Comparative Evolutionary Psychology (Shackelford & Vonk, Eds.). Online Publication Date:Sep 2012DOI:10.1093/oxfordhb/9780199738182.013.0025

Subiaul F, Cantlon JF, Holloway RL, Terrace HS (2004). Cognitive Imitation in Rhesus Macaques. Science, 305(5682), 407-10.

Subiaul, F. (2016). What's Special about Human Imitation? A Comparison with Enculturated Apes. *Behavioral Sciences*, *6*(3), 13. doi: 10.3390/bs6030013

Subiaul F. (2019) Cognitive Imitation. In: Vonk J., Shackelford T. (eds) *Encyclopedia of Animal Cognition and Behavior*. Springer, Cham. Doi: 10.1007/978-3-319-47829-6_1893-1

Subiaul, F., & and Schilder, B. (2014). Working memory constraints on imitation and emulation. *Journal of Experimental Child Psychology*. 128:190–200. https://doi.org/10.1016/j.jecp.2014.07.005

Subiaul, F., & Stanton, M. (2020). Intuitive invention by summative imitation in children and adults. *Cognition*, *202*, 104320. doi: 10.1016/j.cognition.2020.104320

Subiaul, F., Krajkowski, E., Price, E. E., & Etz, A. (2015). Imitation by combination: Preschool age children evidence summative imitation in a novel problem-solving task. *Frontiers in Psychology, 6,* Article 1410. https://doi.org/10.3389/fpsyg.2015.01410

Subiaul, F., Patterson, E., Schilder, B., Renner, E., & Barr, R. (2014). Becoming a high-fidelity -super- imitator: what are the contributions of social and individual learning?. *Developmental Science*, *18*(6), 1025-1035. doi: 10.1111/desc.12276

Stoltzfus, J. C. (2011). Logistic regression: a brief primer. *Academic Emergency Medicine*, *18*(10), 1099-1104.

Szczepanski, S., & Knight, R. (2014). Insights into Human Behavior from Lesions to the Prefrontal Cortex. *Neuron*, *83*(5), 1002-1018. doi: 10.1016/j.neuron.2014.08.011

Tennie, C., Call, J., & Tomasello, M. (2006). Push or Pull: Imitation vs. Emulation in Great Apes and Human Children. *Ethology, 112*(12), 1159–1169. https://doi.org/10.1111/j.1439-0310.2006.01269.x

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Philosophical Transactions Of The Royal Society B: Biological Sciences*, *364*(1528), 2405-2415. doi: 10.1098/rstb.2009.0052

Terrace, H. S., Son, L. K., & Brannon, E. M. (2003). Serial expertise of rhesus macaques. *Psychological Science*, *14*(1), 66-73.

Tomasello, M. (1999). *The cultural origins of human cognition.* Harvard University Press.

Tomasello, M. (2019). *Becoming Human: A Theory of Ontogeny*. Harvard University Press.

van Essen DC. Cerebral cortical folding patterns in primates: why they vary and what they signify. In: Preuss TM, Kaas JH, editors. *Evolution of Nervous Systems. Vol. 4. The Evolution of Primate Nervous Systems.* Oxford: Academic Press; 2007. pp. 267–276.

van Leeuwen, M., van Baaren, R., Martin, D., Dijksterhuis, A., & Bekkering, H. (2009). Executive functioning and imitation: Increasing working memory load facilitates

behavioural imitation. *Neuropsychologia*, *47*(14), 3265-3270. doi: 10.1016/j.neuropsychologia.2009.06.005

Wasielewski, H. (2014). Imitation Is Necessary for Cumulative Cultural Evolution in an Unfamiliar, Opaque Task. *Human Nature*, *25*(1), 161-179. doi: 10.1007/s12110-014-9192-5

Whiten, A., & Ham, R. (1992). On the Nature and Evolution of Imitation in the Animal Kingdom: Reappraisal of a Century of Research. *Advances in The Study of Behavior, 21*, 239-283.

