Intelligence and Connectivity

Group 6: Jenny Hu, Coco Chen, Daniel Meier, Edward Taur, Ziv Schwartz, Jeremy Milkes

Table of Contents

I. Abstract

There are two imaging approaches to measuring the brain: MRI, which captures the brain structure in high resolution, and fMRI, which captures the measures of neuronal activation (oxygen level changes). We are looking to determine the type of relationship between intelligence and connectivity, specifically looking at indicators of working memory and their interaction with MRI and fMRI measures. Our data is obtained from a UCLA study (CNP) which included control patients and those diagnosed with schizophrenia, bipolar disorder, and ADHD.

First, we conduct exploratory analysis to identify variables that are of high significance to our two initial response variables: Verbal Working Memory (VWM_G) and Spatial Working Memory (SWM_G). We use several analysis techniques from random forest to scatterplot matrices in order to determine the important predictors to use in the final model, and find the following predictors to be significant: Age, Ethnicity, School_Yrs, Limbic_Global_Efficiency, Right.Amygdala, Left.Putamen, and non.WM.hypointesities.

Next, we determine from the exploratory data analysis that the two initial outcomes are highly correlated so we proceed to combine the two variables into a single measure: Average Working Memory score (average of VWM_G and SWM_G). With the new calculated outcome variable, Average Working Memory, we run a multiple linear regression model and reduce the model by identifying the significant predictors. We then look at the relationship between age, diagnosis and working memory through various plots.

Further, comparing the full model to the final reduced model, a P-value of 0.2582 indicates that our full model does just as good a job at predicting Average Working Memory as the reduced model. As age increases, working memory decreases with each diagnosis. In the order of most rapid decline with working memory with respect to age: bipolar disorder, control group, schizophrenia, and ADHD. With respect to the control group, the pattern of working memory does not change significantly over the different diagnoses although diagnosis can affect the starting value of working memory.

While we conclude that the there is no significant difference in the relationship between working memory and age across the different diagnoses. Additionally, we see that the diagnoses do affect the starting value of working memory. We can say that while there is no significant difference in how working memory declines across the different diagnoses, the working memory of a patient is more likely to be higher or lower depending on their diagnosis. This is certainly worthy of additional analysis in exploring the relationship between these diagnoses and working memory.

II. Statement of Problems

We aim to investigate how working memory and response inhibition change with fMRI by taking a look at how the brain and cognitive deficits change with age and disease. More specifically, we want to know if we can predict someone's working memory using their MRI measures, fMRI activity, and demographic covariates. Some questions we want to ask are: Does the relationship among cognitive ability, MRI, and fMRI behavior change as a person ages? Do issues with working memory become exaggerated with age and diagnosis? Do these patterns of cognitive ability change with psychiatric diagnosis?

III. Variables of the Study

Table I

fMRI Measures	MRI Measures
Visual_Global_Efficiency Somatomotor_Global_Efficiency Dorsal_Attention_Global_Efficiency Ventral_Attention_Global_Efficiency Limbic_Global_Efficiency Frorntoparietal_Global_Efficiency Default_Mode_Global_Efficiency	Left.Amygdala Right.Amygdala Left.Caudate Right.Caudate Left.Accumbens.area TotalGrayVol CortexVol CorticalWhiteMattervol Left.Putamen Right.Putamen Left.Pallidum Right.Pallidum Left.Hippocampus Right.Hippocampus WM.hypointensities non.WM.hypointensities

