**Intelligence and the Brain: Size Matters**

**Introduction**

Sir Francis Galton was the first to propose a relationship between brain size and general mental ability (Rushton & Ankney, 2009). He determined in the late 19th century that Cambridge University graduates who attained the highest honors had brain sizes 2-5% larger than their less accomplished colleagues. Up until ten years ago, brain size measurements had to be done during autopsy. Since brain size tends to decrease with age (Willerman, Schultz, Rutledge & Bigler, 1991), post mortem brain measurements did not always yield accurate results. With modern brain imaging techniques such as MRI, brain sizes of living younger subjects can be measured reliably. Since childhood malnutrition and early environmental exposure to toxins such as lead can decrease human brain size (Sukel, 2009), it is important to know whether there is a link between brain size and cognitive ability. Major public health consequences could result from such a connection. Another consideration is the fact that women are generally smaller than men. If their brain sizes are in proportion to their body sizes, it may be possible that men and women’s performance on the intelligence tests will differ. Therefore, the questions of interest in this study are whether brain size and gender are significant predictors of intelligence test performance, and whether there is any difference among the mean scores on the various types of intelligence tests.

For the test type main effect, the null hypothesis is that the mean intelligence test scores for all three tests are the same:

The alternative hypothesis is that at least one of three test’s mean score is different:

For the gender main effect, the null hypothesis is that the mean intelligence test scores for men and women are the same:

The alternative hypothesis is that mean intelligence test scores for men and women are different:

**Methods**

This study was conducted at the University of Texas at Austin (Willerman, Schultz, Rutledge & Bigler, 1991). In it, 40 healthy college students with a mean age of 18.9 years were selected from a population of students whose SAT scores were either ≤ 940 or ≥ 1350. These students completed the Vocabulary (VIQ), Similarities, Block Design and Picture Completion (PIQ) subtests of the Wechsler Adult Intelligence Scale-Revised. Their scores on the Vocabulary and Picture Completion subtests were recorded, along with their score on a Full Scale intelligence test (FSIQ). They underwent an MRI which recorded their total brain size, which was obtained by summing all nonzero pixels in the 18 images from each MRI scan (MRI\_Count). Participants’ height and weight were also recorded to adjust for body size, to better reflect the proportion of the brain used for intellectual functioning (Jerison, as cited in Willerman, Schultz, Rutledge & Bigler). Subjects were then equally divided by IQ and gender, however analyses were done blindly with respect to these variables.

This experiment has two factors, Test\_Type and Gender. These are both treated as fixed effects. MRI\_Count is treated as a continuous covariate. The response variable is IQ\_Score. Additional analysis was done to determine whether Height and Weight should also be treated as covariates.

Each student is an experimental unit treated as a random effect, and three measures are made on each student, one for each type of intelligence test. This is a randomized complete block design.

As part of an exploratory data analysis, a simple linear regression was run for IQ\_Score versus MRI\_Count for each test type separately, grouped by gender. Each of the regressions was significant. Graphical output for the regression and tests of significance are included in the Appendix, along with the means and standard errors of the test types.

The ANCOVA used was a compound symmetry model. SAS results of this analysis can be found in the Appendix, along with mean comparisons of the test types using the Tukey method.

A check was made on whether Height and Weight have any effect on IQ\_Scores, warranting their inclusion as additional covariates in the model The fundamental idea of including these as covariates is to ‘control for’ Height and Weight . In other words, including the covariate in the ANOVA allows the comparison between test types to be made without the complicating factors Height or Weight. The relevant scatterplots and regressions are in the Appendix.

**Results**

An initial ANOVA model containing only Test\_Type Gender and an interaction term shows that none of these effects is significant. The slopes of the regression lines for MRI\_Count versus IQ\_Score grouped by Test\_Type were non-zero and all essentially parallel. This indicates that MRI\_Count has an effect on IQ\_Score, but that there are no significant differences in its effect among the different test types. An ANCOVA model containing the continuous covariant, MRI\_Count and its interaction terms with the fixed effects showed no significant interactions. The MRI\_Count interaction terms were left out of the compound symmetry ANCOVA model, which had repeated measures of Test\_Type on each student. Again, the only significant (p < 0.05) effect was MRI\_Count, with a p-value of 0.0079. Figure 1 below shows the mean comparisons of the IQ\_Scores for each of the three test types. Based on the results of the ANOVA and the Tukey mean comparison, there are no significant difference in the mean IQ\_Scores among the different test types.

