NEURAL CORRELATES OF TRAINING AND TRANFSER

BY

AKI NIKOLAIDIS

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Neuroscience
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Arthur Kramer, Chair, Director of Research
Associate Professor Aron Barbey
Assistant Professor Paris Smaragdis
Associate Professor Brad Sutton

ABSTRACT:

Cognitive training holds promise to improve cognitive ability in many people, young, old, both healthy, and those with psychiatric or neurological illness, but this field largely lacks a mechanistic understanding of the process by which training demonstrates transfer to improve underlying cognitive abilities. In Chapter 1, we examine how mapping the neural correlates of training and transfer is critical for developing a mechanistic explanation of how training drives transfer. In the current study, we trained 45 young adults with Mind Frontiers, an adaptive cognitive training game that targets executive function, attention, and reasoning. We investigate how both brain structure and resting state networks are associated with training gain and transfer. In Chapter 2, we investigate how both pre-existing and training-induced differences in brain structure are predictive of training and transfer. In Chapter 3, we assess how both pre-existing, and training-induced differences in resting state network connectivity in the default mode, cingulo-opercular, frontal-parietal, and subcortical networks predict training gain and transfer. In Chapter 4, we examine the relationship of the structural and resting state data in predicting training and transfer. We assess the extent to which these predictors overlap and dissociate with one another over predictions of training gain and transfer. To make our predictions, we utilize a simple machine learning paradigm that we developed to maximize the reliability and interpretability of our findings. We found extensive overlap in structural predictions of training gain and transfer in low level visual and auditory areas, suggesting that greater fidelity in low level sensory systems may contribute to greater signal to noise ratios during training, enabling better training quality and transfer. Furthermore, our resting state results also highlight the importance of training quality through demonstrating the importance of the cingulo-opercular network, which is critical for both the regulation of the default mode network and deployment of sustained attention during training. These results suggest that greater training fidelity through lessened distraction may play an important role in maximizing the benefits of an intervention.

Table of Contents

Chapter 1

Investigations into Cognitive Training

1.1 Introduction

Cognitive neuroscience is entering into a new era of discovery science, and this transition is exemplified by the recent promise to apply this advancing understanding of the brain and mind towards the development of training paradigms that aim to improve general cognitive performance. Cognitive enhancement has captured the attention and imagination of researchers, companies, lawmakers, and the public at large, and with good reason. These tools could have an enormous impact not only on clinical populations, but also on healthy individuals in a wide range of scenarios, prompting calls for funding from not only the NIH and NSF, but also many branches of military research, such as the Office of Naval Research (ONR), Intelligence Advanced Research Projects Association (IARPA), and the Department of Defense (DOD). In the past decade, many studies have successfully demonstrated that cognitive training can lead to improvement of cognitive performance in untrained tasks. The largest cognitive training study, ACTIVE, with over 2800 older adults, demonstrated lasting benefits from training with either reasoning, processing speed, or memory focused training paradigms (Ball et al. 2002), and a recent 10-year follow-up study found these individuals still had significantly less age-related decline in reasoning, processing speed, and memory compared to age matched controls (Rebok et al. 2014). In a more recent study, with a time-intensive cognitive training intervention, with better measurements of cognitive improvement, 100 young adults and 100 older adults trained with several perceptual speed, working memory, and episodic memory training tasks (Schmiedek et al. 2010). Both the young adults and older adults demonstrated significant improvements to several cognitive constructs, though the younger adults benefited from the training in some domains which the older adults showed no improvement. Although these large high quality studies have demonstrated the success of cognitive training, the field has been rife with failures to replicate as well. For example, one of the most controversial findings is that training with the dual n-back, a common working memory task, leads to improvement in fluid intelligence (Jaeggi et al. 2008); we and others have failed to replicate this effect in studies not facing the same methodological constraints (Redick et al. 2012; Shipstead et al. 2012; Oelhafen et al. 2013).

Irrespective of the fact that the cognitive training field is relatively new and so far has been unable to create effects that replicate reliably, a booming cognitive training industry has emerged seeking to take advantage of the limited science that has been done on the topic. This has led to a proliferation of cognitive training products with both varying degrees of claims and scientific backing. Recently, a large contingent of the cognitive neuroscience and cognitive training community released a statement that the industry is grossly overstating the claims that their products have beneficial effects for their customers (http://longevity3.stanford.edu/blog/2014/10/15/the-consensus-on-the-brain-training-industry-from-the-scientific-community-2/). In 2015, one of the largest cognitive training companies, Lumosity, was ordered by the FTC to cease and desist these inaccurate claims on the effect of their products, and they have been made to pay a 2 million USD fine as well, with 48 million USD in fines to be filed in the future if they fail to comply (https://www.ftc.gov/news-events/press-releases/2016/01/lumosity-pay-2-million-settle-ftc-deceptive-advertising-charges). In an era when the scientific community has not yet been able to develop reliably effective training, and the public is becoming increasingly invested in unscientific and unvetted training programs, it has become increasingly urgent to resolve the reliability issues facing cognitive training.

The most important reason that cognitive training studies have faced poor reliability is methodological constraints, such as underpowered samples, lack of adequate control groups, measurements of transfer using single tasks as opposed to cognitive constructs, and lack of adequate theoretical guidance. One way to contribute to the development of a robust theory is to investigate the neural correlates of training and transfer, which can be used to identify how brain characteristics are manifested in individual differences in training and transfer outcomes. Recent studies have successfully identified structural, metabolic, and functional correlates of training and transfer. For example, we found that both the volume and level of iron deposition in the striatum before training was a significant predictor of training-induced improvement in a complex skill task (Erickson et al. 2010; Vo et al. 2011). Others have found that pre-post changes in activity levels in the striatum also predicted transfer to working memory after updating training (Dahlin et al. 2008), and that such training has significant impact on levels of dopamine released in the striatum as well (Bäckman et al. 2011). We found that training-induced changes in activity in the postcentral gyrus predicted the individual differences in improvement in an untrained working memory task (Nikolaidis et al. 2014), and most recently we found that training-induced changes in functional connectivity of the subcortical and motor networks predicted individual differences in learning of a complex training task (Nikolaidis et al. 2015). Nevertheless, a complete picture of how cognitive training and transfer are related to underlying brain structure and function has yet to be established.

In the skill acquisition literature, detailed anatomical, functional, and network characteristics of training-induced learning are available in both humans and animals (Hikosaka et al. 2002; Graybiel 2005; Doyon et al. 2009; Patel et al. 2013). For example, in motor learning it has been well established that the acquisition of skills depends on the interaction between the basal ganglia and cortex (Hikosaka et al. 2002). At the beginning of training, these two networks interact very closely, and the basal ganglia, especially the striatum, play a critical role in formulating the appropriate representations of the skill. As learning progresses and the skill becomes automated, the basal ganglia becomes less important for the performance of the skill. This effect has been strongly demonstrated by studies showing that after removing the striatum, automated responses can still be performed, but new ones cannot be acquired (Eichenbaum and Cohen 2008). To develop a fuller theory of cognitive training and transfer, it will be important to place cognitive training into the context of skill acquisition. In this way, the critical question for cognitive training changes from assessing the relationship of training to transfer, to exploring the relationship between skill acquisition and skill generalization. Previous work on skill acquisition and generalization has demonstrated that modifications of the training that impact the acquisition of skills can have significant effects on their generalization to new contexts, stimuli, and skills as well (Singley and Anderson 1989; Schmidt et al. 1992; Bjork and Linn 2006; Anguera et al. 2007; Boot et al. 2010; Seidler 2010; Soderstrom and Bjork 2015).

In the current set of studies, we investigate the neural correlates of training and transfer by recruiting 45 individuals to perform 30 hours of training with Mind Frontiers, pre and post intervention neuropsychological testing, as well as structural and resting state functional MRI. Mind Frontiers (MF) is a cognitive training paradigm that was developed through a collaboration between our lab at the University of Illinois at Urbana Champaign and Aptima Inc. This paradigm utilizes six training tasks that have previous success in cognitive training studies, and specifically tap into complementary aspects of working memory, executive function, and reasoning (Baniqued et al. 2015). These six tasks have been gamified and unified into a composite “Wild West” themed adaptive cognitive training game. The behavioral results from this study have been published previously (Baniqued et al. 2015), and here Baniqued et al. demonstrated that training with MF significantly improves performance on composite scores of working memory, fluid intelligence, and perceptual speed. They also found that training gain in three of the training tasks predicted transfer to working memory, but not transfer to perceptual speed and fluid intelligence. This is important because it shows both dose dependent transfer of training to working memory, as well as transfer to perceptual speed and fluid intelligence that was independent of training gain. This suggests that multiple, and possibly dissociable learning mechanisms are driving training-induced transfer to these three cognitive domains.

In the current study, we investigate contributions of pre-existing brain structural volumes, thicknesses, and resting state connectivity to predictions of both training gain and transfer. Furthermore, we assess how training-induced changes in these neural characteristics predict training and transfer as well. By examining how pre-existing and training-induced differences in brain structure and connectivity contribute to prediction, we may gain a more detailed picture of the mechanisms by which training and transfer occur, as well as the relationship between characteristics that drive the skill learning and generalization that occurred during training. In the first study, we investigate how structural volumes and thicknesses predict transfer to working memory, perceptual speed, and fluid intelligence, as well as training gain in the three training tasks that were positively correlated with working memory transfer. Our second study assesses how resting state network connectivity in the frontal parietal network, cingulo-opercular network, default mode network, and subcortical network predict training gain and transfer. In the final study, we assess the overlap and dissociations in the structural and resting state network predictions of training and transfer. Each of these studies will help characterize the neural correlates of training and transfer in the context of skill acquisition and generalization. While pre-existing differences in structure and network connectivity have been demonstrated to predict training gain and transfer, how these characteristics relate to training-induced changes in structure and connectivity has not been well established. In the current study we plan to compare the neural correlates that most reliably predict training and transfer in both of these cases. Furthermore, by performing predictions with both structural and resting state characteristics, the current study will offer a unique insight into the processes by which pre-existing and training-induced differences drive learning and transfer. Most previous studies have focused on one category of neural correlate or the other, yet multimodal assessments of structure and function, such as those performed in the current study, are essential to establishing a firmer understanding of cognitive training.

We employ common techniques from machine learning to make these predictions. While Ordinary Least Squares (OLS) is the most common optimization method used in general linear models (GLMs), OLS suffers from likelihood to overfit solutions to a given sample, which reduces reproducibility of its findings (Hastie et al. 2009). Furthermore, OLS is more likely to produce unreliable models in the face of collinear input variables compared to machine learning methods, especially in the case of small samples, which are common in cognitive neuroscience research. Recently, problems with poor reproducibility in cognitive training and psychology have become an issue for which many researchers are seeking solutions (Alexander et al. 2012; Hulme and Melby-Lervåg 2012; Melby-Lervåg and Hulme 2013). It is possible that the common application of OLS based GLM methods to highly collinear datasets with small sample sizes may be a significant contributor to this reproducibility issue. To address these specific issues of variable collinearity and model reproducibility, techniques such as ridge regression have been developed (Marquardt and Snee 1975). While OLS selects models solely based on model coefficients that minimize squared error, ridge regression also adds a penalty for using large model coefficients. By minimizing the model's dependence on large coefficients, ridge regression makes models more conservative in estimating the importance of individual variables. Many other machine learning techniques offer similar advantages of resilience to collinear variables and reduced overfitting, such as lasso, support vector machines (SVMs), and random forests (Kotsiantis et al. 2007; Hastie et al. 2009). While these advantages have led to a surge in the use of machine learning methods in scientific fields, and recently in neuroscience (Hinton 2011; Varoquaux and Thirion 2014), the application of these methods has inherited computer science traditions that tend to value predictive accuracy over other traits that are critically important to scientists, such as interpretability and replicability of findings. It is both important and impressive that machine learning methods are able to predict and detect, with a high degree of accuracy, subtle disease states and characteristics in neuroimaging data that would be impossible for even an expert radiologist (Shen et al. 2010; Dosenbach et al. 2011; Shirer et al. 2011; Liu et al. 2013). However, by focusing on this feature we miss the fact that the real significance of machine learning to the cognitive neuroscience community is not its predictive power, but its ability to extract statistically reliable structure from the overwhelmingly complex data we acquire in neuroimaging. In the current set of studies, we introduce a simple, though computationally expensive, methodological change to common machine learning practice that focuses on maximizing our finding's interpretability and reliability.

Typically, machine learning approaches use a technique called grid search to cross validate over the feature and hyperparameter (lambda) sets and choose the set that gives the highest performance. Cross validation (CV) is a method by which the data is split into training sets which create models and test sets to which the models are applied. Grid search is performed through nested CV, meaning that within each CV loop, the training data is split again into inner CV training and test data. In this inner CV loop, hyperparameters and feature sets are tested and the best performing sets are then applied to the outer CV loop. This means that depending on what the innermost CV loop chooses as the optimal point in model space, the outer CV loop will have a different set of features and lambda values. In a similar way that OLS is more likely to induce overfitting because it prioritizes reducing the squared error term above any other considerations of model composition or sparsity, we can think about grid search optimization of model parameters as being more likely to overfit because it maximizes out of sample accuracy of a single model above any other considerations of model reliability or reproducibility. Previous work has shown that when there is noise present in the outcome variables, as is inevitably the case for training and transfer scores, cross validation can lead to overfitting (Ng 1997). As a conceptual example, in a model space with multiple local minima, the deepest minima could by chance represent a narrow range of models that are predictive of the outcome variable while other shallower minima contain large groups of models more reliably predictive of the outcome variable. Grid search optimization is formulated such that it chooses the deeper minima (Mosteller and Tukey 1968), and this can occur at the expense of reproducibility of the data's predictive power (Ng 1997). Given that data are then often compared based on the relative predictive accuracy of this deepest minima, one may conclude that one dataset is more or less informative based on incomplete and inaccurate information offered by grid search optimization. Obtaining the highest predictive accuracy is an important and worthwhile endeavor in many scientific projects that contain unwieldy or extremely large datasets, but in the current work, where feature sets have been created, cleaned, and preprocessed through well-established methods and guided by prior knowledge, we are more concerned with establishing reliable relationships than we are with achieving the highest accuracy possible. Furthermore, a field currently undergoing a reproducibility crisis needs more emphasis placed on testing the reliability of its own findings rather than reporting large effect sizes.

Another important issue with grid search optimized parameters is that it reduces the interpretability of findings and unnecessarily complicates the detailed investigation of model performance. In grid search there are two sets of CV loops. In the inner loop, models are trained on a range of parameters and the combination of parameters that, on average, performs best on the inner loop test sets is chosen as the model to apply to the outer CV loop. There are no constraints placed on characteristics of different models that are 'optimal' for a given outer CV loop, and thus these models can vary not only in terms of hyperparameter value but also which features are chosen to build the model. Therefore, after looping through all outer CV loops, each test set has been predicted with a variety of combinations of feature sets and hyperparameter values, but a single accuracy and error rate is reported. Given that different hyperparameter values imply dramatically different treatment to the data, it may be inaccurate to combine sets of models with a range of hyperparameter values. While this approach is likely to achieve higher prediction accuracy due to its focus on performance, we suggest that grid search optimization leads to unnecessary obfuscation of the process by which the data is able to predict its outcome variable. If a dataset contains a reliable signal with regards to a specific question, the signal should not be able to be masked or dramatically changed by small modifications in the model parameters and feature sets. To avoid these issues of reproducibility and interpretability, we explore a large coverage of model space and report of significant regions of model space where models are more predictive of the outcome variable than a distribution of null models. We then apply a cluster correction algorithm to the model space map to extract regions where large groups of models can predict the question with high reliability. Exploring model space in this way, enables us not only to assess the relationship between model complexity and predictive accuracy, but also how these characteristics interact with regularization strength. Furthermore, within each significant cluster of models we identify features that have been chosen in >90% of the models in the clusters, which we call consensus features. An important advantage of this approach is that we are able to find features that are included in models that are all reliable contributors to our question, and therefore we are more confident in interpreting the relevance of these features.

The most important step in the application of machine learning is called feature engineering. This involves using both statistical insight and domain expertise to collect, clean, and curate datasets that are relevant to the question of interest (O'Neil and Schutt 2013). In the case of the current projects, we leverage vetted methods of collecting information from structural and resting state MRI to extract variables that are of highest quality and relevance to training and transfer. In our structural data, we use the longitudinal pipeline from Freesurfer, a method of accurately and robustly extracting brain volume and thickness information that has been specifically designed to account for variance at multiple time points (Destrieux et al. 2010; Fischl 2012). Extensive previous research has demonstrated such regional brain volumes and thicknesses are tied to skill acquisition and cognitive training (Kennedy and Raz 2005; Erickson et al. 2010; Basak et al. 2011; Lövdén et al. 2013; Metzler-Baddeley et al. 2016). In our functional data, we apply a well-established functional connectivity preprocessing stream that has been developed for the analysis of large scale resting state data (http://fcp-indi.github.io/). We then extract several highly reliable resting state networks that were previously identified in a sample of 512 individuals, and were then replicated in a second sample of 512 (Power et al. 2011). Furthermore, we choose networks specifically known to be involved in working memory, reasoning, and learning, such as the frontal parietal network (FPN), cingulo-opercular network (CON), default mode network

(DMN), and subcortical network (SUB) (Graybiel 2005; Dosenbach et al. 2006, 2008; Uddin et al. 2009; Sambataro et al. 2010; Elton and Gao 2014; Geerligs et al. 2015). Feature engineering not only allows for greater confidence in the results of a given prediction, but critically, it is the most important step to increasing the interpretability of machine learning results. By creating feature sets that represent dissociable or competing hypotheses and implications, the significant prediction of an outcome with one dataset but not another yields important information regarding the relevance of each competing hypothesis for the question at hand. For example, in our structural project, Chapter 2, we predict training and transfer using both structural volumes and thickness variables. Both of these variable types have been used extensively in the learning and cognitive training literatures (Lustig et al. 2009; Erickson et al. 2010; Basak et al. 2011; Herholz and Zatorre 2012; Metzler-Baddeley et al. 2016), but how thickness and volume offer complementary versus redundant information on these questions is poorly understood. By using these data types both independently and in combination, we may gain insight into this methodological question as well. In our resting state analyses, Chapter 3, we utilize two complementary networks known to be involved in dissociable aspects of cognitive control, the FPN and CON (Dosenbach et al. 2008; Zanto and Gazzaley 2013; Sestieri et al. 2014). Describing which of these networks is more informative and which components of these networks play important roles in the predictive success of our models, will offer important insight into how both dissociable aspects of cognitive control and the unique role of these networks drive learning and transfer.

Aside from the collection, cleaning, and curating of features, the selection of a machine learning method is one of the most important decisions when applying a machine learning pipeline. In the current work, we apply one of the simplest linear models available, ridge regression. This method is not as high performing as other more commonly applied methods, such as support vector machines (SVMs), but what ridge regression loses in predictive power it gains in interpretability, as models from other more powerful methods, such as random forests, neural networks, or SVMs are considerably more difficult to interpret (Kotsiantis et al. 2007). Furthermore, these other approaches have access to more sophisticated methods of signal extraction, and while the neuroimaging data we use in the current set of studies would benefit from more robust signal extraction methods that are sensitive to higher order relationships in the data's covariance structure, these methods may be less likely to replicate findings from previous analyses which have overwhelmingly relied on GLMs. By the same token, the findings from a neural network or other advanced analysis of these data are less likely to be replicated by follow-up GLM analyses, limiting the impact of the current work. Ridge regression on the other hand, while one of the simpler and weaker of tools in the machine learning toolbox, may be the tool most apt for application in this context, given that its simplicity makes it not only more approachable, but also more likely to replicate previous work, and to be replicated by follow up analyses as well.

1.2 References

Alexander A, Barnett-Cowan M, Bartmess E, Bosco F, Brandt M, Carp J, Chandler J, Clay R, Cleary H, Cohn M, others. 2012. An Open, Large-Scale, Collaborative Effort to Estimate the Reproducibility of Psychological Science. Perspect Psychol Sci. 7:657–660.

Anguera J a, Russell C a, Noll DC, Seidler RD. 2007. Neural correlates associated with intermanual transfer of sensorimotor adaptation. Brain Res. 1185:136–151.

Bäckman L, Nyberg L, Soveri A, Johansson J, Andersson M, Dahlin E, Neely AS, Virta J, Laine M, Rinne JO. 2011. Effects of Working-Memory Training on Striatal Dopamine Release. Science (80-). 9–9.

Ball K, Berch DB, Helmers KF, Jobe JB, Leveck MD, Marsiske M, Morris JN, Rebok GW, Smith DM, Tennstedt SL, Unverzagt FW, Willis SL. 2002. Effects of Cognitive Training Interventions With Older Adults. Jorunal Am Med Assoc. 288:2271–2281.

Baniqued PL, Allen CM, Kranz MB, Johnson K, Sipolins A, Dickens C, Ward N, Geyer A, Kramer AF. 2015. Working memory, reasoning, and task switching training: Transfer effects, limitations, and great expectations? PLoS One. 10:1–29.

Basak C, Voss MW, Erickson KI, Boot WR, Kramer AF. 2011. Regional differences in brain volume predict the acquisition of skill in a complex real-time strategy videogame. Brain Cogn. 76:407–414.

Bjork BRA, Linn MC. 2006. The Science of Learning and the Learning of Science Introducing Desirable Difficulties. Science (80-). 19:6–7.

Boot WR, Basak C, Erickson KI, Neider M, Simons DJ, Fabiani M, Gratton G, Voss MW, Prakash R, Lee H, Low K a, Kramer AF. 2010. Transfer of skill engendered by complex task training under conditions of variable priority. Acta Psychol (Amst). 135:349–357.

Dahlin E, Neely AS, Larsson A, Bäckman L, Nyberg L. 2008. Transfer of learning after updating training mediated by the striatum. Science. 320:1510–1512.

Destrieux C, Fischl B, Dale A, Halgren E. 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage. 53:1–15.

Dosenbach NUF, Fair D a, Cohen AL, Schlaggar BL, Petersen SE. 2008. A dual-networks architecture of top-down control. Trends Cogn Sci. 12:99–105.

Dosenbach NUF, Nardos B, Cohen AL, Fair D a, Power D, Church J a, Nelson SM, Wig GS, Vogel AC, Lessov-schlaggar CN, Barnes KA, Dubis JW, Feczko E, Coalson RS, Jr JRP, Barch DM, Petersen SE, Schlaggar BL. 2011. Prediction of Individua Brain Maturity Using fMRI. Science (80-). 329:1358–1361.

Dosenbach NUF, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HC, Burgund ED, Grimes AL, Schlaggar BL, Petersen SE. 2006. A core system for the implementation of task sets. Neuron. 50:799–812.

Doyon J, Bellec P, Amsel R, Penhune V, Monchi O, Carrier J, Lehéricy S, Benali H. 2009. Contributions of the basal ganglia and functionally related brain structures to motor learning. Behav Brain Res. 199:61–75.

Eichenbaum H, Cohen NJ. 2008. From Conditioning to Conscious Recollection: Memory systems of the brain, From Conditioning to Conscious Recollection: Memory systems of the brain.

Elton A, Gao W. 2014. Divergent task-dependent functional connectivity of executive control and salience networks. Cortex. 51:56–66.

Erickson KI, Boot WR, Basak C, Neider MB, Prakash RS, Voss MW, Graybiel AM, Simons DJ, Fabiani M, Gratton G, Kramer AF. 2010. Striatal volume predicts level of video game skill acquisition. Cereb Cortex. 20:2522–2530.

Fischl B. 2012. FreeSurfer. Neuroimage. 62:774–781.

Geerligs L, Renken RJ, Saliasi E, Maurits NM, Lorist MM. 2015. A Brain-Wide Study of Age-Related Changes in Functional Connectivity. Cereb Cortex. 25:1987–1999.

Graybiel AM. 2005. The basal ganglia: learning new tricks and loving it. Curr Opin Neurobiol. 15:638–644.

Hastie T, Tibshirani R, Friedman J. 2009. The elements of statistical learning.

Herholz SC, Zatorre RJ. 2012. Musical Training as a Framework for Brain Plasticity: Behavior, Function, and Structure. Neuron. 76:486–502.

Hikosaka O, Nakamura K, Sakai K, Nakahara H. 2002. Central mechanisms of motor skill learning. Curr Opin Neurobiol. 12:217–222.

Hinton GE. 2011. Machine learning for neuroscience. Neural Syst Circuits. 1:12.

Hulme C, Melby-Lervåg M. 2012. Current evidence does not support the claims made for CogMed working memory training. J Appl Res Mem Cogn. 1:197–200.

Jaeggi SM, Buschkuehl M, Jonides J, Perrig WJ. 2008. Improving fluid intelligence with training on working memory. Proc Natl Acad Sci U S A. 105:6829–6833.

Kennedy KM, Raz N. 2005. Age, sex and regional brain volumes predict perceptual-motor skill acquisition. Cortex. 41:560–569.

Kotsiantis SB, Zaharakis ID, Pintelas PE. 2007. Machine learning: a review of classification and combining techniques. Artif Intell Rev. 26:159–190.

Liu F, Guo W, Fouche J-P, Wang Y, Wang W, Ding J, Zeng L, Qiu C, Gong Q, Zhang W, Chen H. 2013. Multivariate classification of social anxiety disorder using whole brain functional connectivity. Brain Struct Funct. 101–115.

Lövdén M, Wenger E, Mårtensson J, Lindenberger U, Bäckman L. 2013. Structural brain plasticity in adult learning and development. Neurosci Biobehav Rev. 37:2296–2310.

Lustig C, Shah P, Seidler R, Reuter-Lorenz P a. 2009. Aging, training, and the brain: a review and future directions. Neuropsychol Rev. 19:504–522.

Marquardt DW, Snee RD. 1975. Ridge Regression in Practice. Source Am Stat. 29:3–20.

Melby-Lervåg M, Hulme C. 2013. Is working memory training effective? A meta-analytic review. Dev Psychol. 49:270–291.

Metzler-Baddeley C, Caeyenberghs K, Foley S, Jones DK. 2016. Task complexity and location specific changes of cortical thickness in executive and salience networks after working memory training. Neuroimage. 130:48–62.

Mosteller F, Tukey JW. 1968. Data analysis, including statistics. In: Handbook of Social Psychology, Vol. 2.

Ng AY. 1997. Preventing“ overfitting” of cross-validation data. Icml.

Nikolaidis A, Goatz D, Smaragdis P, Kramer A. 2015. Predicting Skill-Based Task Performance and Learning with fMRI Motor and Subcortical Network Connectivity. 2015 Int Work Pattern Recognit NeuroImaging. 93–96.

Nikolaidis A, Voss MW, Lee H, Vo LTK, Kramer AF. 2014. Parietal plasticity after training with a complex video game is associated with individual differences in improvements in an untrained working memory task. Front Hum Neurosci. 8:1–11.

O’Neil C, Schutt R. 2013. Doing Data Science: Straight Talk from the Frontline.

Oelhafen S, Nikolaidis A, Padovani T, Blaser D, Koenig T, Perrig WJ. 2013. Increased parietal

activity after training of interference control. Neuropsychologia. 51:2781–2790.
Patel R, Spreng RN, Turner GR. 2013. Functional brain changes following cognitive and motor skills training: a quantitative meta-analysis. Neurorehabil Neural Repair. 27:187–199.
Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church J a, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, Petersen SE. 2011. Functional network organization of the human brain. Neuron. 72:665–678.
Rebok GW, Ball K, Guey LT, Jones RN, Kim HY, King JW, Marsiske M, Morris JN, Tennstedt SL, Unverzagt FW, Willis SL. 2014. Ten-year effects of the advanced cognitive training for independent and vital elderly cognitive training trial on cognition and everyday functioning in older adults. J Am Geriatr Soc. 62:16–24.
Redick TS, Shipstead Z, Harrison TL, Hicks KL, Fried DE, Hambrick DZ, Kane MJ, Engle RW. 2012. No Evidence of Intelligence Improvement After Working Memory Training: A Randomized, Placebo-Controlled Study. J Exp Psychol Gen. 142:359–379.
Sambataro F, Murty VP, Callicott JH, Tan HY, Das S, Weinberger DR, Mattay VS. 2010. Age-related alterations in default mode network: Impact on working memory performance. Neurobiol Aging. 31:839–852.
Schmidt R a., Bjork R a., Schmidt BRA, Bjork R a., Schmidt R a., Bjork R a. 1992. New Conceptualizations of Practice: Common Principles in Three Paradigms Suggest New Concepts for Training. Psychol Sci. 3:207–217.
Schmiedek F, Lövdén M, Lindenberger U. 2010. Hundred Days of Cognitive Training Enhance Broad Cognitive Abilities in Adulthood: Findings from the COGITO Study. Front Aging Neurosci. 2:1–10.
Seidler RD. 2010. Neural correlates of motor learning, transfer of learning, and learning to learn. Exerc Sport Sci Rev. 38:13.
Sestieri C, Corbetta M, Spadone S, Romani GL, Shulman GL. 2014. Domain-general Signals in the Cingulo-opercular Network for Visuospatial Attention and Episodic Memory. J Cogn Neurosci. 26:551–568.
Shen H, Wang L, Liu Y, Hu D. 2010. Discriminative analysis of resting-state functional connectivity patterns of schizophrenia using low dimensional embedding of fMRI. Neuroimage. 49:3110–3121.
Shipstead Z, Redick TS, Engle RW. 2012. Is working memory training effective? Psychol Bull. 138:628–654.
Shirer WR, Ryali S, Rykhlevskaia E, Menon V, Greicius MD. 2011. Decoding Subject-Driven Cognitive States with Whole-Brain Connectivity Patterns. Cereb Cortex. 22:158–165.
Singley MK, Anderson JR. 1989. The transfer of cognitive skill. Harvard University Press.
Soderstrom NC, Bjork RA. 2015. Learning Versus Performance: An Integrative Review. Perspect Psychol Sci. 10:176–199.
Uddin LQ, Kelly AMC, Biswal BB, Castellanos FX, Milham MP. 2009. Functional Connectivity of Default Mode Network Components: Correlation, Anticorrelation, and Causality. Hum Brain Mapp. 30:625–637.
Varoquaux G, Thirion B. 2014. How machine learning is shaping cognitive neuroimaging. 1–7.
Vo LTK, Walther DB, Kramer AF, Erickson KI, Boot WR, Michelle W, Prakash RS, Lee H, Fabiani M, Gratton G, Simons DJ, Sutton BP, Wang MY. 2011. Predicting individuals' learning success from patterns of pre-learning MRI activity. PLoS One. 6:e16093.
Zanto TP, Gazzaley A. 2013. Fronto-parietal network: Flexible hub of cognitive control. Trends Cogn Sci. 17:602–603.

Chapter 2

Brain Volume and Thickness Predictions of Cognitive Skill Training and Transfer to Working Memory, Fluid Intelligence, and Perceptual Speed Using Machine Learning

Aki Nikolaidis[1], Drew Goatz[1], Pauline Baniqued[2], Michael Kranz[1,5], Aron Barbey[1,5], Paris Smaragdis[1,3,4], Arthur F. Kramer[1,5]

Beckman Institute[1], Department of Computer Science[3], Department of Electrical and Computer Engineering[4],
Department of Psychology[5]
University of Illinois. Urbana Champaign
Urbana, Illinois. USA

Helen Wills Neuroscience Institute[2]
University of California. Berkeley
Berkeley, California. USA

Abstract: Research in cognitive training has demonstrated transfer of training to general cognitive ability, but there have been many failed attempts to replicate significant findings. At this point, a mechanistic understanding of how training drives transfer is not well characterized, and the neural underpinnings of training and transfer even less so. In the current study, we trained 45 young adults with a multimodal training game that targets complimentary aspects of reasoning, attention, and working memory. We use both pre-existing and training-induced differences in structural volumes and thicknesses to predict both training gains and transfer. To make these predictions we utilize a machine learning pipeline that we designed to extract the most reliable findings across models. We find that transfer is best predicted by a broad range of visual and auditory sensory processing and attention regions, suggesting that individual differences in these low level areas, as opposed to regions involved in higher level cognitive processes, may predispose individuals for greater benefits from training. We discuss how these findings support and contrast the neural overlap hypothesis of transfer.

2.1 Introduction

Cognitive neuroscience holds the potential not only to reveal how the brain gives rise to the mind, but also how to train the mind to improve its function through the recent emergence of a field known as cognitive training. While there have been many successful demonstrations of improvement of cognitive performance through training (Ball et al. 2002; Green and Bavelier 2003; Dahlin, Neely, et al. 2008; Dye et al. 2009; Thorell et al. 2009; Boot et al. 2010; Rebok et al. 2014), we and others have failed to replicate some of these findings (Melby-Lervåg and Hulme 2012; Redick et al. 2012; Shipstead et al. 2012; Oelhafen et al. 2013). The reliability of the successful demonstrations of transfer is further called into question due to methodological limitations such as small sample sizes and the use of single tasks to represent transfer. To create reliably effective training regimes, it is critical to not only overcome these methodological restraints, but also to investigate the mechanisms that drive transfer of training. Rather than exploring group differences, a more mechanistic explanation of training and transfer is available through the investigation of individual differences in neural correlates of training and transfer. While recent work has attempted to characterize the neural correlates of transfer by exploring how individual differences in brain structure and function contribute to training and transfer (Olesen et al. 2004; Anguera et al. 2007, 2013; Dahlin, Neely, et al. 2008; Vo et al. 2011; Voss et al. 2011; Oelhafen et al. 2013; Nikolaidis et al. 2014, 2015), the neural correlates of cognitive training have not yet been well characterized. In the skill acquisition literature, highly detailed descriptions of training-induced learning are available in both humans and animals (Hikosaka et al. 2002; Graybiel 2005, 2008; Doyon et al. 2009; Patel et al. 2013), and developing a more accurate description of the mechanisms of cognitive training and transfer will require a description of the phenomenon of human learning that encompasses both cognitive training and skill learning. Cognitive training is often described as a process of transferring training towards the improvement of underlying latent abilities, but another description that is more in line with the skill learning literature is also possible. In this characterization, cognitive training can also be seen as a process of acquiring cognitive skills, and transfer is the phenomenon of those acquired skills generalizing to other contexts, domains, and stimuli (Singley and Anderson 1989; Schmidt et al. 1992; Bjork and Linn 2006; Anguera et al. 2007; Seidler 2010). In this way, a critical question for cognitive training is establishing the relationship and distinctions between skill acquisition and generalization. Some work in the motor and verbal acquisition and transfer literatures have identified specific steps of the learning process, such as the associative learning phase, as well as experimental manipulations, such as modified feedback or spaced learning, which if properly probed causes knowledge to generalize to new contexts (Landauer and Bjork 1978; Bahrick 1979; Singley and Anderson 1989; Schmidt et al. 1992; Hikosaka et al. 2002; Seidler 2010). In order to develop mechanisms of transfer, it will be critical to characterize the neural similarities and distinctions between the process of skill acquisition and skill generalization.