Table II

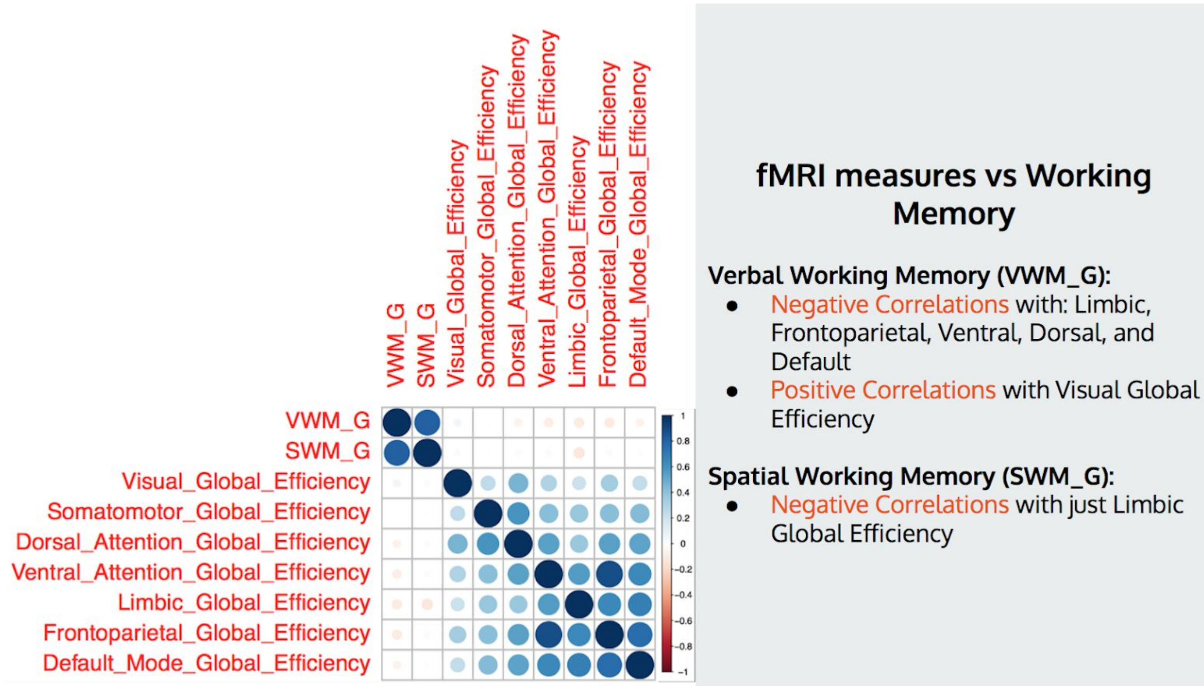
Categorical Variables	Description
DX	Diagnoses and control
Age (>30 and <=30)	Originally numeric. Due to skewness, we split age into two groups by at 30 as the divider as we saw a large divide in patients' ages near this age.

