**Homework 8**

Due date: Wednesday, November 1st, 1:30pm

Submission instructions: Submit both your Rmarkdown file and knitted HTML file to Canvas

**Reading Questions**

1. According to Judd and colleagues (2014):
   1. What are the four assumptions of mediation analyses with unmanipulated mediator variables?
   2. When there is an experimental manipulation of the independent variable X, what other conditions must also be met so that mediation analyses will yield unbiased effects?
2. What does it mean to have "redundancy between predictors" in multiple regression? What is the consequence of having redundant predictors? Answer in 2-3 sentences.
3. A psychologist wants to examine how workers' capacity to concentrate is affected by auditory nuisances in a factory. Imagine she decided to run 6 different experimental “noise” conditions and one control condition (“silence”). In the 6 different experimental conditions, she exposes participants to different types of noise (e.g. in one condition she bangs a hammer on a table, in another condition she instructs a research confederate to yell and scream outside of the room where participants are tested during the study, etc.). She randomly assigned the participants to the 6 experimental conditions and the control group. To operationalize capacity to concentrate, the psychologist asked the participants to first remember a list of words and then asked them to recall the words after a 10-minute delay. Recall performance (the percentage of correctly recalled words) was the DV. After collecting the data, she individually tested whether the recall performance of the participants in each experimental condition was different from the participants in the control condition. Each of the 6 tests yield a *p* value > .05, suggesting no significant effect of noise on the capacity to concentrate. She then decided that she could try adding biological sex as a covariate, since she had read studies in other areas of psychology that statistically controlled for biological sex. After adding biological sex to each of her 6 models, she found that one of her experimental conditions (the one where a confederate yells outside the room) differed from her control condition at a significance level of *p* < .05 (after controlling for sex). Overjoyed, the psychologist wrote up her results for submission to a scientific journal, reporting that she manipulated noise by introducing a screaming confederate, and that this experimental manipulation led to worse memory as compared to a silent control condition (including relevant test statistics). She did not report the 5 other experimental conditions.
   1. Describe two critical problems in the psychologist’s research practices.
   2. How do these problems affect the psychologist’s Type I error rate?
   3. What are the two requirements proposed by Simmons et al. (2011) that would aid peer-reviewers in assessing the psychologist’s work?

**Data Analysis**

Does childhood cognitive ability predict the size of developing brains? Is the effect of childhood cognitive ability on adult brain size mediated by the personality factors which are related to curiosity and intellectual exploration? A team of neuroscientists recruited a longitudinal sample of 40 participants to answer these questions. When the participants reached age 12, the researchers administered the fifth edition of Wechsler Intelligence Scale for Children (WISC-V) to assess childhood cognitive ability. When the participants aged 21, the researchers obtained structural MRI scans of the participants’ brains and calculated total MRI voxel count as a measure of brain size. The researchers also measured the participants’ personalities at age 21 using the Big-Five Inventory. Because body size and biological sex are known to relate to brain size, the neuroscientist also asked the participants to report their height and weight at age 21, as well as their biological sex assigned at birth.

You were asked to conduct the statistical analyses to test the hypothesis that childhood cognitive ability is positively associated with brain size in adulthood. You were also asked to test whether the above association between cognitive ability and brain size is mediated by openness, one of the big-five personalities which is related to curiosity and intellectual exploration. Tragically, the data files came in a total mess. They do not have proper file extensions, so you do not know their exact file type, apart from the fact that they are regular delimited files. The data files also do not have proper variable names labeled. Luckily, your colleague was able to provide you with the codebooks of the two datafiles: “iq” and “brain”.

**Codebook for “iq”**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column** | **Variable** | **Description** | **Values** |
| 1 | subject\_id | Subject ID number | 1 - 40 |
| 2 | simi | WISC-V subtest score: Similarities | 5 - 20 |
| 3 | voca | WISC-V subtest score: Vocabulary | 5 - 20 |
| 4 | bloc | WISC-V subtest score: Block Design | 5 - 20 |
| 5 | matr | WISC-V subtest score: Matrix Reasoning | 5 - 20 |
| 6 | figr | WISC-V subtest score: Figure Weights | 5 - 20 |
| 7 | digi | WISC-V subtest score: Digit Span | 5 - 20 |
| 8 | code | WISC-V subtest score: Coding | 5 - 20 |

**Codebook for “brain”**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column** | **Variable** | **Description** | **Values** |
| 1 | subject\_id | Subject ID number | 1 - 40 |
| 2 | sex | Biological sex assigned at birth. 0 = female, 1 = male | 0, 1 |
| 3 | weight | Weight in pounds | 106 - 192 |
| 4 | height | Height in inches | 62 - 77 |
| 5 | mri\_count | Total voxel count from 18 slices, in thousands | 790.62 – 1079.55 |
| 6 | openness | Big-five inventory: Openness (in adulthood) | -1.65 - 2.17 |

1. Read in both data files (iq and brain) using a function we learned in lab that automatically detects file format and type.
2. For both dataframes, rename the columns according to the codebook so that the column names are meaningful variable names.
3. Use a function we learned in lab to combine the two dataframes, so that you end up with one dataframe that contains all subjects and all variables needed for the analyses.
4. The raw data contain 7 subtest scores from WISC-V, measuring different domains of cognitive functioning. These subtest scores are meant to be added up into a Full-Scale Intelligence Quotient (fsiq) that measures overall cognitive ability.
   1. Obtain the reliability metric of the seven subtest scores.
   2. Use the varScore() function to compute fsiq. If done correctly, fisq would have a mean of around 100 and a standard deviation of roughly 15.
5. Make a correlation/scatterplot matrix of the following variables: mri\_count, sex, weight, height and fsiq. (Hint: look at the package GGally)
6. Report the three strongest correlations among these variables.
7. Fit a 2-parameter model, where fsiq alone predicts total voxel count. Report the statistical test and interpret the parameter estimate for fsiq. Also compute η2.
8. Now fit a 5-parameter model predicting total voxel count (mri\_count) from fsiq, sex, height, and weight. Report the statistical test and interpret the parameter estimate corresponding to the focal predictor. Also compute its partial-η2. How did the partial-η2 of fsiq change from the η2 in the 2-parameter model? Why?
9. Use a function we learned in lab to check for multicollinearity between your predictors in the 5-parameter model. Is multicollinearity a problem? How do you know?
10. Create a publication-quality graph that represents total voxel count (in thousands) as a function of fsiq, when holding height and weight constant at their means and sex constant at the center of its coding scheme. Your plot should include standard error bands around the point estimates, raw data points, and informative axis labels.
11. Fit a mediation model in lavaan to test if the effect of childhood cognitive ability on brain size in adulthood is mediated by openness. Comment on whether each of the four conditions of mediation is satisfied. Use set.seed(123).

*Bonus*: can you figure out how to include sex, height, and weight as covariates in your mediation model, just like what you did for question 8?

1. Indicate the number of hours spent on this assignment.