To explore the neural correlates of training gain and transfer, we recruited 45 individuals to perform pre and post intervention neuropsychological testing, structural MRI, and 30 hours of training with Mind Frontiers. Mind Frontiers (MF) is a recently developed cognitive training program that utilizes six training tasks that each tap into aspects of working memory, executive function, and reasoning (Baniqued et al. 2015). In MF, participants perform six tasks that have demonstrated previous success in cognitive training, but which have been gamified and unified into a "Wild West" themed adaptive training game. The behavioral data for the current study has been published previously (Baniqued 2015), and in this study, the authors found that training with

MF demonstrated transfer to composite scores of working memory, fluid intelligence, and perceptual speed. In the current study, we investigate the unique and overlapping contributions of pre-existing brain structural volumes and thicknesses to both training and transfer. By examining how individuals are predisposed to greater learning and greater transfer, we may gain a more detailed picture of the cognitive and neural mechanisms utilized to maximize training gain and transfer. We also will assess how training-induced changes in brain structural volumes and thicknesses predict to both training and transfer, which enables us to gain a better understanding of how training may target different types of cognitive processing to lead to either training gain or transfer.

To make these predictions, we employ common techniques from machine learning. These techniques, have important advantages over the Ordinary Least Squares (OLS) method that is used in general linear models (GLMs). OLS suffers from problems of low reproducibility due to its likelihood to overfit solutions to a given sample (Hastie et al. 2009). In the face of collinear input variables, OLS is likely to maximize model fit in such a way that reduces model generalizability, thus reducing the likelihood of reproducing such findings, especially in the case of the relatively small samples that are common in cognitive neuroscience research. To address these difficulties, techniques such as ridge regression, which we use here, have been developed (Marquardt and Snee 1975). Rather than exclusively minimize the squared error term, this method minimizes a second term, which penalizes the model for using large model coefficients. This makes models more conservative in estimating the importance of individual variables, which makes the model both less likely to overfit and more likely to generalize to other datasets. The advantage of reduced overfitting is not unique to ridge regression, but rather is common among many machine learning methods such as lasso, support vector machines (SVMs), and random forests (Hastie et al. 2009).

This and other advantages have led to a dramatic increase in the use of machine learning techniques across scientific disciplines, and recently in both cognitive and clinical neuroscience (Hinton 2011; Varoquaux and Thirion 2014). Researchers have used these techniques to explore the structural and functional correlates of psychiatric illness (Shen et al. 2010), aging (Dosenbach et al. 2010; Meier et al. 2012; Vergun et al. 2013), and Alzheimer's disease (Ye et al. 2011). While these techniques have been informative in various domains of cognitive and clinical neuroscience, they have not yet been applied to understanding how regional brain volumes and thicknesses contribute to skill acquisition and generalization.

In the current study we applied machine learning to explore how both pre-existing structural characteristics and training-induced changes in these brain regions contribute to both training gain and transfer to working memory, fluid intelligence, and perceptual speed. To save computational time and maximize predictive accuracy, most applications of machine learning search for the most predictive set of features and regularization values using a technique called grid search. In the current work our goal was to maximize the reliability and interpretability of our findings rather than prediction accuracy. Therefore, we applied ridge regression across a range of model sizes and regularization values, allowing us to fully explore the model space of each analysis. We applied cluster correction to this model space to find regions where all neighboring models were significant predictors, and extracted predictors most reliably found across all significant models. Using this method, we were able to examine the extent to which pre-existing differences in brain structure that predispose individuals to demonstrate greater task independent learning across both trained and untrained domains. Furthermore, by assessing how training-induced changes predict training and transfer we can identify how training-induced changes in the brain may reflect specific skill versus general skill learning. Investigating these questions allows us to extend prior work by performing a more detailed examination of how pre existing and training-induced differences in brain structure reflect on the neural overlap hypothesis of training and transfer.

We expected to find regions that were predictive of both WM transfer and training gain, but which did not predict transfer to GFRT or PS. The neural overlap hypothesis of training and transfer asserts that in order for training to transfer to other domains, brain mechanisms trained must be shared with the transfer task (Jonides 2004; Dahlin et al. 2009; Buschkuehl and Jaeggi 2010; Klingberg 2010; Nikolaidis et al. 2014). Baniqued et al. demonstrated dissociation in the positive relationship of training gain to WM transfer but no relationship to GFRT or PS transfer, and we expected that a similar dissociation in the shared neural predictors should occur as well. Given the sparse literature on the neural correlates of transfer, it is difficult to predict which regions this would involve. On the one hand, overlapping predictors may be in regions unique to WM and the training tasks, or they may represent regions involved in cognitive processing bottlenecks that prevent transfer. Furthermore, given that regional brain volumes are calculated using both cortical surface area and thickness, and both volume and thickness, but not surface area have been widely associated with learning, we expected thickness to be the dataset more predictive of training and transfer. While both of these datasets have been investigated separately, they have rarely been compared to one another in association with skill learning. Finally, given previous demonstrations of the importance of subcortical regions in the acquisition and generalization of skills, we expected that predictions with subcortical regions should out-perform those without. We find that a broad range of regions associated with working memory and basic visual and auditory processing are associated with both transfer to working memory and improvement in training, but that the training task that most closely mirrors the WM transfer tasks shares the least amount of predictors in common. In contrast, training-induced plasticity in regions associated with working memory was reliably predictive of gains in the inductive reasoning training and working memory transfer. We discuss how these findings compare with the neural overlap hypothesis of transfer.

2.2 Methods

2.2.1 Participants

We recruited participants from the University of Illinois and surrounding community through flyers and online advertisements seeking participants for a "cognitive training study." Online pre-screening for demographic information and videogame experience was completed over email. Subsequent phone screening required participants: 1) between 18 and 30 years old, 2) 75% right handed according to the Edinburgh Handedness Inventory, 3) normal or corrected-to-normal hearing and vision, 4) no major psychological or medical conditions, 5) no non-removable metal in the body, and 6) played no more than 5 hours of video games per week in the past six months. A total of 45 participants trained with Mind Frontiers and had complete behavioral and neural data.

2.2.2 Study Design

A complete description of the study design can be found in Baniqued et al. 2015. Briefly, participants completed three initial cognitive testing sessions, and an MRI session before beginning the intervention, in which they completed four to five training sessions per week for four to five weeks, for a total of 20 sessions. After training, subjects returned to lab for a final MRI session and another round of cognitive testing.

Participants trained with Mind Frontiers, which was comprised of six computerized adaptive training tasks that were selected based on their known psychometric properties with reasoning/Gf, working memory, visuospatial reasoning, inductive reasoning, and task switching. See Table 2.1 for a description of each training task.

2.2.3 Cognitive Assessment Protocol

For a full listing of all neuropsychological tasks administered, see Baniqued et al. 2015. Here we briefly describe each of the tasks that were included to create the reasoning, working memory, and perceptual speed composite variables we predicted. The tests below were obtained from the Virginia Cognitive Aging Project Battery (Salthouse 2005), and two different versions were used for pre- and post-testing, with the sequence counterbalanced across subjects.

2.2.3.1 GFRT

The GFRT composite was comprised of standardized reaction times in these four tasks. Descriptions for these and other tasks in section 2.3 are reproduced from Baniqued et al. 2015.

Matrix Reasoning: Select the pattern that completes a missing space on a 3 x 3 grid. Number of correctly answered items is the primary measure. Reaction time on correct trials was also analyzed. Source: Ravens 1962; Crone et al. 2009.

Paper Folding: Identify pattern of holes that results from a punch through folded paper. Number of correctly answered items is the primary measure. Reaction time on correct trials was also analyzed. Source: Ekstrom et al. 1976.

Spatial Relations: Identify 3D object that would match a 2D object when folded. Number of correctly answered items is the primary measure. Reaction time on correct trials was also analyzed. Source: Bennett et al. 1947.

Letter Sets: Determine which letter set is different from the other four. Number of correctly answered items is the primary measure. Reaction time on correct trials was also analyzed. Source: Ekstrom et al. 1976.

2.2.3.2 Perceptual Speed

The Perceptual Speed composite was comprised of standardized accuracy scores in these three tasks.

Digit Symbol Substitution: Write corresponding symbol for each digit using a coding table. The primary measure is the number of correctly answered items within two minutes. Source: Wechsler 1997.

Pattern Comparison: Determine whether pairs of line patterns are the same or different. The primary measure is the number of correctly answered items within 30 seconds, averaged across two sets of problems. Source: Salthouse and Babcock 1991.

Letter Comparison: Determine whether pairs of letter strings are the same or different. The primary measure is the number of correctly answered items within 30 seconds, averaged across two sets of problems. Source: Salthouse and Babcock 1991.

2.2.3.3 Working memory

The WM composite was comprised of the d' scores of the 2-back, 3-back, Dual 2-back, and Dual 3-back.

Single N-back: Determine whether the current letter presented matches the letter presented two or three items back. The primary measure of d-prime was computed separately for the 2-back and 3-back conditions. Reaction times on correct trials were also analyzed. Source: Kirchner 1958; Kane et al. 2007.

Dual N-back (administered in the MRI): Determine whether simultaneously presented auditory and visual stimuli match stimuli presented one, two, or three items ago. The primary measure of d-prime was computed separately for the two-back and three-back conditions. Reaction times on correct trials were also analyzed. Source: Jaeggi et al. 2003, 2007.

2.2.4 Analysis of Behavioral Results

Training with Mind Frontiers produce significant gains in performance not only across all training tasks, but also to three cognitive constructs, working memory (WM), reasoning reaction time (GFRT), and perceptual speed (PS). The gain scores for the training and transfer tasks were calculated by taking the difference between post- and pre-test scores, and dividing by the standard deviation of the pre-test score. In the current study, we use these transfer scores as predictors, as well as training-induced gain scores from three of the training games: Supply Run (SR), Sentry Duty (SD), and Safe Cracker (SC). These three training games were chosen because improvement during training was positively correlated with transfer to the working memory but not perceptual speed or fluid intelligence (Baniqued 2015), indicating that training-induced gains in these tasks may play a role in WM transfer.

2.2.5 Structural MRI Analysis Details

2.2.5.1 MPRAGE Acquisition Parameters

All structural, high-resolution, T1-weighted MPRAGE images were collected on a 3 T Siemens Trio scanner, collected in ascending order parallel to the anterior and posterior commissures with the following parameters: Dataset 1: Echo time (TE) = 2.32 ms, repetition time (TR) = 1900 ms, field of view (FOV) = 230 mm, voxel size $0.9mm^3$, and flip angle = 8°.

2.2.5.2 Structural MRI Preprocessing

All preprocessing of structural images was performed with FSL and Freesurfer. After being skull stripped, each structural image was linearly registered to the MNI 2mm brain template with 12 degrees of freedom (Smith 2004). We used the Freesurfer longitudinal stream to improve the robustness of the longitudinal analysis of cortical volume and thickness (Fischl 2012, Jovich 2009). This scheme creates within-subject anatomical templates that allow for less bias of analysis towards a particular time point (Thomas & Baker 2013). This process has been described elsewhere in detail (Metzler-Baddeley 2016, Fischl 2012), but briefly, this involves Talairach transformation, white matter and gray matter segmentation, automated topology correction, and intensity normalization. After this semi-automated procedure, images are all inspected manually to assess performance of the Freesurfer pipeline. Cortical thickness and volumes and subcortical volumes were calculated using Freesurfer using the Destrieux atlas, which parcellates the cortex into 74 gyral-sulcal regions per hemisphere (Destrieux et al. 2010). Previous studies have explored the association between either cortical thickness or brain volume with learning, but the relative information available in these two metrics with regards to individual differences in learning is poorly understood. Furthermore, contributions of subcortical and white matter volumes are rarely addressed in the same sample as cortical volumes and thicknesses. To address these questions, we used the brain volumes and cortical thickness data to create five separate datasets: Thickness (TH), in which we only include the 150 cortical thickness variables; Volume (VO), which only includes the 148 cortical volume features; Whole Brain (WB), which contains the VO dataset as well as 63 additional subcortical, white matter, and cerebellar volumes; Volume + Thickness (VO-TH), which contains the combined VO and TH datasets; Whole Brain + Thickness (WB-TH), which contains both the WB and TH datasets. We extracted thickness and volume data from both the pre-training structural images (Pre), and to assess change in thickness and volume, we subtracted the pre thicknesses and volumes from the post data (Change). To explore differential contributions of variables from both the Pre and Change datasets, we predicted each of the six scores with ten different datasets, five Pre datasets and five Change datasets.

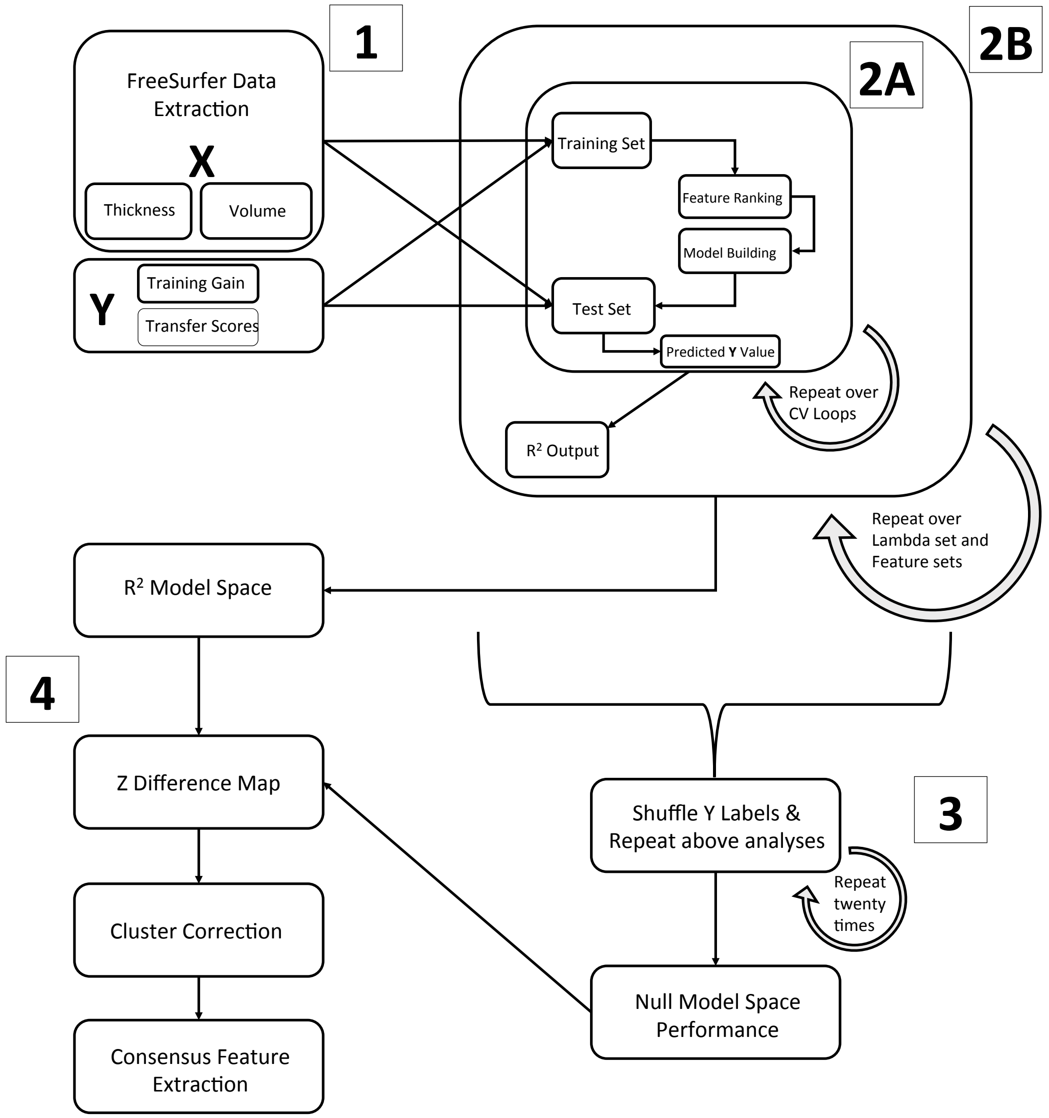

Figure 2.1:
This figure provides a visual flowchart of the methods used in the current work. Analysis begins with the structural features extracted from Freesurfer (1), and then it continues through model building (2A), cross validation (2A), and repeating our analysis across feature and lambda space (2B). Finally, we create the null model distribution by shuffling the Y labels (3), and we assess the difference between the model space and the null model space in the form of a Z-map (4). We apply a cluster correction to this Z map and extract consensus features (4).

2.2.6 Machine Learning Methods

2.2.6.1 Predicting training gain and transfer

We used either pre-training (Pre) or pre-post differences (Change) in cortical thicknesses and brain volumes to make predictions with the Sci-Kit Learn ridge regression (Buitinck et al. 2013) (Figure 2.1: Section 1). While this algorithm does not achieve predictive accuracies as high as other more advanced methods, such as random forests and support vector machines (SVMs), ridge regression offers an important advantage in terms of ease of interpretability and comparability to other analyses in this domain using linear models (Kotsiantis, Zaharakis, & Pintelas, 2007).

Over each leave-one-subject-out cross-validation (LOSO-CV) loop, we trained models of a range of sizes, and we used these models to predict scores of the left-out subject (Figure 2.1: Section 2A). Within each cross validation loop, we ranked all features in the training group by absolute correlation strength. We created feature sets of the top 1 to *n* features, by groups of 5 to save processing time, and predicted each left out subject using each of these feature sets. We calculated R^2 values by squaring the correlation of predicted and actual score values.

In order to assess the true predictive accuracy and overfitting in our models, we created a distribution of null models (Figure 2.1: Section 3). To create these null models, we randomized the score values for the subjects, and repeated the procedure outlined above to predict these values. We repeated this procedure 20 times to create a distribution of null models, and we compared the mean and standard deviation of this null model distribution to our actual model to determine whether our model predicted significantly better than chance.

2.2.6.2 Assessing Reliability of Prediction

Rather than using cross validation to select an optimal hyperparameter (lambda) value for each model, we repeated our predictions and null model testing at a range of lambda values (Figure 2.1: Section 2B). While computationally intensive, this method allows us to more fully explore the space of possible models, extract clusters of predictive models, and assess both the reliability of our models, and the consistency of parameters selected in different models.

We assessed how our model performed over the full model space compared to the distribution of null models, and we transformed the difference in prediction accuracy between the model and null model distribution to a Z score distribution, representing the number of standard deviations that the model outperforms chance (Figure 2.1: Section 4). To get a more accurate assessment of average model prediction in a region and account for variation in neighboring models, we first smoothed the data using a 2 x 3 FWHM Gaussian kernel before applying cluster correction. We set a cluster threshold of one standard deviation so that all the weakest models that predicted less than 1 standard deviation above the null distribution were removed, and then to remove small noisy predicting clusters we removed all clusters of models with less than 12 models. Finally, in order to only assess datasets that were most reliably predictive of the outcome variable, we removed clusters for datasets for which at least one cluster did not occupy at least 3% of model space. We performed cluster correction to not only assess the relative predictive accuracy and reliability of each dataset across a range of models, but also to extract individual features that were reliably included in 90-100% of models in a given cluster, also known as consensus features (CFs).

In this way, we extract only the features that most reliably belong to models that significantly predict our outcome variable. This is an important step because it allows us to look for common individual predictors across outcome variables. While a mass univariate analysis, such as the feature ranking commonly performed in machine learning pipelines (including this one), detects individual variables that are each correlated with the outcome variable, it does not offer any information on conditional relationships between each feature and the outcome variable. This leads to variables that are correlated with the outcome variable but also redundant with one another (Guyon 2003). By exploring the model space more fully, and identifying sets of CFs that are the most reliable predictors of our outcome variable, we are able to identify features that may not be good predictors on their own, but that reliably predict the residual variance that is left after the strongest predictors have been included in the model. CFs allow us to analyze the feature selection process and make inferences based on the features that are selected. For a review on feature selection and how redundant and superficially 'uninformative' features can play an important role in predicting residual variance, see Guyon 2003.

Game	Description	Source
Supply Run (working memory)	Townspeople request items that belong to a certain category. There are five objects in each category, which correspond to stereotypical occupations of the "Wild West." Once the store is reached, the last item from each category must be selected. Difficulty level is manipulated by the number of requests and the number of categories.	Updating WM (Dahlin, Nyberg, et al. 2008)
Sentry Duty (working memory)	Sentries lift their lanterns while saying a word of the phonetic alphabet. The current word spoken and lantern lifted is compared to the word spoken/lantern lifted *n* times previously. There may be an audio match, a visual match, an audio and visual match, or no match between the current sentry and the one who spoke *n* times ago. Difficulty is manipulated by how far back the comparison is (e.g., 1-back, 2-back, 3-back).	Dual n-back(Jaeggi et al. 2008; Redick et al. 2012)
Safe Cracker (reasoning)	Safe combinations are determined by completing the next item in a series. Series may be letter-, number-, or day/month- based, and all are governed by some pattern or rule that must be determined and applied to select the next item in the series. Difficulty is manipulated by the difficulty of the patterns and the number of problems to solve within the given time limit.	Inductive Reasoning (Willis and Schaie 1986)

Table 2.1:
Description of the three training tasks investigated in the current study. Partially reproduced from Baniqued et al. 2015.

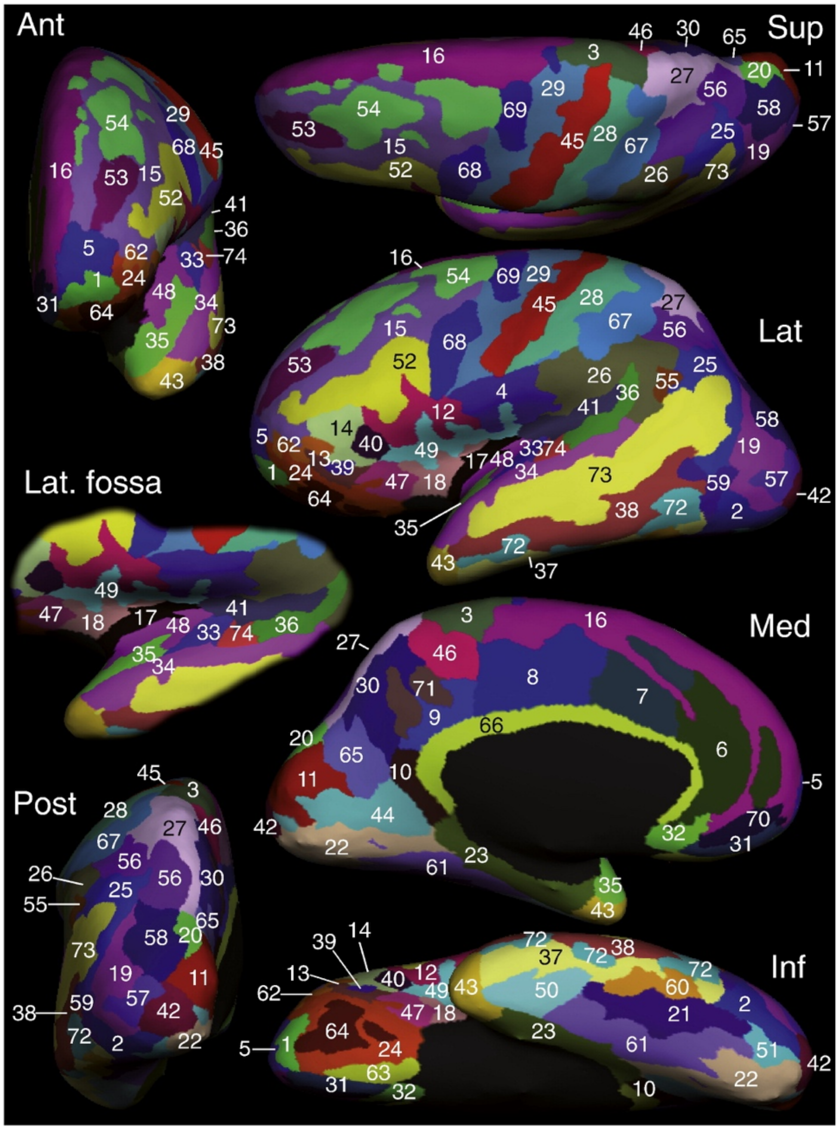

Figure 2.2:
A view of the 74 regions in an inflated hemisphere, displaying both sulcal and gyral regions. Numerical indices refer to anatomical regions defined in Table 2.2. Views from top left to bottom right: Anterior (Ant), Superior (Sup), Lateral (Lat), Lateral Fossa (Lat fossa), Medial (Med), Posterior (Post), and Inferior (Inf). The Lateral Fossa view offers a better view of the gyral and sulcal regions in the insula and temporal cortices. Reproduced from Destrieux et al. 2010.

#	Label	#	Label	#	Label
1	Fronto-marginal gyrus and sulcus	47	Anterior segment of the of the insula	93	Right Lateral Ventricle
2	Inferior occipital gyrus (O3) and sulcus	48	Inferior segment of the insula	94	Right Inferior Lateral Vent
3	Paracentral lobule and sulcus	49	Superior segment of the insula	95	Right Cerebellum WM
4	Central operculum	50	Anterior transverse collateral sulcus	96	Right Cerebellum Cortex
5	Transverse frontopolar gyri and sulci	51	Posterior transverse collateral sulcus	97	Right Thalamus Proper
6	Anterior cingulate cortex (ACC)	52	Inferior frontal sulcus	98	Right Caudate
7	Middle-anterior cingulate cortex and sulcus (aMCC)	53	Middle frontal sulcus	99	Right Putamen
8	Middle-posterior part of the cingulate gyrus and sulcus (pMCC)	54	Superior frontal sulcus	100	Right Pallidum
9	Posterior-dorsal cingulate gyrus (dPCC)	55	Sulcus intermedius primus (of Jensen)	101	Right Hippocampus
10	Posterior-ventral part of the cingulate gyrus (vPCC)	56	Intraparietal sulcus (interparietal sulcus) and transverse parietal sulci	102	Right Amygdala
11	Cuneus (O6)	57	Middle occipital sulcus and lunatus sulcus	103	Right Accumbens area
12	Opercular part of the inferior frontal gyrus	58	Superior occipital sulcus and transverse occipital sulcus	104	Right Ventral DC
13	Orbital part of the inferior frontal gyrus	59	Anterior occipital sulcus and preoccipital	105	Right vessel
14	Triangular inferior frontal gyrus	60	Lateral occipito-temporal sulcus	106	Right choroid plexus
15	Middle frontal gyrus (F2)	61	Medial occipito-temporal sulcus (collateral sulcus) and lingual sulcus	107	5th Ventricle
16	Superior frontal gyrus (F1)	62	Lateral orbital sulcus	108	WM hypointensities
17	Long insular gyrus and central sulcus of the insula	63	Medial orbital sulcus (olfactory sulcus)	109	Left WM hypointensities
18	Short insular gyri	64	Orbital sulci (H-shaped sulci)	110	Right WM hypointensities
19	Middle occipital gyrus (O2, lateral occipital gyrus)	65	Parieto-occipital sulcus (or fissure)	111	Non-WM hypointensities
20	Superior occipital gyrus (O1)	66	Pericallosal sulcus (S of corpus callosum)	112	Left non-WM hypointensities
21	Lateral occipito-temporal gyrus (fusiform gyrus, O4-T4)	67	Postcentral sulcus	113	Right non-WM hypointensities
22	Lingual gyrus, ligual part of the medial occipito-temporal gyrus, (O5)	68	Inferior part of the precentral sulcus	114	Optic Chiasm
23	Parahippocampal gyrus	69	Superior part of the precentral sulcus	115	CC Posterior
24	Orbital gyri	70	Suborbital sulcus (sulcus rostrales, supraorbital sulcus)	116	CC Mid Posterior
25	Angular gyrus	71	Subparietal sulcus	117	CC Central
26	Supramarginal gyrus	72	Inferior temporal sulcus	118	CC Mid Anterior
27	Superior parietal lobule (lateral part of P1)	73	Superior temporal sulcus (parallel sulcus)	119	CC Anterior
28	Postcentral gyrus	74	Transverse temporal sulcus	120	Brain Seg Volume
29	Precentral gyrus	75	Left Lateral Ventricle	121	Brain Seg Volume No Vent
30	Precuneus (medial part of P1)	76	Left Inferior Lateral Vent.	122	Brain Seg Volume No Vent Surface
31	Straight gyrus, Gyrus rectus	77	Left Cerebellum White Matter	123	Left Hem. Cortex Volume
32	Subcallosal area, subcallosal gyrus	78	Left Cerebellum Cortex	124	Right Hem. Cortex Volume
33	Anterior transverse temporal gyrus (of Heschl)	79	Left Thalamus Proper	125	Cortex Volume
34	Lateral aspect of the superior temporal gyrus	80	Left Caudate	126	Left Hem. Cortical White Matter Volume
35	Planum polare of the superior temporal gyrus	81	Left Putamen	127	Right Hem. Cortical White Matter Volume
36	Planum temporale or temporal plane of the superior temporal gyrus	82	Left Pallidum	128	Cortical White Matter Volume
37	Inferior temporal gyrus (T3)	83	3rd Ventricle	129	Subcortical Gray Volume
38	Middle temporal gyrus (T2)	84	4th Ventricle	130	Total Gray Volume
39	Horizontal ramus of the anterior segment of the lateral sulcus (or fissure)	85	Brain Stem	131	Supra Tentorial Volume
40	Vertical ramus of the anterior segment of the lateral sulcus (or fissure)	86	Left Hippocampus	132	Supra Tentorial Volume No Vent
41	Posterior ramus (or segment) of the lateral sulcus (or fissure)	87	Left Amygdala	133	Supra Tentorial Volume No Vent Volume
42	Occipital pole	88	CSF	134	Mask Volume
43	Temporal pole	89	Left Accumbens area	135	Brain Seg Volume to eTIV
44	Calcarine sulcus	90	Left Ventral DC	136	Mask Volume to eTIV
45	Central sulcus (Rolando's fissure)	91	Left vessel	137	Estimated Total Intra Cranial Vol
46	Marginal branch (or part) of the cingulate sulcus	92	Left choroid plexus		

Table 2.2:
Items 1-74 correspond to the cortical regions labeled in Figure 2.2. Items 75-137 correspond to the volumes for white matter, subcortical, whole brain variables, as well as indicators of brain underdevelopment or atrophy, such as ventricle volume and white matter hyperintensities. Columns 1, 3, 5, labeled '#' correspond to the structural assignment, and columns 2, 4, 6, labeled 'Label' are the shortened structural labels from the Destrieux atlas.

2.3 Results

2.3.1 Prediction Accuracies

We found that transfer to fluid intelligence reaction times (GFRT) was significantly predicted by clusters of models in both the Pre VO and Pre WB datasets (Table 2.1). The addition of subcortical regions not only nearly tripled the number of models that were significant predictors (VO Model % Sig = 6.0%; WB Model % Sig = 16.0%), but it also doubled the average predictive accuracy of the cluster of models (VO Avg. R^2 = 8.6%; VO Upper 95% CI R^2 = 12.2%; WB Avg. R^2 = 16.7%; WB Upper 95% CI R^2 = 30.0%), supporting the notion that subcortical regions may be particularly informative in predicting transfer. The thickness variables were not predictive of transfer to GFRT in any dataset, and may have been prevented successful prediction in the case of the VO-TH datasets. Transfer to GFRT was not predicted by any of the Change datasets. On the other hand, transfer to Perceptual Speed (PS) was significantly predicted by both Pre and Change WB-TH datasets. Clusters in both Pre and Change datasets were equally predictive (Table 2.3 & 2.4). The Change VO-TH dataset was the best predictor compared to the Change WB-TH and Pre WB-TH datasets (Change VO-TH Model % Sig = 7.2%; Avg. R^2 = 9.6%; Upper 95% CI R^2 = 14.6%; Table 2.4). Similarly, the WM transfer score was predicted by both the Pre and Change WB-TH (Pre WB-TH Model % Sig = 3.1%; Avg. R^2 = 6.2%; Upper 95% CI R^2 = 13.9%; Table 2.3; Change WB-TH Model % Sig = 6.4%; Avg. R^2 = 5.5%; Upper 95% CI R^2 = 9.4% Table 2.4).

Training gain in Safe Cracker (SC) was predicted by the change dataset only, with predictions in the Change TH (Model % Sig = 3.1%; Avg. R^2 = 5.6%; Upper 95% CI R^2 = 7.2%; Table 2.4) being more than doubled in the Change WB-TH datasets (Model % Sig = 6.5%; Avg. R^2 = 10.3%; Upper 95% CI R^2 = 15.8% Table 2.4), demonstrating the importance of unified volumetric and thickness changes in this prediction. Training gain in Sentry Duty (SD) was best predicted by Pre TH (Model % Sig = 6.8%; Avg. R^2 = 11.2%; Upper 95% CI R^2 = 17.7% Table 2.3). The addition of the volume data decreased the relative size of the cluster passing significance threshold, indicating a relatively smaller set of models passed threshold. This suggests that the volume data may not have been additionally informative for predicting SD gain. SD was not predicted by any Change dataset. Gain in the Supply Run (SR) task was predicted by the Pre TH dataset (Model % Sig = 5.3%; Avg. R^2 = 8.9%; Upper 95% CI R^2 = 13.4% Table 2.3), but not by any of the Pre datasets containing volume information, nor any Change dataset.

Outcome	Pre Structure	Total Model Space	Cluster Model Space	Model Avg R^2	Model Upper 95% R^2	CFs Per Cluster	Null Avg R^2	Null Upper 95% Avg R^2
GFRT								
	VO	5.95%	3.44%	8.61%	12.24%	11, 5	0.17%	0.40%
	WB	16.01%	5.46%	16.70%	30.04%	2, 21, 23	0.16%	0.55%
PS								
	WB-TH	3.13%	3.13%	6.80%	8.71%	3	0.01%	0.02%
WM								
	VO-TH	3.51%	3.51%	6.45%	10.05%	8	0.09%	0.19%
	WB-TH	3.11%	3.11%	6.17%	13.86%	16	0.07%	0.23%
SC	-	-	-	-	-	-	-	-
SD								
	TH	6.77%	6.77%	11.20%	17.70%	0	0.17%	0.47%
	VO-TH	3.70%	3.70%	11.02%	19.81%	3	0.07%	0.31%
SR								
	TH	5.27%	3.14%	8.94%	13.36%	20, 22	0.05%	0.14%

Table 2.3:
This table summarizes the performance of the Pre Structure datasets that were significantly predictive of each training gain and transfer score (Column 1- 'Outcome' Corresponds to the Training gain/Transfer Score, Column 2- 'Pre Structure' corresponds to the dataset used). The third column, 'Total Model Space', displays the percentage of model space that is populated by all clusters of models that significantly predict the outcome variable, and the fourth column, 'Cluster Model Space', shows the percentage of model space covered by the largest cluster. The fifth column, 'Model Avg R^2', shows the average percentage of out-of-sample variance in the outcome variable that is predicted by the largest cluster, which should be compared to column eight, 'Null Avg R^2', which shows the out-of-sample variance predicted by the null models over the same area in model space. Column six 'Model Upper 95% R^2', displays the upper 95% confidence interval of percentage of out-of-sample variance predicted by the cluster, and this should be compared to column nine, 'Null Upper 95% Avg R^2', which contains the upper 95% confidence interval for the null models. Column seven, 'CFs Per Cluster', displays the number of consensus features per significant clusters. Outcome and dataset abbreviations follow the convention of the text. Scores: GFRT- fluid intelligence reaction time, PS- perceptual speed, WM- working memory, SC- Safe Cracker, SD- Sentry Duty, SR- Supply Run. Datasets: VO- volume, WB- whole brain volume, TH- thickness, VO-TH- combined volume and thickness datasets, WB-TH combined whole brain volume and thickness datasets.