Table III

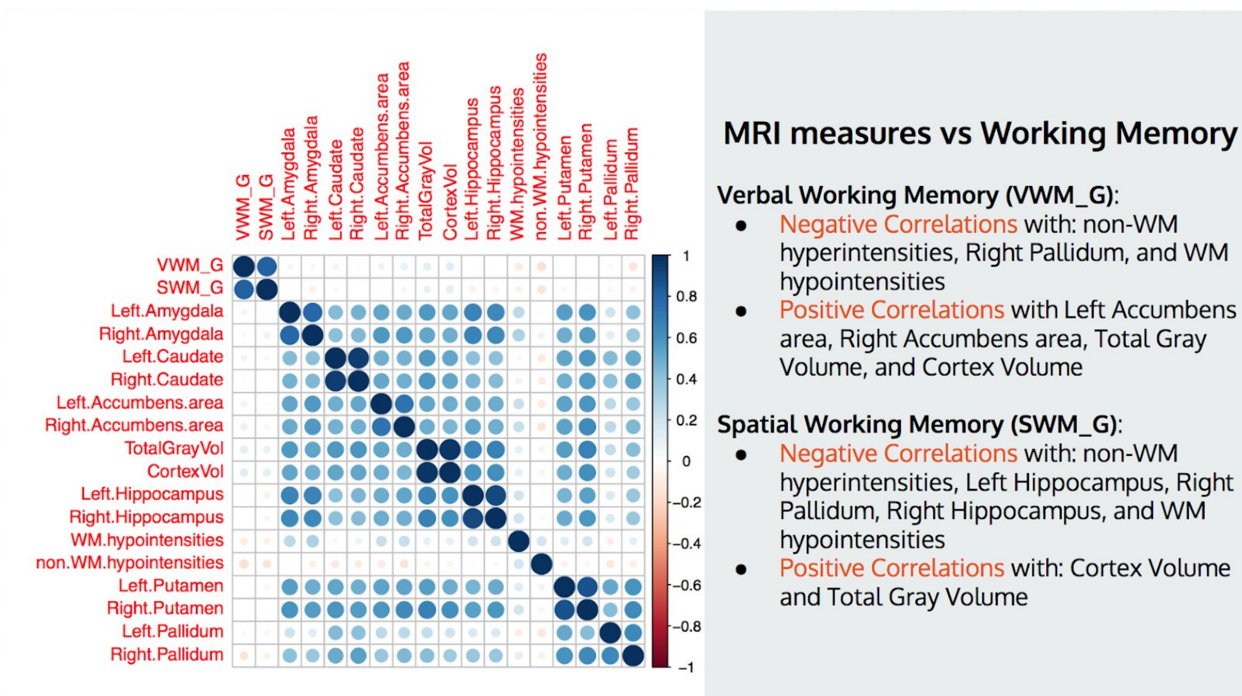
Numeric Variables	Description
PTID	Patient ID
GENDER	1=Male 2=Female
ETHNICITY	1=Hispanic origin 2=Not of Hispanic origin
SCHOOL_YRS	Amount of years of school completed
HOPKINS_SOMTIZATION	Average of items 1, 4, 12, 14, 27, 42, 48, 49, 52, 53, 56, 58; Somatic Symptoms
HOPKINS_OBSCOMP	Average of items 9, 10, 28, 38, 45, 46, 51, 55; Obstetrics Symptoms
HOPKINS_INTSENSITIVITY	Average of items 6, 11, 24, 34, 36, 37, 41; Insensitivity Symptoms
HOPKINS_DEPRESSION	Average of items 5, 15, 19, 20, 22, 26, 29, 30, 31, 32, 54; Depression Symptoms
HOPKINS_ANXIETY	Average of items 2, 17, 23, 33, 39, 50 Anxiety Symptoms
RI_G	Response Inhibition
VWM_G	Verbal Working Memory
SWM_G	Spatial Working Memory
CHAPPER_TOTAL	Total of 35 questions asked using Chapman Scales - **Perceptual Aberrations**
CHAPSOC_TOTAL	Total of 40 questions asked using Chapman Scales - **Social Anhedonia**
CHAPPHY_TOTAL	Total of 61 questions asked using Chapman Scales - **Physical Anhedonia**

