Analysis of the Impact of ScenicView's Neurofeedback program on students' OQ scores and Attendance Rates

Andre Rossi de Oliveira, Duncan Christensen, Shealee Calder,
Jacob Sorenson, Ezra Liau, and Kyle Burch

1. Introduction

This report is a result of the collaboration between UVU's Center for Social Impact and ScenicView Academy, a nonprofit school for young adults with autism and learning disabilities located in Provo, Utah. The research was carried out by faculty and students of the Social Impact Metrics Lab – SIMLAB, with data provided by ScenicView.

In this study, we investigate the impact of ScenicView Academy's neurofeedback program on two measures of student performance, namely attendance rates and their scores on an Outcome Questionnaire Test (OQ). Our analysis was restricted to these two outcomes due to data limitations, but it can be expanded to other metrics if the appropriate data becomes available.

In order to measure that impact, we utilize a series of statistical methods, ranging from visualization aids to statistical tests and regression analysis. In addition, we employ matching techniques to improve sample balance. Our main finding is that there is qualified evidence that participation in the neurofeedback program increases attendance rates.

In the next section, we provide a brief description of a neurofeedback intervention. Section 3 consists of a series of plots that helps the reader visualize the main features of the data set, including how the sample of students is distributed by gender, age and ethnicity and how outcome variables are related to those characteristics. It also includes tests that check whether the mean and median values of the outcome variables differ by student characteristics. Section 4 looks at how the number of training sections affects attendance, OQ scores, and some mental evaluation indicators, while section 5 estimates treatment effects, that is, the impact of the neurofeedback program on the outcome measures, if any. These statistical findings, as well as some data issues, are discussed in section 6. Section 7 concludes.

2. The Neurofeedback program

Neurofeedback is a type of biofeedback, a process by which an indicator is used to measure a bodily function such as a heartbeat, with the goal of increasing an individual's ability to regulate and improve her brain function (Sitaram et al., 2017). In neurofeedback specifically, participants are subjected to an audio or visual stimulus while their brain activity is measured through electroencephalography (EEG). This activity is recorded and monitored, usually by a therapist, who can then select the desired range of acceptable function.

In a typical treatment program for a variety of mental disorders or diseases, the subject either watches a film or plays a game in which either the screen or the sound fades in and out based on the EEG readings. When the readings are within the established zone designated by the therapist overseeing the treatment, the movie or game proceeds normally. Only when the readings drift outside the set boundaries will the changes occur, and then the patient must actively engage to draw the signals back within the training range.

There are seven main types of neurofeedback training: frequency/power, slow cortical potential, low-energy (LENS), Hemoencephalographic (HEG), live Z-score, low-resolution electromagnetic tomography (LORE-TA), and Functional magnetic resonance imaging (fMRI), with the latter being the most recently developed. Each type is aimed at addressing specific issues facing the brain or motor function. Each of these different trainings are then further broken down into brain areas, with different tests and issues requiring different areas of activation. Overall, a specific training is assigned based on diagnosis and a training plan is established after consultations with a therapist (Marzbani, Marateb, and Mansourian, 2016).

Given this setup, "...brain activation is volitionally regulated through learning; as the activation acts as an independent variable, it allows causal inferences to be made between brain activity and behaviour" (Sitaram et al., 2017). This effectively allows for the neurofeedback program to run as a long-term treatment experiment in which the brain activation, independent variable, and mental or behavioral improvement, dependent variable, can be measured and "trained" to be improved.

3. Data visualization and preliminary analysis for the entire sample

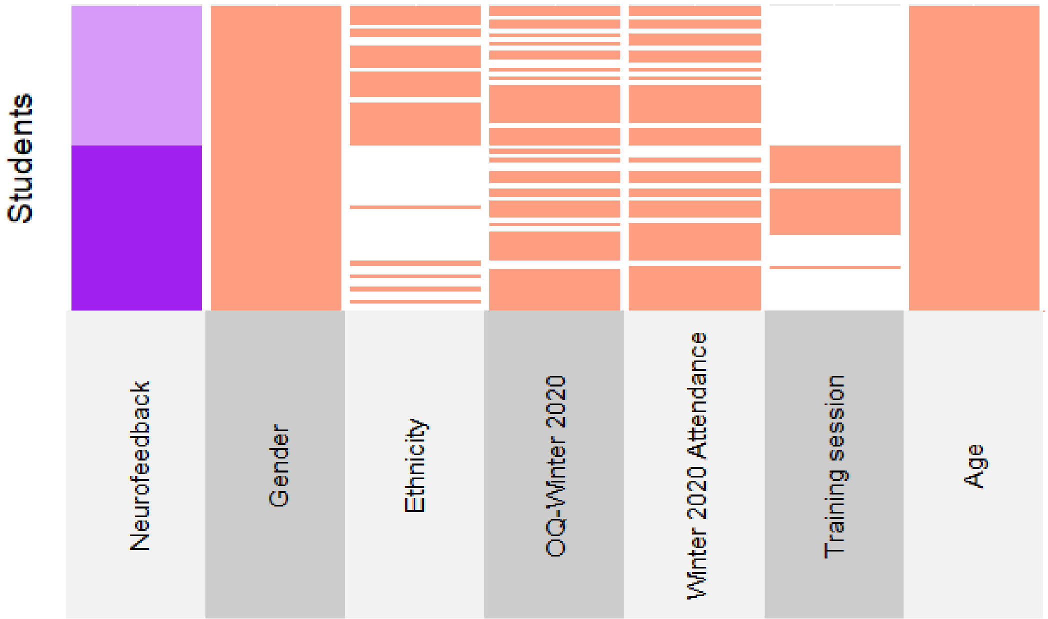
3.1. Visualization and analysis of student characteristics

In this section, we provide plots and descriptive statistics to paint a picture of the sample of students in our dataset. We do that for a select number of variables that are of interest to us. Three of them are characteristics of the students, namely gender, ethnicity, and age. Two of the variables, OQ scores and attendance in the Winter of 2020, are used in our study to measure the impact of the neurofeedback program on students, and for this reason are called outcome variables. The last variable, the number of training sessions, is only available for students who participated in the neurofeedback program.

Our dataset contains many other variables, but we were unable to use them in this study because of rampant missing values. In fact, there is a missing values problem as well with the variables we did include in our study, as can be seen in the heat map below.

Notice that the "Neurofeedback" column has different colors than the rest of the plot. Rows in dark purple are students who have participated in the neurofeedback program, and those in light purple have not. For the other columns, missing values are in white. We can see, for instance, that there is complete information for age and gender, but that the ethnicity of many students is missing. In addition, there is evidently no data on training sessions for non-participants, but even the data available for participants is spotty. Data on OQ test scores and attendance in the Winter of 2020 also have many gaps. This problem will have an impact on our analysis, as we discuss throughout the report.

Figure 1: Heat map



In the next few figures, we visualize the distribution of the sample by student characteristics. The first one shows that a little over 70% of the students are male, and a little less than 30% are female.