Outcome	Change Structure	Total Model Space	Cluster Model Space	Model Avg R^2	Model Upper 95% R^2	CFs per Cluster	Null Avg R^2	Null Upper 95% Avg R^2
GFRT	-	-	-	-	-	-	-	-
PS								
	VO-TH	7.22%	5.04%	9.56%	14.60%	3, 62	0.04%	0.09%
	WB-TH	3.56%	3.56%	8.90%	16.00%	3	0.01%	0.04%
WM								
	WB-TH	6.42%	3.05%	5.54%	9.37%	9, 31	0.04%	0.09%
SC								
	TH	3.05%	3.05%	5.64%	7.17%	16	0.03%	0.06%
	WB-TH	6.48%	4.12%	10.28%	15.83%	0	0.21%	0.51%
SD	-	-	-	-	-	-	-	-
SR	-	-	-	-	-	-	-	-

Table 2.4:
This table summarizes the performance of the Change Structure datasets that were significantly predictive of each training gain and transfer score (Column 1- ‘Outcome’ Corresponds to the Training gain/Transfer Score, Column 2- ‘Pre Structure’ corresponds to the dataset used). The third column, ‘Total Model Space’, displays the percentage of model space that is populated by all clusters of models that significantly predict the outcome variable, and the fourth column, ‘Cluster Model Space’, shows the percentage of model space covered by the largest cluster. The fifth column, ‘Model Avg R^2’, shows the average percentage of out-of-sample variance in the outcome variable that is predicted by the largest cluster, which should be compared to column eight, ‘Null Avg R^2’, which shows the out-of-sample variance predicted by the null models over the same area in model space. Column six ‘Model Upper 95% R^2’, displays the upper 95% confidence interval of percentage of out-of-sample variance predicted by the cluster, and this should be compared to column nine, ‘Null Upper 95% Avg R^2’, which contains the upper 95% confidence interval for the null models. Column seven, ‘CFs Per Cluster’, displays the number of consensus features per significant clusters. Outcome and dataset abbreviations follow the convention of the text. Scores: GFRT- fluid intelligence reaction time, PS- perceptual speed, WM- working memory, SC- Safe Cracker, SD- Sentry Duty, SR- Supply Run. Datasets: VO- volume, WB- whole brain volume, TH- thickness, VO-TH- combined volume and thickness datasets, WB-TH combined whole brain volume and thickness datasets.

2.3.2 Consensus Features

Among the CFs in the Pre VO dataset that predicted GFRT, we see that the most predictive features are the volumes of the left inferior precentral sulcus, and opercular portion of the left inferior frontal gyrus, which both demonstrate weak negative trending correlations with transfer to GFRT (r = -0.24, p = 0.11; r = -0.22, p = 0.14 respectively). In the CFs from the Pre WB dataset, the strongest predictors of GFRT transfer were again the left inferior precentral sulcus, as well as the right cerebellum white matter volume, which was positively correlated with transfer to GFRT (r = 0.35, p = 0.01). Transfer to PS was predicted by both the Pre and Change VO-TH datasets. Among the CFs of the Pre VO-TH dataset, we found that the lingual gyrus demonstrated a positive correlation with transfer to PS (r = 0.35, p = 0.020), and volume in the right medial collateral and lingual sulcus was positively correlated with transfer (r = 0.35, p = 0.017). In the Change VO-TH dataset, the change in thickness of left triangular inferior frontal gyrus demonstrated a positive trend with PS (r = 0.24, p = 0.11), and the change in volume of the left opercular inferior frontal gyrus was correlated with PS (r = 0.31, p = 0.037). WM transfer was predicted by the Pre VO-TH dataset, with thickness of the right dorsal posterior cingulate cortex negatively correlated with WM (dPCC; r = -0.31, p = 0.040), and volume of the left pericallosal sulcus (r = -0.25, p = 0.080) trending negatively with WM transfer. The Pre WB-TH dataset was also able to predict transfer to WM, and both the volume of the left vertical ramus of the lateral sulcus and thickness of the left paracentral lobule and sulcus were negatively correlated with WM transfer (r = -0.30, p = 0.044; r = -0.27, p = 0.072 respectively). In the Change WB-TH dataset, the change in both the thickness and volume of the left vertical ramus of the lateral sulcus (r = 0.41, p = 0.0046; r = 0.43, p = 0.002), and volume of the right posterior collateral sulcus (r = 0.41, p = 0.0043) were positively correlated with WM transfer.

Training gain in SC was predicted by the Change TH and Change WB-TH datasets. In the Change TH dataset, we found that changes in thickness of the left precentral gyrus and right postcentral gyrus were positively correlated with improvement in SC (r = 0.26, p = 0.083; r = 0.38, p = 0.0099 respectively). The Change WB-TH dataset did not yield any CFs. Training gain in SD was predicted by the Pre TH and Pre VO-TH datasets, but the Pre TH dataset did not contain any CFs. In the Pre VO-TH dataset, no region demonstrated either significant or trending correlation with SD. Training gain in SR was only predicted by the Pre TH dataset, and both thickness in the left superior parietal lobe (r = 0.26, p = 0.081) and the right posterior ramus of the lateral sulcus (r

= 0.29, p = 0.050) were positively associated with SR training gain. These correlations did not pass Bonferroni correction for multiple comparisons.

2.3.3 Consensus Similarities Across Training and Transfer

CFs allow us to assess the reliability of a feature's role in prediction. Every consensus feature is one which not only was selected for all cross validation loops for a given feature set size and lambda value, but it also was a consensus feature for 90-100% of the models in a given cluster. Given that transfer to WM was positively correlated to training gain in the SD, SR, and SC tasks, we first tested the extent to which we would see overlap here. While we saw extensive overlap in CFs between WM and the SC and SR training gain, we saw the least overlap between WM and SD, regardless of the fact that they are both n-back based. The only overlap was in the initial thickness of the intermedius primus portion of the right intraparietal sulcus, which is known to play a role in visuospatial working memory (Coull and Frith 1998; Silk et al. 2010). We could not assess any overlap between change predictions in these two scores' CFs because training gain in SD was not predicted by change in any set of volumes or cortical thickness.

Transfer to WM and gain in SR were extensively overlapping in their feature sets, with both sharing a predominantly left hemisphere set of cortical thicknesses. Many of these regions have been previously identified by several studies to demonstrate functional activation changes after either motor or cognitive training, including the right insula (Arima 2011; Puttemans 2005; Ronsse 2011; Zhang 2011), left middle occipital gyrus (Petersson 2001; Thomas 2009), left inferior frontal gyrus (Arima 2011; Thomas 2009; Zhang 2011), and left fusiform gyrus (Chein 2005; Eickhoff 2009). We found that cortical thickness of auditory processing regions, such as the supramarginal gyrus and transverse temporal gyrus, which are involved in phonological storage of working memory (Celsis et al. 1999), were predictive of gain in both SR and WM. Furthermore, cortical thickness of several visual processing regions, such as the cuneus, occipital pole, and calcarine sulcus, as well as the insula and inferior frontal gyrus, which are involved in attention and executive processing, were also commonly predictive in both SR and WM (Aron et al. 2003; Menon and Uddin 2010). Training gain in SR was not predicted by change in any set of volumes or cortical thickness, so we were not able to compare the extent of overlap in these features; however, transfer to WM and SC were both successfully predicted by the Change WB-TH dataset.

We found that several CFs in the combined change in cortical thicknesses and volumes of sensory processing were shared between clusters of models predicting both SC and WM. Changes in the cortical thickness of the left superior frontal gyrus, which is involved in visuospatial working memory (Rypma et al. 1999; Boisgueheneuc et al. 2006), left anterior transverse temporal gyrus, which contains the primary auditory cortex (Destrieux et al. 2010), and right collateral sulcus and lingual sulcus, which are involved in visual processing (Newman-Norlund et al. 2006; Ilg et al. 2008; Ronsse et al. 2011) were identified as CFs in models that predicted both transfer to WM and training-induced gains in SC. Similarly, models that predicted transfer to WM and gains in SC both contained changes in cortical volume in sensory processing and attention regions such as the left and right collateral sulci, which are involved in object recognition and visual attention (Awh et al. 2000), and the left cingulate sulcus which is critical for attention (Vogt et al. 1992; Nebel et al. 2005; Colla et al. 2008). Several of these regions have been implicated in general skill acquisition by demonstrating training-induced changes in functional activation after both motor and cognitive skill training. For example, both right and left lingual gyri, left posterior cingulate cortex, and right fusiform gyrus have demonstrated changes in activity both after motor training (Penhune and Doyon 2002; Kim et al. 2004; Ronsse et al. 2011; Zhang et al. 2011), and cognitive training (Puttemans 2005; Newman-Norlund et al. 2006; Wong et al. 2007; Ilg et al. 2008; Raboyeau et al. 2010; Schmidt-Wilcke et al. 2010). We could not assess overlapping predictions between initial volumes and cortical thickness because SC was not predicted by any Pre dataset. Furthermore, as GFRT and WM were predicted by completely non-overlapping sets of data, we could not assess overlapping CFs between these predictions either, nor could we assess overlapping predictions between GFRT and the training gain scores (SR, SD, SC).

2.3.4 Comparing Predictions

Given that previous work has both examined how regional cortical thickness and regional brain volumes relate to cognitive performance and learning, we wanted to compare how our predictions would fare when each of these datasets was used independently, and whether predictions would be improved by combining thickness and volume information. On average we saw that the thickness data was about twice as informative as the volume data. Significant clusters of models emerged in 3/12 (25%) of the TH dataset predictions, while only 3/24 (12.5%) of VO dataset analyses contained significant predictors. Furthermore, only 7/48 (14.5%) of all analyses using the VO data contained significantly predictive clusters of models, while 11/36 (30.5%) of the analyses using the TH data contained significant clusters. Only transfer to GFRT is uniquely predicted by the volume datasets, whereas all other transfer and training variables are predicted by TH and combinations of TH and VO. It is important to note that cortical volumes are calculated using cortical thickness and surface area, so these data are necessarily collinear. Regardless of the collinearity of these data, the WB-TH and VO-TH datasets are the most informative, predicting of41.7% and 25% of their analyses respectively.

2.4 Discussion

We assessed the extent to which the CFs in the training and transfer tasks demonstrated overlap, and while we did not find any CFs that were common between training gain and transfer to either PS or GFRT, an interesting pattern of common CFs emerged in the models predicting WM and training gain. Behaviorally, WM transfer is positively associated with gain scores in these three training tasks, but WM transfer shares pre CFs with SR and change CFs with SC, and shares no CFs with SD. Training in these three tasks all tap into working memory in different ways. SR is largely a running span task requiring participants to repeatedly add and remove an item in working memory. SD is a dual n-back task, which taps into both the updating and storage component of working memory, as well as the executive component due to the presence of lures. This task is the most similar to the WM transfer tasks, which are also n-back tasks. SC is an inductive reasoning task, in a fashion similar to the Ravens Advanced Progressive Matrices task. This task requires subjects to take in a large amount of information all at once. Participants must use their short term memory storage ability while quickly and repeatedly comparing distinct pieces of information in the larger set in order to extract patterns, and compare the answer choices on screen to the pattern extracted. The cognitive training and transfer literature suggests that transfer occurs through a neural, cognitive, or behavioral "overlap" between the transfer and training tasks (Jonides 2004; Dahlin, Neely, et al. 2008; Buschkuehl and Jaeggi 2010). Therefore, we expected to see extensive commonalities in the CFs of SD gain and WM transfer, and to a less extent between WM transfer and SR and SC gain. We found the clusters of models that predicted SR and SC had many of the same CFs in sensory and attention processing regions as the clusters of models predicting WM. Surprisingly, we found little overlap between WM transfer and gain in SD, the training task that most resembles the n-back task. These results support a refinement of the neural overlap hypothesis of transfer such that training that shares neural correlates with processing limitations in the transfer tasks will drive improvement in these untrained domains. This suggests that changes predicting SD may reflect changes that were skill specific and did not drive transfer to WM. On the other hand, changes predicting WM transfer may reflect the training's widening of the sensory and attention processing bottlenecks that allowed for WM transfer. For example, given that training gain in SC and SR may be more dependent on changes in rapid and accurate processing of information in memory, rather than improvement in the size of memory capacity primarily targeted by SD, our results suggest that training gain in SR and SC targeted the same low level sensory and attention processing that would generalize to the WM transfer tasks as well.

SR and WM both share a large set of mostly left lateralized Pre TH CFs which have been reliably associated with motor and cognitive training (see Patel 2013 for a review), and that have been implicated in auditory and visual processing and attention. These results suggest that pre-existing differences in cortical thickness in regions involved in low level sensory processing and awareness may play an important role in the outcome of an intervention, and that this effect extends not only to the training-induced gain, but also to the transfer of that training to untrained domains. Previous work has suggested that improving these low level sensory processes may be key to generating transfer to higher cognition in older adults (Mahncke et al. 2006), and the current work suggests that young adults may benefit from this approach as well. Finally, training-induced changes in cortical volume and thickness share a common set of CFs in predictions of gain in SC and WM. Similarly to regions shared between SR and WM, SC and WM share many regions involved in working memory, such as the superior frontal gyrus and left lingual gyrus (Courtney et al. 1996; Rypma and D'Esposito 1999), as well as sensory processing such as the transverse temporal gyrus and right lingual gyrus, which are involved in auditory and visual processing respectively (Paulesu et al. 1993; Macaluso et al. 2000). Rather than sharing regions associated with high level reasoning, such as the dorsolateral prefrontal cortex and superior parietal lobe, these regions may be reflective of changes in structure that are due to improvements in rapid access and comparison of information.

We found that cortical volumes were predictive of transfer to GFRT, but that the inclusion of white matter and subcortical regions doubled not only the average predictive accuracy of the model-cluster, but also the size of the model cluster. Brain volume has been extensively studied in intelligence, and there are several convergent pieces of evidence that suggest that 'bigger is better' concerning brain volumes, but to our knowledge this the first result demonstrating that brain volumes can also predict future transfer to fluid intelligence. We found that individuals with larger right cerebellum white matter were better able to derive improvements in reasoning speed from the training. While the cerebellum used to be considered mostly a motor processing center, more recently it has demonstrated to play a role in working memory and other cognitive functions (Desmond and Fiez 1998), as well as being a central hub of the skill acquisition process (Hikosaka et al. 2002). Furthermore, both the striatum and cerebellum play critical roles in the acquisition of cognitive skills, with previous work demonstrating that greater striatal volumes predict greater skill learning (Erickson et al. 2010). Given that the cerebellum plays an especially important role in the speed of motor processing (Ghez and Fahn 1985), it may be possible that greater cerebellar volumes also improve the capacity of individuals to decrease decision making reaction times with training, and to transfer those speeded response patterns to other tasks. In this way GFRT transfer may be more reflective of improvements in gains in the speed of processing in decision making and pattern recognition, rather than gain in capacity or accuracy in these higher cognitive functions. As participants trained for 30 hours with the six training tasks, it is likely that significant gains were made in the low level processing of a variety of types of information, and our results in WM and PS suggest that speed of processing in low level sensory domains may have been critical across all transfer scores.

Transfer to PS was best predicted by Pre CFs in the right collateral and lingual sulci. This region is important for visual processing and plays a key role in the ventral attention stream (Newman-Norlund et al. 2006; Ilg et al. 2008; Ronsse et al. 2011). Individuals with greater increases in the left opercular inferior frontal gyrus, a region important for attention (Aron et al. 2003; Menon and Uddin 2010), also demonstrated the greatest gains in PS. These results suggest that individuals that benefited from better initial visual processing and for whom the training had more of an impact on attention processing also demonstrated greater transfer. Transfer to WM was best predicted by the pre training and training-induced change in the volume of the left lateral sulcus, the location of the primary auditory cortex. This may suggest that individuals with the smallest primary auditory cortical volume would transfer the most benefit from auditory working memory training to WM, and they may be also the same individuals that demonstrated greatest volumetric increases with training. The initial volume of the right dPCC, a region implicated in attention and a central hub of the DMN, was surprisingly negatively correlated with transfer to WM. Previous work has demonstrated that individuals with MCI and AD have atrophied PCC volumes compared to healthy age-matched controls (Han et al. 2010), and as in the case of the primary auditory cortex, individuals with the smallest PCC may see the most benefit from cognitive training and transfer to WM. Finally, we found that WM was positively correlated with training-induced increases in right transverse collateral sulcus volume, which encompasses both the fusiform and lingual gyri, both of which are critical for visual processing, attention, and object recognition (Macaluso et al. 2000; Weiner and Zilles 2015). Which suggests that individuals for whom the training-induced the greatest increases in the volume of visual processing regions, it also demonstrated the greatest generalization of their training to untrained tasks.

Taken together, these results suggest that training-induced changes to low level sensory (auditory and visual) processing may have allowed for higher quality training and skill acquisition and subsequent greater information flow between cognitive systems, which enabled the skills trained in MF to generalize to untrained tasks. Previous work has suggested that improved low level sensory processing results in improvements in higher cognitive performance through the strengthening of the signal-to-noise ratio of relevant cortical activity (Mahncke et al. 2006). These clearer signals may support the subcortical regions in unitization of skill representation and may thus be beneficial both for the acquisition and transfer of skills. In the face of the sensory degradation that occurs with aging, the brain adapts by lengthening the time window of processing, which not only makes processing slower, but also prevents the accurate representation of spatiotemporally complex signals (Mahncke et al. 2006). In this way, individuals with better sensory processing may benefit from a secondary benefit of faster reaction times, and the current work suggests that these individuals may not only demonstrate greater training-induced benefits to processing speed, but also gains to higher cognitive function such as working memory.

While we identified sets of regions that were reliably associated with learning and transfer across the brain, the novel analysis pipeline we applied may raise a few questions that are important to address. We did not find significant clusters of models in many datasets, suggesting that the majority of model-clusters did not have high predictive accuracy and were sparsely found in model space. There are many reasons that may affect both the percentage of variance our models account for and the proportion of model-space for which our models are significant predictors. For example, ridge regression, is both a linear method and it employs one of the simplest regularization methods, and thus is not able to capture complex interdependencies in the data's covariance structure, as is more common with advanced methods such as support vector machines (SVMs), random forests, and neural networks (Smola and Schölkopf 2004; Hastie et al. 2009). As predictive accuracy was not a high priority in the current project, but rather reliability and interpretability of findings, we decided it was better to pursue a simple linear model that would underperform compared to more advanced models that have been described as both non-transparent and difficult to interpret (Kotsiantis 2007). Freesurfer's longitudinal pipeline has been demonstrated to be sensitive to longitudinal changes, and this pipeline is known to have high precision, and reproducibility (Destrieux et al. 2010; Fischl 2012; Metzler-Baddeley et al. 2016). We chose the Destrieux atlas because of its sensitivity to individual differences in gyral and sulcal folding, which have been demonstrated to be superior to volumetric approaches (Fischl 1995, Van Essen 2005). Nevertheless, most of the literature finding large effects in investigating the relationships between brain structure and skill and motor learning uses voxel based morphometry (VBM), but this method has limitations as well. Multiple reviews of the skill learning literature have discussed methodological problems with the use of VBM in assessing structure-cognition relationships, especially in low reproducibility of effects (Lövdén et al. 2013; Thomas and Baker 2013). It is important to note that the reason we were able to assess the reliability of our models across cross validation loops, many model sizes, and lambda values is because our feature set was not exceedingly large. Using the full dataset from VBM would be far too computationally expensive to have performed the current analysis, therefore while the Destrieux atlas may not have been the most common choice for our data, it was the most practical choice.

2.5 Conclusion

In the current work we demonstrated that individual differences in brain structure, and training-induced changes in these regions contribute significantly to greater gains in training tasks and transfer to untrained domains. We found support and further refinement of the neural overlap hypothesis in the current work, and we demonstrated the utility of reliability-maximizing machine learning in establishing associations between brain structure and learning. Future work investigating the structural correlates of training and transfer should take advantage of multivariate methods, such as voxel-wise PCA, for identifying structural networks of the brain for prediction. Furthermore, it would be highly informative to follow this work with resting state analyses comparing how networks involved in attention, executive function, and skill learning may be involved in the skill learning and generalization process.

Acknowledgements: This work was supported by a Beckman Fellowship to AN. The authors would like to thank Brad Sutton, Tracey Wszalek for technical guidance, and Nancy Dodge, Holly Tracy, and members of the Lifelong Brain and Cognition Laboratory for assistance with performing the experiments.

2.6 References

Anguera J a, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, Kong E, Larraburo Y, Rolle C, Johnston E, Gazzaley a. 2013. Video game training enhances cognitive control in older adults. Nature. 501:97–101.

Anguera J a, Russell C a, Noll DC, Seidler RD. 2007. Neural correlates associated with intermanual transfer of sensorimotor adaptation. Brain Res. 1185:136–151.

Aron AR, Fletcher PC, Bullmore ET, Sahakian BJ, Robbins TW. 2003. Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. Nat Neurosci. 6:115–116.

Awh E, Baylis GC, Cantwell D, Clark VP, Courtney SM, Davidson MC, Digirolamo GJ, Greenwood PM, Haxby J V, Johnson MH, Joseph JS, Knight RT, Luck SJ, Marrocco RT, Motter BC, Nestor PG, Donnell BFO. 2000. The Attentive Brain.

Bahrick HP. 1979. Maintenance of knowledge: Questions about memory we forgot to ask. J Exp Psychol Gen. 108:296–308.

Ball K, Berch DB, Helmers KF, Jobe JB, Leveck MD, Marsiske M, Morris JN, Rebok GW, Smith DM, Tennstedt SL, Unverzagt FW, Willis SL. 2002. Effects of Cognitive Training Interventions With Older Adults. Jorunal Am Med Assoc. 288:2271–2281.

Baniqued PL, Allen CM, Kranz MB, Johnson K, Sipolins A, Dickens C, Ward N, Geyer A, Kramer AF. 2015. Working memory, reasoning, and task switching training: Transfer effects, limitations, and great expectations? PLoS One. 10:1–29.

Bennett G, Seashore H, Wesman A. 1947. Differential aptitude tests, Form A. San Antonio: The Psychological Corporation.

Bjork BRA, Linn MC. 2006. The Science of Learning and the Learning of Science Introducing Desirable Difficulties. Science (80-). 19:6–7.

Boisgueheneuc F Du, Levy R, Volle E, Seassau M, Duffau H, Kinkingnehun S, Samson Y, Zhang S, Dubois B. 2006. Functions of the left superior frontal gyrus in humans: A lesion study. Brain. 129:3315–3328.

Boot WR, Basak C, Erickson KI, Neider M, Simons DJ, Fabiani M, Gratton G, Voss MW, Prakash R, Lee H, Low K a, Kramer AF. 2010. Transfer of skill engendered by complex task training under conditions of variable priority. Acta Psychol (Amst). 135:349–357.

Buitinck L, Louppe G, Blondel M. 2013. API design for machine learning software: experiences from the scikit-learn project. arXiv Prepr arXiv …. 1–15.

Buschkuehl M, Jaeggi SM. 2010. Improving intelligence: A literature review. Swiss Med Wkly. 140:266–272.

Celsis P, Boulanouar K, Doyon B, Ranjeva JP, Berry I, Nespoulous JL, Chollet F. 1999. Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. Neuroimage. 9:135–144.

Colla M, Ende G, Alm B, Deuschle M, Heuser I, Kronenberg G. 2008. Cognitive MR spectroscopy of anterior cingulate cortex in ADHD: elevated choline signal correlates with slowed hit reaction times. J Psychiatr Res. 42:587–595.

Coull JT, Frith CD. 1998. Differential activation of right superior parietal cortex and intraparietal sulcus by spatial and nonspatial attention. Neuroimage. 8:176–187.

Courtney SM, Ungerleider LG, Keil K, Haxby J V. 1996. Object and Spatial Visual Working Memory Activate Separate Neural Systems in Human Cortex. Cereb Cortex. 6:39–49.

Crone EA, Wendelken C, Van Leijenhorst L, Honomichl RD, Christoff K, Bunge SA. 2009.

Neurocognitive development of relational reasoning. Dev Sci. 12:55–66.

Dahlin E, Bäckman L, Neely AS, Nyberg L. 2009. Training of the executive component of working memory: subcortical areas mediate transfer effects. Restor Neurol Neurosci. 27:405–419.

Dahlin E, Neely AS, Larsson A, Bäckman L, Nyberg L. 2008. Transfer of learning after updating training mediated by the striatum. Science. 320:1510–1512.

Dahlin E, Nyberg L, Bäckman L, Neely AS. 2008. Plasticity of executive functioning in young and older adults: immediate training gains, transfer, and long-term maintenance. Psychol Aging. 23:720–730.

Desmond JE, Fiez JA. 1998. Neuroimaging studies of the cerebellum: Language, learning and memory. Trends Cogn Sci. 2:355–362.

Destrieux C, Fischl B, Dale A, Halgren E. 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage. 53:1–15.

Doyon J, Bellec P, Amsel R, Penhune V, Monchi O, Carrier J, Lehéricy S, Benali H. 2009. Contributions of the basal ganglia and functionally related brain structures to motor learning. Behav Brain Res. 199:61–75.

Dye MWG, Green CS, Bavelier D. 2009. Increasing Speed of Processing With Action Video Games. Curr Dir Psychol Sci. 18:321–326.

Ekstrom RBR, French JJW, Harman HH, Dermen D. 1976. Manual for kit of factor-referenced cognitive tests. Princet NJ Educ Test Serv. 102:117.

Erickson KI, Boot WR, Basak C, Neider MB, Prakash RS, Voss MW, Graybiel AM, Simons DJ, Fabiani M, Gratton G, Kramer AF. 2010. Striatal volume predicts level of video game skill acquisition. Cereb Cortex. 20:2522–2530.

Fischl B. 2012. FreeSurfer. Neuroimage. 62:774–781.

Ghez C, Fahn S. 1985. The cerebellum. In: Principles of neural science. p. 502–522.

Graybiel AM. 2005. The basal ganglia: learning new tricks and loving it. Curr Opin Neurobiol. 15:638–644.

Graybiel AM. 2008. Habits, rituals, and the evaluative brain. Annu Rev Neurosci. 31:359–387.

Green CS, Bavelier D. 2003. Action video game modifies visual selective attention. Nature. 423:534–537.

Guyon I. 2003. An Introduction to Variable and Feature Selection 1 Introduction. 3:1157–1182.

Han IL, Young D, Oh JS, Sung J, Soo D, Chan I, Choul J, Gyeom S, Woong K, Hyeong J, Inn J. 2010. Posterior cingulate cortex atrophy and regional cingulum disruption in mild cognitive impairment and Alzheimer ' s disease. Neurobiol Aging. 31:772–779.

Hastie T, Tibshirani R, Friedman J. 2009. The elements of statistical learning.

Hikosaka O, Nakamura K, Sakai K, Nakahara H. 2002. Central mechanisms of motor skill learning. Curr Opin Neurobiol. 12:217–222.

Ilg R, Wohlschlager AM, Gaser C, Liebau Y, Dauner R, Woller A, Zimmer C, Zihl J, Muhlau M. 2008. Gray matter increase induced by practice correlates with task-specific activation: A combined functional and morphometric magnetic resonance Imaging study. J Neurosci. 28:4210–4215.

Jaeggi SM, Buschkuehl M, Etienne A, Ozdoba C, Perrig WJ, Nirkko AC. 2007. On how high performers keep cool brains in situations of cognitive overload. Cogn Affect Behav Neurosci. 7:75–89.

Jaeggi SM, Buschkuehl M, Jonides J, Perrig WJ. 2008. Improving fluid intelligence with training on working memory. Proc Natl Acad Sci U S A. 105:6829–6833.

Jaeggi SM, Seewer R, Nirkko AC, Eckstein D, Schroth G, Groner R, Gutbrod K. 2003. Does excessive memory load attenuate activation in the prefrontal cortex? Load-dependent processing in single and dual tasks: Functional magnetic resonance imaging study. Neuroimage. 19:210–225.

Jonides J. 2004. How does practice makes perfect? Nat Neurosci.

Kane MJ, Conway AR, Miura TK, Colflesh GJ. 2007. Working memory, attention control, and the N-back task: a question of construct validity. J Exp Psychol Learn Mem Cogn. 33:615–622.

Kim DE, Shin MJ, Lee KM, Chu K, Woo SH, Kim YR, Song EC, Lee JW, Park SH, Roh JK. 2004. Musical training-induced functional reorganization of the adult brain: Functional magnetic resonance imaging and transcranial magnetic stimulation study on amateur string players. Hum Brain Mapp. 23:188.

Kirchner WK. 1958. Age differences in short-term retention of rapidly changing information. J Exp Psychol. 55:352–358.

Klingberg T. 2010. Training and plasticity of working memory. Trends Cogn Sci. 14:317–324.

Landauer TK, Bjork RA. 1978. Optimum rehearsal patterns and name learning. Pract Asp Mem. 625–632.

Lövdén M, Wenger E, Mårtensson J, Lindenberger U, Bäckman L. 2013. Structural brain plasticity in adult learning and development. Neurosci Biobehav Rev. 37:2296–2310.

Macaluso E, Frith C, Driver J. 2000. Modulation of Human Visual Cortex by Crossmodal Spatial Attention. Science (80-). 289:1206–1208.

Mahncke HW, Bronstone A, Merzenich MM. 2006. Brain plasticity and functional losses in the aged : scientific bases for a novel intervention. Brain. 157:81–109.

Melby-Lervåg M, Hulme C. 2012. Is Working Memory Training Effective? A Meta-Analytic Review. Dev Psychol.

Menon V, Uddin LQ. 2010. Saliency, switching, attention and control: a network model of insula function. Brain Struct Funct. 214:1–13.

Metzler-Baddeley C, Caeyenberghs K, Foley S, Jones DK. 2016. Task complexity and location specific changes of cortical thickness in executive and salience networks after working memory training. Neuroimage. 130:48–62.

Nebel K, Wiese H, Stude P, de Greiff A, Diener H-C, Keidel M. 2005. On the neural basis of focused and divided attention. Brain Res Cogn Brain Res. 25:760–776.

Newman-Norlund RD, Frey SH, Petitto L-A, Grafton ST. 2006. Anatomical substrates of visual and auditory miniature second-language learning. J Cogn Neurosci. 18:1984–1997.

Nikolaidis A, Goatz D, Smaragdis P, Kramer A. 2015. Predicting Skill-Based Task Performance and Learning with fMRI Motor and Subcortical Network Connectivity. 2015 Int Work Pattern Recognit NeuroImaging. 93–96.

Nikolaidis A, Voss MW, Lee H, Vo LTK, Kramer AF. 2014. Parietal plasticity after training with a complex video game is associated with individual differences in improvements in an untrained working memory task. Front Hum Neurosci. 8:1–11.

Oelhafen S, Nikolaidis A, Padovani T, Blaser D, Koenig T, Perrig WJ. 2013. Increased parietal activity after training of interference control. Neuropsychologia. 51:2781–2790.

Olesen PJ, Westerberg H, Klingberg T. 2004. Increased prefrontal and parietal activity after training of working memory. Nat Neurosci. 7:75–79.

Patel R, Spreng RN, Turner GR. 2013. Functional brain changes following cognitive and motor skills training: a quantitative meta-analysis. Neurorehabil Neural Repair. 27:187–199.

Paulesu E, Frith CD, Frackowiak RS. 1993. The neural correlates of the verbal component of working memory. Nature. 362:342–345.

Penhune VB, Doyon J. 2002. Dynamic cortical and subcortical networks in learning and delayed recall of timed motor sequences. J Neurosci. 22:1397–1406.

Puttemans V. 2005. Changes in Brain Activation during the Acquisition of a Multifrequency Bimanual Coordination Task: From the Cognitive Stage to Advanced Levels of Automaticity. J Neurosci. 25:4270–4278.

Raboyeau G, Marcotte K, Adrover-Roig D, Ansaldo AI. 2010. Brain activation and lexical learning: The impact of learning phase and word type. Neuroimage. 49:2850–2861.

Ravens J. 1962. Advanced Progressive Matrices: Set II. London.

Rebok GW, Ball K, Guey LT, Jones RN, Kim HY, King JW, Marsiske M, Morris JN, Tennstedt SL, Unverzagt FW, Willis SL. 2014. Ten-year effects of the advanced cognitive training for independent and vital elderly cognitive training trial on cognition and everyday functioning in older adults. J Am Geriatr Soc. 62:16–24.

Redick TS, Shipstead Z, Harrison TL, Hicks KL, Fried DE, Hambrick DZ, Kane MJ, Engle RW. 2012. No Evidence of Intelligence Improvement After Working Memory Training: A Randomized, Placebo-Controlled Study. J Exp Psychol Gen. 142:359–379.

Ronsse R, Puttemans V, Coxon JP, Goble DJ, Wagemans J, Wenderoth N, Swinnen SP. 2011. Motor learning with augmented feedback: Modality-dependent behavioral and neural consequences. Cereb Cortex. 21:1283–1294.

Rypma B, D'Esposito M. 1999. The roles of prefrontal brain regions in components of working memory: effects of memory load and individual differences. Proc Natl Acad Sci U S A. 96:6558–6563.

Rypma B, Prabhakaran V, Desmond JE, Glover GH, Gabrieli JD. 1999. Load-dependent roles of frontal brain regions in the maintenance of working memory. Neuroimage. 9:216–226.

Salthouse TA. 2005. Relations between cognitive abilities and measures of executive functioning. Neuropsychology. 19:532–545.

Salthouse TA, Babcock RL. 1991. Decomposing adult age differences in working memory. Dev Psychol. 27:763–776.

Schmidt R a., Bjork R a., Schmidt BRA, Bjork R a., Schmidt R a., Bjork R a. 1992. New Conceptualizations of Practice: Common Principles in Three Paradigms Suggest New Concepts for Training. Psychol Sci. 3:207–217.

Schmidt-Wilcke T, Rosengarth K, Luerding R, Bogdahn U, Greenlee MW. 2010. Distinct patterns of functional and structural neuroplasticity associated with learning Morse code. Neuroimage. 51:1234–1241.

Seidler RD. 2010. Neural correlates of motor learning, transfer of learning, and learning to learn. Exerc Sport Sci Rev. 38:13.

Shipstead Z, Redick TS, Engle RW. 2012. Is working memory training effective? Psychol Bull. 138:628–654.

Silk TJ, Bellgrove MA, Wrafter P, Mattingley JB, Cunnington R. 2010. Spatial working memory and spatial attention rely on common neural processes in the intraparietal sulcus. Neuroimage. 53:718–724.

Singley MK, Anderson JR. 1989. The transfer of cognitive skill. Harvard University Press.