Whiten (2019). Cultural Evolution in Animals. *Annu. Rev. Ecol. Evol. Syst*. 50:27–48

Whiten, A., & van de Waal, E. (2017). Social learning, culture and the 'socio-cultural brain' of human and non-human primates. *Neuroscience and biobehavioral reviews*, *82*, 58–75. doi: 10.1016/j.neubiorev.2016.12.018

Whiten, A., Allan, G., Devlin, S., Kseib, N., Raw, N., & McGuigan, N. (2016). Social Learning in the Real-World: 'Over-Imitation' Occurs in Both Children and Adults Unaware of Participation in an Experiment and Independently of Social Interaction. *PLOS ONE*, *11*(7), e0159920. doi: 10.1371/journal.pone.0159920

Yuan, P., & Raz, N. (2014). Prefrontal cortex and executive functions in healthy adults: A meta-analysis of structural neuroimaging studies. *Neuroscience &Amp; Biobehavioral Reviews*, *42*, 180-192. doi: 10.1016/j.neubiorev.2014.02.005

Zmyj, N., Buttelmann, D., Carpenter, M., & Daum, M. (2010). The reliability of a model influences 14-month-olds' imitation. *Journal Of Experimental Child Psychology*, *106*(4), 208-220. doi: 10.1016/j.jecp.2010.03.002

Appendix A: Criteria for Binomial Logistic Regressions

Data does not have to be normally distributed, however, it needs to satisfy the following criteria: (1) the dependent variable must be binary, (2) the observations to be independent of each other (in this analysis, the individual scores, summary scores, and composite scores were separated for this reason), (3) little to no multicollinearity among the independent variables, (4) linearity of independent variables and log odds (for this analysis, the log odds for the Cognitive and Motor-Spatial tasks were fairly linear; the Tower Task was less so), (5) few to no outliers (observed using graphs of Cook's Distance; outliers were preserved, however, to avoid losing valuable data), (6) a large sample size (Stolzfus, 2011).

Appendix B: Expanded Bootstrap for Cognitive Imitation Task

The SE was highest for WM over Plan and WM over Inhibit (0.382 and 0.372 respectively). For the complete dominance analysis, WM dominated Plan with 0.455 reproducibility. Age was dominant over Plan and Inhibit with a reproducibility of 0.455 and 0.451 respectively, but its dominance over Shift and Emotion was 0.695 and 0.735 reproducibility respectively. WM was the only measure that Age did not dominate, with 0.618 reproducibility. The rest of the results were inconclusive dominance. These results were similar for both the conditional and general dominance. For the summary scores, Age had incomplete dominance over EMI with a reproducibility of 0.494, and dominated ISCI and FI with 0.506 and 0.622 reproducibility respectively. EMI was dominant over ISCI and FI with 0.766 and 0.508 reproducibility respectively. The results were similar for the conditional and general conditions. The SE ranged from 0.2-0.35 and the bias was low. Finally, Age was dominant over GEC with a SE of 0.483, low bias, and reproducibility of 0.632.

Appendix C: Cognitive Interaction Analysis

Binomial Logistic Regression (Interaction Terms Removed). The analysis was repeated removing Age and FI. In the generalized linear model, none of the variables had significant p-values. In the Analysis of Deviance table, Plan had the largest Deviance Residual (6.493) and a significant p-value of 0.011 and Shift had the next largest (3.696) and a p-value of 0.055. The McFadden fit was poorer than when age was included, at 0.093. For the summary scores, in the generalized linear model, EMI did not have a significant p-value (0.074). In the Analysis of Deviance Table, however, the Deviance Residual for EMI was 7.968 and its p-value was significant at 0.005. The McFadden Index again was lower than when interaction terms were included (0.057). GEC was not considered since it is the composite of the other scores and no longer had Age to be compared against.