Figure 2: Distribution by gender

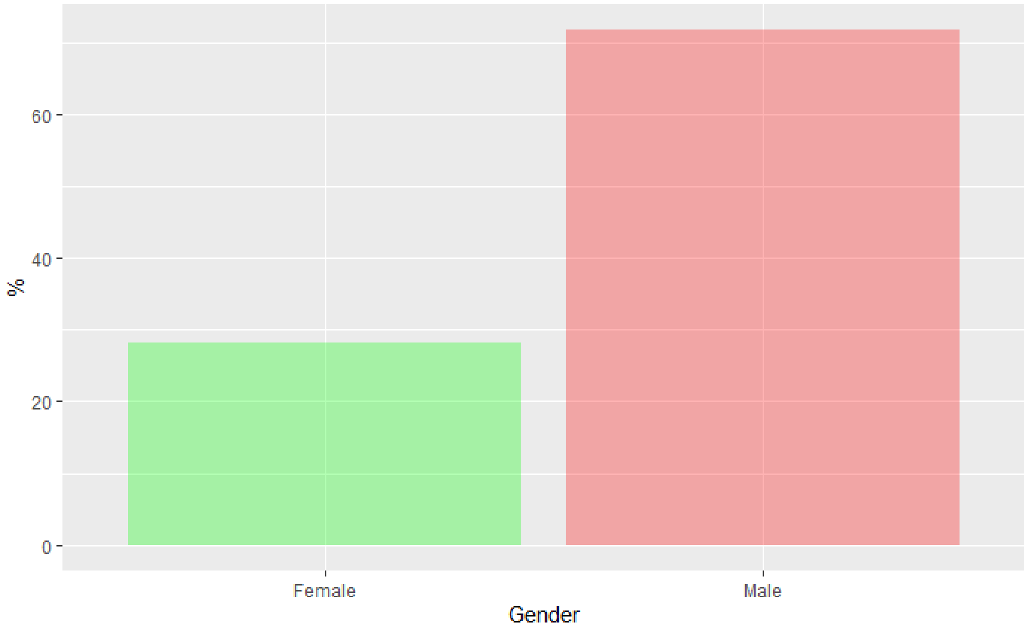
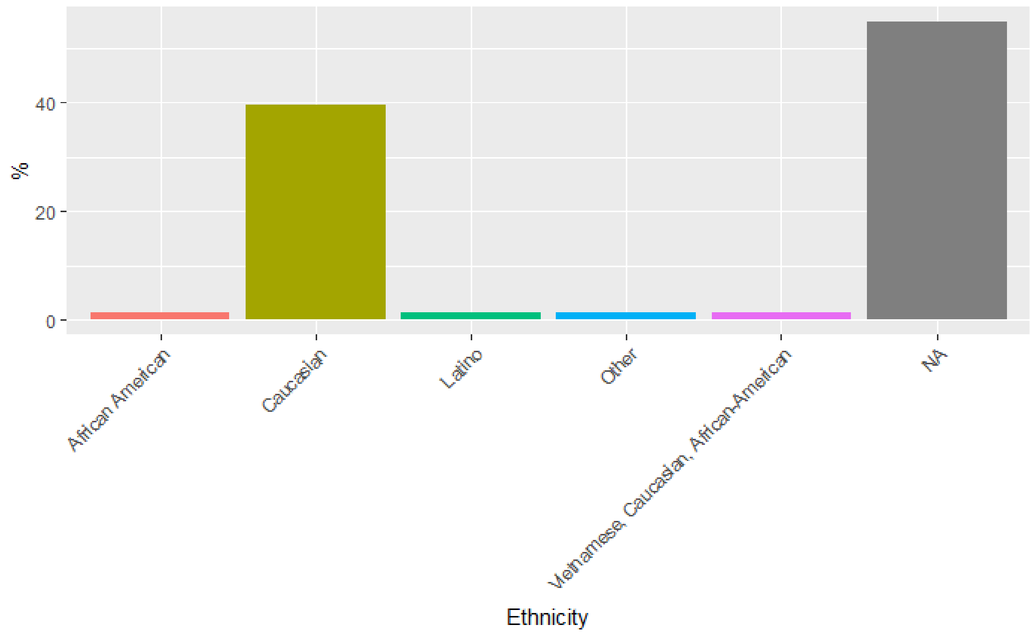


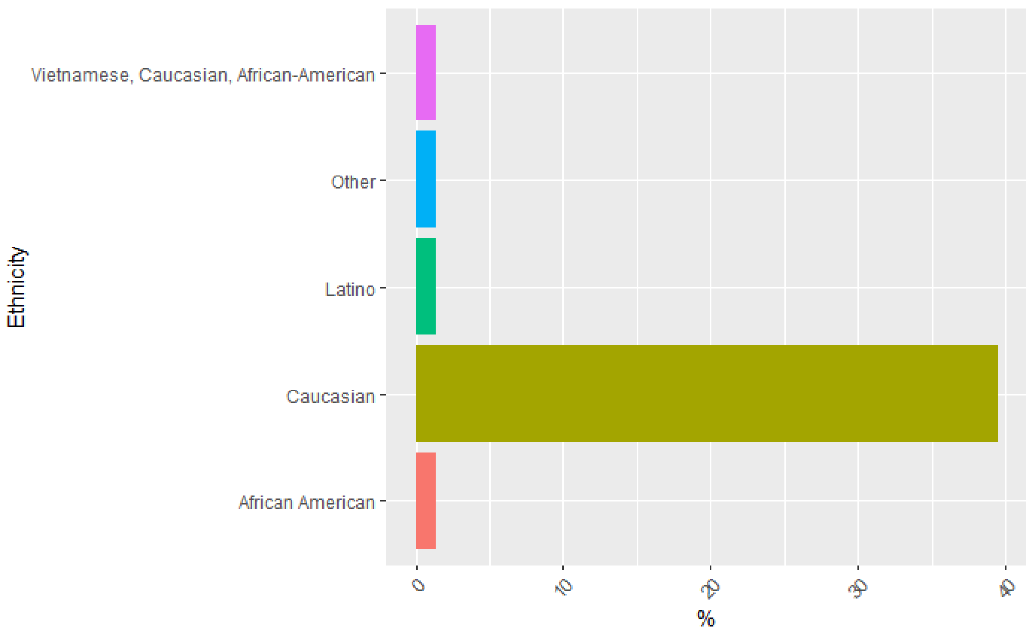
Figure 3 shows the distribution by ethnicity. It is evident that the large majority of students are Caucasian. Notice as well that we don't have information on this feature for almost 60% of the sample.