Figure 1. Mean Comparisons of IQ Scores for Different Test Types.

Based on the regressions of Height and Weight versus IQ\_Score, the slopes are not all zero. However they do appear to be the same for all three test types. When grouped by Gender, the slopes of the regression lines for Height and Weight versus IQ\_Score do not appear to be the same, suggesting an interaction with Gender. An ANCOVA was run to determine the significance of the interaction terms. A significant interaction existed between Weight and Test\_Type. No significant interactions existed between Height and any of the other effects. Results of the ANCOVA are in the Appendix.

A final model was run incorporating Weight and the Weight\*Test\_Type interaction term. The significant effects now include Test\_Type, MRI\_Count, the Weight and Test\_Type interaction, and the Test\_Type and Gender interaction. The result of the mean comparison of the IQ\_Scores for Test\_Type and Gender interaction is shown in Figure 2 below. Based on the results of the Tukey mean comparison, there is a significant difference in means between the men’s IQ\_Score on the Vocabulary IQ test and the men’s IQ\_Score on the Picture Completion IQ test.

Figure 2. Mean Comparisons of IQ Scores for Different Test Types and Genders.

We can conclude that MRI\_Count, hence brain size, has a significant effect on intelligence test scores, but not the IQ test type or gender of the student. We can conclude that men score significantly lower in the Picture Completion IQ test compared to the Vocabulary IQ test. We can also conclude that there are significant differences in IQ scores on one of the three test types based on Weight.

**Conclusion**

Through evolution there have been many factors that contributed to the intelligence and brain size of humans (UCLA, 2012).  Now, with the availability of MRI technologies and IQ testing, we are able to more confidently compare the differences in brain size and intelligence among people and genders.  The data show that, in general, people with higher MRI counts have higher IQ scores.  However, when comparing differences between genders, MRI scores, as expected, are related to body size.  Yet IQ scores are fairly uniform across genders.  In closing, it should be noted that there are many other factors that contribute to intelligence besides brain size.  This data set examined physical characteristics alone, though environment must also be taken into consideration when investigating intelligence.

**References**

Rushton, J.P. & Ankney, C.D. (2009). Whole brain size and general mental ability. *Journal of International Neuroscience*. 119(5): 692-732. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2668913/#!po=94.5652

Sukel, K. (2009). Fact or Fiction: When It Comes to Intelligence, Does Brain Size Matter? *Scientific American*. April 14, 2009. http://www.scientificamerican.com/article.cfm?id=does-brain-size-matter

University of California, Los Angeles (UCLA), Health Sciences (2012, April 15). New genes linked to brain size, intelligence. *ScienceDaily*. http://www.sciencedaily.com­/releases/2012/04/120415150123.htm

Willerman, L., Schultz, R., Rutledge, J.N., & Bigler, E.D. (1991). In vivo brain size and intelligence. *Intelligence*, 15, 223-228.

**Appendix**

**Scatterplot of IQ scores versus MRI, grouped by test type:**

The slopes are not zero, and are all roughly the same regardless of test type.



**Scatterplot of IQ scores versus MRI, grouped by gender**



**Regression Analyses of each test score type versus MRI**

All of the regressions are significant, verifying that the regression line slopes are not zero.



**Regression Analysis: FSIQ versus MRI\_Count**

The regression equation is

FSIQ = 5.2 + 0.000119 MRI\_Count

Predictor Coef SE Coef T P

Constant 5.17 46.01 0.11 0.911

MRI\_Count 0.00011915 0.00005047 2.36 0.023

S = 22.7833 R-Sq = 12.8% R-Sq(adj) = 10.5%

Analysis of Variance

Source DF SS MS F P

Regression 1 2893.0 2893.0 5.57 0.023

Residual Error 38 19724.9 519.1

Total 39 22617.9

Unusual Observations

Obs MRI\_Count FSIQ Fit SE Fit Residual St Resid

12 1079549 141.00 133.80 9.34 7.20 0.35 X

X denotes an observation whose X value gives it large leverage.



**Regression Analysis: MRI\_Count versus VIQ**

The regression equation is

MRI\_Count = 792706 + 1033 VIQ

Predictor Coef SE Coef T P

Constant 792706 53630 14.78 0.000

VIQ 1032.9 467.4 2.21 0.033

S = 68931.0 R-Sq = 11.4% R-Sq(adj) = 9.1%

Analysis of Variance

Source DF SS MS F P

Regression 1 23206826203 23206826203 4.88 0.033

Residual Error 38 1.80556E+11 4751480608

Total 39 2.03763E+11

Unusual Observations

Obs VIQ MRI\_Count Fit SE Fit Residual St Resid

12 150 1079549 947644 20699 131905 2.01R

23 129 790619 925953 13392 -135334 -2.00R

28 96 1062462 891867 13311 170595 2.52R

R denotes an observation with a large standardized residual.