Summary. These results suggest that Plan contributes the most to the model in agreement with the previous section's results. For the summary scores, again in agreement, EMI contributed the most to the model, unsurprisingly, as it is the composite of Plan and WM.

Dominance Analysis (Interaction Terms Removed). For the complete dominance analysis matrix, WM was dominant over Plan and Inhibit; the rest were inconclusive. In the conditional dominance analysis matrix, WM was dominant over Inhibit, Emotion and Plan, and Inhibit was dominant over Plan. Finally, in the general dominance condition, WM was dominant over every other variable and Inhibit over Plan, Shift, and Emotion. The average contribution by predictor was the following: WM (0.032)> Inhibit (0.021)> Plan (0.019)> Shift (0.012)> Emotion (0.008). For the complete, conditional, and general

dominance analysis matrix for the summary scores, EMI was dominant over ISCI. The average contribution by predictor was: EMI (0.04)> ISCI (0.017).

Summary. WM again rather than Plan appeared as most contributive in the dominance analysis; consistent with previous results, EMI contributed the most to the model for the summary scores.

Bootstrapping (Interaction Terms Removed). The results of the bootstrapping showed a small bias, the SE was slightly higher than when interaction terms were included (ranging from 0.191-0.408). For the complete dominance results, WM dominating other variables had poor reproducibility except over Shift and Emotion, where it had 0.698 and 0.666 reproducibility respectively. These results were consistent for the conditional and general dominance conditions as well. Again, for the summary results, there was a small bias and a standard error of 0.314. The reproducibility for complete dominance of EMI over ISCI was 0.889, which is more robust than the analysis with the interaction term included. These results were consistent for the conditional and general dominance conditions as well. Overall, the reproducibility for these dominance relationships was greater than when interaction terms were included.

Summary. These results show that WM dominating Plan and Inhibit had poor reproducibility; and looking at their general dominance rank, Plan and Inhibit are very close in contribution. EMI being dominant over ISCI was a highly robust result. EMI then is confirmed as the most robust outcome of these analyses.

Appendix D: Motor-Spatial Interaction Analysis

Binomial Logistic Regression (Interaction Terms Removed). Removing Age, WM, and Inhibit as predictors in the binomial logistic regression analysis yielded a poor fit (0.026 and 0.017 respectively) for both individual and summary scores and no significant results. Although the p-value for Plan was 0.095 in the generalized linear model, in the Analysis of Deviance table, Emotion had the greatest Deviance Residual, and decreased the Residual Deviance by the greatest amount.

Summary. The variables that ended up having the greatest effects were Plan and Emotion, and correspondingly EMI and ISCI, but none of these results were significant. Again, GEC was removed because it is a composite of the other measures and could not be analyzed independently without Age as comparison.

Dominance Analysis (Interaction Terms Removed). Plan was dominant over all other variables, while Emotion was dominant over Shift for complete, conditional, and general dominance. In the complete, conditional, and general dominance analysis matrices, EMI was dominant over ISCI and FI. Looking at the average contribution by predictor showed the following ranking: Plan (0.015)> Emotion (0.009)>Shift (0.002); for the summary scores: EMI (0.011)> ISCI (0.004)> FI (0.001).

Summary. The results of the dominance analysis showed that Plan and EMI contributed most to the models in agreement with the analysis where interaction terms were included.

Bootstrapping (Interaction Terms Removed). The results of the bootstrap showed small bias, a SE around 0.25. Across conditional, complete, and general dominance the most robust result was Plan dominating Shift, which had a reproducibility of 0.649. The results

for the summary scores showed small bias and a SE around 0.4. In the complete, conditional, and general dominance matrices, the reproducibility for EMI over ISCI and EMI over FI was 0.554 and 0.555 respectively, which were not particularly robust results.

Summary. The only somewhat robust result was Plan being dominant over Shift. While EMI was dominant over ISCI and FI, these results were not very robust.