Smola AJ, Schölkopf B. 2004. A tutorial on support vector regression. Stat Comput. 14:199–222.

Suzuki WA. 2007. Integrating Associative Learning Signals Across the Brain. 850:842–850.

Thomas C, Baker CI. 2013. Teaching an adult brain new tricks: A critical review of evidence for training-dependent structural plasticity in humans. Neuroimage. 73:225–236.

Thorell LB, Lindqvist S, Bergman Nutley S, Bohlin G, Klingberg T. 2009. Training and transfer effects of executive functions in preschool children. Dev Sci. 12:106–113.

Vo LTK, Walther DB, Kramer AF, Erickson KI, Boot WR, Michelle W, Prakash RS, Lee H, Fabiani M, Gratton G, Simons DJ, Sutton BP, Wang MY. 2011. Predicting individuals' learning success from patterns of pre-learning MRI activity. PLoS One. 6:e16093.

Vogt B a, Finch DM, Olson CR. 1992. Functional heterogeneity in cingulate cortex: the anterior executive and posterior evaluative regions. Cereb Cortex. 2:435–443.

Voss MW, Prakash RS, Erickson KI, Boot WR, Basak C, Neider MB, Simons DJ, Fabiani M, Gratton G, Kramer AF. 2011. Effects of training strategies implemented in a complex videogame on functional connectivity of attentional networks. Neuroimage.

Wechsler D. 1997. WAIS-III: Wechsler adult intelligence scale. San Antonio: The Psychological Corporation.

Weiner KS, Zilles K. 2015. The anatomical and functional specialization of the fusiform gyrus. Neuropsychologia. 83:48–62.

Willis SL, Schaie KW. 1986. Trianing the elderly on the ability factors of spatial orientationand inductive reasoning. Psychol Aging.

Wong PCM, Perrachione TK, Parrish TB. 2007. Neural characteristics of successful and less successful speech and word learning in adults. Hum Brain Mapp. 28:995–1006.

Zhang H, Xu L, Wang S, Xie B, Guo J, Long Z, Yao L. 2011. Behavioral improvements and brain functional alterations by motor imagery training. Brain Res. 1407:38–46.

Chapter 3

Machine Learning Predictions of Cognitive Skill Training and Transfer to Working Memory, Fluid Intelligence, and Perceptual Speed Using Resting State Functional Connectivity

Aki Nikolaidis[1], Drew Goatz[1], Pauline Baniqued[2], Michael Kranz[1, 5], Aron Barbey[1, 5], Paris Smaragdis[1, 3,4], Arthur F. Kramer[1, 5]

Beckman Institute[1], Department of Computer Science[3], Department of Electrical and Computer Engineering[4],
Department of Psychology[5]
University of Illinois. Urbana Champaign
Urbana, Illinois. USA

Helen Wills Neuroscience Institute[2]
University of California. Berkeley
Berkeley, California. USA

Abstract: Cognitive training studies have demonstrated transfer of training to underlying cognitive abilities, but many attempts to replicate these effects have failed. Understanding individual differences in the brain networks that contribute to training and transfer is essential to developing a mechanistic understanding of these learning phenomena. We trained 45 young adults with Mind Frontiers, a multimodal training game that taps complimentary aspects of reasoning, attention, and working memory. We predicted gains in training and untrained transfer domains using both pre-existing and training-induced differences in the Default Mode, Cingulo-Opercular, Frontal-Parietal, and Subcortical networks. We utilize a machine learning pipeline that we designed to extract the most reliably predictive models, network connections, and nodes. We find that networks involved in the development and deployment of sustained attention and regulation of internal cognition are reliable predictors of both skill acquisition and generalization of Mind Frontiers training to fluid intelligence, working memory, and perceptual speed.

3.1 Introduction

Cognitive training holds a great deal of potential for offering an ecologically valid and healthy approach to both remediating symptoms of many psychiatric disorders and enhancing cognition in healthy participants. Recent attempts to improve cognitive function through training have been successful (Ball et al. 2002; Green and Bavelier 2003; Dahlin, Neely, et al. 2008; Jaeggi et al. 2008; Dye et al. 2009a; Thorell et al. 2009; Anguera et al. 2013; Rebok et al. 2014; Baniqued et al. 2015), but there have also been large scale failures to replicate many of these effects (Melby-Lervåg and Hulme 2012; Redick et al. 2012; Shipstead et al. 2012; Oelhafen et al. 2013). The low reliability of these effects may be due in part to methodological constraints such as small sample sizes, estimation of enhancement through the use of single tasks, or the use of single training tasks. The investigation of how individual differences are associated with positive training outcomes may offer important insight into the mechanisms by which improvements during training transfer to other domains.

Some previous work has investigated how brain structure and function relate to individual differences in training and transfer (Olesen et al. 2004; Anguera et al. 2007, 2013; Dahlin, Neely, et al. 2008; Vo et al. 2011; Voss et al. 2011; Oelhafen et al. 2013; Nikolaidis et al. 2014, 2015), but the neural correlates of transfer have not yet been well established. In a closely related field, skill learning, neural correlates have been increasingly well defined in both animals and humans (Hikosaka et al. 2002; Graybiel 2005; Doyon et al. 2009; Patel et al. 2013). To gain a fuller mechanistic understanding of the process by which training transfers to new domains, it will be critical to establish a common theory that encompasses the learning that occurs in both skill learning and cognitive training. Therefore, the critical effort in cognitive training may be reframed as establishing both the relationships and distinctions between the neural correlates of skill acquisition and skill generalization. Previous work has demonstrated that skill acquisition and generalization are closely linked, and modifications that impact skill acquisition can have impacts on generalization of learning as well (Singley and Anderson 1989; Schmidt et al. 1992; Bjork and Linn 2006; Boot et al. 2010; Seidler 2010). Creating targeted cognitive training interventions that utilize mechanisms of training and transfer will depend on the characterization of the neural similarities and distinctions between the process of skill acquisition and generalization.

In order to explore the relationship between training and transfer, we recruited 45 subjects for a cognitive training study in which they performed pre and post neuropsychological assessments, resting state fMRI (rsfMRI), and 30 hours of training with Mind Frontiers. Mind Frontiers (MF) is a cognitive training study with six different training games that target complementary aspects of executive function, reasoning, working memory, and attention (Baniqued et al. 2015). The training comprises six adaptive training tasks that have been gamified and united into a common "Wild West" themed game, and which have all shown previous success in cognitive training. We recently published the behavioral results from this study, in which we found that MF training led to significant improvements in constructs of perceptual speed, fluid intelligence, and working memory. In the current study, we investigate how pre-training resting state network connectivity predicts training gain and transfer. By examining how resting state connectivity in specific networks predicts training gain and transfer, we may gain a better understanding of not only the mechanisms by which these behavioral effects occur, but also how training and transfer are affected by common neural mechanisms. We will also assess how training-induced changes in these resting state networks predict both training and transfer, which will allow us to gain a better understanding of the path through which training drives common changes in both resting state connectivity and transfer.

The executive, attention, and reasoning cognitive functions critical for the performance of the training tasks in MF are known to depend critically on the frontal parietal (FPN) and cingulo-opercular networks (CON) (Kelly et al. 2008; Uddin et al. 2009; Sestieri et al. 2014). These networks are not only important for the performance of these cognitive processes, but they also may be important for the acquisition of skills that largely depend on these abilities, and by the same token may be involved in their training-induced improvement. For example, previous work has demonstrated that greater training-induced synchrony between frontal and parietal regions was associated with greater transfer to attention and multitasking (Anguera et al. 2013). Furthermore, while it has been well established that subcortical regions, such as the striatum, are critical for the acquisition of skills (Hikosaka et al. 2002; Graybiel 2005; Erickson et al. 2010), these regions have also been established as critical for the generalization of trained skills to untrained domains (Dahlin, Neely, et al. 2008), providing support for the notion that training-induced transfer is related to skill acquisition and generalization. In order to examine how these networks may play a role in both the acquisition and generalization of trained skills, we assessed functional network connectivity in the FPN, CON, and subcortical (SUB) networks. Motor skills are known to be acquired through the network interaction of the motor and subcortical regions (Hikosaka et al. 2002), and given that nearly all of cortex projects to the striatum (Postuma 2005), it is possible that higher cognitive skills are acquired through a similar pattern of interaction with these subcortical regions. To assess how this may play a role in driving improvements in training tasks and transfer to other domains, we assessed functional connectivity between the SUB and FPN networks (SUBxFPN), as well as between the SUB and CON networks (SUBxCON). Finally, the functional scans we collected were during the resting state, in which the default mode network (DMN) is dominant. Prior work has established a close anti-correlated relationship between the

DMN and the executive control CON and FPN networks (Fox et al. 2005). In fact, less distinction between these networks has been associated with a wide range of cognitive symptomology, ranging from Alzheimer's disease (Brier et al. 2012) and aging (Sambataro et al. 2010; Geerligs et al. 2015), to schizophrenia (Tu et al. 2012) and poor performance in higher cognitive functions (Zanto and Gazzaley 2013). Therefore, it may be the case that the connectivity of the DMN and its associated control networks (CON and FPN) may reflect on the performance of these higher cognitive functions and their responsiveness to training-induced improvement. As in the case of the SUB network, we also assessed functional connectivity in the DMN, as well as functional connectivity between the DMN and FPN (DMNxFPN), and CON (DMNxCON) networks. We use pre-training connectivity in these four networks and four internetworks, as well as training-induced changes in these networks, to predict improvement in three cognitive training tasks and transfer to perceptual speed, fluid intelligence, and working memory.

We employ common techniques from machine learning to achieve these predictions. Machine learning methods, such as the ridge regression we apply, are more likely to develop models that reproduce their effects in new data compared to ordinary least squares (OLS), which is used in most general linear model (GLM) based analyses. OLS suffers from problems of low reproducibility because it exclusively optimizes reducing squared error, which increases the likelihood of creating models that overfit a given sample, and therefore tend to not replicate well to other samples (Hastie et al. 2009). The tendency of OLS to overfit is especially problematic in the case of small samples, or high collinearity between variables, two characteristics that are commonplace in cognitive neuroscience research. Ridge regression addresses this problem by penalizing the model parameters, which prevents a model from making strong assumptions on the importance of a variable for the sake of improving within-sample predictive accuracy, thereby reducing chance for overfitting (Marquardt and Snee 1975). Aside from ridge regression, many other methods, such as support vector machines, lasso, and random forests have the same advantage that they produce models that are less likely to overfit compared to OLS (Hastie et al. 2009), although these methods are more computationally intensive. The statistical advantages offered by these techniques, in concert with increases in computing power has led to a surge in the use of these methods in many scientific fields, and more recently these methods have begun to be applied in neuroscience as well (Varoquaux and Thirion 2010; Hinton 2011). Recent work has investigated how functional network connectivity can be used to detect strokes (Varoquaux and Thirion 2010) and disease states (Craddock et al. 2009), predict skill performance and learning rate (Nikolaidis et al. 2015), and predict brain-age and chronological age (Dosenbach et al. 2011; Meier et al. 2012). In the current study, we apply such techniques to the question of cognitive training and make methodological advancements that enhance the ability of these methods to make reliable predictions.

We compare how predictors of training and transfer are shared across four resting state networks and the connectivity between these networks. Many applications in machine learning focus on selecting model parameters to maximize accuracy, but in the current work, our goal was to maximize the reliability and interpretability of our findings. We applied ridge regression across a wide range of model parameters: lambda regularization values and feature sets to map the predictive ability of our dataset over the entire model space. We applied a cluster correction algorithm to this space to identify regions of model space where clusters of models were reliable predictors of our outcome variables. This enabled us to search for consensus features, which are those that are included not only in every cross validation loop in a given point in model space, but also across all models within the cluster of significant models. By identifying consensus features, we were able to find features that were highly reliable predictors of our outcome variable, and remained predictive even with varied feature sets and regularization values. We found that the CON and its regulation of the DMN were the most reliable and significant predictors of both skill acquisition and generalization. Critically, we found that both the SUB, which is involved in skill learning, and the SUBxCON network, which is involved in the development and deployment of the sustained attention skills of the CON, was predictive of transfer to fluid intelligence, working memory, and perceptual speed. We discuss these and other findings in the greater context of learning, skill acquisition, and cognitive training.

3.2 Methods

3.2.1 Participants

We recruited participants from the University of Illinois and surrounding community through flyers and online advertisements seeking participants for a “cognitive training study.” Online pre-screening for demographic information and videogame experience was completed over email. Subsequent phone screening required participants: 1) between 18 and 30 years old, 2) 75% right handed according to the Edinburgh Handedness Inventory, 3) normal or corrected-to-normal hearing and vision, 4) no major psychological or medical conditions, 5) no non-removable metal in the body, and 6) played no more than 5 hours of video games per week in the past six months. A total of 45 participants trained with Mind Frontiers and had complete behavioral and neural data.

3.2.2 Study Design

A complete description of the study design can be found in Baniqued et al. 2015. Briefly, participants completed three initial cognitive testing sessions, and an MRI session before beginning the intervention, in which they completed four to five training sessions per week for four to five weeks, for a total of 20 sessions. After training, subjects returned to lab for a final MRI session and another round of cognitive testing.

Participants trained with six computerized adaptive training tasks that were selected based on their known psychometric properties with reasoning/Gf, working memory, visuospatial reasoning, inductive reasoning, and task switching. See Table 3.1 for a description of each training task.

3.2.3 Cognitive Assessment Protocol

For a full listing of all neuropsychological tasks administered see Baniqued et al. 2015. Here we briefly describe each of the tasks that were included to create the reasoning, working memory, and perceptual speed composite variables we predicted. The tests below were obtained from the Virginia Cognitive Aging Project Battery (Salthouse 2005), and two different versions were used for pre- and post-testing, with the sequence counterbalanced across subjects.

3.2.3.1 GFRT

The GFRT composite was comprised of standardized reaction times in these four tasks. Descriptions for these and other tasks in section 2.3 are reproduced from Baniqued et al. 2015.

Matrix Reasoning: Select the pattern that completes a missing space on a 3 x 3 grid. Number of correctly answered items is the primary measure. Reaction time on correct trials was also analyzed. Source: Ravens 1962; Crone et al. 2009.

Paper Folding: Identify pattern of holes that results from a punch through folded paper. Number of correctly answered items is the primary measure. Reaction time on correct trials was also analyzed. Source: Ekstrom et al. 1976.

Spatial Relations: Identify 3D object that would match a 2D object when folded. Number of correctly answered items is the primary measure. Reaction time on correct trials was also analyzed. Source: Bennett et al. 1947.

Letter Sets: Determine which letter set is different from the other four. Number of correctly answered items is the primary measure. Reaction time on correct trials was also analyzed. Source: Ekstrom et al. 1976.

3.2.3.2 Perceptual Speed

The Perceptual Speed composite was comprised of standardized accuracy scores in these three tasks.

Digit Symbol Substitution: Write corresponding symbol for each digit using a coding table. The primary measure is the number of correctly answered items within two minutes. Source: Wechsler 1997.

Pattern Comparison: Determine whether pairs of line patterns are the same or different. The primary measure is the number of correctly answered items within 30 seconds, averaged across two sets of problems. Source: Salthouse and Babcock 1991.

Letter Comparison: Determine whether pairs of letter strings are the same or different. The primary measure is the number of correctly answered items within 30 seconds, averaged across two sets of problems. Source: Salthouse and Babcock 1991.

3.2.3.3 Working memory

The WM N-back composite was comprised of the d' scores of the 2-back, 3-back, Dual 2-back, and Dual 3-back.

Single N-back: Determine whether the current letter presented matches the letter presented two or three items back. The primary measure of d-prime was computed separately for the 2-back and 3-back conditions. Reaction times on correct trials were also analyzed. Source: Kirchner 1958; Kane et al. 2007.

Dual N-back (administered in the MRI): Determine whether simultaneously presented auditory and visual stimuli match stimuli presented one, two, or three items ago. The primary measure of d-prime was computed separately for the two-back and three-back conditions. Reaction times on correct trials were also analyzed. Source: Jaeggi et al. 2003, 2007.

3.2.4 Behavioral Results Analysis

Training with Mind Frontiers produced significant gains in performance not only across all training tasks, but also to three cognitive constructs, working memory (WM), reasoning reaction time (GFRT), and perceptual speed (PS). The gain scores for the training and transfer tasks were calculated by taking the difference between post- and pre-test scores, and dividing by the standard deviation of the pre-test score. In the current study, we use these transfer scores as predictors, as well as training-induced gain scores from three of the training games: Supply Run (SR), Sentry Duty (SD), and Safe Cracker (SC). These three training games were chosen because improvement during training was positively correlated with transfer to the working memory construct (Baniqued 2015), indicating that training-induced gains in these tasks may play a role in WM transfer.

3.2.5 MRI Preprocessing & Analysis

3.2.5.1 MPRAGE Acquisition and Processing

All structural, high-resolution, T1-weighted MPRAGE images were collected on a 3 T Siemens Trio scanner, collected in ascending order parallel to the anterior and posterior commissures with the following parameters: Echo time (TE) = 2.32 ms, repetition time (TR) = 1900 ms, field of view (FOV) = 230 mm, voxel size 0.9mm^3, and flip angle = 8°. All preprocessing of structural images was performed with FSL. After being skull stripped, each structural image was linearly registered to the MNI 2mm brain template with 12 degrees of freedom (Smith 2004).

3.2.5.2 RS Acquisition and Processing

For all functional scans, we used an echo-planar imaging (EPI) sequence with Blood Oxygenation Level Dependent (BOLD) contrast (FoV Read: 220, 2.4 × 2.4 × 3.5 mm voxel size, TR = 2000 ms, TE = 25 ms, and flip angle = 80). BOLD repetitions: 180 volumes.

Preprocessing of the functional data was performed with C-PAC (http://fcp-indi.github.io/), which is an open source, configurable, and automated processing pipeline for resting functional MRI. This package utilizes tools from a variety of publicly available packages, such as FSL, AFNI, and ANTS. After being skull stripped with FSL, each structural image was registered to the MNI 2mm brain template and functional images were registered to the anatomical image using ANTS. These transformations were combined to register the functional images to the MNI 2mm brain template. Slice timing and motion correction using Friston's 24 parameter motion model were applied to the functional data to remove effects of motion from the dataset. To account for motion related artifacts further, we also performed motion scrubbing such that TRs with excessive motion (greater than 0.5mm frame displacement) were removed, along with 2 subsequent TRs. We excluded subjects with less than 120 volumes.

We performed nuisance signal regression by regressing out mean signal of the white matter and CSF. We also used CompCorr, with 4 parameters, WM, CSF, and global as a method to remove excessive noise due to signal from the white matter, CSF, and global signal. Finally, we applied a temporal band pass filter from 0.01 to 0.1 Hz, and a spatial smoothing kernel of 6mm FWHM.

3.2.5.3 Connectivity

We chose the nodes for the default mode (DMN), frontal parietal (FPN), cingulo-opercular (CON), and subcortical (SUB) networks from the Petersen 264 node network atlas (Power *et al.* 2011). The CON contained nodes in the supramarginal gyrus, superior frontal gyrus, insular cortex, central opercular cortex, cingulate gyrus, and paracingulate gyrus. The DMN was made up of nodes in the frontal pole, middle frontal gyrus, PCC, lateral occipital cortex, temporal pole, precuneous, superior lateral occipital cortex, middle temporal gyrus, and angular gyrus. The FPN contained nodes in the precentral gyrus, middle frontal gyrus, posterior supramarginal gyrus, inferior temporal gyrus, frontal pole, superior parietal lobe, and superior lateral occipital cortex. The SUB network contained nodes in the thalamus, brain stem, putamen, and pallidum. For a full listing of network regions and their anatomical assignments, see (Power *et al.* 2011).

To create the internetworks, we assessed only the connections between each of the networks and excluded all within network connections. For example, for the SUBxFPN network, we exclusively extracted correlations between the SUB and FPN, and excluded all intranetwork connections. To extract and process each network's connectivity information, we calculated node-averaged timecourses within each subject's network of interest and we used these timecourses to create correlation matrices that would become features for our machine learning pipeline.

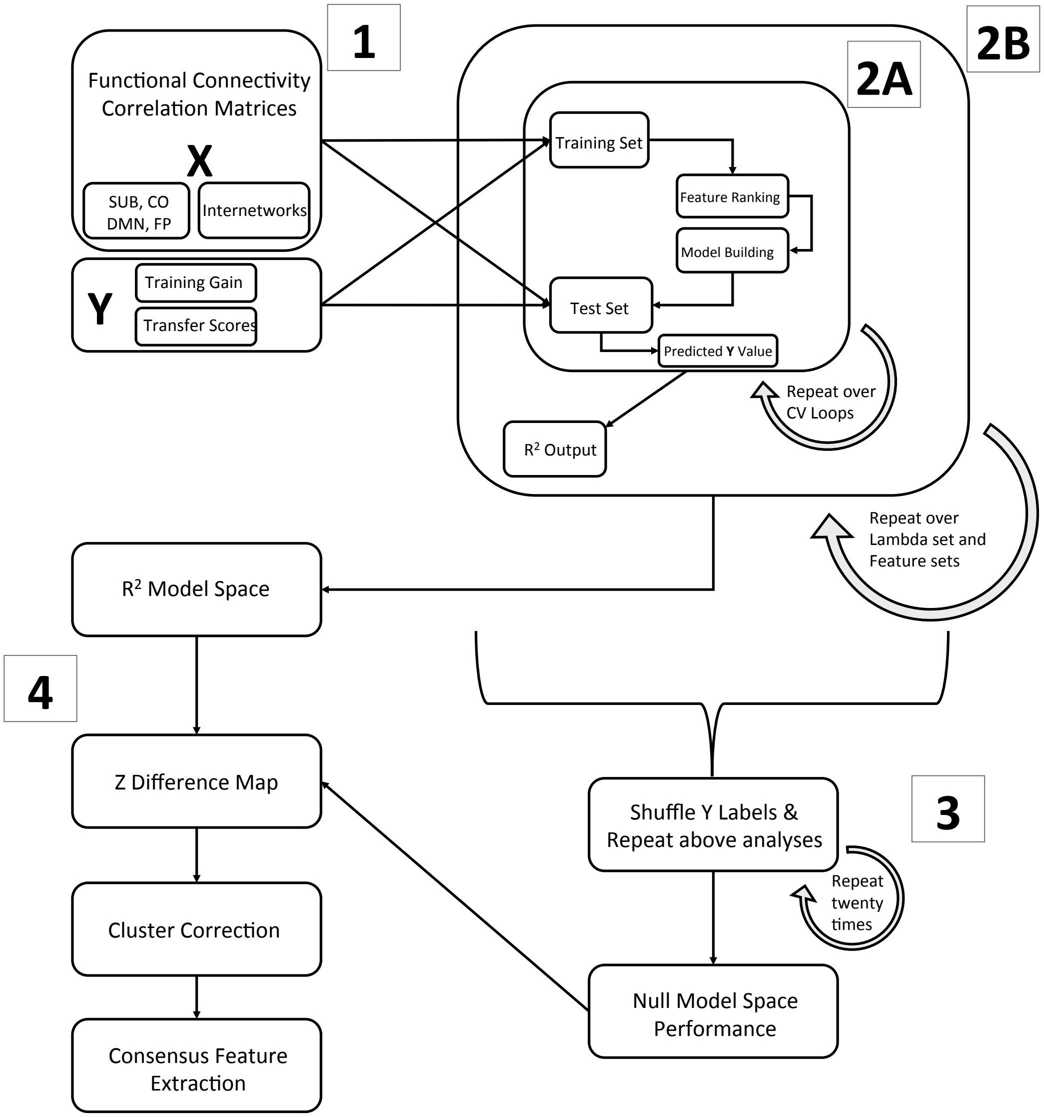

Figure 3.1:
This figure provides a visual flowchart of the methods used in the current work. Analysis begins with the structural features extracted from our resting state scans (1), and then it continues through model building (2A), cross validation (2A), and repeating our analysis across feature and lambda space (2B). Finally we create the null model distribution by shuffling the Y labels (3), and we assess the difference between the model space and the null model space in the form of a Z-map (4). We apply a cluster correction to this Z map and extract consensus features (4).

3.2.6 Machine Learning Methods

3.2.6.1 Predicting training gain and transfer

To make our predictions we used either pre-training or pre-post differences in network connectivity. We used Sci-Kit Learn to perform ridge regression (Figure 3.1: Section 1) (Buitinck et al. 2013). While this algorithm does not achieve predictive accuracies as high as other more advanced methods, such as random forests and support vector machines (SVMs), ridge regression offers an important advantage in terms of ease of interpretability and comparability to other analyses in this domain using linear models (Kotsiantis, Zaharakis, & Pintelas, 2007).

Over each leave-one-subject-out cross-validation (LOSO-CV) loop, we trained models of a range of sizes, and we used these models to predict scores of the left-out subject (Figure 3.1: Section 2A). Within each cross validation loop, we ranked all features in the training group by absolute correlation strength. We created feature sets of the top 1 to *n* features, by groups of 5 to save processing time, and predicted each left out subject using each of these feature sets. We calculated R^2 values by squaring the correlation of predicted and actual score values.

In order to assess the true predictive accuracy and overfitting in our models, we created a distribution of null models (Figure 3.1: Section 3). To create these null models, we randomized the score values for the subjects, and repeated the procedure outlined above to predict these values. We repeated this procedure 20 times to create a distribution of null models. The performance accuracy of our models was compared to the mean performance of our null models.

3.2.6.2 Assessing Reliability of Prediction

Rather than using cross validation to select an optimal hyperparameter (lambda) value for each model, we repeated our predictions and null model testing at a range of hyperparameter values (Figure 3.1: Section 2B). While computationally intensive, this method allows us to more fully explore the space of possible models, extract clusters of predictive models, and assess the reliability of our models independent of lambda value.

We assessed how our model performed over the full model space compared to the distribution of null models, and we transformed the difference in prediction accuracy between the model and null model distribution to a Z score distribution, representing the number of standard deviations that the model outperforms chance (Figure 3.1: Section 4). To get a more accurate assessment of average model prediction in a region and account for variation in neighboring models, we first smoothed the data using a 2 x 3 FWHM Gaussian kernel before applying cluster correction. We then removed all models that predicted less than 1 standard deviation above the null distribution to remove weakly predicting models, and then to remove small noisy predicting clusters we removed all clusters of models with less than 12 models. Finally, in order to only assess datasets that were reliable predictors of the outcome variable, we applied a threshold that the clusters in each dataset must cover at least 2% of the total model space for the clusters to be investigated further. We performed cluster correction to not only assess the relative predictive accuracy and reliability of each dataset across a range of models, but also to extract individual features that were reliably included in 90-100% of models in a given cluster, also known as consensus features. In this way, we extract only the features that most reliably belong to models that significantly predict our outcome variable. This is an important step because it allows us to look for common individual predictors across outcome variables. While a mass univariate analysis, such as the feature ranking commonly performed in machine learning pipelines (including this one), detects individual variables that are each correlated with the outcome variable, it does not offer any information on conditional relationships between each feature and the outcome variable, thereby identifying mostly groups of redundant variables (Guyon 2003). By exploring the model space more fully, and identifying sets of consensus features that are the most reliable predictors of our outcome variable, we are also able to identify features that may not be good predictors on their own, but that reliably predict the residual variance that is left after the strongest predictors have been included in the model. For a review on feature selection and how redundant and superficially 'uninformative' features can play an important role in predicting residual variance, see Guyon 2003.

3.2.6.3 Assessing Consensus Features

Consensus features allow us to analyze the feature selection process and make inferences based on the features that are selected. We generate a cluster-level table that describes the consensus features found in at least 90% of the models of a given cluster (Tables 3.4 & 3.5). Among consensus features, some of the nodes are dramatically over-represented, suggesting that edges leading to these nodes may be more important for prediction than others. We term these high degree nodes 'consensus hubs'. Similar to hubs in the graph theoretical sense, these consensus hubs have a strong influence on the rest of the consensus features.

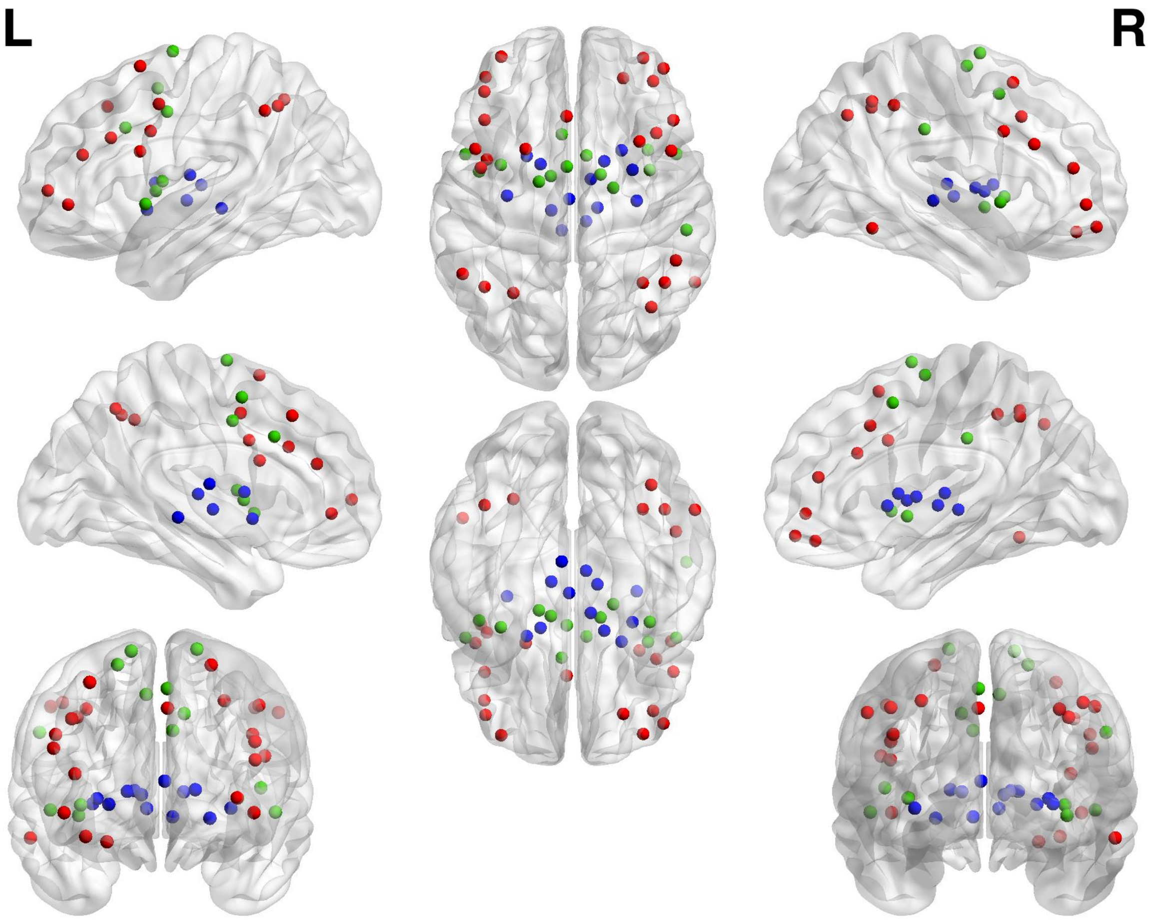

Figure 3.2:
Display of the approximate position of the nodes in 2mm MNI Template Space. SUB (Blue), CON (Green), and FPN (Red). From left to right- The 3 top row views are: Lateral view of Left Hemisphere, Top side view, Lateral View Right Hemisphere. The 3 mid row views are: Medial Left Hemisphere, Bottom side view, Medial Right Hemisphere, and the 2 bottom row views are: Front Side view, Back Side view

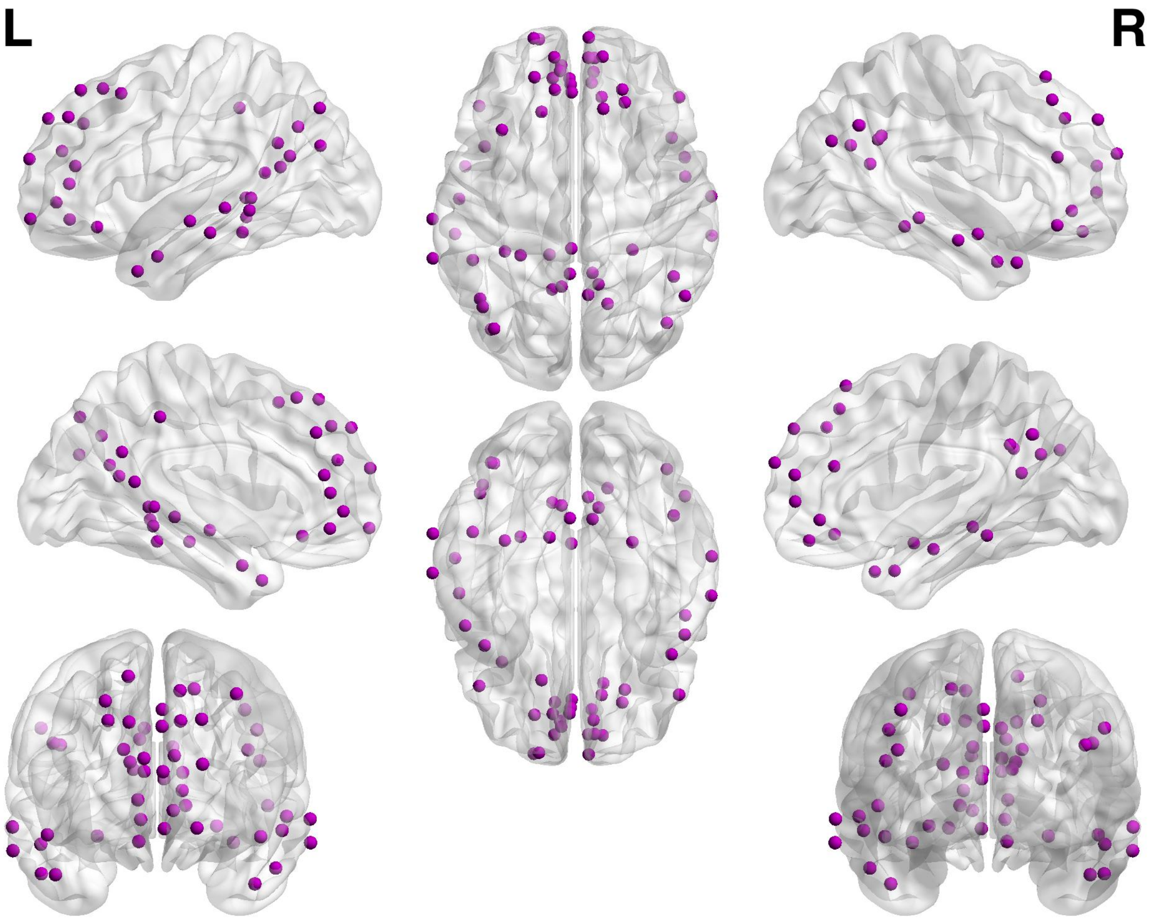

Figure 3.3:
Displays the approximate position and distribution of the Default Mode Network in 2mm MNI Template Space. From left to right- The 3 top row views are: Lateral view of Left Hemisphere, Top side view, Lateral View Right Hemisphere. The 3 mid row views are: Medial Left Hemisphere, Bottom side view, Medial Right Hemisphere, and the 2 bottom row views are: Front Side view, Back Side view

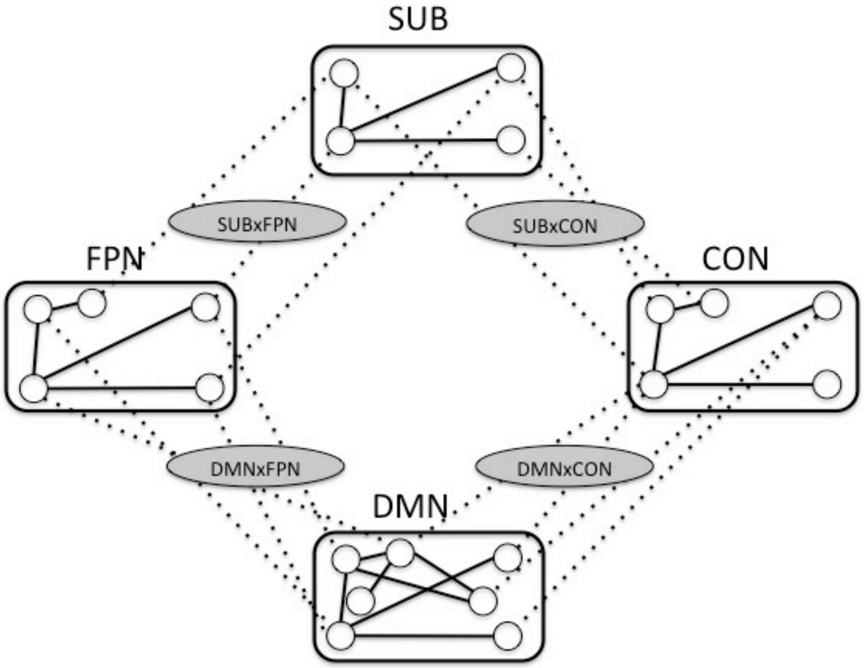

Figure 3.4:
This displays a summary of all eight networks investigated in the current study. Each rounded box corresponds to a single network, which are depicted in Figures 3.2 & 3.3. As shown here, relationships between single networks are assessed to create internetworks, which are represented by the dashed lines between rounded boxes, and the gray ovals label their corresponding internetwork.