Figure 3: Distribution by ethnicity with missing values



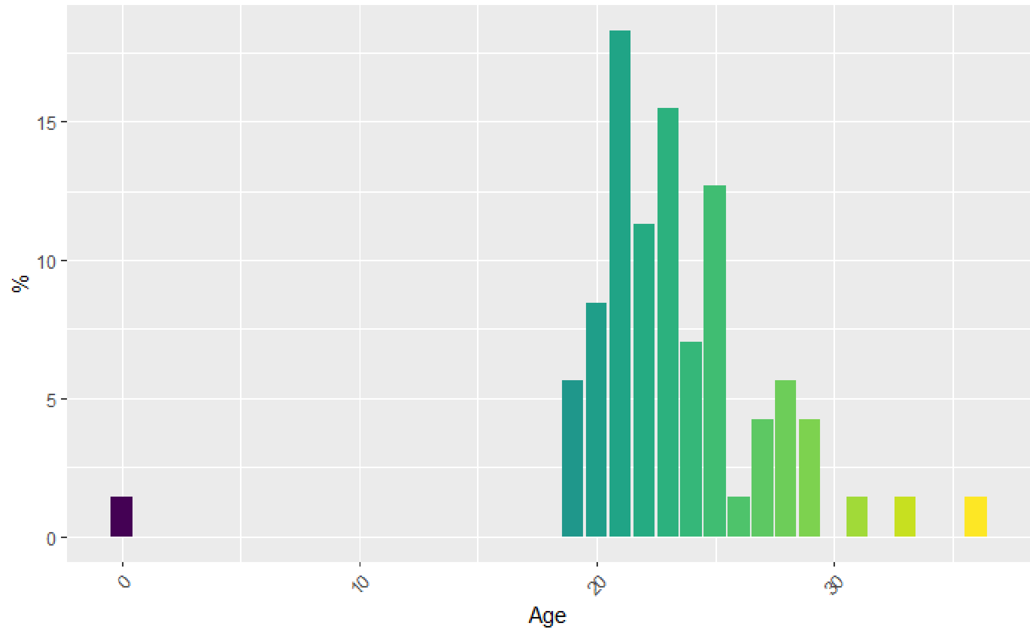
Another way to see the prevalence of Caucasian students is a plot without the missing values, as shown below.

Figure 4: Distribution by ethnicity without missing values



The next characteristic we consider is age. As the figure below makes clear, ages range from 18 to 36, with the majority of the students around 20 years old[1].

Figure 5: Distribution by age



In summary, our sample is comprised of mostly Caucasian, young, male students. In what follows, we investigate how this sample is distributed along the outcome variables. In other words, we want to know if students' performance, measured by their attendance and scores on an OQ test, depend on their characteristics. Since there are so many missing values for ethnicity in our dataset, we limit the analysis to gender and age[2].

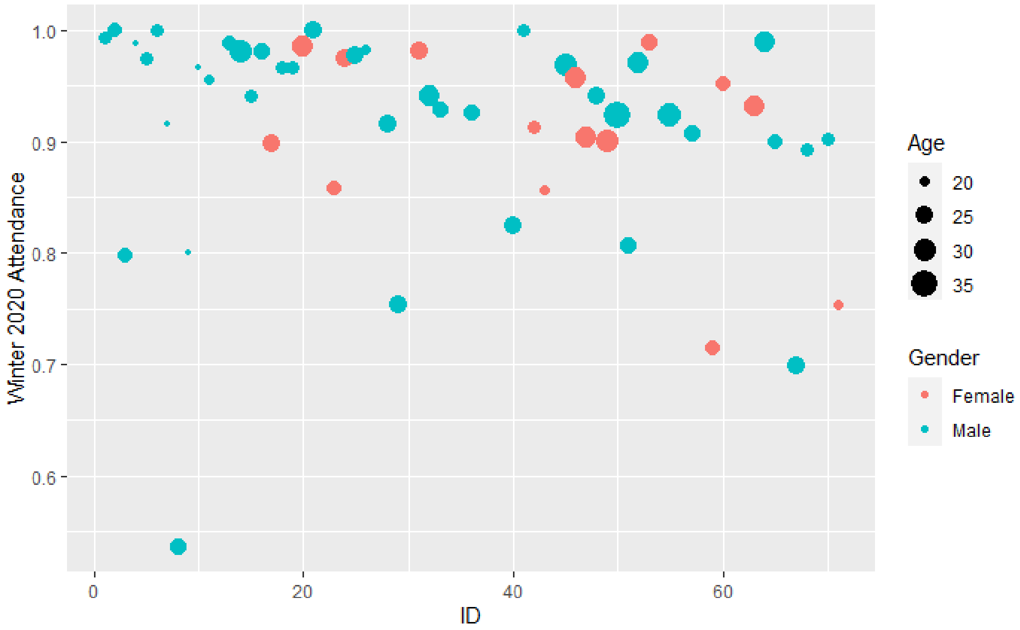
3.2. Visualization and analysis of outcome variables by student characteristics

3.2.1. Attendance

The first outcome we consider is attendance in the Winter 2020 semester. Students at ScenicView do not receive course grades, and whether they pass a class depends on their attendance. Therefore, it makes sense to use attendance as a measure of student performance. As can be seen on the figure below, almost all students have attendance rates above 70% (with the exception of an outlier), and the majority has more than 90% attendance. There are more males than females with relatively low attendance, but that is not very informative, since, as we have pointed out before, males account for more than 70% of our sample. It is also hard to detect any differences in attendance rates by age based on the scatterplot.

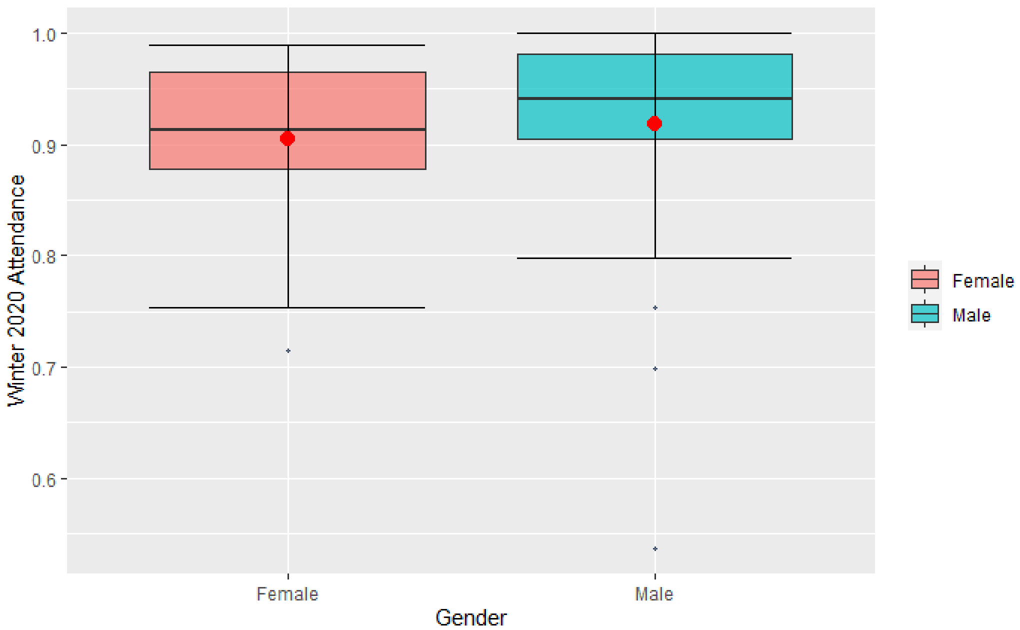
[1] There is an outlier that is believed to be a recording error where the birthdate was recorded as 2020.
[2] We have removed the observation whose birthdate was recorded as 2020.

Figure 6: Scatterplot of attendance by gender and age



The boxplot below provides a more helpful visualization of the attendance rates by gender[3].