Correlation Plots

fMRI Measures vs. Working Memory



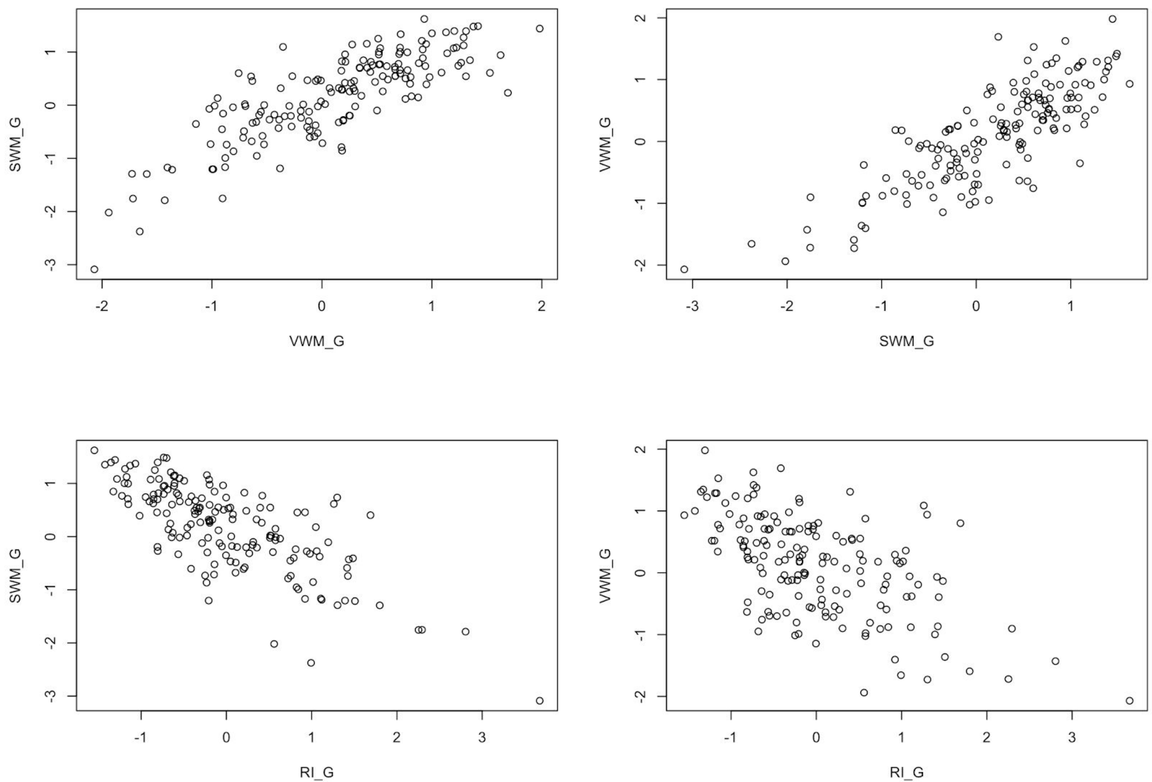
MRI Measures vs. Working Memory



IV. Exploratory Data Analysis

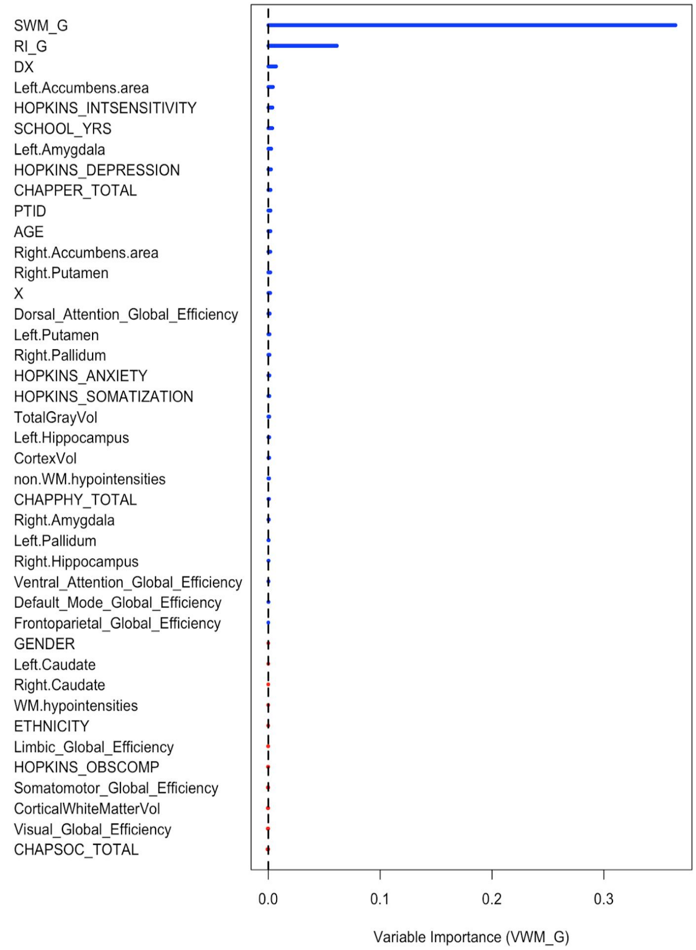
We started off our exploratory data analysis by using scatterplot matrices to visualize the relationships of our two response variables Verbal Working Memory (VWM_G) and Spatial Working Memory (SWM_G) with highly significant variables. In doing this, we were able to identify the variables correlation with our two response variables, as well as the relationship between the two response variables.