Game	Description	Source
Supply Run (working memory)	Townspeople request items that belong to a certain category. There are five objects in each category, which correspond to stereotypical occupations of the "Wild West." Once the store is reached, the last item from each category must be selected. Difficulty level is manipulated by the number of requests and the number of categories.	Updating WM (Dahlin, Nyberg, et al. 2008)
Sentry Duty (working memory)	Sentries lift their lanterns while saying a word of the phonetic alphabet. The current word spoken and lantern lifted is compared to the word spoken/lantern lifted *n* times previously. There may be an audio match, a visual match, an audio and visual match, or no match between the current sentry and the one who spoke *n* times ago. Difficulty is manipulated by how far back the comparison is (e.g., 1-back, 2-back, 3-back).	Dual n-back(Jaeggi et al. 2008; Redick et al. 2012)
Safe Cracker (reasoning)	Safe combinations are determined by completing the next item in a series. Series may be letter-, number-, or day/month- based, and all are governed by some pattern or rule that must be determined and applied to select the next item in the series. Difficulty is manipulated by the difficulty of the patterns and the number of problems to solve within the given time limit.	Inductive Reasoning (Willis and Schaie 1986)

Table 3.1:
Description of the three training tasks investigated in the current study. Partially reproduced from Baniqued et al. 2015.

3.3 Results

3.3.1 Predicting Training and Transfer

We found that transfer to perceptual speed (PS), was best predicted by Pre CON, with both the greatest reliability across model space (Model % Sig: 15.95%), as well as the highest average and upper 95% confidence interval performance (Avg R^2 = 9.95%; Upper 95% CI R^2 = 20.52%). Among the Change networks, PS was best predicted by the SUBxCON network, showing high reliability across model space (Model % Sig: 10.66%) and high predictive accuracy (Avg R^2 = 12.04%; Upper 95% CI R^2 = 21.45%). The importance of the CON network for PS was also evidenced by the fact that the majority of networks significantly predictive of PS contained the nodes of the CON. Transfer to fluid intelligence reaction time (GFRT), was significantly predicted by clusters in a large number of networks, but the Pre DMNxFPN was the most accurate predictor of GFRT (Model % Sig: 3.59%; Avg R^2 = 12.03%; Upper 95% CI R^2 = 24.28%), and the SUB was the most reliable predictor (Model % Sig: 12.04%), but with relatively poorer model accuracy (Avg R^2 = 8.79%; Upper 95% CI R^2 = 14.32%). Among the Change networks, GFRT was best predicted by the Change SUBxCON and SUBxFPN networks, which both showed comparably high model reliability and accuracy (Table 3.3). Transfer to working memory (WM), was significantly predicted in a large number of networks, but the most predictive network was the Pre SUBxFPN, which demonstrated superior reliability and model accuracy (Model % Sig: 16.82%; Avg R^2 = 12.01%; Upper 95% CI R^2 = 25.24%). Among the change networks, WM was also very clearly predicted most successfully by the change in DMNxFPN network, which was both a reliable and accurate predictor of transfer (Model % Sig: 7.58%; Avg R^2 = 10.30%; Upper 95% CI R^2 = 16.13%). The DMN connectivity seemed to have a particularly important role for predicting WM transfer, given that a majority (62.5%) of datasets predictive of WM contained DMN nodes.

The training gain scores to SC were predicted by almost all networks, but the Pre DMN was the most predictive in terms of both model reliability and accuracy (Model % Sig: 13.42%; Avg R^2 = 6.02%; Upper 95% CI R^2 = 8.81%). The Change SUBxCON was the best predictor of SC gain, with both high model reliability and greatest model accuracy (Model % Sig: 9.42%; Avg R^2 = 8.96%; Upper 95% CI R^2 = 13.71%). The training gain scores in SD were best predicted by both the Pre FPN and DMNxCON networks, with significant clusters of models covering equal sized portions of model space and demonstrating comparable accuracy (Table 3.2). Similarly, among the Change networks, there was no single dataset with superior performance, and the SUB, FPN, and DMNxCON networks were all about equally predictive (Table 3.3). Finally, training gain in SR was best predicted by the Pre SUBxCON, demonstrating high model space coverage and model accuracy (Model % Sig: 6.09%; Avg R^2 = 9.90%; Upper 95% CI R^2 = 25.20%). Importantly, among Change networks, the DMN was overwhelmingly the best performer in predictions of SR, demonstrating strong ties between training-induced changes in this network and gain in SR (Model % Sig: 23.30%; Avg R^2 = 10.65%; Upper 95% CI R^2 = 16.57%).

3.3.2 Comparing Predictions with Pre and Change Networks

Investigating how Pre networks predict training and transfer is important to gain an understanding of what types of pre-existing characteristics determine whether someone is likely to improve in a training task and transfer that learning to other domains. On the other hand, providing evidence for relationships between training-induced network plasticity and training gain and transfer allows us to gain a better understanding of the effect of the training on the brain networks involved in training, and the extent to which these effects may be driving training gain and transfer. Investigating the overlap between networks that predict at Pre and those that predict with Change is critical for assessing to what extent pre-existing differences in network connectivity may have an influence on training-induced changes in these networks. We found a high degree of overlap between the networks that were predictive with the Pre training network connectivity and those that were predictive with Change in network connectivity. In predictions of PS transfer, both Pre and Change SUB, CON, FPN and SUBxCON networks were significantly predictive, and in predictions of GRFT, the SUB, FPN, SUBxCON, and DMNxFPN networks were all predictive of transfer with both Pre and Change network values. WM transfer was predicted by both Pre and Change SUB, DMN, SUBxFPN, and DMNxFPN networks. In predictions of SC training gain, the FPN, DMNxCON, and DMNxFPN networks were all predictive, and in predictions of SD training gain, the SUB, FPN, DMN, and DMNxCON networks were significant in both Pre and Change networks. Training gain in SR was predicted by both the SUB, FPN, DMN, SUBxCON, and DMNxCON Pre and Change networks. The overlap between the Pre and Change networks may also be informative in assessing how network overlap may show similarities between training gain and transfer.

3.3.3 Comparing Predictions between Training and Transfer

Training is thought to lead to transfer through the overlapping behavioral, cognitive, or neural mechanisms that are shared between the training and transfer tasks (Jonides 2004; Dahlin, Neely, et al. 2008; Buschkuehl and Jaeggi 2010; Klingberg 2010). To assess the extent to which this is the case in our dataset, we looked for predictive overlap between network predictors of transfer and training tasks. We found that transfer to PS and training gain in SC, SD, and SR were all predicted by Pre SUB, FPN, and DMNxCON networks, while GFRT transfer and training gain in these three games were predicted by SUB and FPN. Transfer to WM and training gain in SC, SD, and SR were predicted by Pre SUB, DMN, and DMNxCON network connectivity. All three transfer scores, along with training gain in both SD and SR, were predicted by the Pre SUBxCON internetwork, and both GFRT and WM, as well as SC and SR were predicted by Pre SUBxFPN connectivity. In comparing predictive overlap between Change networks, we found that transfer to PS and WM and training gain in all three games were predicted by changes in the FPN network, while improvement in GFRT and training games was predicted by changes in the FPN and DMNxCON networks.

In past behavioral results, we found that transfer to WM was significantly correlated with gain in the SC, SD, and SR tasks, but that gains in these tasks was not associated with transfer to either PS or GFRT (Baniqued 2015). Based on the ‘overlap’ theory of transfer, we would expect therefore that there are network predictors of WM that are shared with SC, SD, and SR which are not shared with PS and GFRT. We found that these three training tasks, as well as WM, were all predicted by the Pre DMN network connectivity, but the other transfer constructs, PS and GFRT, were not. Overall, the most reliably predictive pre-training network was the SUB network, which was predictive of all training and transfer scores, and is thought to be directly involved in skill learning. The most reliable predictor among training-induced changes in network connectivity was the FPN, which was predictive of all training gain and transfer scores, and is strongly involved in executive function (Dosenbach et al. 2008). Furthermore, prior work has implied that the FPN is critical for skill generalization that occurs during transfer (Zanto and Gazzaley 2013).

3.3.4 Consensus Features

Applying cluster correction to model space allows us to find clusters of models that are all significant predictors of our outcome variable, demonstrating that the dataset in question is reliably informative regarding individual differences in the outcome variable. A further test of the reliability of this dataset pertains to the individual features identified in the model clusters. If each model in a significant cluster contains a largely non-overlapping set of variables, we can assume that this dataset may not be as reliable of a predictor than one in which the significant cluster contains sets of features that are included in >90% of the models in the cluster. We call these features that are reliable predictors across the model cluster consensus features (CFs). To extract CFs, we identified which features were included in all CV loops within a given point in model space, and then assessed the extent to which that feature was a consensus feature for all other points in model space within a cluster.

A summary of the CFs can be found in Tables 3.4 and 3.5. Exploring CFs gives us a reduced group of networks with reliable features predictors. We found that GFRT had the largest set of CFs in the Pre SUBxFPN and Change DMNxFPN internetworks. While PS had the largest set of CFs in Pre CON, the Change DMN had a much larger group of CFs reliably predicting PS. WM had a large set of CFs in both the Pre and Change DMN, as well as the Change SUBxFPN internetwork. Both SC and SR showed similar prediction patterns in the Pre networks, with the DMN and DMNxCON having a large set of CF predictors. In the Change networks we see that SC, SR, and SD all contain their largest set of CFs in the DMNxCON internetwork.

We compared the sets of CFs and found that while WM transfer had many shared CFs with SC and SR, WM and SD shared very few CFs, which is surprising given that both the SD and WM tasks are n-back based. We found that the Pre DMN network shared 26 CFs between SC and WM, 28 between SR & WM, 23 CFs between SC and SR, but only 1 between SD and WM. In the Change DMN, PS shared 18 CFs with SD, 20 CFs with SR, and SR & SD shared 31 CFs, but PS did note share any CFs with SC. Furthermore, we also found that the Pre DMNxCON network shared 11 CFs between SC and WM, 9 CFs between SR and WM, and 55 CFs between SC and SR, but no CFs between SD and WM. We also found some overlap between transfer to GFRT and the training gain scores in the Change DMNxCON network, for example 5 between GFRT and SC, 9 between GFRT and SR, 3 between GFRT and SD, and 31 SR and SD. We found that the training gain in SC and SD had 54 overlapping CFs, SC and SR 17, and SD and SR had 8 overlapping CFs.

3.3.5 Consensus Hubs

In a functional connectivity matrix, the strength of each edge represents the variance shared by the timecourses in two nodes. We propose that if many edges that are CFs emanate from the same node, this node's position within the functional network may play a particularly important role in the prediction, and we call these nodes consensus hubs (CHs). Therefore, following the identification of CFs, we found network nodes that were included in either more than 10 CFs. For networks with less of CFs, nodes present in over 40% of the CFs were deemed CHs. We present a summary of all the CHs found in both the Pre and Change network analyses in Tables 3.6 & 3.7, and we display the percentage of CFs emanating from that node that were negatively correlated with the predicted score over the total number of CFs present in that node. We found a majority of the Pre network CHs were nodes in the DMN and CON networks, and that these CHs were predominantly informative of training gain in SC and SR. In the Change network CHs, the DMN and CON networks had an even larger majority, but unlike Pre, these CHs were evenly distributed among each of the training scores.

We found considerable overlap in the consensus hubs that were predictive in the PS and training gain scores. We found that the five consensus hubs where pre-training connectivity was predictive of PS were also predictive of training gain in SC, SR, or SD. It is worth noting that these consensus hubs most often played roles in different networks. For example, both the connectivity of the insula and superior frontal gyrus were predictive within the CON for PS, but these same regions were consensus hubs in the DMNxCON network for SC. It is also important to note that some of these consensus hubs had opposing correlations with their predictor variable. For example, while 88.9% of the CFs that make up the superior frontal gyrus consensus hub were negatively correlated with PS, 90.9% of the CFs for this hub in the DMNxCON were positively correlated with SC. We also found consensus hubs where change in that node's connectivity was commonly found in the CFs across multiple networks. Predictions from the Change DMN identified the middle temporal gyrus and temporal pole as both predictive of PS and SD, suggesting that the training may have had a common impact on these regions to drive gains in both PS and SD. The right insula connectivity in the Pre CON and Pre DMNxCON networks was predictive of GFRT transfer and SC gain respective

Outcome	Pre Network	Total Model Space	Cluster Model Space	Model Average R^2	Model Upper 95% R^2	Null Average R^2	Null Upper 95% R^2
GFRT	SUB	12.04%	5.81%	8.79%	14.32%	0.13%	0.30%
	CON	5.56%	2.87%	6.64%	11.35%	0.07%	0.26%
	FPN	3.49%	3.49%	7.18%	12.10%	0.20%	0.40%
	DMN	-	-	-	-	-	-
	SUBxCON	6.09%	3.41%	5.64%	11.18%	0.05%	0.13%
	SUBxFPN	9.13%	5.31%	5.60%	11.41%	0.03%	0.08%
	DMNxCON	-	-	-	-	-	-
	DMNxFPN	3.59%	3.59%	12.03%	24.28%	0.02%	0.07%
PS	SUB	2.80%	2.80%	7.66%	12.35%	0.17%	0.44%
	CON	15.95%	6.27%	9.95%	20.52%	0.02%	0.05%
	FPN	13.17%	6.45%	6.34%	12.48%	0.12%	0.30%
	DMN	-	-	-	-	-	-
	SUBxCON	11.20%	8.42%	7.62%	16.50%	0.20%	0.57%
	SUBxFPN	-	-	-	-	-	-
	DMNxCON	4.50%	4.50%	2.65%	4.67%	0.01%	0.02%
	DMNxFPN	-	-	-	-	-	-
WM	SUB	4.09%	4.09%	12.02%	22.01%	0.16%	0.50%
	CON	-	-	-	-	-	-
	FPN	-	-	-	-	-	-
	DMN	7.71%	4.83%	6.56%	10.81%	0.07%	0.18%
	SUBxCON	4.48%	2.24%	5.97%	11.25%	0.03%	0.09%
	SUBxFPN	16.82%	9.33%	12.01%	25.24%	0.07%	0.16%
	DMNxCON	8.42%	5.06%	5.62%	10.37%	0.01%	0.04%
	DMNxFPN	9.59%	9.59%	7.34%	13.08%	0.07%	0.22%
SC	SUB	11.61%	6.88%	5.06%	10.21%	0.34%	0.72%
	CON	-	-	-	-	-	-
	FPN	12.53%	8.98%	4.40%	9.31%	0.09%	0.20%
	DMN	13.42%	5.16%	6.02%	8.81%	0.10%	0.32%
	SUBxCON	-	-	-	-	-	-
	SUBxFPN	5.51%	3.03%	6.05%	12.01%	0.19%	0.48%
	DMNxCON	5.50%	3.48%	5.48%	7.94%	0.02%	0.05%
	DMNxFPN	3.35%	3.35%	4.58%	7.61%	0.02%	0.06%
SD	SUB	2.15%	2.15%	3.88%	6.94%	0.04%	0.13%
	CON	6.27%	3.94%	2.86%	5.43%	0.03%	0.07%
	FPN	7.85%	7.85%	6.99%	14.74%	0.10%	0.24%
	DMN	2.41%	2.41%	3.54%	6.37%	0.07%	0.24%
	SUBxCON	2.51%	2.51%	4.14%	8.47%	0.06%	0.17%
	SUBxFPN	-	-	-	-	-	-
	DMNxCON	5.68%	3.09%	7.92%	14.68%	0.07%	0.23%
	DMNxFPN	-	-	-	-	-	-
SR	SUB	6.02%	6.02%	3.94%	8.32%	0.01%	0.05%
	CON	-	-	-	-	-	-
	FPN	6.13%	6.13%	3.75%	7.83%	0.09%	0.30%
	DMN	2.89%	2.89%	6.55%	10.29%	0.01%	0.02%
	SUBxCON	6.09%	3.41%	9.90%	25.20%	0.05%	0.17%
	SUBxFPN	3.77%	3.77%	1.88%	4.09%	0.04%	0.12%
	DMNxCON	2.23%	2.23%	4.94%	8.23%	0.03%	0.09%
	DMNxFPN	-	-	-	-	-	-

Table 3.2: This table summarizes the performance of the Pre network datasets in predictions of each training gain and transfer score (Column 1- ‘Outcome’ Training gain/Transfer Score, Column 2- ‘Pre Network’ Dataset). The third column, ‘Total Model Space’, displays the percentage of model space that is populated by all clusters of models that significantly predict the outcome variable, and the fourth column, ‘Cluster Model Space’, shows the percentage of model space covered by the largest cluster. The fifth column, ‘Model Average R^2’, shows the average percentage of out-of-sample variance in the outcome variable that is predicted by the largest cluster, which should be compared to column seven, ‘Null Average R^2’, which shows the out of sample variance predicted by the null models over the same area in model space. Column six, ‘Model Upper 95% R^2’, displays the upper 95% confidence interval of percentage of out-of-sample variance predicted by the cluster, and this should be compared to column eight, ‘Null Upper 95% R^2’, which contains the upper 95% confidence interval for the null models. Outcome and dataset abbreviations follow the convention of the text. Scores: GFRT- fluid intelligence reaction time, PS- perceptual speed, WM- working memory, SC- Safe Cracker, SD- Sentry Duty, SR- Supply Run. Datasets: SUB- subcortical, CON- cingulo-opercular network, FPN- frontal-parietal network, DMN- default mode network, SUBxCON- subcortical and cingulo-opercular internetwork, SUBxFPN- subcortical and frontal-parietal internetwork, DMNxCON- default mode and cingulo-opercular internetwork, DMNxFPN- default mode and frontal-parietal internetwork.

Outcome	Change Network	Total Model Space	Cluster Model Space	Model Average R^2	Model Upper 95% R^2	Null Average R^2	Null Upper 95% R^2
GFRT	SUB	4.09%	4.09%	6.10%	12.65%	0.08%	0.21%
	CON	-	-	-	-	-	-
	FPN	7.53%	7.53%	4.39%	10.05%	0.08%	0.25%
	DMN	9.88%	9.88%	3.14%	5.93%	0.08%	0.23%
	SUBxCON	5.65%	3.23%	7.74%	15.12%	0.08%	0.24%
	SUBxFPN	6.40%	3.97%	7.39%	11.81%	0.04%	0.12%
	DMNxCON	8.80%	4.62%	5.69%	11.47%	0.02%	0.06%
	DMNxFPN	5.36%	3.31%	5.23%	8.72%	0.08%	0.31%
PS	SUB	6.88%	4.30%	3.97%	8.06%	0.03%	0.07%
	CON	4.12%	4.12%	9.66%	18.38%	0.12%	0.36%
	FPN	5.27%	5.27%	7.37%	13.93%	0.18%	0.34%
	DMN	12.15%	8.61%	4.74%	10.39%	0.05%	0.10%
	SUBxCON	10.66%	7.08%	12.04%	21.45%	0.05%	0.17%
	SUBxFPN	-	-	-	-	-	-
	DMNxCON	-	-	-	-	-	-
	DMNxFPN	-	-	-	-	-	-
WM	SUB	3.66%	3.66%	5.54%	11.75%	0.03%	0.12%
	CON	-	-	-	-	-	-
	FPN	3.12%	3.12%	4.35%	8.60%	0.17%	0.51%
	DMN	6.61%	4.22%	6.10%	9.00%	0.01%	0.03%
	SUBxCON	-	-	-	-	-	-
	SUBxFPN	4.57%	2.48%	5.07%	12.28%	0.07%	0.14%
	DMNxCON	-	-	-	-	-	-
	DMNxFPN	7.58%	7.58%	10.30%	16.13%	0.01%	0.05%
SC	SUB	-	-	-	-	-	-
	CON	-	-	-	-	-	-
	FPN	3.28%	3.28%	8.17%	12.14%	0.04%	0.11%
	DMN	-	-	-	-	-	-
	SUBxCON	7.26%	5.11%	8.96%	13.71%	0.02%	0.07%
	SUBxFPN	-	-	-	-	-	-
	DMNxCON	9.42%	7.41%	2.87%	6.67%	0.02%	0.06%
	DMNxFPN	3.59%	3.59%	6.73%	10.09%	0.01%	0.06%
SD	SUB	6.88%	6.88%	9.17%	14.59%	0.07%	0.18%
	CON	-	-	-	-	-	-
	FPN	8.71%	8.71%	6.75%	12.89%	0.07%	0.20%
	DMN	3.12%	3.12%	8.35%	18.26%	0.03%	0.08%
	SUBxCON	-	-	-	-	-	-
	SUBxFPN	-	-	-	-	-	-
	DMNxCON	7.59%	5.00%	6.88%	14.00%	0.04%	0.14%
	DMNxFPN	-	-	-	-	-	-
SR	SUB	4.95%	4.95%	4.47%	8.82%	0.08%	0.20%
	CON	7.89%	2.69%	8.63%	20.15%	0.05%	0.14%
	FPN	2.15%	2.15%	9.92%	14.64%	0.14%	0.31%
	DMN	23.30%	18.15%	10.65%	16.57%	0.04%	0.15%
	SUBxCON	9.95%	9.95%	6.42%	12.27%	0.07%	0.26%
	SUBxFPN	-	-	-	-	-	-
	DMNxCON	2.95%	2.95%	6.37%	10.37%	0.11%	0.25%
	DMNxFPN	3.24%	3.24%	4.48%	8.35%	0.03%	0.05%

Table 3.3: This table summarizes the performance of the Change network datasets in predictions of each training gain and transfer score (Column 1- ‘Outcome’ Training gain/Transfer Score, Column 2- ‘Pre Network’ Dataset). The third column, ‘Total Model Space’, displays the percentage of model space that is populated by all clusters of models that significantly predict the outcome variable, and the fourth column, ‘Cluster Model Space’, shows the percentage of model space covered by the largest cluster. The fifth column, ‘Model Average R^2’, shows the average percentage of out-of-sample variance in the outcome variable that is predicted by the largest cluster, which should be compared to column seven, ‘Null Average R^2’, which shows the out of sample variance predicted by the null models over the same area in model space. Column six, ‘Model Upper 95% R^2’, displays the upper 95% confidence interval of percentage of out-of-sample variance predicted by the cluster, and this should be compared to column eight, ‘Null Upper 95% R^2’, which contains the upper 95% confidence interval for the null models. Outcome and dataset abbreviations follow the convention of the text. Scores: GFRT- fluid intelligence reaction time, PS- perceptual speed, WM- working memory, SC- Safe Cracker, SD- Sentry Duty, SR- Supply Run. Datasets: SUB- subcortical, CON- cingulo-opercular network, FPN- frontal-parietal network, DMN- default mode network, SUBxCON- subcortical and cingulo-opercular internetwork, SUBxFPN- subcortical and frontal-parietal internetwork, DMNxCON- default mode and cingulo-opercular internetwork, DMNxFPN- default mode and frontal-parietal internetwork.

Network- Pre	GFRT	PS	WM	SC	SD	SR
SUB-# CFs	2	5	1	1, 15	9	1
SUB-Consensus Hubs	0	0	0	1	1	0
CON-# CFs	1, 8	5, 6, 26	-	-	0	-
CON-Consensus Hubs	1	2	-	-	0	-
FPN-# CFs	2	5, 15, 1	-	2, 22	4	0
FPN-Consensus Hubs	0	1	-	1	0	0
DMN-# CFs	-	-	24, 120	4, 142, 66	3	130
DMN-Consensus Hubs	-	-	4	4	0	6
SUBxCON-# CFs	5	7	0	-	0	15, 14
SUBxCON-Consensus Hubs	0	1	0	-	0	1
SUBxFPN-# CFs	5, 13	-	1, 1	40	-	27
SUBxFPN-Consensus Hubs	1	-		0	-	0
DMNxCON-# CFs	-	7	27, 26	58, 155	1	190
DMNxCON-Consensus Hubs	-	1	0	6	0	8
DMNxFPN-# CFs	11	-	4	23	-	-
DMNxFPN-Consensus Hubs	0	-	0	0	-	-

Table 3.4: This summarizes, in alternating order, the number of consensus features found across each analysis, and the number of consensus hubs found within those consensus features for all Pre datasets. Rows labeled '# CFs' display the number of consensus features in each cluster. Rows labeled 'Consensus Hubs' display the number of consensus hubs the overall analysis. Cells marked with a '-' were analyses for which the dataset did not successfully predict the outcome variable, and those marked with '0' are analyses that did not contain any consensus features.

Network- Change	GFRT	PS	WM	SC	SD	SR
SUB-# CFs	1, 1	1, 1	3	-	0	4
SUB-Consensus Hubs	0	0	0	-	0	0
CON-# CFs	-	8	-	-	-	10, 11, 5
CON-Consensus Hubs	-	1	-	-	-	0
FPN-# CFs	1	4	0	40	1	6
FPN-Consensus Hubs	0	0	0	1	0	0
DMN-# CFs	0	3, 177	0, 22	0	158	143, 74
DMN-Consensus Hubs	0	13	0	0	12	7
SUBxCON-# CFs	7, 2	3, 2	0	3, 9	0	0
SUBxCON-Consensus Hubs	0	0	0	0	0	0
SUBxFPN-# CFs	12	0	2, 38	0	0	0
SUBxFPN-Consensus Hubs	0	0	3	0	0	0
DMNxCON-# CFs	29	0	0	2, 151	208	70
DMNxCON-Consensus Hubs	2	0	0	8	14	2
DMNxFPN-#	0, 208	0	1	62	0	7
DMNxFPN-Consensus Hubs	10	0	0	0	0	0

Table 3.5: This summarizes, in alternating order, the number of consensus features found across each analysis, and the number of consensus hubs found within those consensus features for all Change datasets. Cells marked with a '-' were analyses for which the dataset did not successfully predict the outcome variable, and those marked with '0' are analyses that did not contain any consensus features.

Network- Pre	Score	Degree	Node #	Region	% Neg. Corr
SUB	SC	7	Node #233	Right Pallidum	57.1%
	SD	7	Node #230	Right Putamen	85.7%
CON	GFRT	4	Node #60	Right Insular Cortex	100%
	PS	4	Node #52	Right Insular Cortex	100%
	PS	9	Node #49	Right Superior Frontal Gyrus	88.9%
FPN	PS	6	Node #195	Left Angular Gyrus	100%
	SC	6	Node #175	Right Middle Frontal Gyrus	66.7%
DMN	WM	12	Node #120	Left Middle Temporal Gyrus	75.0%
	WM	13	Node #105	Right Paracingulate Gyrus	23.1%
	WM	11	Node #118	Left Middle Temporal Gyrus	45.5%
	WM	12	Node #75	Right Frontal Pole	8.3%
	SC	11	Node #74	Left Lateral Occipital Cortex	0.0%
	SC	10	Node #89	Right Precuneous Cortex	20.0%
	SC	10	Node #96	Right Lateral Occipital Cortex	60.0%
	SC	11	Node #117	Left Middle Temporal Gyrus	9.1%
	SR	11	Node #109	Left Paracingulate Gyrus	27.2%
	SR	14	Node #105	Right Paracingulate Gyrus	0.0%
	SR	17	Node #121	Right Superior Frontal Gyrus	100%
	SR	11	Node #104	Left Frontal Pole	0.0%
	SR	13	Node #116	Right Middle Temporal Gyrus	0.0%
	SR	15	Node #129	Left Middle Temporal Gyrus	0.0%
SUBxCON	PS	6	Node #53	Right Superior Frontal Gyrus	100%
	SR	4	Node #223	Left Thalamus	100%
SUBxFPN	GFRT	6	Node #202	Left Paracingulate Gyrus	16.7%
DMNxCON	PS	3	Node #56	Right Central Opercular Cortex	33.3%
	SC	13	Node #60	Right Insular Cortex	61.5%
	SC	10	Node #58	Left Central Opercular Cortex	20.0%
	SC	13	Node #55	Left Central Opercular Cortex	15.4%
	SC	11	Node #48	Right Ant. Supramarginal Gyrus	27.3%
	SC	11	Node #49	Right Superior Frontal Gyrus	9.1%
	SC	11	Node #52	Right Insular Cortex	63.6%
	SR	16	Node #52	Right Insular Cortex	12.5%
	SR	22	Node #53	Right Superior Frontal Gyrus	18.2%
	SR	21	Node #54	Right Juxt. Lobule Cortex	9.5%
	SR	15	Node #55	Left Central Opercular Cortex	53.3%
	SR	16	Node #56	Right Central Opercular Cortex	68.8%
	SR	20	Node #57	Left Insular Cortex	40.0%
	SR	12	Node #59	Left Cingulate Gyrus	58.3%
	SR	10	Node #60	Right Insular Cortex	30.0%

Table 3.6:
This table displays all consensus hubs (CHs) in each analysis. The first column, 'Network-Pre', displays the name of the corresponding Pre network of the CH, and the second column, 'Score', shows the score predicted. The third column, 'Degree', corresponds to the number of consensus features emanating from the CH. Column four and five, 'Node #' and 'Region', give the node number and anatomical label from the Power et al. 2011 functional network parcellation. The sixth column, '% Neg. Corr', shows the percentage of edges emanating from the CH that are negatively correlated with the score variable in column 2.

Network-Change	Score	Degree	Node #	Region	% Neg. Corr
CON	PS	4	Node #57	Left Insular Cortex	75.0%
FPN	SC	11	Node #181	Right Frontal Pole	36.4%
DMN	PS	11	Node #75	Right Frontal Pole	58.3%
	PS	14	Node #78	Left Frontal Pole	85.7%
	PS	15	Node #86	Left Lateral Occipital Cortex	46.7%
	PS	16	Node #89	Right Precuneous Cortex	50.0%
	PS	14	Node #91	Left Cingulate Gyrus	78.6%
	PS	11	Node #92	Right Cingulate Gyrus	81.2%
	PS	11	Node #98	Left Superior Frontal Gyrus	90.1%
	PS	14	Node #114	Left Frontal Pole	57.1%
	PS	11	Node #116	Right Middle Temporal Gyrus	72.7%
	PS	10	Node #118	Middle Temporal Gyrus	90%
	PS	13	Node #124	Left Lingual Gyrus	76.9%
	PS	16	Node #125	Right Parahippocampal Gyrus	31.3%
	PS	10	Node #128	Right Temporal Pole	60.0%
	SD	10	Node #81	Left Temporal Pole	100%
	SD	11	Node #92	Right Cingulate Gyrus	100%
	SD	14	Node #101	Right Frontal Pole	100%
	SD	11	Node #102	Right Frontal Pole	100%
	SD	18	Node #104	Left Frontal Pole	100%
	SD	11	Node #107	Left Paracingulate Gyrus	100%
	SD	10	Node #116	Right Middle Temporal Gyrus	100%
	SD	15	Node #122	Right Cingulate Gyrus	100%
	SD	15	Node #126	Left Temporal Fusiform Cortex	100%
	SD	13	Node #128	Right Temporal Pole	100%
	SD	14	Node #129	Left Middle Temporal Gyrus	100%
	SD	12	Node #130	Right Angular Gyrus	100%
	SR	15	Node #121	Right Superior Frontal Gyrus	6.7%
	SR	14	Node #98	Left Superior Frontal Gyrus	50.0%
	SR	13	Node #83	Left Middle Temporal Gyrus	69.2%
	SR	12	Node #123	Right Superior Temporal Gyrus	91.7%
	SR	12	Node #125	Right Parahippocampal Gyrus	75.0%
	SR	13	Node #116	Right Middle Temporal Gyrus	100%
	SR	10	Node #105	Right Paracingulate Gyrus	100%
SUBxFPN	WM	7	Node #226	Left Brain-Stem	71.4%
	WM	8	Node #227	Left Putamen	50.0%
	WM	5	Node #179	Right Inferior Temporal Gyrus	100%
DMNxCON	GFRT	6	Node #48	Right Supramarginal Gyrus	50.0%
	GFRT	6	Node #51	Left Juxt Lobule Cortex	16.7%
	SC	33	Node #53	Right Superior Frontal Gyrus	75.8%
	SC	29	Node #57	Left Insular Cortex	82.8%
	SC	15	Node #49	Right Superior Frontal Gyrus	86.7%
	SC	15	Node #50	Left Superior Frontal Gyrus	100%
	SC	15	Node #55	Left Central Opercular Cortex	93.3%
	SC	12	Node #113	Left Anterior Cingulate Cortex	100%
	SC	11	Node #59	Left Anterior Cingulate Cortex	81.2%
	SC	10	Node #60	Right Insular Cortex	90.0%

Table 3.7

	SD	18	Node #51	Left Juxt Lobule Cortex	94.4%
	SD	14	Node #52	Right Insular Cortex	100%
	SD	12	Node #54	Right Juxt Lobule Cortex	100%
	SD	12	Node #55	Left Central Opercular Cortex	100%
	SD	11	Node #56	Right Central Opercular Cortex	100%
	SD	16	Node #57	Left Insular Cortex	93.8%
	SD	24	Node #58	Left Central Opercular Cortex	100%
	SD	22	Node #59	Left Cingulate Gyrus	95.5%
	SD	15	Node #60	Right Insular Cortex	100%
	SD	12	Node #99	Left Superior Frontal Gyrus	100%
	SD	11	Node #109	Left Paracingulate Gyrus	100%
	SD	11	Node #110	Right Paracingulate Gyrus	100%
	SD	13	Node #111	Left Anterior Cingulate Gyrus	100%
	SD	12	Node #113	Left Anterior Cingulate Gyrus	100%
	SR	14	Node #50	Left Superior Frontal Gyrus	100%
	SR	13	Node #57	Left Insular Cortex	76.9%
DMNxFPN	GFRT	24	Node #197	Left Frontal Pole	8.3%
	GFRT	20	Node #180	Right Frontal Pole	20.0%
	GFRT	14	Node #187	Left Middle Frontal gyrus	28.6%
	GFRT	13	Node #188	Left Frontal Pole	38.5%
	GFRT	11	Node #189	Right Frontal Pole	45.5%
	GFRT	10	Node #104	Left Frontal Pole	10.0%
	GFRT	10	Node #123	Right Superior Temporal Gyrus	50.0%
	GFRT	10	Node #181	Right Frontal Pole	40.0%
	GFRT	10	Node #196	Right Middle Frontal Gyrus	50.0%
	GFRT	10	Node #202	Left Paracingulate Gyrus	10.0%

Table 3.7 (cont.): This table displays all consensus hubs (CHs) in each analysis. The first column, 'Network-Change', displays the name of the corresponding Change network of the CH, and the second column, 'Score', shows the score predicted. The third column, 'Degree', corresponds to the number of consensus features emanating from the CH. Column four and five, 'Node #' and 'Region', give the node number and anatomical label from the Power et al. 2011 functional network parcellation. The sixth column, '% Neg. Corr', shows the percentage of edges emanating from the CH that are negatively correlated with the score variable in column 2.