Figure 7: Boxplot of attendance by gender



[3] The lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles). The upper whisker extends from the hinge to the largest value no further than 1.5*IQR from the hinge (where IQR is the inter-quartile range, or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5*IQR of the hinge. Data beyond the end of the whiskers are called "outlying" points and are plotted individually.

In order to check if these differences are statistically significant, we run a t-test, whose results are found in the table below:

Table 1: t-test attendance by gender

Value of t statistic	-0.53622
Degrees of freedom	29.862
P-value	0.5958
Mean female	0.9047333
Mean male	0.9187436

Since the p-value of 0.5958 is higher than 0.05, we do not reject the null hypothesis that the difference in means is equal to zero, that is, we cannot say that the two means are statistically different.

Let's now run Mood's Median test to check if the medians are statistically different.

Table 2: Mood's Median test for attendance by gender

Value of z statistic	-0.57011
P-value	0.5686

The p-value is greater than 0.05, so we fail to reject the null hypothesis. We do not have sufficient evidence to say that there is a statistically significant difference in the median attendance between the two groups. We conclude that, even though males have higher mean and median attendances, the differences are not statistically significant.

The next figure shows the boxplot for the variable age. It seems to indicate that older students have higher attendance, with the exception of 21-year-olds. This is not very informative, so we will carry out an ANOVA test to investigate further.

Figure 8: Boxplot of attendance by age

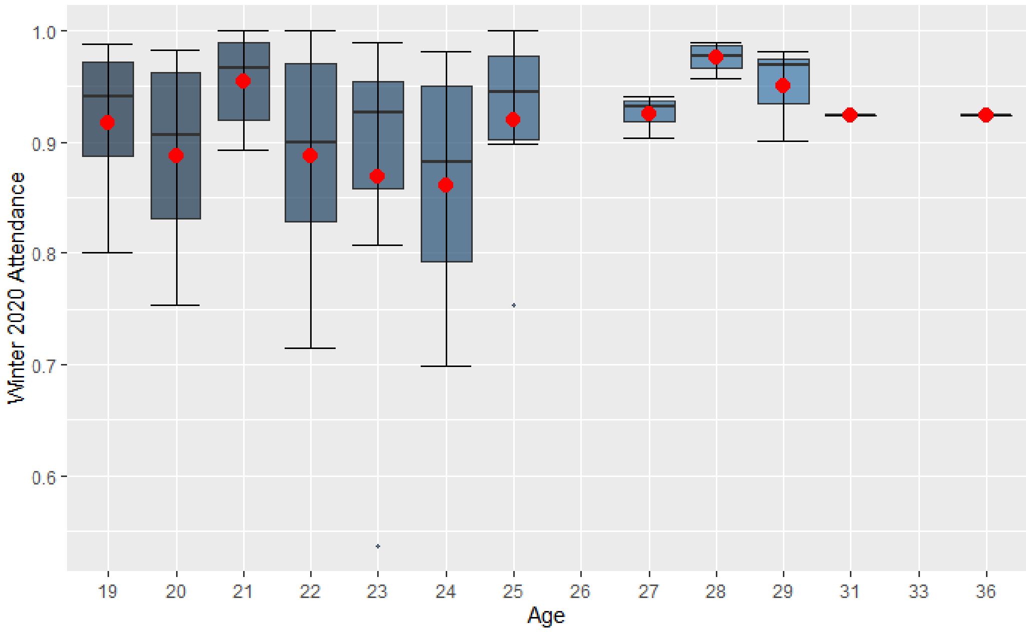


Table 3: ANOVA for attendance by age

Value of F statistic	0.523
P-value	0.473

Since the p-value is greater than 0.05, we cannot reject the null hypothesis that the means of the different groups are the same. A similar conclusion stems from Mood's median test, as seen below.

Table 4: Mood's Median test for attendance by age

Value of chi-squared	11.71
Degrees of freedom	11
P-value	0.3858

The p-value of 0.3858 tells us that null hypothesis that the medians of the different groups are the same cannot be rejected. In summary, there is no evidence that attendance varies significantly with age.

3.2.2. OQ scores

We now turn to the second outcome variable, OQ scores in the Winter of 2020. Outcome Questionnaire Tests (OQs) are tests or surveys issued to subjects after they have received a treatment in order to assess whether the treatment had any significant impact.

A generally accepted questionnaire is "Outcome Questionnaire 45". It is recognized in the psychological community and is used by clinicians. The information below is specific to OQ45, but gives a view of Outcome Questionnaires in general.

The subjects answer questions regarding mental and emotional key indicators at the beginning of the treatment and at specified intervals throughout. The OQ is meant to measure short term changes in subject behavior. The lower the score, the better the mental state of the subjects. The OQ is divided into 3 subsets. The Symptom Distress (SD), Interpersonal Relations (IR), and Social Score (SR) are all used to measure different factors regarding the subject's mental and emotional state.

The SD measures the symptoms of any mental or emotional disorders like anxiety or stress related disorders. This category ranges on a scale from 0-100, with a score of 36 or higher being considered an indication that the subject is showing clinically significant symptoms of a disorder. A score of 0 represents a subject without any symptoms.

The IR score measures the subject's ability to interact with other people and maintain a relationship with them. This score ranges on a scale of 0-44, with a score of 15 or more considered clinically significant. It not only measures potential disorders regarding interpersonal relationships, but also satisfaction with the relationships they have.

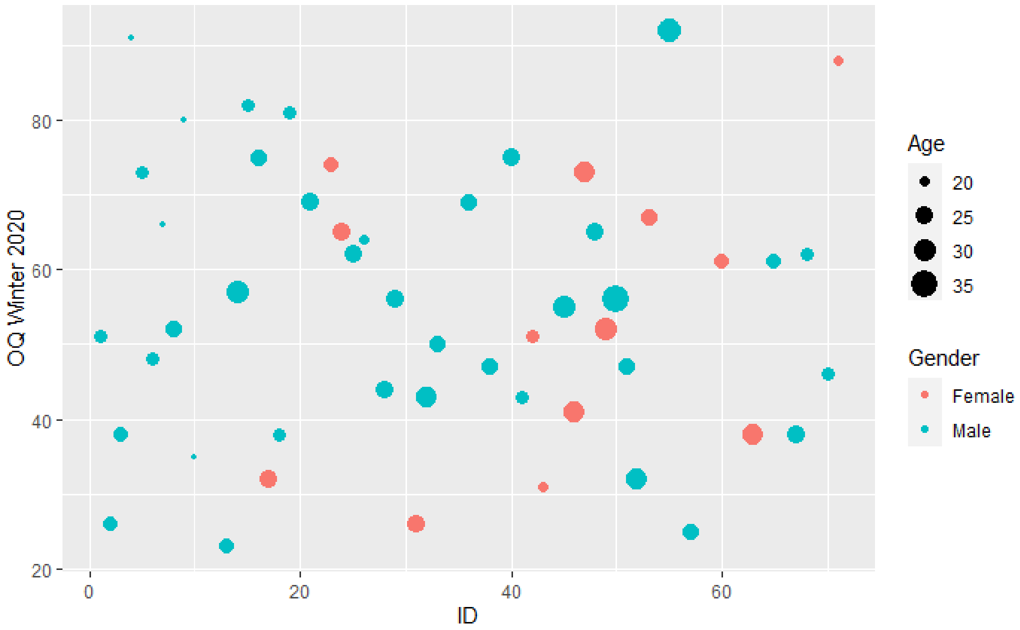
The SR score represents the subject's aptitude in filling social roles such as employee, husband/wife, student, and so forth. This scale ranges on a scale of 0-36, with a score of 12 or more considered clinically significant. It is important to note that low scores in this category can result from either aptitude or lack of social roles to fill (unemployed for example).