Appendix E: Tower Task Analysis without Age

Binomial Logistic Regression (No Age). The initial generalized linear model showed no significant p-values. For the Analysis of Deviance Table, again nothing was significant but Shift had a Deviance Residual of 2.994 and a p-value of 0.087. The McFadden index was 0.038. For the summary scores, the generalized linear model showed nothing significant with EMI having a p-value of 0.148 and FI of 0.080. In the Analysis of Deviance table, FI had a deviance residual of 3.350 and a p-value of 0.067; EMI had a deviance residual of 1.620 and a p-value of 0.203. The McFadden index was 0.035, suggesting a very poor fit.

Summary. The results of this analysis showed, that Shift was the most predictive of outcome for the individual scores. For the summary scores, FI and EMI contributed the most to the model. None of these results, however, were significant.

Dominance Analysis (No Age). For the complete dominance analysis, WM overcame Plan and Inhibit overcame Plan. For the conditional dominance analysis, WM overcame Plan, Inhibit overcame Plan, and Shift overcame Emotion. For the general dominance analysis, Shift overcame WM, Inhibit, Emotion, and Plan; WM overcame Inhibit, Emotion, and Plan; Inhibit overcame Emotion and Plan; and Emotion overcame Plan. For the summary scores, EMI overcame ISCI in the complete and conditional dominance matrices. For the general dominance analysis, EMI overcame FI and ISCI, and FI overcame ISCI. For the average contribution by predictor, WM (0.009)> Shift (0.009)> Inhibit (0.008)> Emotion (0.008)> Plan (0.004). For the summary scores, EMI (0.015)> FI (0.014)> ISCI (0.006).

Summary. The results of the dominance analysis suggest that WM and Shift contributed similarly to the model, and EMI and FI contributed similarly to the model.

Bootstrapping (No Age). For the individual scores, the bias was small and the SE ranged from 0.3-0.4. In the complete dominance analysis, WM dominated Plan with a reproducibility of 0.395; and Inhibit dominated Plan with a reproducibility of 0.322. These results were consistent for conditional and general dominance analyses. For the summary scores, EMI overcame ISCI with a reproducibility of 0.588 in all dominance analyses. The SE ranged from 0.35-0.4 and the bias was small.

Summary. The bootstrap suggested that none of the results were robust; the best result was EMI being dominant over ISCI, which had a reproducibility of 0.588.

Table: Summarized results of bootstrap analysis, where 1 indicates dominance, 0 indicates being dominated, and 0.5 indicates no dominance was established.

i	k	SE	Dominance i>k	REP
WM	Plan	0.361	1	0.395
WM	Inhibit	0.384	0.5	0.411
WM	Shift	0.307	0.5	0.624
WM	Emotion	0.248	0.5	0.754
Plan	Inhibit	0.332	0	0.322
Plan	Shift	0.258	0.5	0.702
Plan	Emotion	0.228	0.5	0.783
EMI	ISCI	0.397	1	0.588
EMI	FI	0.393	0.5	0.368
ISCI	FI	0.345	0.5	0.453

Appendix F: Explanation of Interaction Plot

An interaction occurs if the relationship between one predictor, and the response variable, depends on another independent variable (Fisher, 1926). For two predictors for and one response, if the slopes of an interaction plot are parallel, the effect of one predictor on the response is the same at all levels of the other variable, and there is no interaction (Aiken & West, 1991; Allison, 2014; Fisher, 1926; Hardy, 1993; Hayes & Matthes, 2009; Jaccard, 2001; Johnson & Neyman, 1936). To understand the interactions, the logit or estimated probabilities can be used for logistic regressions. In R, first the results of a binary logistic regression (with the response as binary) were expressed in log-odds metric; and then the conditional effects of the focal predictor at values of the moderators were used to visualize the data (Hayes, 2017). The y-axis of the plot is in logit metric. Just as a typical interaction plot for linear regressions, the convergence of the lines indicates interaction.