3.4 Discussion

Several converging sets of evidence in the current study suggest that the cingulo-opercular network (CON) plays a critical role not only in the acquisition of cognitive skills, but also in their transfer to untrained tasks. The CON plays a key role in executive function, maintaining on task awareness and suppressing internal off-task distractions (Dosenbach et al. 2007, 2008). Furthermore, while this network is largely separate from the FPN, CON plays an important role in the flexible control of goal directed behavior through stable task maintenance (Dosenbach et al. 2008; Zanto and Gazzaley 2013; Sestieri et al. 2014). In this way, CON provides the sustained attention that is critical during long and challenging cognitive training sessions. This type of sustained attention is not an inflexible innate ability, but rather a cognitive skill that can be enhanced with training (Tang et al. 2007; Tang and Posner 2009; Xue et al. 2011).

The development and deployment of all skills depends critically on the interaction between the basal ganglia and relevant cortical regions. While skill learning is thoroughly established in the motor learning literature (Karni et al. 1995; Driemeyer et al. 2008; Doyon et al. 2009), many cognitive skill training studies have been performed as well (Basak et al. 2008; Erickson et al. 2010; Vo et al. 2011; Nikolaidis et al. 2014, 2015; Knowlton et al. 2015), and a recent quantitative meta analysis found remarkably similar sets of neural correlates across both motor and cognitive skill acquisition (Patel et al. 2013). Just as the performance and acquisition of other skills depend on an interaction between the basal ganglia and the relevant cortical regions, the performance and acquisition of sustained attention skills depend not only on the CON, but also its relationship to the subcortical regions (SUB). We found the SUB network was also predictive of all three transfer domains, suggesting that prior to learning, brain mechanisms of skill learning offer insight into which individuals will show the greatest skill generalization; however, the development and deployment of these sustained attention skills does not rely only on the SUB network, but also on its relationship to CON. Critically, we found that the Pre SUBxCON internetwork was predictive of all transfer scores. This may suggest that the pre-training sustained attention processing was critical for benefiting the most from training. The importance of the pre-training connectivity of these two networks in generalizing training to untrained tasks suggests that there are individual differences in the process of skill learning that predispose trainees towards greater transfer; however, to address the full implications of this finding, we must also consider the role of the DMN in the process of training and transfer.

From a network perspective, the CON plays an important role as the primary regulator of the DMN (Fox et al. 2005; Sestieri et al. 2014), which is involved in internally oriented cognition and distraction (Buckner 2012). It is important to note that the current analyses were performed during resting state, in which the DMN should be prominent. During resting state, activity in the DMN rises and activity in its anti-correlated networks, such as the FPN and CON, are suppressed. Previous studies found that DMN abnormalities are associated with many disorders, such as autism (Kennedy et al. 2006), Alzheimer's Disease (Greicius et al. 2004; Buckner et al. 2005; Brier et al. 2012), schizophrenia (Liang et al. 2006), and attention deficit hyperactivity disorder (Tian et al. 2006). The role of the DMN in these disorders is specifically manifested as decreased attention and executive function due primarily to the decreased coordination between the DMN and its anti-networks (Uddin et al. 2009). Direct evidence of the relationship between these networks and cognition has demonstrated that individual differences in the variability of the negative correlation between DMN and task positive attentional networks, such as the CON, were predictive of individual differences in behavioral variability in executive attention as well (Kelly et al. 2008).

This suggests that behavioral performance is mediated by the relationship between the CON and DMN networks.

In the current work, we found that pre-training DMN connectivity was predictive of transfer to WM, and training gain in SC, SD, and SR, but was not predictive of transfer to PS or GFRT. WM transfer relies on the quick and accurate storage and maintenance of information, and individual differences in DMN connectivity are likely to be critical for both WM training and transfer because internally oriented thinking could be particularly disruptive for WM storage. The ability to both learn training tasks that rely on working memory and to generalize those skills to WM may depend on the DMN not playing a suppressive role on WM performance during learning. Given that skills are acquired through the interaction of negative feedback loops between cortex and striatum (Hikosaka et al. 2002; Doyon et al. 2003), and that the DMN and executive regions are anti-correlated, the DMN could also play an obstructive role in learning by prohibiting feedback from relevant cortical regions to the striatum. This notion is further supported by our findings that SR, the training task most dependent on accurate storage of information and thus most likely to be disrupted by DMN processing, showed strong relationships between consensus hub connectivity and training gain, with each CH demonstrating either 100% positive or negative correlations with SR gain. CHs in the Pre DMN were uniquely present in predictions of WM, SC, and SR, and Change DMN CHs were all 100% negatively associated with training gain in SD, supporting the notion that DMN connectivity may play a role in obstructing working memory training. Finally, we found that a very large group of CHs in the Change DMNxCON internetwork, and the greater decreases in connectivity of these CON regions were all associated with greater improvements in SC and SD. These results provide evidence that training-induced plasticity in the

DMN and its relationship with the CON, the DMN's regulator network, are important contributors to successful training and transfer.

The compilation of results from the SUB, CON, and DMN networks suggest that the regulation of distraction and the development and deployment of sustained attention plays a critical role for not only skill learning but also for skill generalization. One prominent theory in this domain is that attention processes and their associated brain regions provide a "scaffold" by which other cognitive processes are able to perform the task at hand and unitize the skill based stimulus response associations required for task performance (Petersen et al. 1998). In this way, the acquisition of skills depends critically on the deployment of sustained attention, and the current results suggest that skill generalization may depend on this ability as well. Sustained attention need not necessarily play a direct role in transfer, but rather may aid in generalization through improving the learning process. For example, previous cognitive training studies have demonstrated the importance of non-specific training that maximizes user engagement (Green and Bavelier 2003; Dye et al. 2009b), and highly engaging learning tasks are thought to offer superior learning opportunities as well (Kirriemuir and McFarlane 2003). Given the impact of these engaging states on learning, it is possible that sustained attention skills impact training gain and generalization through the regulation of the default mode network and subsequent facilitation of flow states during learning.

Based on previous work, we might have expected that the FPN would play the most critical role in the acquisition and transfer of skills. The FPN is thought to be the network most critical for the processing of intelligence (Jung and Haier 2007), and the interaction of regions in the FPN, such as the superior parietal lobe and dorsal lateral prefrontal cortex, is critical for processing WM as well (Koenigs et al. 2009; Andrews et al. 2011; Beatty et al. 2015). Given that the FPN is known to be critical for many of the executive functions performed in the training here, and some have suggested it plays a particularly important role in the transfer of training (Zanto and Gazzaley 2013), we thought that the FPN may play a key role not only in the acquisition of the training tasks but also in transfer to WM and GFRT. While we found that the Pre FPN was a reliably strong predictor of each of the three training tasks, surprisingly it was only a strong predictor of PS transfer, weakly predicting GFRT transfer and failing to predict WM transfer. Critically, Change FPN was the only network which significantly predicted every training and transfer score. Thus the beneficial impact of the intervention may have manifested most clearly in resting state changes to the FPN, which largely agrees with the cognitive functions targeted and improved by the training. Given that the Pre FPN was mostly predictive of training gain, but that Change was predictive of both training and transfer, changes in the FPN connectivity can be an important marker of intervention progress. Although the Change FPN demonstrated reliable predictive accuracy across all training and transfer tasks, the exceedingly low numbers of CHs and complete lack of shared CHs between predictions raises questions regarding its mechanistic importance in training and transfer.

While the neural correlates of training and skill acquisition have been well developed, a similar understanding of transfer and skill generalization are not yet well established. The predominant theory on the neural correlates of transfer, the neural overlap hypothesis, while being relatively non-specific and non-mechanistic, is a good starting point and has received some experimental support (Dahlin, Neely, et al. 2008; Dahlin et al. 2009; Anguera et al. 2013; Nikolaidis et al. 2014). This theory states that the neural correlates of transfer correspond to brain regions or networks that are commonly recruited in both the training and transfer tasks, and that these correlate regions will play a role in transferring the learning from the training task to the transfer domain. In the current study, several sets of results offer interesting insight into the nature of the neural overlap between training and transfer, and how this hypothesis is represented at different levels of analysis. We found considerable overlap in the predictive abilities of the pre and change datasets, demonstrating that for most networks in which pre-intervention connectivity was predictive of the behavioral effect of the training, the impact of training in those same networks would reflect these same learning phenomena.

Finally, there are several important clarifications and additional considerations that must be discussed. First, while resting state connectivity and task based connectivity are largely reflections of one another (Kelly et al. 2008; Smith et al. 2009), we should be cautious when making extrapolations and comparisons from one to the other. In the current work we demonstrate that specific networks and regions play important roles in the training and transfer process, but we cannot strongly assert that these roles carried over into the actual performance of cognitive tasks or future tasks. While the current results suggest that training may have induced changes in task processing, we would need to acquire scans during task performance to make this specific claim; however, we can be confident in concluding that general pre-existing conditions of the resting state, and training-induced changes in this resting state are tied to learning and transfer. We argue that measurement is advantageous because it may reflect task independent processing and network relationships. Furthermore, while examining the neural correlates of training and transfer offers an important insight into mechanisms of these learning phenomena, it is important to remember that cross sectional or pre-post longitudinal analyses offer a coarse representation of the complexity of the neural correlates of training. The process we are attempting to characterize is highly variable, not only across time, but also across individuals. Each training study therefore includes individuals at different stages of a learning process that is not yet well characterized in time. Repeated MRI over the course of learning of complex cognitive skills would be highly informative in charting the neural correlates of this learning process. Finally, it is important for us to express that while several of the results herein demonstrate the involvement of common neural mechanisms for training and transfer, these results do not necessitate common cognitive mechanism as well. For example, both visual and motor training may utilize the same attention networks, and both motor and attention networks may be upregulated in a motor training task.

3.5 Conclusion

We found that both pre-existing and training-induced differences in resting state connectivity are informative of both improvements in cognitive training games and transfer to underlying cognitive abilities. Our results highlight the importance of sustained attention via the CON in regulating the DMN, resulting in greater generalization of training to cognitive function through enhancing task engagement and training quality. Future attempts to unite cognitive training with other interventions shown to improve attention, such as exercise and meditation, may be especially promising in enhancing cognition.

Acknowledgements: This work was supported by a Beckman Fellowship to AN. The authors would like to thank Brad Sutton, Tracey Wszalek for technical guidance, and Nancy Dodge, Holly Tracy, and members of the Lifelong Brain and Cognition Laboratory for assistance with performing the experiments.

3.6 References

Andrews SC, Hoy KE, Enticott PG, Daskalakis ZJ, Fitzgerald PB. 2011. Improving working memory: the effect of combining cognitive activity and anodal transcranial direct current stimulation to the left dorsolateral prefrontal cortex. Brain Stimul. 4:84–89.

Anguera J a, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, Kong E, Larraburo Y, Rolle C, Johnston E, Gazzaley a. 2013. Video game training enhances cognitive control in older adults. Nature. 501:97–101.

Anguera J a, Russell C a, Noll DC, Seidler RD. 2007. Neural correlates associated with intermanual transfer of sensorimotor adaptation. Brain Res. 1185:136–151.

Ball K, Berch DB, Helmers KF, Jobe JB, Leveck MD, Marsiske M, Morris JN, Rebok GW, Smith DM, Tennstedt SL, Unverzagt FW, Willis SL. 2002. Effects of Cognitive Training Interventions With Older Adults. Jorunal Am Med Assoc. 288:2271–2281.

Baniqued PL, Allen CM, Kranz MB, Johnson K, Sipolins A, Dickens C, Ward N, Geyer A, Kramer AF. 2015. Working memory, reasoning, and task switching training: Transfer effects, limitations, and great expectations? PLoS One. 10:1–29.

Basak C, Boot WR, Voss MW, Kramer AF. 2008. Can training in a real-time strategy video game attenuate cognitive decline in older adults? Psychol Aging. 23:765–777.

Beatty EL, Jobidon M, Bouak F, Nakashima A, Smith I, Lam Q, Blackler K, Cheung B, Vartanian O. 2015. Transfer of training from one working memory task to another: behavioural and neural evidence. Front Syst Neurosci. 9:86.

Bennett G, Seashore H, Wesman A. 1947. Differential aptitude tests, Form A. San Antonio: The Psychological Corporation.

Bjork BRA, Linn MC. 2006. The Science of Learning and the Learning of Science Introducing Desirable Difficulties. Science (80-). 19:6–7.

Boot WR, Basak C, Erickson KI, Neider M, Simons DJ, Fabiani M, Gratton G, Voss MW, Prakash R, Lee H, Low K a, Kramer AF. 2010. Transfer of skill engendered by complex task training under conditions of variable priority. Acta Psychol (Amst). 135:349–357.

Brier MR, Thomas JB, Snyder AZ, Benzinger TL, Zhang D, Raichle ME, Holtzman DM, Morris JC, Ances BM. 2012. Loss of intranetwork and internetwork resting state functional connections with Alzheimer's disease progression. J Neurosci. 32:8890–8899.

Buckner R, Snyder A, Shannon B, LaRossa G, Sachs R, Fotenos A, Sheline Y, Klunk W, Mathis C, Morris J, Mintun M. 2005. Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. J Neurosci. 25:7709–7717.

Buckner RL. 2012. The serendipitous discovery of the brain's default network. Neuroimage. 62:1137–1145.

Buitinck L, Louppe G, Blondel M. 2013. API design for machine learning software: experiences from the scikit-learn project. arXiv Prepr arXiv …. 1–15.

Buschkuehl M, Jaeggi SM. 2010. Improving intelligence: A literature review. Swiss Med Wkly. 140:266–272.

Clare Kelly AM, Uddin LQ, Biswal BB, Castellanos FX, Milham MP. 2008. Competition between functional brain networks mediates behavioral variability. Neuroimage. 39:527–537.

Craddock RC, Holtzheimer PE, Hu XP, Mayberg HS. 2009. Disease state prediction from resting state functional connectivity. Magn Reson Med. 62:1619–1628.

Crone EA, Wendelken C, Van Leijenhorst L, Honomichl RD, Christoff K, Bunge SA. 2009. Neurocognitive development of relational reasoning. Dev Sci. 12:55–66.

Dahlin E, Bäckman L, Neely AS, Nyberg L. 2009. Training of the executive component of working memory: subcortical areas mediate transfer effects. Restor Neurol Neurosci. 27:405–419.

Dahlin E, Neely AS, Larsson A, Bäckman L, Nyberg L. 2008. Transfer of learning after updating training mediated by the striatum. Science. 320:1510–1512.

Dahlin E, Nyberg L, Bäckman L, Neely AS. 2008. Plasticity of executive functioning in young and older adults: immediate training gains, transfer, and long-term maintenance. Psychol Aging. 23:720–730.

Dosenbach NUF, Fair D a, Cohen AL, Schlaggar BL, Petersen SE. 2008. A dual-networks architecture of top-down control. Trends Cogn Sci. 12:99–105.

Dosenbach NUF, Fair D a, Miezin FM, Cohen AL, Wenger KK, Dosenbach R a T, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen SE. 2007. Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci U S A. 104:11073–11078.

Dosenbach NUF, Nardos B, Cohen AL, Fair D a, Power D, Church J a, Nelson SM, Wig GS, Vogel AC, Lessov-schlaggar CN, Barnes KA, Dubis JW, Feczko E, Coalson RS, Jr JRP, Barch DM, Petersen SE, Schlaggar BL. 2011. Prediction of Individua Brain Maturity Using fMRI. Science (80-). 329:1358–1361.

Doyon J, Bellec P, Amsel R, Penhune V, Monchi O, Carrier J, Lehéricy S, Benali H. 2009. Contributions of the basal ganglia and functionally related brain structures to motor learning. Behav Brain Res. 199:61–75.

Doyon J, Penhune V, Ungerleider LG. 2003. Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. Neuropsychologia. 41:252–262.

Driemeyer J, Boyke J, Gaser C, Büchel C, May A. 2008. Changes in gray matter induced by learning--revisited. PLoS One. 3:e2669.

Dye MWG, Green CS, Bavelier D. 2009a. Increasing Speed of Processing With Action Video Games. Curr Dir Psychol Sci. 18:321–326.

Dye MWG, Green CS, Bavelier D. 2009b. The development of attention skills in action video game players. Neuropsychologia. 47:1780–1789.

Ekstrom RBR, French JJW, Harman HH, Dermen D. 1976. Manual for kit of factor-referenced cognitive tests. Princet NJ Educ Test Serv. 102:117.

Erickson KI, Boot WR, Basak C, Neider MB, Prakash RS, Voss MW, Graybiel AM, Simons DJ, Fabiani M, Gratton G, Kramer AF. 2010. Striatal volume predicts level of video game skill acquisition. Cereb Cortex. 20:2522–2530.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci U S A. 102:9673–9678.

Geerligs L, Renken RJ, Saliasi E, Maurits NM, Lorist MM. 2015. A Brain-Wide Study of Age-Related Changes in Functional Connectivity. Cereb Cortex. 25:1987–1999.

Graybiel AM. 2005. The basal ganglia: learning new tricks and loving it. Curr Opin Neurobiol. 15:638–644.

Green CS, Bavelier D. 2003. Action video game modifies visual selective attention. Nature. 423:534–537.

Greicius M, Srivastava G, Reiss A, Menon V. 2004. Default- mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI. Proc Natl Acad Sci USA. 101:4637–4642.

Guyon I. 2003. An Introduction to Variable and Feature Selection 1 Introduction. 3:1157–1182.

Hastie T, Tibshirani R, Friedman J. 2009. The elements of statistical learning.

Hikosaka O, Nakamura K, Sakai K, Nakahara H. 2002. Central mechanisms of motor skill learning. Curr Opin Neurobiol. 12:217–222.

Hinton GE. 2011. Machine learning for neuroscience. Neural Syst Circuits. 1:12.

Jaeggi SM, Buschkuehl M, Etienne A, Ozdoba C, Perrig WJ, Nirkko AC. 2007. On how high performers keep cool brains in situations of cognitive overload. Cogn Affect Behav Neurosci. 7:75–89.

Jaeggi SM, Buschkuehl M, Jonides J, Perrig WJ. 2008. Improving fluid intelligence with training on working memory. Proc Natl Acad Sci U S A. 105:6829–6833.

Jaeggi SM, Seewer R, Nirkko AC, Eckstein D, Schroth G, Groner R, Gutbrod K. 2003. Does excessive memory load attenuate activation in the prefrontal cortex? Load-dependent processing in single and dual tasks: Functional magnetic resonance imaging study. Neuroimage. 19:210–225.

Jonides J. 2004. How does practice makes perfect? Nat Neurosci.

Jung RE, Haier RJ. 2007. The Parieto-Frontal Integration Theory (P-FIT) of intelligence: converging neuroimaging evidence. Behav Brain Sci. 30:135–154; discussion 154–187.

Kane MJ, Conway AR, Miura TK, Colflesh GJ. 2007. Working memory, attention control, and the N-back task: a question of construct validity. J Exp Psychol Learn Mem Cogn. 33:615–622.

Karni A, Meyer G, Jezzard P, M. Adams M, Tuner R, G. Ungerleider L, Adams M, Turner R, Ungerleider LG, M. Adams M, Tuner R, G. Ungerleider L, Adams M, Turner R, Ungerleider LG. 1995. Functional MRI evidence for adult motor cortex plasticity during motor skill learning. Lett to Nat.

Kelly AMC, Uddin LQ, Biswal BB, Castellanos FX, Milham MP. 2008. Competition between functional brain networks mediates behavioral variability. Neuroimage. 39:527–537.

Kennedy DP, Redcay E, Courchesne E. 2006. Failing to deactivate: resting functional abnormalities in autism. Proc Natl Acad Sci U S A. 103:8275–8280.

Kirchner WK. 1958. Age differences in short-term retention of rapidly changing information. J Exp Psychol. 55:352–358.

Kirriemuir J, McFarlane a. 2003. Literature Review in Games and Learning. A Rep NESTA Futur.

Klingberg T. 2010. Training and plasticity of working memory. Trends Cogn Sci. 14:317–324.

Knowlton BJ, Foerde K, Poldrack RA, Knowlton BJ, Sabb FW, Bookheimer SY, Bilder RM, Guthrie D, Granholm E. 2015. Selective corticostriatal dysfunction in schizophrenia : Examination of motor and cognitive skill learning . Selective Corticostriatal Dysfunction in Schizophrenia : Examination of Motor and Cognitive Skill Learning.

Koenigs M, Barbey AK, Postle BR, Grafman J. 2009. Superior parietal cortex is critical for the manipulation of information in working memory. J Neurosci. 29:14980–14986.

Liang M, Zhou Y, Jiang T, Liu Z, Tian L, Liu H, Hao Y. 2006. Widespread functional disconnectivity in schizophrenia with resting-state functional magnetic resonance imaging. Neuroreport. 17:209–213.

Marquardt DW, Snee RD. 1975. Ridge Regression in Practice. Source Am Stat. 29:3–20.

Meier TB, Desphande AS, Vergun S, Nair V a., Song J, Biswal BB, Meyerand ME, Birn RM, Prabhakaran V. 2012. Support vector machine classification and characterization of age-related reorganization of functional brain networks. Neuroimage. 60:601–613.

Melby-Lervåg M, Hulme C. 2012. Is Working Memory Training Effective? A Meta-Analytic Review. Dev Psychol.

Nikolaidis A, Goatz D, Smaragdis P, Kramer A. 2015. Predicting Skill-Based Task Performance and Learning with fMRI Motor and Subcortical Network Connectivity. 2015 Int Work Pattern Recognit NeuroImaging. 93–96.

Nikolaidis A, Voss MW, Lee H, Vo LTK, Kramer AF. 2014. Parietal plasticity after training with a complex video game is associated with individual differences in improvements in an untrained working memory task. Front Hum Neurosci. 8:1–11.

Oelhafen S, Nikolaidis A, Padovani T, Blaser D, Koenig T, Perrig WJ. 2013. Increased parietal activity after training of interference control. Neuropsychologia. 51:2781–2790.

Olesen PJ, Westerberg H, Klingberg T. 2004. Increased prefrontal and parietal activity after training of working memory. Nat Neurosci. 7:75–79.

Patel R, Spreng RN, Turner GR. 2013. Functional brain changes following cognitive and motor skills training: a quantitative meta-analysis. Neurorehabil Neural Repair. 27:187–199.

Petersen SE, van Mier H, Fiez JA, Raichle ME. 1998. The effects of practice on the functional anatomy of task performance. Proc Natl Acad Sci. 95:853–860.

Postuma RB. 2005. Basal Ganglia Functional Connectivity Based on a Meta-Analysis of 126 Positron Emission Tomography and Functional Magnetic Resonance Imaging Publications. Cereb Cortex. 16:1508–1521.

Ravens J. 1962. Advanced Progressive Matrices: Set II. London.

Rebok GW, Ball K, Guey LT, Jones RN, Kim HY, King JW, Marsiske M, Morris JN, Tennstedt SL, Unverzagt FW, Willis SL. 2014. Ten-year effects of the advanced cognitive training for independent and vital elderly cognitive training trial on cognition and everyday functioning in older adults. J Am Geriatr Soc. 62:16–24.

Redick TS, Shipstead Z, Harrison TL, Hicks KL, Fried DE, Hambrick DZ, Kane MJ, Engle RW. 2012. No Evidence of Intelligence Improvement After Working Memory Training: A Randomized, Placebo-Controlled Study. J Exp Psychol Gen. 142:359–379.

Salthouse TA. 2005. Relations between cognitive abilities and measures of executive functioning. Neuropsychology. 19:532–545.

Salthouse TA, Babcock RL. 1991. Decomposing adult age differences in working memory. Dev Psychol. 27:763–776.

Sambataro F, Murty VP, Callicott JH, Tan HY, Das S, Weinberger DR, Mattay VS. 2010. Age-related alterations in default mode network: Impact on working memory performance. Neurobiol Aging. 31:839–852.

Schmidt R a., Bjork R a., Schmidt BRA, Bjork R a., Schmidt R a., Bjork R a. 1992. New Conceptualizations of Practice: Common Principles in Three Paradigms Suggest New Concepts for Training. Psychol Sci. 3:207–217.

Seidler RD. 2010. Neural correlates of motor learning, transfer of learning, and learning to learn. Exerc Sport Sci Rev. 38:13.

Sestieri C, Corbetta M, Spadone S, Romani GL, Shulman GL. 2014. Domain-general Signals in the Cingulo-opercular Network for Visuospatial Attention and Episodic Memory. J Cogn Neurosci. 26:551–568.

Shipstead Z, Redick TS, Engle RW. 2012. Is working memory training effective? Psychol Bull. 138:628–654.

Singley MK, Anderson JR. 1989. The transfer of cognitive skill. Harvard University Press.

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR, Beckmann CF. 2009. Correspondence of the brain ' s functional architecture during activation and rest.

Tang Y-Y, Ma Y, Wang J, Fan Y, Feng S, Lu Q, Yu Q, Sui D, Rothbart MK, Fan M, Posner MI. 2007. Short-term meditation training improves attention and self-regulation. Proc Natl Acad Sci U S A. 104:17152–17156.

Tang Y-Y, Posner MI. 2009. Attention training and attention state training. Trends Cogn Sci. 13:222–227.

Thorell LB, Lindqvist S, Bergman Nutley S, Bohlin G, Klingberg T. 2009. Training and transfer effects of executive functions in preschool children. Dev Sci. 12:106–113.

Tian L, Jiang T, Wang Y, Zang Y, He Y, Liang M, Sui M, Cao Q, Hu S, Peng M, Zhuo Y. 2006. Altered resting-state functional connectivity patterns of anterior cingulate cortex in adolescents with attention deficit hyperactivity disorder. Neurosci Lett. 400:39–43.

Tu PC, Hsieh JC, Li CT, Bai YM, Su TP. 2012. Cortico-striatal disconnection within the cingulo-opercular network in schizophrenia revealed by intrinsic functional connectivity analysis: A resting fMRI study. Neuroimage. 59:238–247.

Uddin LQ, Kelly AMC, Biswal BB, Castellanos FX, Milham MP. 2009. Functional Connectivity of Default Mode Network Components: Correlation, Anticorrelation, and Causality. Hum Brain Mapp. 30:625–637.

Varoquaux G, Thirion B. 2010. Detection of Brain Functional-Connectivity Difference in Post-stroke Patients Using Group-Level Covariance Modeling. Brain. 200–208.

Vo LTK, Walther DB, Kramer AF, Erickson KI, Boot WR, Michelle W, Prakash RS, Lee H, Fabiani M, Gratton G, Simons DJ, Sutton BP, Wang MY. 2011. Predicting individuals' learning success from patterns of pre-learning MRI activity. PLoS One. 6:e16093.

Voss MW, Prakash RS, Erickson KI, Boot WR, Basak C, Neider MB, Simons DJ, Fabiani M, Gratton G, Kramer AF. 2011. Effects of training strategies implemented in a complex videogame on functional connectivity of attentional networks. Neuroimage.

Wechsler D. 1997. WAIS-III: Wechsler adult intelligence scale. San Antonio: The Psychological Corporation.

Willis SL, Schaie KW. 1986. Trianing the elderly on the ability factors of spatial orientationand inductive reasoning. Psychol Aging.

Xue S, Tang Y-Y, Posner MI. 2011. Short-term meditation increases network efficiency of the anterior cingulate cortex. Neuroreport. 22:570–574.

Zanto TP, Gazzaley A. 2013. Fronto-parietal network: Flexible hub of cognitive control. Trends Cogn Sci. 17:602–603.

Chapter 4

Comparing Resting State and Brain Structural Predictions of Cognitive Skill Training and Transfer to Working Memory, Fluid Intelligence, and Perceptual Speed Using Machine Learning

Aki Nikolaidis[1], Drew Goatz[1], Pauline Baniqued[2], Michael Kranz[1,5], Aron Barbey[1,5], Paris Smaragdis[1,3,4], Arthur F. Kramer[1,5]

Beckman Institute[1], Department of Computer Science[3], Department of Electrical and Computer Engineering[4],
Department of Psychology[5]
University of Illinois. Urbana Champaign
Urbana, Illinois. USA

Helen Wills Neuroscience Institute[2]
University of California. Berkeley
Berkeley, California. USA

Abstract:
Cognitive training holds great potential for targeted improvement and remediation of cognitive ability, and while there have been many strong demonstrations of the success of these approaches, attempts to replicate some of these effects have failed. One necessary requirement for the successful and reliable application of cognitive training is the development of a mechanistic theory of how training drives transfer, yet this is currently not well established. Furthermore, the neural correlates that drive these cognitive mechanisms are even less well characterized. In the current study, we address both the cognitive and neural mechanisms of transfer by training 45 young adults with a multimodal training game that targets complimentary aspects of working memory, attention, and reasoning. Using both pre existing and training-induced changes in brain structure and resting state network connectivity, we assess the extent to which these two neuroimaging modalities contribute to predictions of training and transfer, and compare the reliable effects across structure and function. We find that transfer to working memory shares many structural and functional predictors with training improvement in an inductive reasoning and working memory updating game, but not with a working memory n-back based game, even though the transfer to working memory was measured with the n-back. We discuss these findings in the context of skill learning, skill generalization, and the neural overlap hypothesis of training and transfer.

4.1 Introduction

The possibility of improving cognition through guided training may hold potential for improving quality of life for many people either afflicted with neurological or psychiatric illness, or healthy individuals who could benefit from improved cognitive performance. The cognitive training field has focused on the study and development of these training programs and many studies have found significant improvements in cognitive performance after training (Ball et al. 2002; Green and Bavelier 2003; Dahlin, Neely, et al. 2008; Jaeggi et al. 2008; Dye et al. 2009; Thorell et al. 2009; Anguera et al. 2013; Rebok et al. 2014; Baniqued et al. 2015), but we and others have failed to replicate some of these effects in follow-up studies (Melby-Lervåg and Hulme 2012; Redick et al. 2012; Shipstead et al. 2012; Oelhafen et al. 2013). A major problem facing this field is not only methodological constraints in the design and implementation of these studies, such as small sample sizes, but also in the underdeveloped theory of mechanisms by which these training paradigms lead to improvement in cognitive performance. The study of cognitive training in the context of individual differences in brain structure and function leads to important insight into the mechanisms by which training may impact cognition. While several studies have investigated the neural correlates of cognitive training and transfer (Olesen et al. 2004; Anguera et al. 2007, 2013; Dahlin, Neely, et al. 2008; Vo et al. 2011; Voss et al. 2011; Oelhafen et al. 2013; Nikolaidis et al. 2014, 2015), a comprehensive theory of the neural underpinnings of training and transfer has yet to become established. In contrast, the skill learning literature has established a deep mechanistic understanding of training in both the human and animal literature (Hikosaka et al. 2002; Graybiel 2005; Doyon et al. 2009; Patel et al. 2013). In order to understand how cognitive training leads to transfer it is critical that cognitive training is explained within the greater context of skill learning. In this way, an important question for cognitive training changes from how training leads to transfer, to how skill acquisition leads to skill generalization. Previous work on establishing the relationship between skill acquisition and generalization have demonstrated that these phenomena are closely linked (Landauer and Bjork 1978; Bahrick 1979; Singley and Anderson 1989; Schmidt et al. 1992; Bjork and Linn 2006; Anguera et al. 2007; Seidler 2010), and modifications to training that affect acquisition can have subsequent effects of skill generalization as well. In order to develop a robust mechanistic theory of cognitive training and transfer, it is critical to establish the neural mechanisms that explain the relationship between skill acquisition and skill generalization.

In order to contribute to the establishment of such a theory, we recruited 45 young adults who trained with Mind Frontiers, an adaptive cognitive training program that targets complementary aspects of executive function, attention, working memory, and reasoning (Baniqued et al. 2015). Participants recruited to our cognitive training study completed pre and post training neuropsychological testing, structural MRI and resting state fMRI, as well as 30 hours of cognitive training. Mind Frontiers (MF) includes a comprehensive battery of six adaptive cognitive tasks that have demonstrated previous success in cognitive training, and which were gamified and united into a common “Wild West” theme videogame. The behavioral results from this study have been published previously (Baniqued et al. 2015), in which we found that training with MF induced transfer to constructs of Perceptual Speed (PS), fluid intelligence reaction times (GFRT), and working memory (WM). More recently we demonstrated that individual differences in regional brain volumes and cortical thickness in sensory processing and attention regions predict both training and transfer to these constructs (Nikolaidis et al. 2016a). We have also found that network connectivity of the cingulo-opercular network, default mode network, and subcortical network are highly relevant for both training and transfer, suggesting that the deployment of higher quality sustained attention may play an important role in deriving maximum benefits from training (Nikolaidis et al. 2016b).