These three subsection scores are added up to generate the subject's individual OQ score. As with the three distinct categories, a higher score generally means that a subject has more mental or emotional distress and is potentially in need of further treatment.

Some of the individual questions have their own separate scoring and are to be considered individually. These questions regard suicide, substance abuse, and violence. If any of these scores meet the individual threshold requirements, immediate individual care is necessary.

The figure below shows a scatterplot of students' OQ scores in the Winter of 2020, where points are differentiated by color (gender) and size (age).

Figure 9: Scatterplot of OQ scores by gender and age



We notice a large variation in test scores, but apparently there is no connection between scores and either gender or age. In order to further investigate this, we create some plots and run some tests for the gender and age covariates separately.

Figure 10: Boxplot of OQ scores by gender

As we can see, the distribution of OQ scores for males is a little wider than that for females. The median and mean scores for males are slightly higher than those for females. However, these differences do not seem to be significant. The next two tables show the results of testing these hypotheses.

Table 5: t-test OQ scores by gender

Value of t statistic	-0.31805
Degrees of freedom	19.554
P-value	0.7538
Mean female	53.76923
Mean male	55.71053

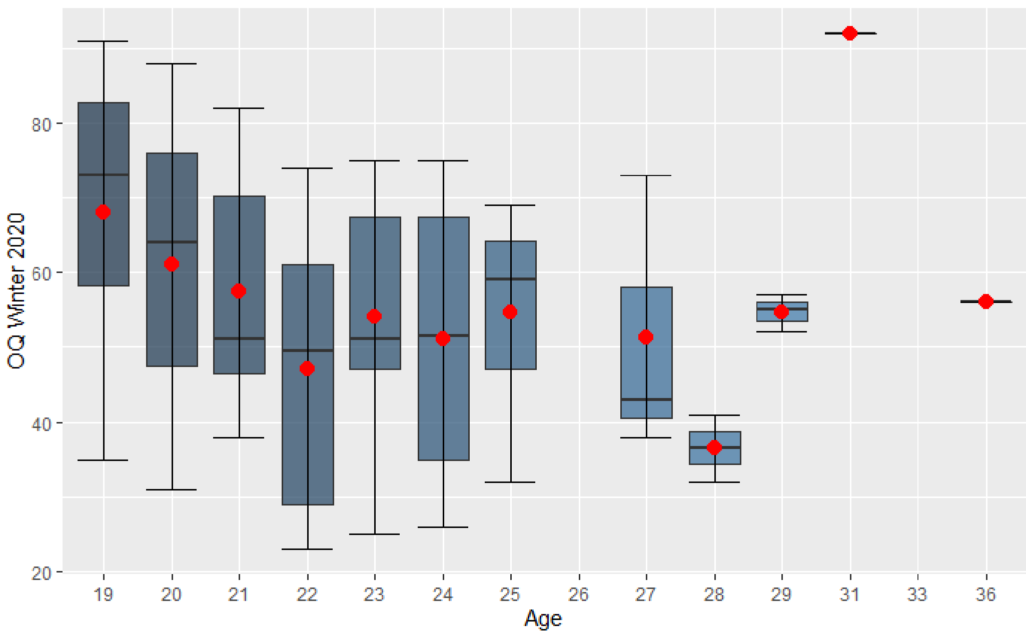
Table 6: Mood's Median test for OQ scores by gender

Value of z statistic	-0.23709
P-value	0.8126

Both tests show p-values higher than 0.05, so we cannot reject the hypotheses that there is no difference between the mean and median values of OQ scores for males and females.

We turn our attention now to the relationship between OQ scores and age. The boxplot in the figure below doesn't show any consistent pattern. Scores first decrease with age, then increase somewhat, decrease again and go back up for students 29 and older.

Figure 11: Boxplot of OQ scores by age



An ANOVA test will help determine if there is any difference in mean scores between age groups.

Table 7: ANOVA for OQ scores by age

Value of F statistic	0.238
P-value	0.628

The p-value of 0.628 tells us that the null hypothesis that the means of the different groups are the same cannot be rejected. Mood's median test, shown below, leads to the same conclusion for medians. Since the p-value of 0.7657 is higher than 0.05, there is no evidence to reject the null hypothesis that the medians of the different groups are the same. In summary, there is no evidence that OQ scores vary significantly with age.

Table 8: Mood's Median test for OQ scores by age

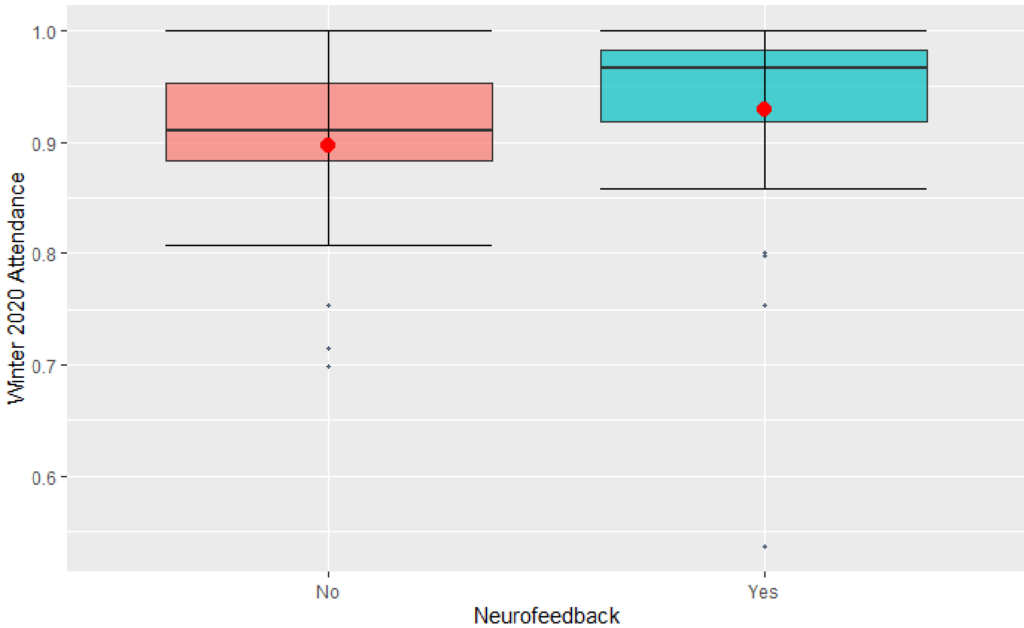
Value of chi-squared	7.4019
Degrees of freedom	11
P-value	0.7657

3.3. Distribution of outcome variables by participation in Neurofeedback program

The next order of business is to perform an initial evaluation of the impact of the Neurofeedback program on the two outcome variables we are considering: OQ scores and attendance. We split the sample into two groups, called treatment and control. The treatment group consists of the students who participated in the Neurofeedback program, and the control group, of those who didn't. We then compare the distributions of the outcome variables in these two groups.