Scatterplots of Significant Response Variable Relationships



From the scatterplots, we can see that the variable RI_G has a negative correlation with our two response variables. Additionally, we can see that our two response variables are positively correlated with one another with a correlation of ~0.8. Following up on the positive correlation between our response variables, we decide to perform a Random Forest analysis to see if our response variables also have the same high significance predictors.

Random Forest Analysis



From our Random Forest analysis, we are able to obtain the significant predictors for predicting working memory for both of our response variables. For spatial working memory, VMG_G and RI_G are variables of high significance and for verbal working memory, SWM_G and RI_G have high significance. Moreover, because the response variables are highly correlated, we decide to combine VWM and SWM into a single measure: average working memory score. We did this by taking the average of the two variables.

Combining Verbal (VWM) and Spatial (SWM) Working Memory

```
dat <- read.csv("~/Desktop/Stats 141SL/Final Project -
Intelligence & Connectivity/intelligence and connectivity by
age.csv")
dat$workingmem <- rowMeans(dat[,13:14])
```

We are now ready to perform statistical analysis in the form of a multiple regression model.

V. Statistical Analysis

We decide to run a multiple regression model to model our data using our newly combined variable average working memory as the response variable and fMRI measures, MRI measures, and demographic covariates as the explanatory variables. We first ran our full model lm(working memory ~ fmri+mri measures) and used vif() to identify and remove variables with high multicollinearity.

VIF() Output

AGE	79.357045
GENDER	100.680340
ETHNICITY	98.851369
SCHOOL_YRS	42.499165
AGE:GENDER	26.424306
AGE:ETHNICITY	34.969738
AGE:SCHOOL_YRS	86.763583
GENDER:ETHNICITY	26.457767
GENDER:SCHOOL_YRS	96.601515
ETHNICITY:SCHOOL_YRS	101.556018

Next, we used anova() to identify the significant predictors. In doing this, we were able to reduce our model from 34 predictor variables to 17.

Model 1 (Reduced Model) vs. Model 2 (Full Model)

```
Analysis of Variance Table

Model 1: workingmem ~ Visual_Global_Efficiency + Somatomotor_Global_Efficiency +
    Dorsal_Attention_Global_Efficiency + Limbic_Global_Efficiency +
    Default_Mode_Global_Efficiency + Left.Amygdala + Right.Amygdala +
    Left.Accumbens.area + Right.Accumbens.area + CorticalWhiteMatterVol +
    Left.Pallidum + Right.Pallidum + WM.hypointensities + non.WM.hypointensities +
    AGE + GENDER + SCHOOL_YRS
Model 2: workingmem ~ Visual_Global_Efficiency + Somatomotor_Global_Efficiency +
    Dorsal_Attention_Global_Efficiency + Ventral_Attention_Global_Efficiency +
    Limbic_Global_Efficiency + Frontoparietal_Global_Efficiency +
    Default_Mode_Global_Efficiency + Left.Amygdala + Right.Amygdala +
    Left.Caudate + Right.Caudate + Left.Accumbens.area + Right.Accumbens.area +
    TotalGrayVol + CortexVol + CorticalWhiteMatterVol + Left.Putamen +
    Right.Putamen + Left.Pallidum + Right.Pallidum + Left.Hippocampus +
    Right.Hippocampus + WM.hypointensities + non.WM.hypointensities +
    AGE + GENDER + ETHNICITY + SCHOOL_YRS + AGE * GENDER + AGE *
    ETHNICITY + AGE * SCHOOL_YRS + GENDER * ETHNICITY + GENDER *
    SCHOOL_YRS + ETHNICITY * SCHOOL_YRS
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1    210 74.493
2    193 67.309 17    7.1837 1.2117 0.2582
```

From the anova of our reduced model with our full model, we notice that the high p-value of 0.2582 indicates that our reduced model can predict the data just as well as the full model.