In the current study, we assess how the functional network and structural correlates of training gain and transfer are related. By assessing how pre-existing differences in functional and structural correlates of training gain and transfer are related, we gain a fuller understanding of the brain mechanisms by which skill acquisition and generalization are obtained. Brain structure and resting state networks are two highly interactive systems, with structural volume measurements changing due to many functional characteristics, like blood flow to a region. Most analyses have treated these modalities separately due to the difficulty of processing both the functional and structural data in a given study for one project. Structural and network topography may reflect on one another. Given that resting state networks are both functionally and metabolically linked over time and may respond in sync to external stimuli, their structural volumes may be linked as well. We plan on assessing both how training-induced changes in functional networks and structural characteristics relate to one another and to training gain and transfer. The extent to which training-induced changes in networks and structure overlap has not been well established. It is known that structural regions can exhibit volumetric changes in as little as two hours of training (Kwok et al. 2011), and this raises the important point that structural volume measurements are a combination of many different functionally and structurally relevant signals, such as synaptogenesis, neuronal swelling, dentritic arborization, and enhanced blood flow (Lövdén et al. 2013; Zatorre et al. 2013). It has been well established that structural regions undergo volumetric changes according to the partial renormalization hypothesis (Lövdén et al. 2013). This hypothesis states that with learning, relevant cortical regions will first undergo an expansion, which is soon to be followed by a process of pruning that occurs as training continues. Given the time-course of these structural changes, it is likely that they may follow similar functional changes, such as expansion followed by winnowing of activated regions as skills are unitized and become automated. It is possible that training-induced changes in functional networks may be reflected in concomitant changes in structural regions as well, and therefore it is critical to investigate these two sets of data concurrently, as the reciprocal relationship between structure and function directly reflect the learning process.

To assess how both pre-existing and training-induced differences in brain structure and functional connectivity contribute to individual differences in training gain and transfer, we apply common machine learning techniques to structural MRI volumes and thicknesses, as well as resting state network connectivity. Unlike machine learning methods, ordinary least squares (OLS) is likely to produce models that overfit and fail to replicate to new samples (Hastie et al. 2009). This is especially true in cases of collinear variables or low sample sizes, which are common characteristics in cognitive neuroscience. OLS optimizes the model to reduce the squared error term, and does not place any constraints on the models it creates to do so. To overcome these issues, more advanced statistical techniques such as ridge regression have been developed (Marquardt and Snee 1975). Ridge regression places constraints on the size of the largest model parameters, which prevents the model from taking strong positions on the importance of a given variable in predicting an outcome. This allows ridge regression to not only produce models that are more likely to replicate than OLS, but it also produces better results in scenarios with collinear models. The benefit of greater replicability is not unique to ridge regression, but common to many machine learning methods such as lasso, elastic net, support vector machines, and random forests (Hastie et al. 2009). This has led to increased use of these methods in the scientific literature, and more recently these methods have made their way to neuroscience as well (Hinton 2011;

Varoquaux and Thirion 2014). Researchers have used machine learning and neuroimaging to classify disease states (Shen et al. 2010; Ye et al. 2011), predict aging (Dosenbach et al. 2010; Meier et al. 2012), and complex skill performance and learning (Nikolaidis et al. 2015). Recently, we have applied such models to both structural (Nikolaidis et al. 2016a) and functional network characteristics (Nikolaidis et al 2016b) to predict training gain and transfer as well. In the current work, we attempt to consolidate these findings and explore the relationship between the two in depth.

In the current study, we focus on investigating the most reliable structural and functional network predictors of training gain and transfer. Typically machine learning approaches optimize model parameters to maximize performance, but in the current study, we explore how well the data predict the outcome independent of model parameters, allowing us to assess the reliability of our predictions. We focus on the reliability of our findings in the current study by running our models over a wide range of model parameters, and assessing the range of models over which our data is predictive of our outcome variables. This allows us to define the full model space, and assess the extent to which small permutations in the feature set or regression method affect the predictive ability of our data. Furthermore, this allows us to find clusters of models that are nearby in model space which are all predictive of our outcome variable. We parse these clusters looking for features that all significant models include as predictors of the outcome variable, called consensus features. We find that structural and functional predictors of transfer to working memory are greatly overlapping with predictors of gain in Safe Cracker, a reasoning game, and Supply Run, a working memory updating game, but no overlap with Sentry Duty. Given that both Sentry Duty and our working memory transfer tasks were based in the n-back, the neural overlap hypothesis would have predicted the greater similarity in predictors between Sentry Duty gain and working memory transfer. We address these findings in the context of skill learning and assess refinements to the neural overlap hypothesis.

4.2 Methods

4.2.1 Participants

We recruited participants from the University of Illinois and surrounding community through flyers and online advertisements seeking participants for a "cognitive training study." Online pre-screening for demographic information and videogame experience was completed over email. Subsequent phone screening required participants: 1) between 18 and 30 years old, 2) 75% right handed according to the Edinburgh Handedness Inventory, 3) normal or corrected-to-normal hearing and vision, 4) no major psychological or medical conditions, 5) no non-removable metal in the body, and 6) played no more than 5 hours of video games per week in the past six months. A total of 45 participants trained with Mind Frontiers and had complete behavioral and neural data.

4.2.2 Study Design

A complete description of the study design can be found in Baniqued et al. 2015. Briefly, participants completed three initial cognitive testing sessions, and an MRI session before beginning the intervention, in which they completed four to five training sessions per week for four to five weeks, for a total of 20 sessions. After training, subjects returned to lab for a final MRI session and another round of cognitive testing.

Participants trained with six computerized adaptive training tasks that were selected based on their known psychometric properties with reasoning/Gf, working memory, visuospatial reasoning, inductive reasoning, and task switching. See Table 1.1 for a description of each training task.

4.2.3 Cognitive Assessment Protocol

For a full listing of all neuropsychological tasks administered, see Baniqued et al. 2015. Here we briefly describe each of the tasks that were included to create the reasoning, working memory, and perceptual speed composite variables we predicted. The tests below were obtained from the Virginia Cognitive Aging Project Battery (Salthouse 2005), and two different versions were used for pre- and post-testing, with the sequence counterbalanced across subjects.

4.2.3.1 GFRT

The fluid intelligence reaction time (GFRT) composite was comprised of standardized reaction times in these four tasks. Descriptions for these and other tasks in section 2.3 are reproduced from Baniqued et al. 2015.

Matrix Reasoning: Select the pattern that completes a missing space on a 3 x 3 grid. Number of correctly answered items is the primary measure. Reaction time on correct trials was also analyzed. Source: Ravens 1962; Crone et al. 2009.

Paper Folding: Identify pattern of holes that results from a punch through folded paper. Number of correctly answered items is the primary measure. Reaction time on correct trials was also analyzed. Source: Ekstrom et al. 1976.

Spatial Relations: Identify 3D object that would match a 2D object when folded. Number of correctly answered items is the primary measure. Reaction time on correct trials was also analyzed. Source: Bennett et al. 1947.

Letter Sets: Determine which letter set is different from the other four. Number of correctly answered items is the primary measure. Reaction time on correct trials was also analyzed. Source: Ekstrom et al. 1976.

4.2.3.2 Perceptual Speed

The Perceptual Speed (PS) composite was comprised of standardized accuracy scores in these three tasks.

Digit Symbol Substitution: Write corresponding symbol for each digit using a coding table. The primary measure is number of correctly answered items within two minutes. Source: Wechsler 1997.

Pattern Comparison: Determine whether pairs of line patterns are the same or different. The primary measure is number of correctly answered items within 30 seconds, averaged across two sets of problems. Source: Salthouse and Babcock 1991.

Letter Comparison: Determine whether pairs of letter strings are the same or different. The primary measure is number of correctly answered items within 30 seconds, averaged across two sets of problems. Source: Salthouse and Babcock 1991.

4.2.3.3 Working memory

The Working memory (WM) N-back composite was comprised of the d' scores of the 2-back, 3-back, Dual 2-back, and Dual 3-back.

Single N-back: Determine whether the current letter presented matches the letter presented two or three items back. The primary measure of d-prime was computed separately for the 2-back and 3-back conditions. Reaction times on correct trials were also analyzed. Source: Kirchner 1958; Kane et al. 2007.

Dual N-back (administered in the MRI): Determine whether simultaneously presented auditory and visual stimuli match stimuli presented one, two, or three items ago. The primary measure of d-prime was computed separately for the two-back and three-back conditions. Reaction times on correct trials were also analyzed. Source: Jaeggi et al. 2003, 2007.

4.2.4 Behavioral Results Analysis

Training with Mind Frontiers produced significant gains in performance not only across all training tasks, but also to three cognitive constructs, working memory (WM), reasoning reaction time (GFRT), and perceptual speed (PS). The gain scores for the training and transfer tasks were calculated by taking the difference between post- and pre-test scores, and dividing by the standard deviation of the pre-test score. In the current study, we use these transfer scores as predictors, as well as training-induced gain scores from three of the training games: Supply Run (SR), Sentry Duty (SD), and Safe Cracker (SC). These three training games were chosen because improvement during training was positively correlated with transfer to the working memory construct (Baniqued et al. 2015), indicating that training-induced gains in these tasks may play a role in WM transfer.

4.2.5 MRI Processing and Analysis Details

4.2.5.1 MPRAGE Acquisition Parameters

All structural, high-resolution, T1-weighted MPRAGE images were collected on a 3 T Siemens Trio scanner, collected in ascending order parallel to the anterior and posterior commissures with the following parameters: Dataset 1: Echo time (TE) = 2.32 ms, repetition time (TR) = 1900 ms, field of view (FOV) = 230 mm, voxel size 0.9mm^3, and flip angle = 8°.

4.2.5.2 Structural MRI Preprocessing

All preprocessing of structural images was performed with FSL and Freesurfer. After being skull stripped, each structural image was linearly registered to the MNI 2mm brain template with 12 degrees of freedom (Smith 2004). We used the Freesurfer longitudinal stream to improve the robustness of the longitudinal analysis of cortical volume and thickness (Fischl 2012, Jovich 2009). This scheme creates within-subject anatomical templates that allow for less bias of analysis towards a particular time point (Thomas & Baker 2013). This process has been described elsewhere in detail (Metzler-Baddeley 2016, Fischl 2012), but briefly, this involves Talairach transformation, white matter and gray matter segmentation, automated topology correction, and intensity normalization. After this semi-automated procedure, images are all inspected manually to assess performance of the Freesurfer pipeline. Cortical thickness and volumes and subcortical volumes were calculated using Freesurfer using the Destrieux atlas, which parcellates the cortex into 74 gyral-sulcal regions per hemisphere (Destrieux et al. 2010). Previous studies have explored the association between either cortical thickness or brain volume with learning, but the relative information available in these two metrics with regards to individual differences in learning is poorly understood. Furthermore, contributions of subcortical and white matter volumes are rarely addressed in the same sample as cortical volumes and thicknesses. To address these questions, we used the brain volumes and cortical thickness data to create five separate datasets: Thickness, in which we only include the 150 cortical thickness variables; Volume, which only includes the 148 cortical volume features; Whole Brain, which contains the Volume dataset as well as 63 additional subcortical, white matter, and cerebellar volumes; Thickness + Volume, which contains the combined Thickness and Volume datasets; Thickness + Whole Brain, which contains both the Thickness and Whole Brain datasets. We extracted thickness and volume data from both the pre-training structural images (Pre), and to assess change in thickness and volume, we subtracted the pre thicknesses and volumes from the post data (Change). To explore differential contributions of variables from both the Pre and Change datasets, we predicted each of the six scores with ten different datasets, five Pre datasets and five Change datasets.

4.2.5.3 RS Acquisition and Processing

For all functional scans, we used an echo-planar imaging (EPI) sequence with Blood Oxygenation Level Dependent (BOLD) contrast (FoV Read: 220, 2.4 × 2.4 × 3.5 mm voxel size, TR = 2000 ms, TE = 25 ms, and flip angle = 80). BOLD repetitions: 180 volumes.

Preprocessing of the functional data was performed with C-PAC (http://fcp-indi.github.io/), which is an open source, configurable, and automated processing pipeline for resting functional MRI. This package utilizes tools from a variety of publicly available packages, such as FSL, AFNI, and ANTS. After being skull stripped with FSL, each structural image was registered to the MNI 2mm brain template and functional images were registered to the anatomical image using ANTS. These transformations were combined to register the functional images to the MNI 2mm brain template. Slice timing and motion correction using Friston's 24 parameter motion model were applied to the functional data to remove effects of motion from the dataset. To account for motion related artifacts further, we also performed motion scrubbing such that TRs with excessive motion (greater than 0.5mm frame displacement), were removed, along with 2 subsequent TRs. We excluded subjects with less than 120 volumes.

We performed nuisance signal regression by regressing out the mean signal of the white matter and CSF. We also used CompCorr, with 4 parameters, WM, CSF, and global as a method to remove excessive noise due to signal from the white matter, CSF, and global signal. Finally, we applied a temporal band pass filter from 0.01 to 0.1 Hz, and a spatial smoothing kernel of 6mm FWHM.

4.2.5.4 Connectivity

We chose the nodes for the default mode (DMN), frontal parietal (FPN), cingulo-opercular (CON), and subcortical (SUB) networks from the Petersen 264 node network atlas (Power *et al.* 2011). The CON contained nodes in the supramarginal gyrus, superior frontal gyrus, insular cortex, central opercular cortex, cingulate gyrus, and paracingulate gyrus. The DMN was made up of nodes in the frontal pole, middle frontal gyrus, PCC, lateral occipital cortex, temporal pole, precuneous, superior lateral occipital cortex, middle temporal gyrus, and angular gyrus. The FPN contained nodes in precentral gyrus, middle frontal gyrus, posterior supramarginal gyrus, inferior temporal gyrus, frontal pole, superior parietal lobe, and superior lateral occipital cortex. The SUB network contained nodes in the thalamus, brain stem, putamen, and pallidum. For a full listing of network regions and their anatomical assignments, see Power *et al.* 2011.

To create the internetworks, we assessed only the connections between each of the networks and excluded all within network connections. For example, for the SUBxFPN network, we exclusively extracted correlations between the SUB and FPN, and excluded all intranetwork connections. To extract and process each network's connectivity information, we calculated node-averaged timecourses within each subject's network of interest and we used these timecourses to create correlation matrices that would become features for our machine learning pipeline.

4.2.5.5 Structural and Resting State Network Overlap.

The structural and functional data in our sample were transformed to two different atlases, each with a corresponding standard space. We used correspondence between the regional labeling of the two atlases, the Harvard-Oxford cortical atlas locations and labeling, and visual inspection of both atlas maps to identify whether a given structural region corresponded to the same region in a given network from Power *et al.* 2011. Power *et al.'s* resting state atlas was computed at the voxel level (Figure 4.4), and clustering was applied to every voxel depending on its global connectivity pattern. The clustering that emerged matched the node clustering that was applied as well, demonstrating the strength of reliability of these functional connectivity networks. Furthermore, given that these networks were computed at both the voxel and nodal level, we can be confident that structural regions that correspond to a given functional atlas region will play a role in that same resting state network. This is the basis upon which we can compare the structural and resting state network results in the current work.

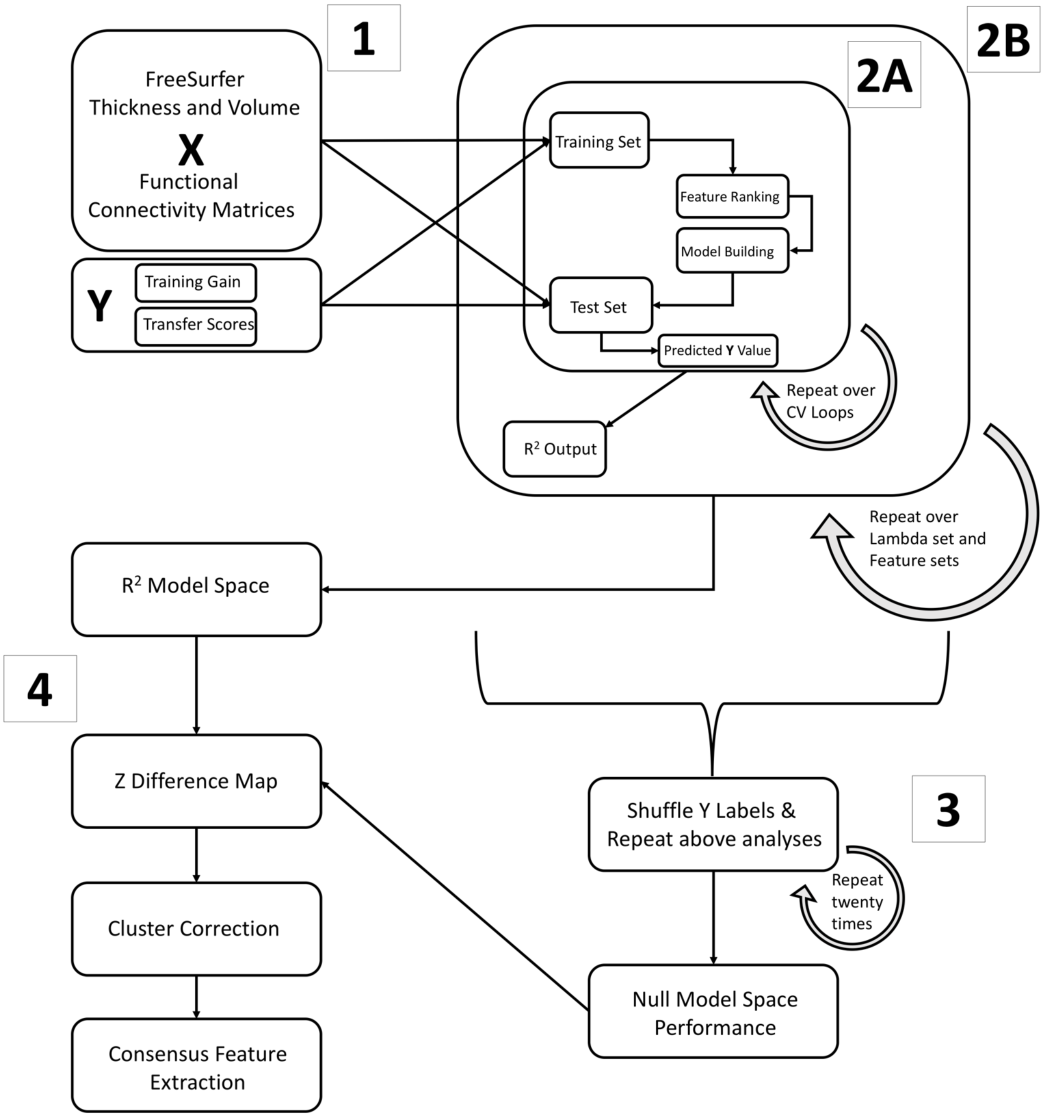

Figure 4.1:
This figure provides a visual flowchart of the methods used in the current work. Analysis begins with the structural features extracted from Freesurfer, or the resting state connectivity features (1), and then it continues through model building (2A), cross validation (2A), and repeating our analysis across feature and lambda space (2B). Finally, we create the null model distribution by shuffling the Y labels (3), and we assess the difference between the model space and the null model space in the form of a Z-map (4). We apply a cluster correction to this Z map and extract consensus features (4).

4.2.6 Machine Learning Methods

4.2.6.1 Predicting training gain and transfer

To make our predictions we used either pre-training or pre-post differences in cortical thicknesses, brain volumes, or resting state connectivity. We used Sci-Kit Learn to perform ridge regression (Figure 4.1: Section 1)(Buitinck et al. 2013). While this algorithm does not achieve predictive accuracies as high as other more advanced methods, such as random forests and support vector machines (SVMs), ridge regression offers an important advantage in terms of ease of interpretability and comparability to GLMs (Kotsiantis, Zaharakis, & Pintelas, 2007).

Over each leave-one-subject-out cross-validation (LOSO-CV) loop, we trained models of a range of sizes, and we used these models to predict scores of the left-out subject (Figure 4.1: Section 2A). Within each cross validation loop, we ranked all features in the training group by absolute correlation strength. We created feature sets of the top 1 to *n* features, by groups of 5 to save processing time, and predicted each left out subject using each of these feature sets. We calculated R^2 values by squaring the correlation of predicted and actual score values.

In order to assess the true predictive accuracy and overfitting in our models, we created a distribution of null models (Figure 4.1: Section 3). To create these null models, we randomized the score values for the subjects, and repeated the procedure outlined above to predict these values. We repeated this procedure 20 times to create a distribution of null models. The performance accuracy of our models were compared to the mean performance of our null models.

4.2.6.2 Assessing Reliability of Prediction

Rather than using cross validation to select an optimal lambda value for each model, we repeated our predictions and null model testing at a range of lambda values (Figure 4.1: Section

2B). While computationally intensive, this method allows us to more fully explore the space of possible models, extract clusters of predictive models, and assess the reliability of our models independent of hyperparameter value.

We assessed how our model performed over the full model space compared to the distribution of null models, and we transformed the difference in prediction accuracy between the model and null model distribution to a Z score distribution, representing the number of standard deviations that the model outperforms chance (Figure 4.1: Section 4). To get a more accurate assessment of average model prediction in a region and account for variation in neighboring models, we first smoothed the data using a 2 x 3 FWHM Gaussian kernel before applying cluster correction. We then removed all models that predicted less than 1 standard deviation above the null distribution to remove weakly predicting models, and then to remove small noisy predicting clusters we removed all clusters of models with less than 12 models. Finally, we removed clusters for datasets for which at least one cluster did not occupy at least 2% of the total model space, in order to only assess datasets that were most reliably predictive of the outcome variable. We performed cluster correction to not only assess the relative predictive accuracy and reliability of each dataset across a range of models, but also to extract individual features that were reliably included in 90-100% of models in a given cluster, also known as consensus features. In this way, we extract only the features that most reliably belong to models that significantly predict our outcome variable. This is an important step because it allows us to look for common individual predictors across outcome variables. While a mass univariate analysis, such as the feature ranking commonly performed in machine learning pipelines (including this one), detects individual variables that are each correlated with the outcome variable, it does not offer any information on conditional relationships between each feature and the outcome variable, thereby identifying mostly groups of redundant variables (Guyon 2003). By exploring the model space more fully, and identifying sets of consensus features that are most the reliable predictors of our outcome variable, we are also able to identify features that may not be good predictors on their own, but that reliably predict the residual variance that is left after the strongest predictors have been included in the model. For a review on feature selection and how redundant and superficially 'uninformative' features can play an important role in predicting residual variance, see Guyon 2003.

4.2.6.3 Assessing Consensus Features

Consensus features (CFs) allow us to analyze the feature selection process and make inferences based on the features that are selected. CFs represent features that were reliably included in many models across both feature set sizes and model regularization parameters (lambda). For more detailed results of the consensus features in the structural and functional datasets, refer to Nikolaidis et al. 2016a, 2016b. Among consensus features in the resting state datasets, some of the nodes are dramatically over-represented, suggesting that edges leading to these nodes may be more important for prediction than others. We term these high degree nodes 'consensus hubs'. Similar to hubs in the graph theoretical sense, these consensus hubs (CHs) have a strong influence on the rest of the consensus features.

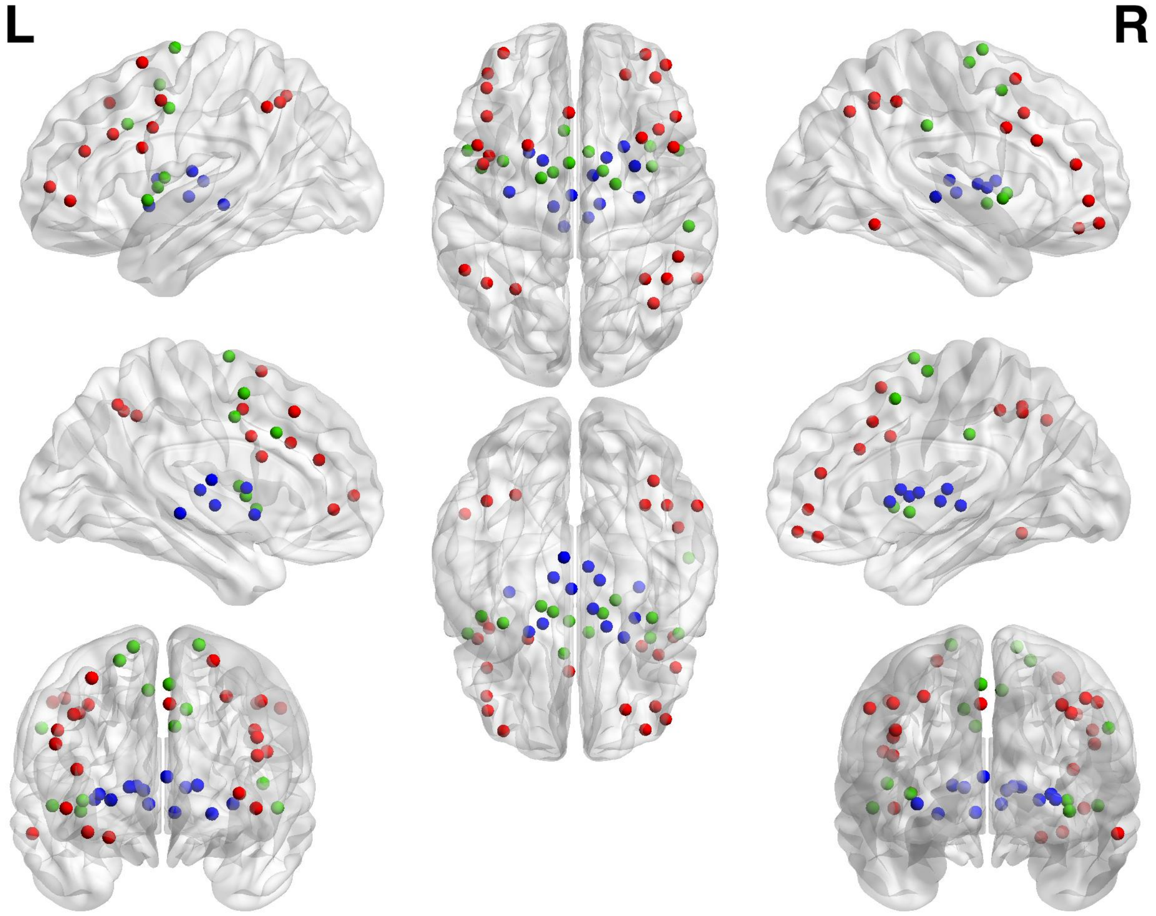

Figure 4.2:
Display of the approximate position of the nodes in 2mm MNI Template Space. SUB (Blue), CON (Green), and FPN (Red). From left to right- The 3 top row views are: Lateral view of Left Hemisphere, Top side view, Lateral View Right Hemisphere. The 3 mid row views are: Medial Left Hemisphere, Bottom side view, Medial Right Hemisphere, and the 2 bottom row views are: Front Side view, Back Side view.

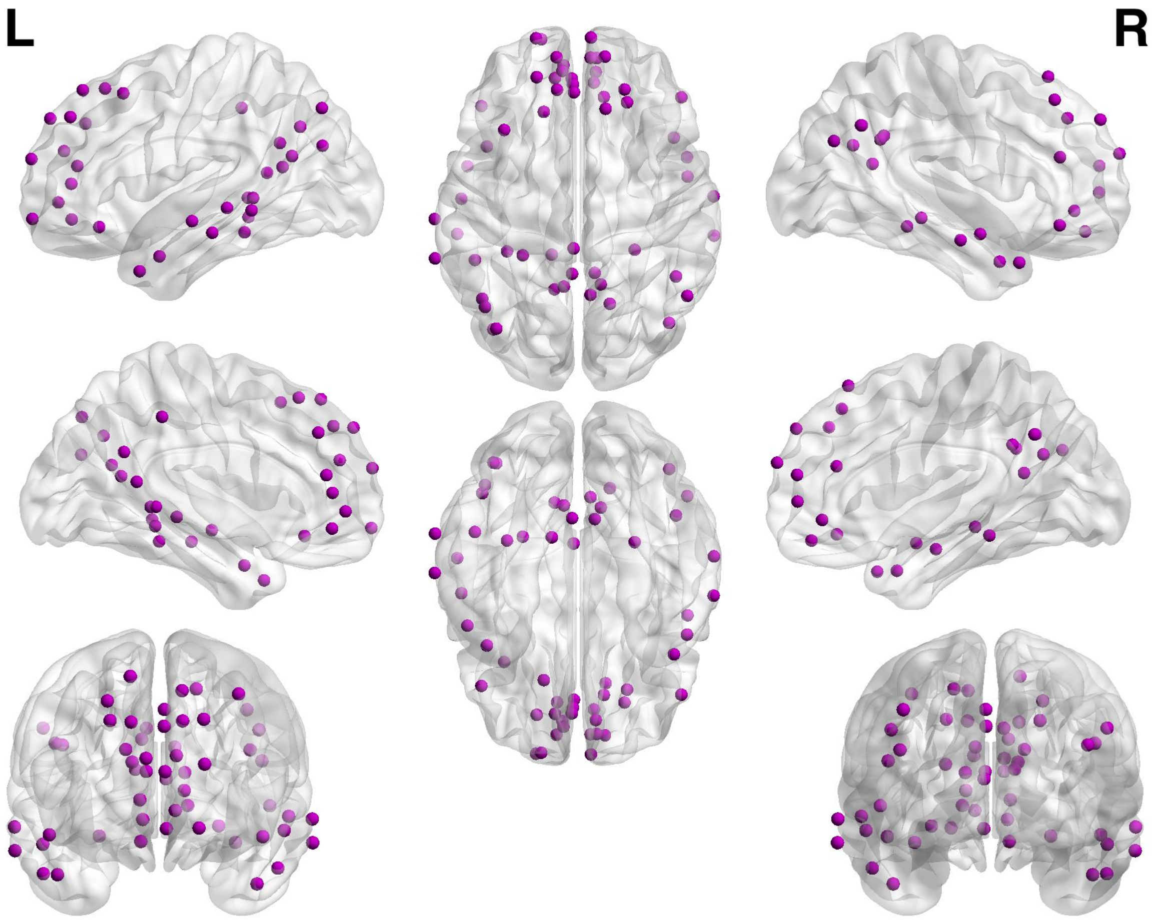

Figure 4.3:
Displays the approximate position and distribution of the Default Mode Network in 2mm MNI Template Space. From left to right- The 3 top row views are: Lateral view of Left Hemisphere, Top side view, Lateral View Right Hemisphere. The 3 mid row views are: Medial Left Hemisphere, Bottom side view, Medial Right Hemisphere, and the 2 bottom row views are: Front Side view, Back Side view.

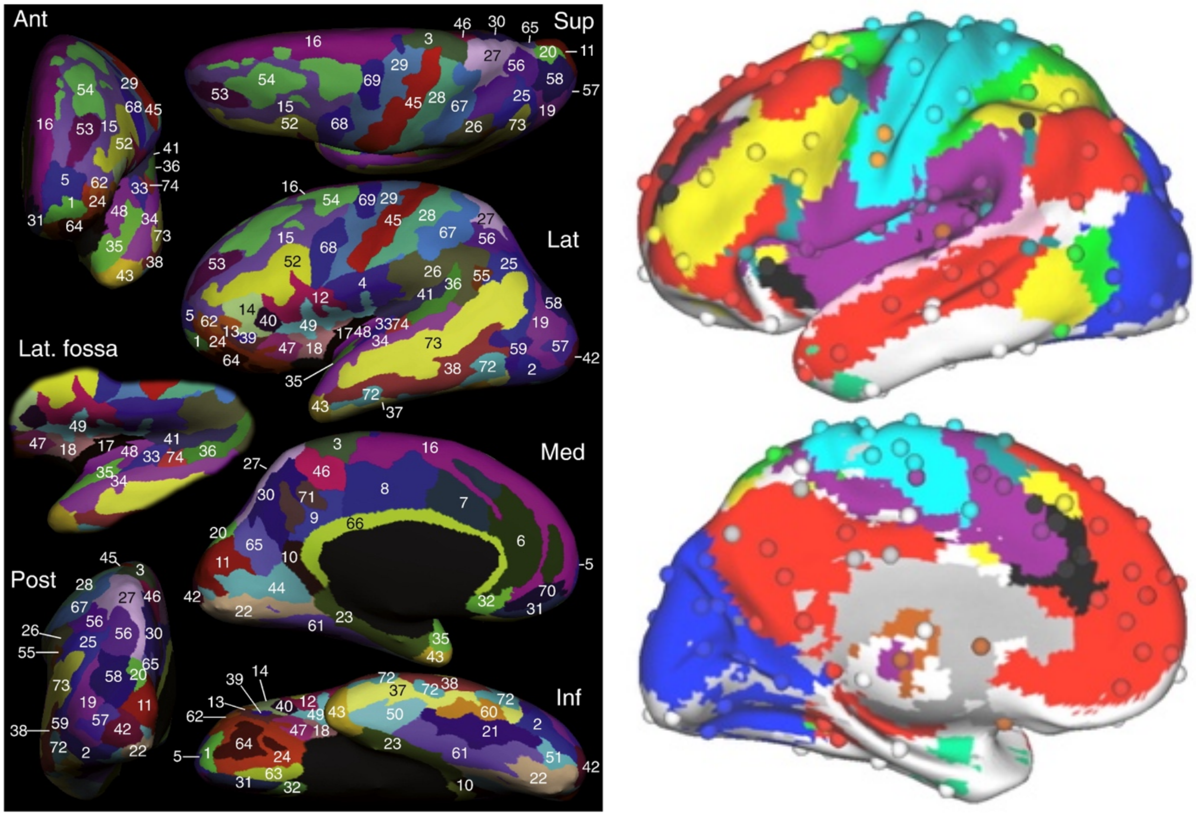

Figure 4.4:
Left: A view of the 74 regions in an inflated hemisphere, displaying both sulcal and gyral regions. Numerical indices refer to anatomical regions defined in Table 4.1. Views from top left to bottom right: Anterior (Ant), Superior (Sup), Lateral (Lat), Lateral Fossa (Lat fossa), Medial (Med), Posterior (Post), and Inferior (Inf). The Lateral Fossa view offers a better view of the gyral and sulcal regions in the insula and temporal cortices. Reproduced from (Destrieux et al. 2010). Right: A view of the all networks and most nodes from the resting state analysis. The top and bottom images correspond to lateral and medial views of a single hemisphere respectively. Red: DMN, Yellow: FPN, Purple: CON, Brown: SUB. Other networks included in Table 4.2: White: UNCERTAIN, Blue: VISUAL, Light Green: DAN (Dorsal Attention Network), Cyan: MOTOR, Black: SALIENCE, Teal: VAN (Ventral Attention Network). Reproduced from Power et al. 2011.