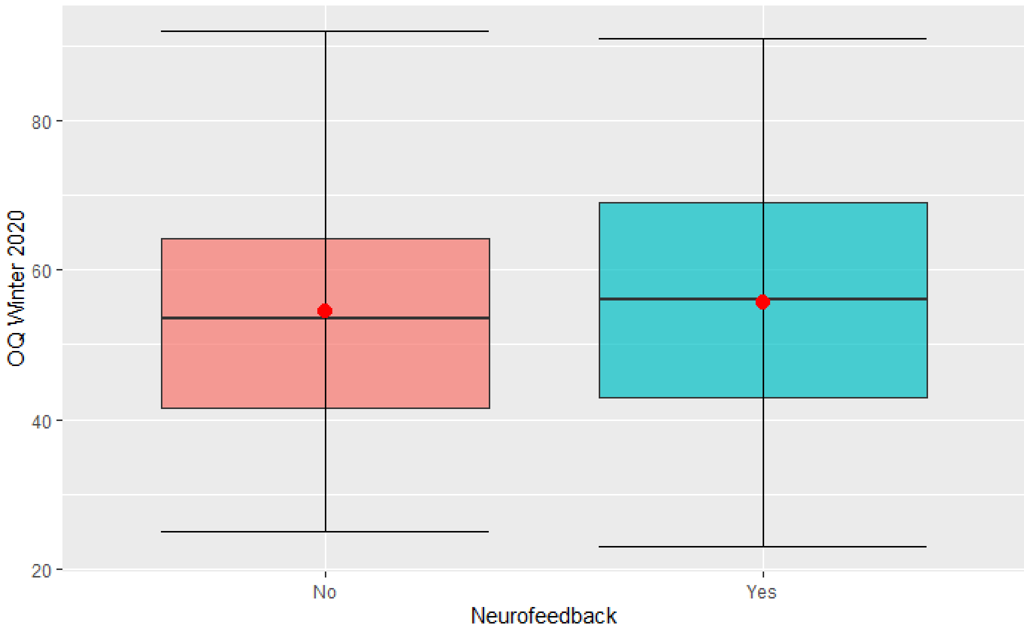
The figure below shows a boxplot of attendance by participation. Notice that the distribution is considerably higher for those who participated in the program, that is, those in the treatment group.

Figure 12: Boxplot of Attendance by participation in the Neurofeedback program



The boxplot of OQ scores by participation does not show a distinctive impact of the neurofeedback program, as can be seen below. The mean and median, and in fact the entire distribution of OQ scores for students who participated in the program are similar to those for students who did not.

Figure 13: Boxplot of OQ scores by participation in the Neurofeedback program



We now run a few tests to determine if there is any difference between the means and medians of the two groups. The next table shows the results of t-tests for means.

Table 9: t-test attendance and OQ scores by participation in Neurofeedback program

	Attendance	OQ scores
Value of t statistic	-1.3135	-0.24491
Degrees of freedom	51.793	46.48
P-value	0.1948	0.8076
Mean non-participant	0.8969	54.5
Mean participant	0.9292	55.7586

Even though the mean attendance of participants is 3 percentage points higher than that of non-participants, the t-test indicates that this difference is not statistically significant, for the p-value of 0.1948 is higher than the 5% significance level. As for the OQ scores, the p-value is a lot higher at 0.8076, so we conclude there is no significant difference between the means of participants and non-participants in this case either.

The results of the tests for differences in medians are shown in the table below.

Table 10: Mood's Median test for OQ scores and attendance by participation in Neurofeedback program

	Attendance	OQ scores
Value of z statistic	-2.2369	-0.43922
P-value	0.02529	0.6605

The test provides strong evidence that the medians of the OQ scores for participants and non-participants are the same. On the other hand, with a p-value of 0.02529, we reject the null hypothesis that the median attendances are the same. Combined with the t-test for mean attendance, this provides evidence, albeit weak, that the neurofeedback program has a positive effect on attendance rates. We will explore this possibility further in one of the next sections.

4. Data visualization for the treatment group (Neurofeedback participants)

In this section, we investigate if there is a relationship between neurofeedback training sections and student performance. Our hypothesis is that a larger numbers of training sections should translate into higher scores and attendance rates. In order to do so, we restrict our sample to students who participated in the program, which reduces the sample size to 39 observations. We should also point out that there is data on the number of training sessions for only 21 of the 39 students who participated in the neurofeedback program.

The scatterplots in the next two figures don't show any monotonic relationship between the outcome variables and the number of training sessions. The blue smoothing line is a LOESS curve fit to the data and the grey band gives the 95% confidence interval[4]. They were added to the plot to help detect trends.

Figure 14: Scatterplot of OQ scores by training sessions – Treatment group

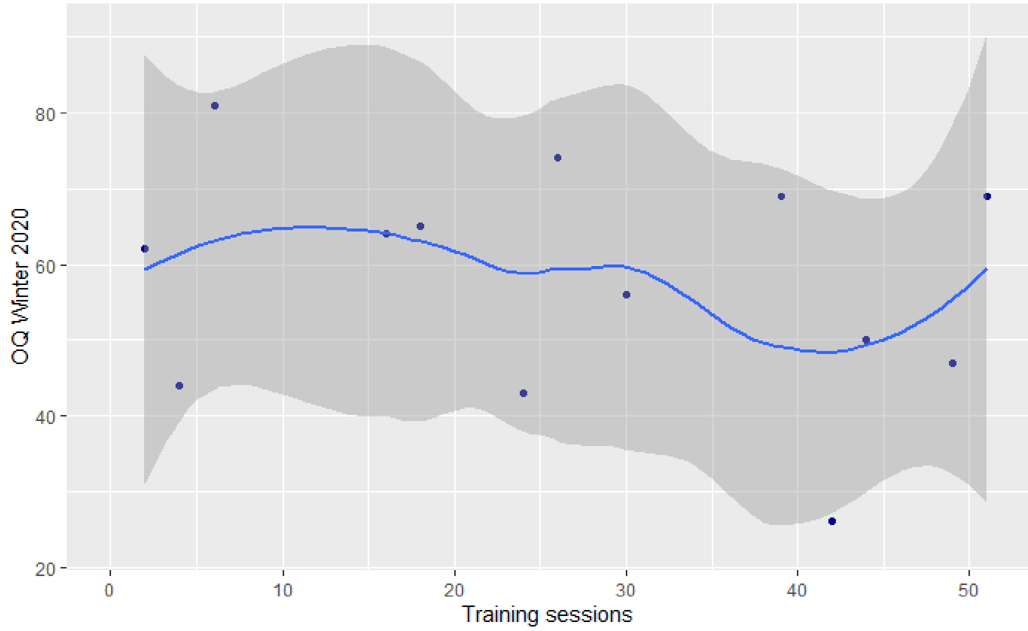
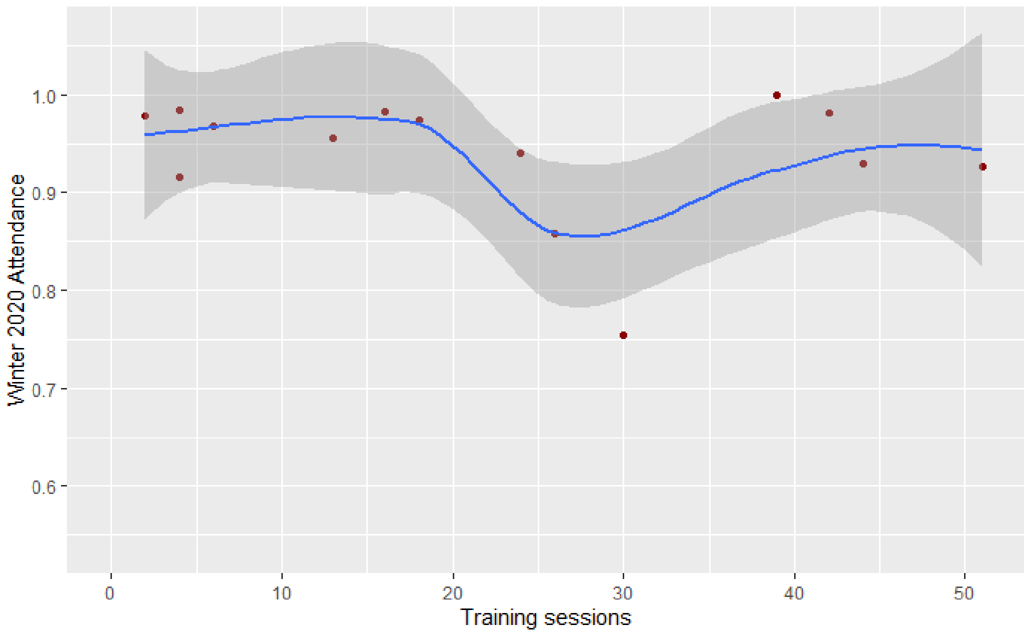