VI. Summary of Results

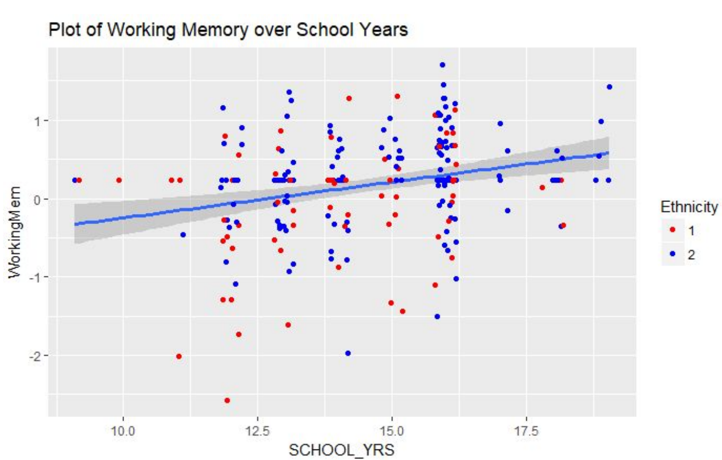
Summary Table

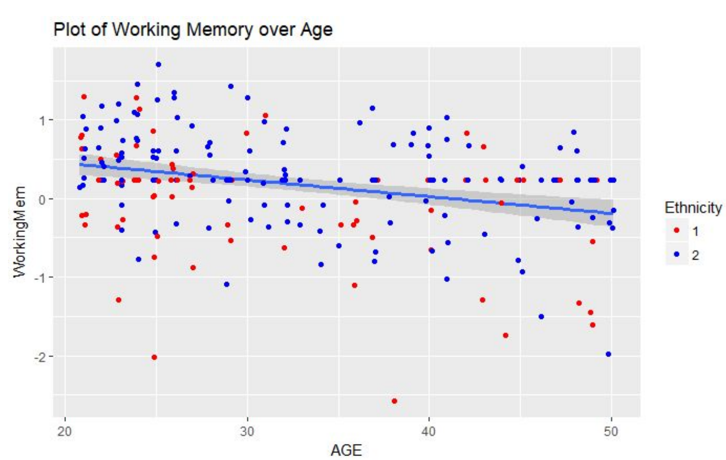
	Dependent variable:
	WorkingMem
Limbic_Global_Efficiency	-0.346
	(0.434)
Right.Amygdala	-0.230
	(0.743)
Left.Putamen	-0.417
	(0.728)
non.WM.hypointensities	-0.990
	(0.628)
AGE	-0.020^{***}
	(0.005)
ETHNICITY	0.304^{***}
	(0.088)
SCHOOL_YRS	0.069^{***}
	(0.021)
Constant	-0.557
	(0.415)
Observations	228
R^2	0.210
Adjusted R^2	0.185
Residual Std. Error	0.594 (df = 220)
F Statistic	8.377^{***} (df = 7; 220)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

VII. Interpretation of Results

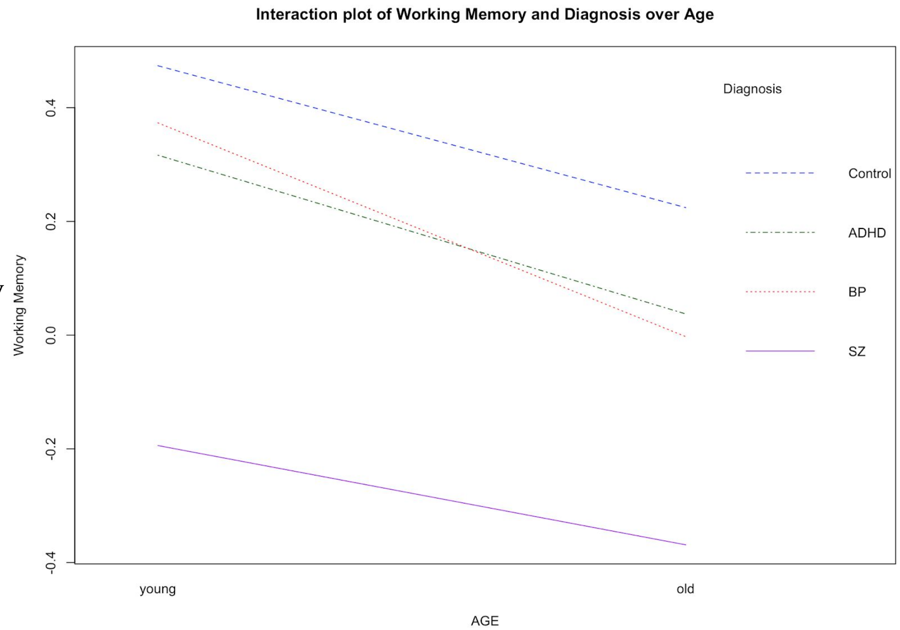
The results table of the model shows that of the predictors, Age, Ethnicity, and SCHOOL_YRS are most significant. The model has an F statistic of 8.377, resulting in a near 0 p-value of 4.524e-09. While the p-value is very low, the R^2 value (0.210) is also low. The low p-value indicates that the predictors have a significant relationship with working memory, but the low R^2 indicates that there is a high amount of variability in this relationship.