#	List of anatomical parcellations	Network	#	List of anatomical parcellations	Network
1	Fronto-marginal gyrus (of Wernicke) and sulcus	DMN	38	Middle temporal gyrus (T2)	DMN
2	Inferior occipital gyrus (O3) and sulcus	VISUAL	39	Horizontal ramus of anterior lateral sulcus	DMN
3	Paracentral lobule and sulcus	MOTOR	40	Vertical ramus of anterior lateral sulcus	SALIENCE
4	Subcentral gyrus (central operculum) and sulci	CON	41	Posterior ramus of lateral sulcus	CON
5	Transverse frontopolar gyri and sulci	FPN	42	Occipital pole	VISUAL
6	Anterior cingulate cortex (ACC)	MULTIP	43	Temporal pole	DMN
7	Middle anterior cingulate cortex (mACC)	CON	44	Calcarine sulcus	VISUAL
8	Middle posterior cingulate cortex (pMCC)	CON	45	Central sulcus (Rolando's fissure)	MOTOR
9	Posterior dorsal cingulate cortex (dPCC)	DMN	46	Marginal branch of cingulate sulcus	DMN
10	Posterior-ventral cingulate gyrus (vPCC)	DMN	47	Anterior insula circular sulcus	CON
11	Cuneus (O6)	VISUAL	48	Inferior insula circular sulcus	FPN
12	Opercular part of the inferior frontal gyrus	CON	49	Superior insula circular sulcus	CON
13	Orbital part of the inferior frontal gyrus	DMN	50	Anterior transverse collateral sulcus	UNCERTA
14	Triangular part of the inferior frontal gyrus	FPN	51	Posterior transverse collateral sulcus	VISUAL
15	Middle frontal gyrus (F2)	FPN	52	Inferior frontal sulcus	FPN
16	Superior frontal gyrus (F1)	MULTIP	53	Middle frontal sulcus	SALIENCE
17	Long insular gyrus and central sulcus of insula	CON	54	Superior frontal sulcus	MULTIPL
18	Short insular gyri	CON	55	Sulcus intermedius primus (of Jensen)	DMN
19	Middle occipital gyrus (O2 lateral occipital	VISUAL	56	Intraparietal and transverse parietal sulci	FPN
20	Superior occipital gyrus (O1)	DAN	57	Middle occipital sulcus and lunatus sulcus	VISUAL
21	Lateral occipito-temporal gyrus, fusiform gyrus	VISUAL	58	Superior occipital, transverse occipital sulci	VISUAL
22	Ligual medial occipito-temporal gyrus (O5)	VISUAL	59	Anterior occipital sulcus	FPN
23	Parahippocampal medial occipito-temporal gyrus	DMN	60	Lateral occipito-temporal sulcus	UNCERTA
24	Orbital gyri	DMN	61	Medial collateral sulcus and lingual sulcus	UNCERTA
25	Angular gyrus	DMN	62	Lateral orbital sulcus	FPN
26	Supramarginal gyrus	CON	63	Medial orbital sulcus (olfactory sulcus)	DMN
27	Superior parietal lobule (lateral part of P1)	DAN	64	Orbital sulci (H-shaped sulci)	MULTIPL
28	Postcentral gyrus	MOTOR	65	Parieto-occipital sulcus	VISUAL
29	Precentral gyrus	CON	66	Pericallosal sulcus	UNCERTA
30	Precuneus (medial part of P1)	DMN	67	Postcentral sulcus	FPN
31	Straight gyrus, Gyrus rectus	DMN	68	Inferior part of the precentral sulcus	MOTOR
32	Subcallosal area, subcallosal gyrus	DMN	69	Superior part of the precentral sulcus	MOTOR
33	Anterior transverse temporal gyrus (of Heschl)	DMN	70	Suborbital sulcus	DMN
34	Lateral aspect of the superior temporal gyrus	DMN	71	Subparietal sulcus	DMN
35	Planum polare of the superior temporal gyrus	CON	72	Inferior temporal sulcus	VAN
36	Temporal plane of superior temporal gyrus	CON	73	Superior temporal sulcus	DMN
37	Inferior temporal gyrus (T3)	CON	74	Transverse temporal sulcus	CON

Table 4.1:
A listing of the 74 Destrieux atlas regions and their resting state network assignment in the Power *et al.* networks. The first and fourth columns, labeled '#', correspond to each regions unique numbering according to the Destrieux Atlas and Figure 4.4 (Left). The second and fifth columns, labeled 'List of anatomical parcellations', correspond to shortened versions of assigned names of each structural region. The third and sixth columns, labeled 'Network' correspond to the assigned network for that region based on the full listing of 264 nodes from the Power *et al* atlas and their respective anatomical labels, as well as visual inspection and naming conventions from the Harvard-Oxford Cortical Atlas in FSL. DMN- Default mode network, CON- Cingulo-opercular network, FPN-Frontal-parietal network, DAN- Dorsal attention network, VAN- Ventral attention network.

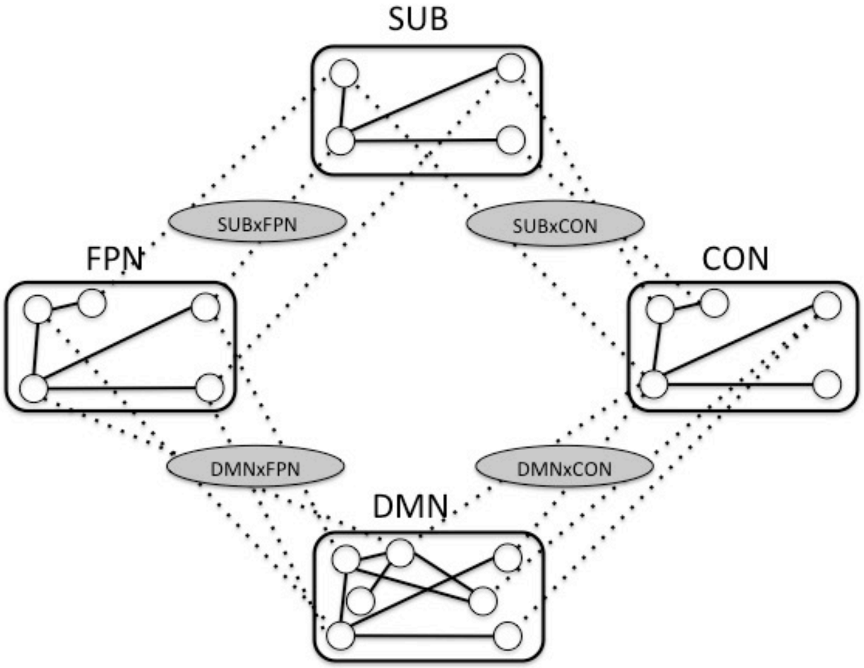

Figure 4.5:
This displays a summary of the eight networks originally investigated in Nikolaidis et al. 2016b. Each rounded box corresponds to a single network, which are depicted in Figures 4.2 & 4.3. As shown here, relationships between single networks are assessed to create internetworks, which are represented by the dashed lines between rounded boxes, and the gray ovals label their corresponding internetwork. Reproduced from Nikolaidis et al. 2016b.

Game	Description	Source
Supply Run (working memory)	Townspeople request items that belong to a certain category. There are five objects in each category, which correspond to stereotypical occupations of the "Wild West." Once the store is reached, the last item from each category must be selected. Difficulty level is manipulated by the number of requests and the number of categories.	Updating WM (Dahlin, Nyberg, et al. 2008)
Sentry Duty (working memory)	Sentries lift their lanterns while saying a word of the phonetic alphabet. The current word spoken and lantern lifted is compared to the word spoken/lantern lifted *n* times previously. There may be an audio match, a visual match, an audio and visual match, or no match between the current sentry and the one who spoke *n* times ago. Difficulty is manipulated by how far back the comparison is (e.g., 1-back, 2-back, 3-back).	Dual n-back (Jaeggi et al. 2008; Redick et al. 2012)
Safe Cracker (reasoning)	Safe combinations are determined by completing the next item in a series. Series may be letter-, number-, or day/month- based, and all are governed by some pattern or rule that must be determined and applied to select the next item in the series. Difficulty is manipulated by the difficulty of the patterns and the number of problems to solve within the given time limit.	Inductive Reasoning (Willis and Schaie 1986)

Table 4.2:
Description of the three training tasks investigated in the current study. Partially reproduced from Baniqued et al. 2015.

4.3 Results

4.3.1 Shared Network Consensus Features and Consensus Hubs

We investigated overlap between the resting state and structural predictors of training and transfer. We assessed the network assignment of each of the consensus features (CFs) and compared the strongest structural CFs to the resting state predictors in those same networks. Using the small set of strong CF predictors from the structural data, we assessed which network each of these CFs belonged to, and compared the CHs in these networks with these structural CFs. In the DMN, initial thickness of the dorsal posterior cingulate cortex, a hub of the DMN critical for internally oriented attention, and the initial volume of the right middle frontal sulcus, a region involved in working memory maintenance (Zhang et al. 2003), were negatively correlated with gain in SD. Initial connectivity of the DMN was also predictive of WM, SC, and SR, with several CHs in each analysis, but none in predicting SD. We found that consensus hubs in the DMN's middle temporal gyrus, a region previously associated with transfer of motor skill (Anguera et al. 2007), were present in prediction of WM, SC, and SR . We also found that the right paracingulate gyrus, a region implicated in attention and executive processing (Fornito et al. 2004; Wang et al. 2009, was a CH in predictions of both SR and WM. The frontal pole, which is also involved in attention and executive function(Wang et al. 2009), was a prominent CH in the predictions of WM. SC predictions also included CHs in both the right and left lateral occipital cortices, which play an important role in object recognition (Grill-spector et al. 1999) and the right precuneous which is a central node of the DMN and plays a role in internal visual-spatial imagery (Cavanna and Trimble 2006).

Pre-training volume of the left opercular inferior frontal gyrus (IFG), a region in the FPN and implicated in inhibition and working memory (Aron et al. 2003; Zhang et al. 2003), was predictive of transfer to GFRT. Similarly, we found that connectivity to the right middle frontal gyrus and inferior frontal gyrus was a CH of SC, the inductive reasoning training task. Furthermore, pre-training connectivity of the node between the right middle frontal gyrus and inferior frontal gyrus, both of which are involved in working memory (Aron et al. 2003; Zhang et al. 2003), was identified as a CH in predictions of SC. In the Pre SUBxFPN network, the left paracingulate gyrus, a node of the FPN involved in executive function (Fornito et al. 2004), was identified as a CH in predicting GFRT and 83.3% of the associations in this node were positively correlated with GFRT.

The pre-training thickness of the right insula, a critical region of the CON that plays a role in regulating the DMN (Menon and Uddin 2010; Uddin et al. 2011), was positively correlated with training gain in SD, and thickness of the posterior ramus of the lateral sulcus, a region of the CON that contains both the primary auditory and secondary somatosensory cortex, was also positively associated with training gain in SR. Pre training connectivity in the CON also demonstrated that the right insula was a consensus hub in predicting both GFRT and PS, with 100% of these node's connections being negatively correlated with transfer. The right superior frontal gyrus, a region involved in sustained attention and sensitive to working memory load was also a CH in predicting PS (Rypma and D'Esposito 1999; Dosenbach et al. 2006), and 88.9% of this CH's consensus features were negatively associated with transfer to PS. Furthermore we found many CON CH's that were associated with training gain to SC and SR (Nikolaidis et al. 2016b). In the Pre DMNxCON internetwork both SC and SR identified left central opercular cortex, right superior frontal gyrus, and two nodes in the right insula as CHs.

We found that training-induced changes in the structural regions of the FPN, in the thickness of the left triangular inferior frontal gyrus and the volume of the left opercular inferior frontal gyrus were both positively correlated with transfer to PS. In assessing corresponding training-induced changes in connectivity, while we did not find CHs in the inferior frontal gyrus, we found many CHs in the two regions on the superior and anterior bound of the inferior frontal gyrus, the middle frontal gyrus and the frontal pole. We found that change in connectivity of the frontal pole, in the FPN and DMNxFPN, were reliably positively correlated with gain in SC and GFRT (Nikolaidis et al. 2016b), suggesting that increases in the connectivity between the frontal pole and the DMN reflected greater generalized learning through training. Furthermore, we found that the middle frontal gyrus nodes in both the right and left hemispheres were CHs predicting GFRT as well, further supporting the notion that training-induced behavioral benefits are reflected in the interaction between frontal executive regions and the DMN.

4.3.2 Assessing CH and CF Overlap between WM Transfer and Training Scores

Given that WM and the three training scores were correlated, we assessed the amount of overlap between their respective predictors in both the structural and resting state datasets. In the structural data, we found a large set of overlapping consensus features at Pre that were included in clusters predicting WM and SC, and several of these regions are part of the DMN or CON. In the DMN, both the left fronto-marginal gyrus, which is involved in higher order pursuit of goal related activity (Ramnani and Miall 2004), and right precuneous, which is involved in internally oriented cognition, were CFs for both WM and SC (Cavanna and Trimble 2006). In the CON the left short insular gyrus, which plays an important role in regulating the DMN (Menon and Uddin 2010), and the right middle cingulate cortex which plays a role in attentional monitoring (Alexander and Brown 2012), were CFs for both WM and SC. Furthermore, WM and SC also overlapped in their predictions in training-induced change in several regions of the DMN, including the left superior frontal gyrus a region critical for high load working memory processing (Rypma and D'Esposito 1999; Boisgueheneuc et al. 2006), the left anterior transverse temporal gyrus which contains the primary auditory cortex, and the marginal branch of the left cingulate, a part of the precuneous and involved in internally oriented processing (Cavanna and Trimble 2006). We also found several CFs common between WM and SR predictions, including regions in the FPN that are important for working memory such as the left supramarginal gyrus (Silk et al. 2010), and left inferior frontal gyrus (Rypma and D'Esposito 1999; Zhang et al. 2003), as well as regions in the DMN and CON such as the left anterior transverse temporal gyrus (DMN), which contains the primary auditory cortex (DMN), and the left inferior frontal sulcus (CON), which is also involved in working memory (Gruber 2001; Barbey et al. 2013). We found very little overlap in CFs between predictions of WM and SD, though SD is the training task that most closely mirrors WM. Only a small section of the right intraparietal sulcus (DMN), which is involved in working memory and attention (Silk et al. 2010) was shared as a CF between the two predictions.

In the resting state datasets, the WM and SC scores shared 26 CFs in the Pre DMN, including several connections to the frontal pole, precuneous, and paracingulate gyrus. The WM and SR scores shared 28 CFs in the Pre DMN which also included several connections to regions involved in executive function such as the frontal pole (Pochon et al. 2002), and the paracingulate gyrus (Fornito et al. 2004), as well as the lateral occipital cortex which is involved in object recognition and visual working memory (Grill-spector et al. 1999; Silk et al. 2010), while WM and SD shared only a single CF. Furthermore, in the DMNxCON internetwork, predictions of WM and SC shared 11 CFs, with several connections to the paracingulate gyrus, insula, and frontal pole, WM and SR shared 9 CFs with several emanating from the insula, and WM and SD had no common CFs. It is important to note that in internetworks such as the DMNxCON, all CHs that are identified are comprised exclusively of connections between that CH and the other network. This means that the paracingulate gyrus, and frontal pole's most reliably predictive connections were those to the CON, while the insula's connectivity was reliably predictive. The CHs identified by the overlap between scores in the DMN and DMNxCON networks are remarkably conserved, and point to the importance of pre existing connectivity in these executive regions of the DMN and CON for gains in both working memory updating and inductive reasoning training, as well as transfer to WM. In a similar fashion as the structural data, we found that WM and SD shared only a single overlapping consensus feature in the DMN. Given the psychological similarity between WM and SD, as well as the correlation between the two improvement scores, the neural overlap hypothesis would have predicted common neural correlates between WM and SD.

4.4 Discussion

While the structural CFs that predicted training and transfer were predominantly regions involved in lower level sensory processing and attention (Nikolaidis et al. 2016a), we found overlap in the structural and resting state correlates of several structural regions associated with higher cognitive function. Pre-existing and training-induced differences in regions of the FPN, including the inferior frontal gyrus, a region implicated in inhibitory processing and working memory (Aron et al. 2003; Zhang et al. 2003; Colom et al. 2013), were predictive of training gain and transfer in the structural and resting state datasets. This suggests that the functional network connectivity and structural size of this region may both reflect propensity for greater training gain and transfer due to more effective integration of this executive function area before training. Furthermore, we found that training and transfer scores were predicted by pre-training thickness and network connectivity of the right insula. This region is a hub of the CON and is known to play a critical role in the regulation of the DMN (Menon and Uddin 2010; Uddin et al. 2011; Sestieri et al. 2014). We proposed that the development and deployment of sustained attention skills may be critical to enhancing learning and transfer during training through an increase in flow states (Nikolaidis et al. 2016b), and the current work suggests that pre-training structural thickness of the insula may contribute to the ability of the CON to regulate the DMN as well. Interestingly, prior work has demonstrated the importance of DMN's middle temporal gyrus in motor transfer (Anguera et al. 2007), and in the current work we found that this region was a DMN CH for WM, SC, and SR, all of which are metrics of higher learning. This region is known to play a role in semantic memory and visual perception (Vandenberghe et al. 1996; Visser et al. 2012), but it also plays a role in multimodal sensory integration (Onitsuka et al. 2004). Given that this region plays a role both in generalization of motor and cognitive skills and in the integration of multimodal sensory information, interactions between this region and the striatum and other cortical regions may play a particularly important role in broadening the initially unitized representations of skills into more generalized skills that can be deployed in other contexts.

We found markedly less overlap in the predictions for SD & WM compared to other training tasks, although SD and WM are both n-back based and thus likely share similar cognitive and neural processes. This raises important questions for the most simplistic interpretation of the neural overlap hypothesis. This simple version states that if initially the training and transfer tasks share the same neural correlates, which is highly likely with SD and WM given the overwhelming similarity between them, then training in one will lead to transfer in the other. It was the case in the current results that although WM and SD were very similar tasks, and likely shared underlying neural architecture, and even behaviorally paralleled one another in training-induced gain, our functional and structural results suggest that the underlying neural mechanisms during WM transfer and gain to SD are quite different. In fact so completely different that one wonders whether if SD had been removed would transfer to WM still manifest? To assess why it may be the case that two tasks could share underlying neural similarity but not drive transfer to one another, it is important to consider the skill acquisition process that occurs during training. In the case of training with SD, there are several complex skills that are being concurrently developed. Participants must take in independent streams of auditory and visual information, store these representations, and repeatedly access them on each trial to both update the stored set, and to the assess whether the items n trials previously match the stimuli of the current trial. There are many levels of processing that may be preventing individuals from performing better at this task, such as a limited capacity for auditory and visual information in working memory, high levels of susceptibility to lure trials due to poor interference resolution, or poor storage accuracy due to low quality sensory representations. As level gain in SD is specifically measured by increasing sizes of the n-back level, training gain in SD is primarily a reflection of increased storage capacity. However, the tasks used to assess WM transfer did not depend on improvement of storage capacity, as they largely were dependent on low levels of storage, 2 and 3 back. Instead, improvements in the d' measurements used as transfer metrics depend critically on both the speed and accuracy of storage and retrieval, as well as the resolution of interferences during lure trials.

A modified instantiation of the neural overlap theory may explain the current result in the context of skill acquisition and generalization, and we have suggested this to be the case previously as well (Nikolaidis et al. 2016a). Rather than transfer and training tasks overlapping in their dominant brain activity, perhaps instead the neural overlap needs to occur in brain regions and networks which form a processing bottleneck for greater performance on transfer tasks. More specifically, perhaps the improvement of skills in those bottleneck domains will lead to improvement of the transfer tasks not through improvement of a specific aspect of WM, such as storage capacity or accuracy in manipulation of information, but rather through the development of faster or more accurate low level processing through training on these tasks. Rather than depending on improvement in the size of memory capacity, perhaps training gains in SC and SR are more dependent on rapid and accurate processing of information in memory, which are in turn dependent on the low level sensory and attention processing correlates found in our previous structural analysis (Nikolaidis et al. 2016a), and in the regulation of default mode activity by the CON, as found in our previous resting state analysis (Nikolaidis et al. 2016b).

4.5 Conclusion

We find that support for overlapping structural and resting state network correlates of training and transfer. Furthermore, both structural and resting state correlates of training gain and transfer demonstrated the same dissociation in predictive overlap between working memory and the training tasks, suggesting that a refinement of the neural overlap hypothesis that takes into consideration skill learning within processing bottlenecks could be critical to training-induced transfer.

Acknowledgements: This work was supported by a Beckman Fellowship to AN. The authors would like to thank Brad Sutton, Tracey Wszalek for technical guidance, and Nancy Dodge, Holly Tracy, and members of the Lifelong Brain and Cognition Laboratory for assistance with performing the experiments.

4.6 References

Alexander WH, Brown JW. 2012. Medial Prefrontal Cortex as an action-outcome predictor. Nat Neurosci. 14:1338–1344.

Anguera J a, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, Kong E, Larraburo Y, Rolle C, Johnston E, Gazzaley a. 2013. Video game training enhances cognitive control in older adults. Nature. 501:97–101.

Anguera J a, Russell C a, Noll DC, Seidler RD. 2007. Neural correlates associated with intermanual transfer of sensorimotor adaptation. Brain Res. 1185:136–151.

Aron AR, Fletcher PC, Bullmore ET, Sahakian BJ, Robbins TW. 2003. Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. Nat Neurosci. 6:115–116.

Bahrick HP. 1979. Maintenance of knowledge: Questions about memory we forgot to ask. J Exp Psychol Gen. 108:296–308.

Ball K, Berch DB, Helmers KF, Jobe JB, Leveck MD, Marsiske M, Morris JN, Rebok GW, Smith DM, Tennstedt SL, Unverzagt FW, Willis SL. 2002. Effects of Cognitive Training Interventions With Older Adults. Jorunal Am Med Assoc. 288:2271–2281.

Baniqued PL, Allen CM, Kranz MB, Johnson K, Sipolins A, Dickens C, Ward N, Geyer A, Kramer AF. 2015. Working memory, reasoning, and task switching training: Transfer effects, limitations, and great expectations? PLoS One. 10:1–29.

Barbey AK, Colom R, Paul EJ, Grafman J. 2013. Architecture of fluid intelligence and working memory revealed by lesion mapping. Brain Struct Funct. 219:485–494.

Bennett G, Seashore H, Wesman A. 1947. Differential aptitude tests, Form A. San Antonio: The Psychological Corporation.

Bjork BRA, Linn MC. 2006. The Science of Learning and the Learning of Science Introducing Desirable Difficulties. Science (80-). 19:6–7.

Boisgueheneuc F Du, Levy R, Volle E, Seassau M, Duffau H, Kinkingnehun S, Samson Y, Zhang S, Dubois B. 2006. Functions of the left superior frontal gyrus in humans: A lesion study. Brain. 129:3315–3328.

Buitinck L, Louppe G, Blondel M. 2013. API design for machine learning software: experiences from the scikit-learn project. arXiv Prepr arXiv …. 1–15.

Cavanna AE, Trimble MR. 2006. The precuneus: a review of its functional anatomy and behavioural correlates. Brain. 129:564–583.

Colom R, Burgaleta M, Román FJ, Karama S, Alvarez-Linera J, Abad FJ, Martínez K, Quiroga MÁ, Haier RJ. 2013. Neuroanatomic overlap between intelligence and cognitive factors: morphometry methods provide support for the key role of the frontal lobes. Neuroimage. 72:143–152.

Crone EA, Wendelken C, Van Leijenhorst L, Honomichl RD, Christoff K, Bunge SA. 2009. Neurocognitive development of relational reasoning. Dev Sci. 12:55–66.

Dahlin E, Neely AS, Larsson A, Bäckman L, Nyberg L. 2008. Transfer of learning after updating training mediated by the striatum. Science. 320:1510–1512.

Dahlin E, Nyberg L, Bäckman L, Neely AS. 2008. Plasticity of executive functioning in young and older adults: immediate training gains, transfer, and long-term maintenance. Psychol Aging. 23:720–730.

Destrieux C, Fischl B, Dale A, Halgren E. 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage. 53:1–15.

Dosenbach NUF, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HC, Burgund ED, Grimes AL, Schlaggar BL, Petersen SE. 2006. A core system for the implementation of task sets. Neuron. 50:799–812.

Dosenbach NUFN, Nardos B, Cohen AL, Fair D a, Power JD, Church J a, Nelson SM, Wig GS, Vogel AC, Lessov-Schlaggar CN, Barnes KA, Dubis JW, Feczko E, Coalson RS, Pruett JR, Barch DM, Petersen SE, Schlaggar BL. 2010. Prediction of Individual Brain Maturity using fMRI. Science (80-). 329:1358–1361.

Doyon J, Bellec P, Amsel R, Penhune V, Monchi O, Carrier J, Lehéricy S, Benali H. 2009. Contributions of the basal ganglia and functionally related brain structures to motor learning. Behav Brain Res. 199:61–75.

Dye MWG, Green CS, Bavelier D. 2009. Increasing Speed of Processing With Action Video Games. Curr Dir Psychol Sci. 18:321–326.

Ekstrom RBR, French JJW, Harman HH, Dermen D. 1976. Manual for kit of factor-referenced cognitive tests. Princet NJ Educ Test Serv. 102:117.

Fornito A, Yücel M, Wood S, Stuart GW, Buchanan J, Proffitt T, Anderson V, Velakoulis D, Pantelis C. 2004. Individual Differences in Anterior Cingulate / Paracingulate Morphology Are Related to Executive Functions in Healthy Males. 424–431.

Graybiel AM. 2005. The basal ganglia: learning new tricks and loving it. Curr Opin Neurobiol. 15:638–644.

Green CS, Bavelier D. 2003. Action video game modifies visual selective attention. Nature. 423:534–537.

Grill-spector K, Kushnir T, Edelman S, Avidan G, Itzchak Y, Malach R. 1999. Differential Processing of Objects under Various Viewing Conditions in the Human Lateral Occipital Complex. 24:187–203.

Gruber O. 2001. Effects of domain-specific interference on brain activation associated with verbal working memory task performance. Cereb Cortex. 11:1047–1055.

Guyon I. 2003. An Introduction to Variable and Feature Selection 1 Introduction. 3:1157–1182.

Hastie T, Tibshirani R, Friedman J. 2009. The elements of statistical learning.

Hikosaka O, Nakamura K, Sakai K, Nakahara H. 2002. Central mechanisms of motor skill learning. Curr Opin Neurobiol. 12:217–222.

Jaeggi SM, Buschkuehl M, Etienne A, Ozdoba C, Perrig WJ, Nirkko AC. 2007. On how high performers keep cool brains in situations of cognitive overload. Cogn Affect Behav Neurosci. 7:75–89.

Jaeggi SM, Buschkuehl M, Jonides J, Perrig WJ. 2008. Improving fluid intelligence with training on working memory. Proc Natl Acad Sci U S A. 105:6829–6833.

Jaeggi SM, Seewer R, Nirkko AC, Eckstein D, Schroth G, Groner R, Gutbrod K. 2003. Does excessive memory load attenuate activation in the prefrontal cortex? Load-dependent processing in single and dual tasks: Functional magnetic resonance imaging study. Neuroimage. 19:210–225.

Kane MJ, Conway AR, Miura TK, Colflesh GJ. 2007. Working memory, attention control, and the N-back task: a question of construct validity. J Exp Psychol Learn Mem Cogn. 33:615–622.

Kirchner WK. 1958. Age differences in short-term retention of rapidly changing information. J Exp Psychol. 55:352–358.

Kwok V, Niu Z, Kay P, Zhou K, Mo L, Jin Z, So K-F, Tan LH. 2011. Learning new color names produces rapid increase in gray matter in the intact adult human cortex. Proc Natl Acad Sci. 108:6686–6688.

Landauer TK, Bjork RA. 1978. Optimum rehearsal patterns and name learning. Pract Asp Mem. 625–632.

Lövdén M, Wenger E, Mårtensson J, Lindenberger U, Bäckman L. 2013. Structural brain plasticity in adult learning and development. Neurosci Biobehav Rev. 37:2296–2310.

Meier TB, Desphande AS, Vergun S, Nair V a., Song J, Biswal BB, Meyerand ME, Birn RM, Prabhakaran V. 2012. Support vector machine classification and characterization of age-related reorganization of functional brain networks. Neuroimage. 60:601–613.

Melby-Lervåg M, Hulme C. 2012. Is Working Memory Training Effective? A Meta-Analytic Review. Dev Psychol.

Menon V, Uddin LQ. 2010. Saliency, switching, attention and control: a network model of insula function. Brain Struct Funct. 214:1–13.

Nikolaidis A, Goatz D, Baniqued PL, Kranz MB, Barbey AK, Smaragdis P, Kramer AF. 2016a. Brain Volume and Thickness Predictions of Cognitive Skill Training and Transfer to Working Memory, Fluid Intelligence, and Perceptual Speed Using Machine Learning. Prep.

Nikolaidis A, Goatz D, Baniqued PL, Kranz MB, Barbey AK, Smaragdis P, Kramer AF. 2016b. Machine Learning Predictions of Cognitive Skill Training and Transfer to Working Memory, Fluid Intelligence, and Perceptual Speed Using Resting State Functional Connectivity. Prep.

Nikolaidis A, Goatz D, Smaragdis P, Kramer A. 2015. Predicting Skill-Based Task Performance and Learning with fMRI Motor and Subcortical Network Connectivity. 2015 Int Work Pattern Recognit NeuroImaging. 93–96.

Nikolaidis A, Voss MW, Lee H, Vo LTK, Kramer AF. 2014. Parietal plasticity after training with a complex video game is associated with individual differences in improvements in an untrained working memory task. Front Hum Neurosci. 8:1–11.

Oelhafen S, Nikolaidis A, Padovani T, Blaser D, Koenig T, Perrig WJ. 2013. Increased parietal activity after training of interference control. Neuropsychologia. 51:2781–2790.

Olesen PJ, Westerberg H, Klingberg T. 2004. Increased prefrontal and parietal activity after training of working memory. Nat Neurosci. 7:75–79.

Onitsuka T, Shenton ME, Salisbury DF, Dickey CC, Kasai K, Toner SK, Frumin M, Kikinis R, Jolesz FA, McCarley RW. 2004. Middle and inferior temporal gyrus gray matter volume abnormalities in chronic schizophrenia: An MRI study. Am J Psychiatry. 161:1603–1611.

Patel R, Spreng RN, Turner GR. 2013. Functional brain changes following cognitive and motor skills training: a quantitative meta-analysis. Neurorehabil Neural Repair. 27:187–199.

Pochon JB, Levy R, Fossati P, Lehericy S, Poline JB, Pillon B, Le Bihan D, Dubois B. 2002. The neural system that bridges reward and cognition in humans: an fMRI study. Proc Natl Acad Sci U S A. 99:5669–5674.

Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church J a, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, Petersen SE. 2011. Functional network organization of the human brain. Neuron. 72:665–678.

Ramnani N, Miall RC. 2004. A system in the human brain for predicting the actions of others. Nat Neurosci. 7:85–90.

Ravens J. 1962. Advanced Progressive Matrices: Set II. London.

Rebok GW, Ball K, Guey LT, Jones RN, Kim HY, King JW, Marsiske M, Morris JN, Tennstedt SL, Unverzagt FW, Willis SL. 2014. Ten-year effects of the advanced cognitive training for independent and vital elderly cognitive training trial on cognition and everyday functioning in older adults. J Am Geriatr Soc. 62:16–24.

Redick TS, Shipstead Z, Harrison TL, Hicks KL, Fried DE, Hambrick DZ, Kane MJ, Engle RW. 2012. No Evidence of Intelligence Improvement After Working Memory Training: A Randomized, Placebo-Controlled Study. J Exp Psychol Gen. 142:359–379.

Rypma B, D'Esposito M. 1999. The roles of prefrontal brain regions in components of working memory: effects of memory load and individual differences. Proc Natl Acad Sci U S A. 96:6558–6563.

Salthouse TA. 2005. Relations between cognitive abilities and measures of executive functioning. Neuropsychology. 19:532–545.

Salthouse TA, Babcock RL. 1991. Decomposing adult age differences in working memory. Dev Psychol. 27:763–776.

Schmidt R a., Bjork R a., Schmidt BRA, Bjork R a., Schmidt R a., Bjork R a. 1992. New Conceptualizations of Practice: Common Principles in Three Paradigms Suggest New Concepts for Training. Psychol Sci. 3:207–217.

Seidler RD. 2010. Neural correlates of motor learning, transfer of learning, and learning to learn. Exerc Sport Sci Rev. 38:13.

Sestieri C, Corbetta M, Spadone S, Romani GL, Shulman GL. 2014. Domain-general Signals in the Cingulo-opercular Network for Visuospatial Attention and Episodic Memory. J Cogn Neurosci. 26:551–568.

Shen H, Wang L, Liu Y, Hu D. 2010. Discriminative analysis of resting-state functional connectivity patterns of schizophrenia using low dimensional embedding of fMRI. Neuroimage. 49:3110–3121.

Shipstead Z, Redick TS, Engle RW. 2012. Is working memory training effective? Psychol Bull. 138:628–654.

Silk TJ, Bellgrove MA, Wrafter P, Mattingley JB, Cunnington R. 2010. Spatial working memory and spatial attention rely on common neural processes in the intraparietal sulcus. Neuroimage. 53:718–724.

Singley MK, Anderson JR. 1989. The transfer of cognitive skill. Harvard University Press.

Thorell LB, Lindqvist S, Bergman Nutley S, Bohlin G, Klingberg T. 2009. Training and transfer effects of executive functions in preschool children. Dev Sci. 12:106–113.

Uddin LQ, Supekar KS, Ryali S, Menon V. 2011. Dynamic reconfiguration of structural and functional connectivity across core neurocognitive brain networks with development. J Neurosci. 31:18578–18589.

Vandenberghe R, Price C, Wise R, Josephs O, Frackowiak RS. 1996. Functional anatomy of a common semantic system for words and pictures. Nature. 383:254–256.

Visser M, Jefferies E, Embleton K V., Lambon Ralph M a. 2012. Both the Middle Temporal Gyrus and the Ventral Anterior Temporal Area Are Crucial for Multimodal Semantic Processing: Distortion-corrected fMRI Evidence for a Double Gradient of Information Convergence in the Temporal Lobes. J Cogn Neurosci. 24:1766–1778.

Vo LTK, Walther DB, Kramer AF, Erickson KI, Boot WR, Michelle W, Prakash RS, Lee H, Fabiani M, Gratton G, Simons DJ, Sutton BP, Wang MY. 2011. Predicting individuals' learning success from patterns of pre-learning MRI activity. PLoS One. 6:e16093.

Voss MW, Prakash RS, Erickson KI, Boot WR, Basak C, Neider MB, Simons DJ, Fabiani M, Gratton G, Kramer AF. 2011. Effects of training strategies implemented in a complex videogame on functional connectivity of attentional networks. Neuroimage.

Wang L, Liu X, Guise KG, Knight RT, Ghajar J, Fan J. 2009. Effective Connectivity of the Fronto-parietal Network during Attentional Control. J Cogn Neurosci. 543–553.

Wechsler D. 1997. WAIS-III: Wechsler adult intelligence scale. San Antonio: The Psychological Corporation.

Willis SL, Schaie KW. 1986. Trianing the elderly on the ability factors of spatial orientationand inductive reasoning. Psychol Aging.

Ye J, Wu T, Li J, Chen K. 2011. Machine Learning Approaches for the Neuroimaging Study of Alzheimer's Disease. IEEE Comput. 44:99–101.

Zatorre RJ, Fields R., Johansen-Berg H. 2013. Plasticity in Gray and White : Neuroimaging changes in brain structure during learning. Nat Neurosci. 15:528–536.

Zhang JX, Leung HC, Johnson MK. 2003. Frontal activations associated with accessing and evaluating information in working memory: An fMRI study. Neuroimage. 20:1531–1539.