Figure 15: Scatterplot of attendance by training sessions – Treatment group

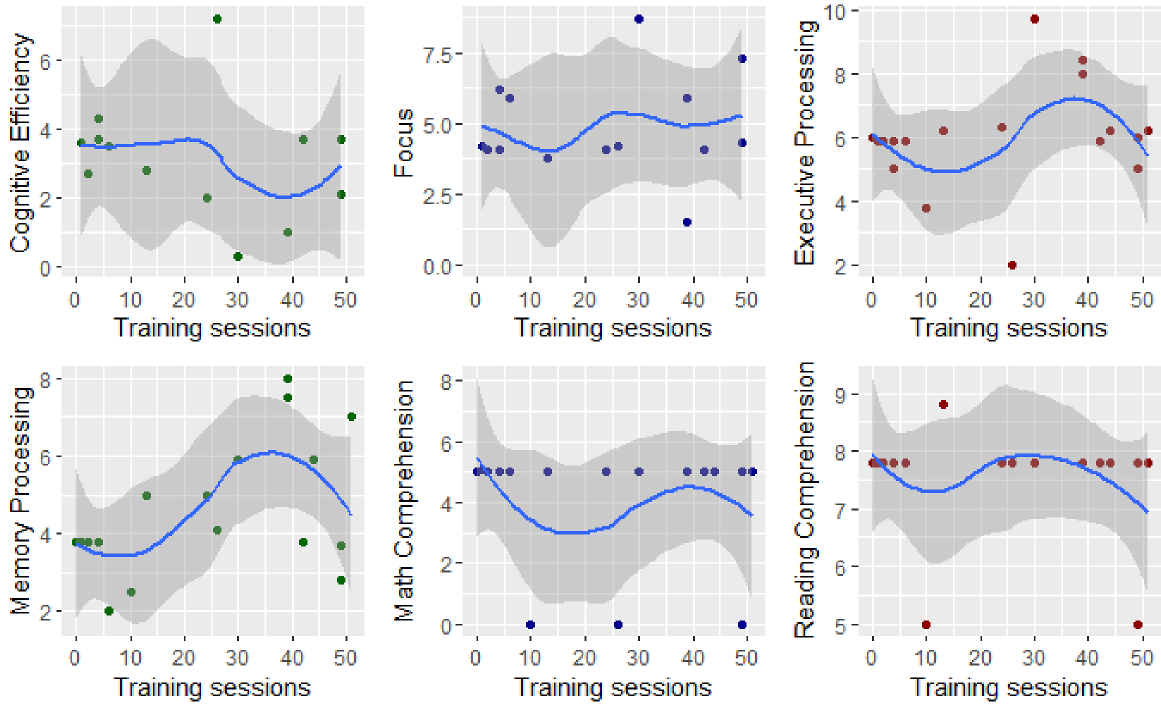


[4] LOESS denotes a method known as locally weighted polynomial regression. At each point in the data set a low-degree polynomial is fit to a subset of the data given the values of the explanatory variables near the point where the dependent variable is being estimated. The polynomial is fit using weighted least squares, with more weight given to points nearby and less weight given to points further away.

We now shift to the relationships between number of training sessions and some mental evaluation metrics that were collected only for students who participated in the Neurofeedback program (and for this reason, were not included in our set of outcome variables that measure student performance), namely cognitive efficiency, focus, executive processing, memory processing, math comprehension, and reading comprehension.

As can be seen in the figure below, there is no upward or downward trend connecting training sections to any of the mental evaluation metrics. In other words, the graphs don't present evidence that training sections have an effect on those indicators.

Figure 16: Scatterplot of cognitive efficiency, focus, executive processing, memory processing, math comprehension, and reading comprehension by training sessions – Treatment group



Based on the graphical analysis conducted above, we conclude that there is no evidence that the number of Neurofeedback training sections a student participates in has an effect on her OQ score, attendance, or mental abilities.

5. Estimation of treatment effects

In this section, we employ more advanced statistical methods to evaluate the potential impact of the Neurofeedback program on student performance. The first method is linear regression. Our model is

$$y = \beta_1 + \beta_2 \times \text{neuro} + \beta_3 \times gender + \beta_4 \times age + \varepsilon \,,$$

where y is the dependent or outcome variable (OQ scores or attendance), β_1 is the intercept, β_2, β_3 and β_4 are the coefficients of the explanatory variables (neuro, gender and age), and ε is a random disturbance. The 'neuro' variable is a dummy equal to 1 when the student participated in the Neurofeedback program and 0 otherwise.

The results of the linear regressions can be found in the tables below.

Table 11: Linear regression results – OQ scores

Variable	Coefficient	St. error	T statistic	P-value
Intercept	61.99313	20.69318	2.996	0.00436
Neuro	0.07679	5.80251	0.013	0.98950
Gender	1.65368	6.29594	0.263	0.79396
Age	-0.34255	0.80714	-0.424	0.67321

Table 12: Linear regression results – Attendance

Variable	Coefficient	St. error	T statistic	P-value
Intercept	0.783968	0.097769	8.019	1.56e-10
Neuro	0.040182	0.027320	1.471	0.148
Gender	0.005378	0.029090	0.185	0.854
Age	0.004437	0.003796	1.169	0.248

Notice that the neuro variable is not significant (p-value greater than 0.05) for either outcome variable, although the p-value for attendance is a lot smaller than that for OQ scores. This means that there is no statistical evidence that participation in the Neurofeedback program has an effect on students' OQ scores or attendance. The other explanatory variables, gender and age, are not significant either.