**Regression Analysis: PIQ versus MRI\_Count**

The regression equation is

PIQ = 1.7 + 0.000120 MRI\_Count

Predictor Coef SE Coef T P

Constant 1.74 42.39 0.04 0.967

MRI\_Count 0.00012025 0.00004651 2.59 0.014

S = 20.9927 R-Sq = 15.0% R-Sq(adj) = 12.7%

Analysis of Variance

Source DF SS MS F P

Regression 1 2946.6 2946.6 6.69 0.014

Residual Error 38 16746.4 440.7

Total 39 19693.0

Unusual Observations

Obs MRI\_Count PIQ Fit SE Fit Residual St Resid

12 1079549 128.00 131.56 8.61 -3.56 -0.19 X

14 856472 147.00 104.74 4.11 42.26 2.05R

R denotes an observation with a large standardized residual.

X denotes an observation whose X value gives it large leverage.

**ANOVA Model with no Covariate MRI\_Count**

| **Type 3 Tests of Fixed Effects** | | | | |
| --- | --- | --- | --- | --- |
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| Test\_Type | 2 | 114 | 0.10 | 0.9020 |
| Gender | 1 | 114 | 0.52 | 0.4708 |
| Test\_Type\*Gender | 2 | 114 | 0.04 | 0.9581 |

None of the effects are significant.

**ANOVA Output Test for Equal Slopes**

Interaction terms with the covariate MRI\_Count are not significant, so interaction terms can be removed from the model.

| **Type 3 Tests of Fixed Effects** | | | | |
| --- | --- | --- | --- | --- |
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| Test\_Type | 2 | 110 | 0.20 | 0.8176 |
| Gender | 1 | 110 | 1.50 | 0.2227 |
| Test\_Type\*Gender | 2 | 110 | 0.21 | 0.8105 |
| MRI\_Count | 1 | 110 | 5.89 | 0.0169 |
| MRI\_Count\*Gender | 1 | 110 | 1.08 | 0.3004 |
| MRI\_Count\*Test\_Type | 2 | 110 | 0.20 | 0.8199 |

**ANCOVA Output Compound Symmetry Model**

| **Type 3 Tests of Fixed Effects** | | | | |
| --- | --- | --- | --- | --- |
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| Test\_Type | 2 | 76 | 0.95 | 0.3894 |
| MRI\_Count | 1 | 37 | 7.90 | 0.0079 |
| Gender | 1 | 37 | 2.12 | 0.1540 |
| Test\_Type\*Gender | 2 | 76 | 0.40 | 0.6742 |

MRI\_Count is the only significant effect.

**Mean Comparisons**

|  |  |
| --- | --- |
| |  | | --- | | **Effect=Test\_Type   Method=Tukey-Kramer(P<.05)   Set=1** | |
| |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | | **Obs** | **Test\_Type** | **Estimate** | **Standard Error** | **Letter Group** | | --- | --- | --- | --- | --- | | 1 | FSIQ | 113.28 | 3.4173 | A | | 2 | VIQ | 111.83 | 3.4173 | A | | 3 | PIQ | 110.93 | 3.4173 | A | | |

There is no significant difference among the mean IQ score for each test type.

**Additional Covariates of Weight and Height**

There appears to be no effect of Weight on IQ\_Scores when they are grouped by Test\_Type



Here are the regression results

**Regression Analysis: IQ\_Scores versus Wt, Test\_Type**

The regression equation is

IQ\_Scores = 120 - 0.0353 Wt - 1.42 Num\_Type

114 cases used, 6 cases contain missing values

Predictor Coef SE Coef T P

Constant 120.26 15.02 8.01 0.000

Wt -0.03534 0.09214 -0.38 0.702

Num\_Type -1.421 2.614 -0.54 0.588

S = 22.7925 R-Sq = 0.4% R-Sq(adj) = 0.0%

Analysis of Variance

Source DF SS MS F P

Regression 2 229.9 114.9 0.22 0.802

Residual Error 111 57664.4 519.5

Total 113 57894.3

Source DF Seq SS

Wt 1 76.4

Num\_Type 1 153.5

There appears to be a slight downward trend to IQ\_Scores versus Height when they are grouped by Test\_Type