This can be observed in the below scatterplots of working memory versus age and school years, two of the most significant predictors. One can observe a positive correlation between working memory and school years, and a negative correlation between working memory and age. While these relationships are observable, there is a high amount of variation in them, as can be seen in how many points stray from the regression lines.

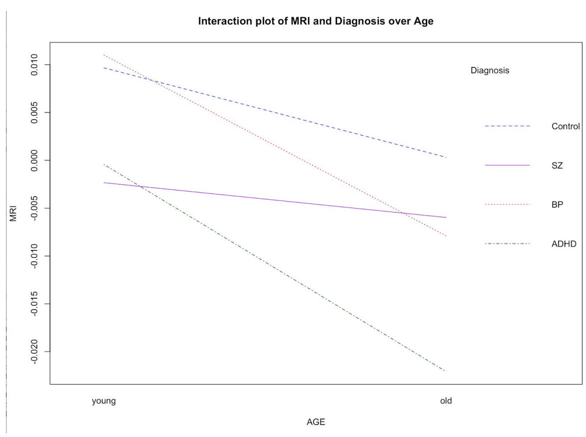


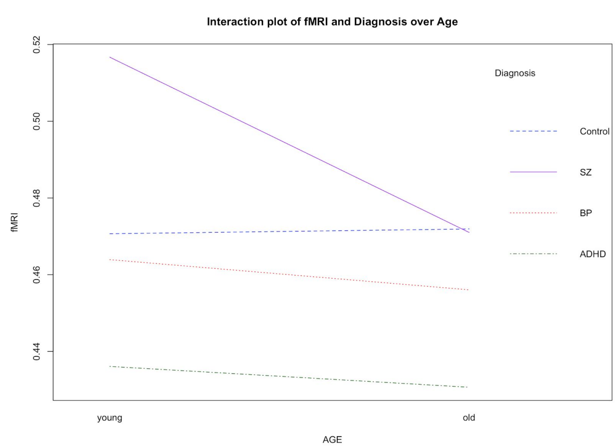


VII. Interpretation of Plots



We see large decrease in overall working memory as age increases. There are not many intersection points, which suggests minimal interaction between working memory and diagnosis. Those diagnosed with schizophrenia, however, start off with much lower working memory than those not diagnosed with schizophrenia. Those diagnosed with schizophrenia see less of a decrease in working memory as they age than those in the other groups, although this could be because they start off with much lower working memory to begin with.





Looking at MRI measures across the different diagnoses, we see that MRI decreases for all 4 groups. Those with bipolar disorder and ADHD saw a steep decline in MRI score as they aged. Those in the control and schizophrenia group do not show as much decline. It should also be noted that the ADHD and schizophrenia group started off with lower MRI scores than those in the control and bipolar group. There are many intersection points on this plot, suggesting heavy interaction between MRI score and diagnosis.