Here are the regression results:

**Regression Analysis: IQ\_Scores versus Ht, Test\_Type**

The regression equation is

IQ\_Scores = 149 - 0.489 Ht - 1.24 Num\_Type

117 cases used, 3 cases contain missing values

Predictor Coef SE Coef T P

Constant 148.74 37.15 4.00 0.000

Ht -0.4893 0.5360 -0.91 0.363

Num\_Type -1.244 2.588 -0.48 0.632

S = 22.8594 R-Sq = 0.9% R-Sq(adj) = 0.0%

Analysis of Variance

Source DF SS MS F P

Regression 2 556.1 278.1 0.53 0.589

Residual Error 114 59571.1 522.6

Total 116 60127.2

Source DF Seq SS

Ht 1 435.5

Num\_Type 1 120.6

There is a general trend for men’s IQ\_Scores to decrease as their weight or height increases. There is a slight increase in women’s IQ\_Scores as their weight increases, and a slight decrease in their IQ\_Scores as their height increases.





**Results of ANCOVA with Weight**

| **Type 3 Tests of Fixed Effects** | | | | |
| --- | --- | --- | --- | --- |
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| Test\_Type | 2 | 68 | 4.88 | 0.0105 |
| MRI\_Count | 1 | 33 | 5.87 | 0.0210 |
| Gender | 1 | 33 | 0.02 | 0.9012 |
| Test\_Type\*Gender | 2 | 68 | 0.80 | 0.4517 |
| Weight | 1 | 33 | 1.23 | 0.2756 |
| Weight\*Test\_Type | 2 | 68 | 4.18 | 0.0193 |
| Weight\*Gender | 1 | 33 | 0.05 | 0.8311 |
| Weight\*Test\_T\*Gender | 2 | 68 | 0.34 | 0.7163 |

The interaction term between Weight and Test\_Type is significant. The interaction term between Weight and Gender is not significant and will be removed from the model.

**Results of ANCOVA with Height**

| **Type 3 Tests of Fixed Effects** | | | | |
| --- | --- | --- | --- | --- |
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| Test\_Type | 2 | 70 | 1.55 | 0.2203 |
| MRI\_Count | 1 | 34 | 10.58 | 0.0026 |
| Gender | 1 | 34 | 0.80 | 0.3768 |
| Test\_Type\*Gender | 2 | 70 | 0.19 | 0.8261 |
| Height | 1 | 34 | 4.00 | 0.0535 |
| Height\*Test\_Type | 2 | 70 | 1.50 | 0.2292 |
| Height\*Gender | 1 | 34 | 0.78 | 0.3830 |
| Height\*Test\_T\*Gender | 2 | 70 | 0.25 | 0.7768 |

None of the interaction terms with Height are significant and so can be left out of the model. Height is not significant and will be removed from the model.

**Final Model Incorporating Weight as a Covariate**

| **Type 3 Tests of Fixed Effects** | | | | |
| --- | --- | --- | --- | --- |
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| Test\_Type | 2 | 70 | 5.01 | 0.0092 |
| MRI\_Count | 1 | 34 | 6.89 | 0.0129 |
| Gender | 1 | 34 | 0.29 | 0.5911 |
| Weight | 1 | 34 | 1.38 | 0.2476 |
| Test\_Type\*Gender | 2 | 70 | 4.73 | 0.0119 |
| Weight\*Test\_Type | 2 | 70 | 4.60 | 0.0133 |

**Mean Comparisons of Gender and Test\_Type Interaction**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | |  | | --- | |  | | | |  |  | | --- | --- | | |  | | --- | | **Effect=Test\_Type\*Gender   Method=Tukey-Kramer(P<.05)   Set=1** | | | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | | **Obs** | **Test\_Type** | **Gender** | **Estimate** | **Standard Error** | **Letter Group** | | --- | --- | --- | --- | --- | --- | | 1 | PIQ | F | 116.23 | 5.9700 | AB | | 2 | FSIQ | F | 115.81 | 5.9700 | AB | | 3 | VIQ | M | 112.77 | 6.4145 | A | | 4 | VIQ | F | 111.85 | 5.9700 | AB | | 5 | FSIQ | M | 110.71 | 6.4145 | AB | | 6 | PIQ | M | 104.24 | 6.4145 | B | | | | |