For fMRI measure we see that those in the control group slightly increase, while those with any diagnosis except schizophrenia only slightly decrease as they age. Schizophrenia starts very high in fMRI score and then rapidly and steeply declines as the person ages.

IX. Conclusion

A. ***Can we predict someone's working memory using their MRI measures, fMRI activity, and all other covariates (education, age, etc.)?***

```
Analysis of Variance Table

Model 1: workingmem ~ Visual_Global_Efficiency + Somatomotor_Global_Efficiency +
    Dorsal_Attention_Global_Efficiency + Limbic_Global_Efficiency +
    Default_Mode_Global_Efficiency + Left.Amygdala + Right.Amygdala +
    Left.Accumbens.area + Right.Accumbens.area + CorticalWhiteMatterVol +
    Left.Pallidum + Right.Pallidum + WM.hypointensities + non.WM.hypointensities +
    AGE + GENDER + SCHOOL_YRS
Model 2: workingmem ~ Visual_Global_Efficiency + Somatomotor_Global_Efficiency +
    Dorsal_Attention_Global_Efficiency + Ventral_Attention_Global_Efficiency +
    Limbic_Global_Efficiency + Frontoparietal_Global_Efficiency +
    Default_Mode_Global_Efficiency + Left.Amygdala + Right.Amygdala +
    Left.Caudate + Right.Caudate + Left.Accumbens.area + Right.Accumbens.area +
    TotalGrayVol + CortexVol + CorticalWhiteMatterVol + Left.Putamen +
    Right.Putamen + Left.Pallidum + Right.Pallidum + Left.Hippocampus +
    Right.Hippocampus + WM.hypointensities + non.WM.hypointensities +
    AGE + GENDER + ETHNICITY + SCHOOL_YRS + AGE * GENDER + AGE *
    ETHNICITY + AGE * SCHOOL_YRS + GENDER * ETHNICITY + GENDER *
    SCHOOL_YRS + ETHNICITY * SCHOOL_YRS
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1    210 74.493
2    193 67.309 17    7.1837 1.2117 0.2582
```

From our anova, the large p-value tells us that our reduced model can predict the data just as well as the full model, even with a significantly less amount of predictors. A multiple regression model is the appropriate model to use given the dataset.

B. ***Does the relationship among cognitive ability, MRI, and fMRI behavior change as a person ages? Do issues with working memory become exaggerated with age and diagnosis?***

The relationships among cognitive ability, MRI, and fMRI behavior changes with age a lot. In general, all working memory and cognitive ability decline with age. As age increases, working memory also decreases with each diagnosis. In the order of most rapid decline with working memory with respect to age: bipolar disorder, ADHD, schizophrenia, and then the control group.

C. ***Finally, do these patterns of cognitive ability change with a psychiatric diagnosis?***

With respect to the healthy control group, the pattern of working memory does not change significantly over the different diagnoses although diagnosis can affect the starting value of working memory.

X. Challenges & Future Recommendations

The initial challenges we had in approaching this project were fully understanding the data, deciding on which statistical method to use for the data, and identifying the significant predictors. At first, we were unsure of which statistical method to employ and we considered running a multivariate regression due to the fact that we had two outcome variables. However, after our exploratory data analysis where we used scatterplots and Random Forest analysis to see that our two outcome variables could be combined into one, the solution became clear that we should use a multiple regression model. Similarly, we overcame the challenge of identifying the significant predictors in our exploratory data analysis stage, where Random Forest analysis and anova() revealed the significant predictors.

The data did not have an even distribution for ages. This prevented us from exploring the relationship between age and working memory as thoroughly as possible. In the future, a larger, random sample of patients would likely lead to a wider variance in age, and allow us to better analyze the relationship between age and the other variables.

While we concluded that the there is not a significant difference in the relationship between working memory and age across the different diagnoses, we did see that the diagnoses had an effect on the starting value of working memory. This means that while there is not a significant difference in how working memory declines across the different diagnoses, the working memory of a patient is more likely to be higher or lower depending on their diagnosis. This is certainly worthy of additional analysis to further explore the relationship between these diagnoses and working memory.