Even though the regression results were not promising, we decided to carry out a propensity score analysis to match students in the treatment group (participants) to students in the control group (non-participants) that have similar characteristics. This reduces the bias that might be present in the treatment assignment process.

We used a greedy one-to-one matching with replacement process based on the propensity score. Here are some statistics comparing the two groups before and after matching for the outcome variable 'Attendance'[5].

Table 13: Balance before and after matching – Attendance[6]

	Full sample			Matched sample			
	Mean treatment	Mean control	Stand. mean difference	Mean treatment	Mean control	Stand. mean difference	% improvement in Stand. mean difference
Distance	0.6216	0.4730	0.8551	0.6216	0.6186	0.0174	97.9594
Female	0.1667	0.4167	-0.6595	0.1667	0.1	0.1759	73.3333
Male	0.8333	0.5833	0.6595	0.8333	0.9	-0.1759	73.3333
Age	22.6667	24.75	-0.7767	22.667	23.2667	-0.2237	71.2

As can be seen in the table above, there were considerable differences between the treatment and control groups before matching. The average age in the treatment group was more than two years lower than in the control group, and the share of female students was 25 percentage points lower in the treatment group. After matching, the balance improved considerably. The percentage improvements in the standardized mean differences were 97.96%, 73.33% and 71.2% for the variables distance, gender and age, respectively.

We ran t-tests and regressions on the matched data. The results for attendance are shown in the next two tables.

Table 14: t-test Attendance by participation in Neurofeedback program

Value of t statistic	-1.9681
Degrees of freedom	22.951
P-value	0.06125
Mean non-participant	0.86554
Mean participant	0.9292

The p-value of the t-test is 0.06125, so we do not reject the null hypothesis at the 5% significance level. Notice, however, that the two means are statistically significant at the 10% level (actually, at the 6.125% level). This relationship is further investigated in a linear regression.

[5] We had to create two matched data sets, one for attendance and one for OQ scores, because the observations with missing data for those variables did not coincide.
[6] The 'distance' variable is the propensity score.

Table 15: Linear regression results for matched data – Attendance

Variable	Coefficient	St. error	T statistic	P-value
Intercept	0.731098	0.139713	5.233	5.99e-06
Neuro	0.067342	0.033319	2.021	0.0502
Gender	0.025608	0.039568	0.647	0.5213
Age	0.004827	0.005374	0.898	0.3746

After we control for gender and age (which are not statistically significant), the p-value for the neurofeedback dummy is barely higher than 5%. If we are willing to accept this significance level, we have evidence that participation in the neurofeedback program has an impact on attendance. Since the coefficient of the 'neuro' variable is 0.067342, we conclude that this impact is positive, that is, participants have, on average, a higher attendance rate than non-participants (in the Winter of 2020).

We replicated this type of analysis for OQ scores. The first table shows how the full and matched samples compare.

Table 16: Balance before and after matching – OQ scores

	Full sample			Matched sample			
	Mean treatment	Mean control	Stand. mean difference	Mean treatment	Mean control	Stand. mean difference	% improvement in Stand. mean difference
Distance	0.6426	0.4711	0.9989	0.6426	0.6425	0.0009	99.9086
Female	0.1379	0.4091	-0.7727	0.1379	0.690	0.1965	74.5665
Male	0.8621	0.5909	0.7727	0.8621	0.9310	-0.1965	74.5665
Age	22.5862	24.7273	-0.8613	22.5862	23.1724	-0.2358	72.6208

Similar to what happened when we constructed a matched sample for the variable 'Attendance', there were significant differences in the covariates between the treatment and control groups before matching. Matching improved the balance substantially, with percentage improvements in standardized mean differences of 99.91%, 74.57% and 72.6% for the variables distance, gender and age, respectively.

A t-test of OQ scores by participation in the Neurofeedback program generated the results shown in the table below. Notice that the p-value is very high, so we do not reject the assumption that the average OQ score is the same in the treatment and control groups.

Table 17: t-test OQ scores by participation in Neurofeedback program

Value of t statistic	-0.079932
Degrees of freedom	20.826
P-value	0.9371
Mean non-participant	55.25
Mean participant	55.75862

The conclusion that participation in the Neurofeedback program does not have an impact on OQ scores is confirmed by a linear regression, whose results are seen in the table below. As a matter of fact, all of the regressors are not statistically significant, which means that our model does not explain OQ scores well.

Table 18: Linear regression results for matched data – OQ scores

Variable	Coefficient	St. error	T statistic	P-value
Intercept	40.1918	27.2464	1.475	0.149
Neuro	0.5320	6.5302	0.081	0.936
Gender	8.76	8.3272	1.052	0.300
Age	0.3313	1.0947	0.303	0.764

6. Data issues and discussion of main findings

As we mentioned throughout this report, there are many missing values in our data. This has affected our analysis in different ways. First, we were not able to provide a complete picture of how the sample is distributed by student characteristics, especially ethnicity. Second, it limited our outcome variables to just two, attendance and OQ scores in the Winter of 2020. The original data set has data on OQ scores and attendance for other periods, but there are so many gaps in the data that we were not able to use them. It also contains information on test scores other than OQ and health indicators, but there is no data available for students who did not participate in the Neurofeedback program, which prevented us from including them in our comparison of the treatment (participants) and control (non-participants) groups. We should also point out that there is no information in the data set about the exact dates of the Neurofeedback training sessions or when tests and health measures were taken. This complicates the analysis, for we can only use outcome indicators that were measured after the intervention. That explains why we decided to work with the latest period available to us, namely the Winter of 2020.

That being said, let's now summarize our main findings. It is clear from the graphical analysis as well as statistical tests and linear regressions that the neurofeedback program did not have a statistically significant impact on students' performance on the OQ test. Boxplots and other plots did not show any consistent relationship between participation in the neurofeedback program and OQ scores, a finding that was reinforced by the high p-values generated by a t-test, a median test and a linear regression using the entire data set (minus observations with missing values). This conclusion was confirmed by a t-test and a linear regression carried out on a matched sample where the treatment and control groups were better balanced.

The initial graphical and statistical analyses of the neurofeedback program's impact on students' rates of attendance provided conflicting evidence. The boxplot showed that the distribution of the outcome variable 'attendance' was higher for those who participated in the program, but the t-test for means indicated that they were not statistically different. This finding was confirmed by a linear regression run on the entire data set (minus observations with missing values) that, in addition to participation in the program, included gender and age as regressors. On the other hand, the Mood's median test showed that the two medians were different, statistically speaking. This statistical controversy was revisited in a propensity score analysis, which included a t-test and a linear regression run on a new sample with treatment and control groups matched with respect to the propensity score. The p-value of the t-test was slightly higher than 5%, while that of the neurofeedback dummy in the regression was practically indistinguishable from 5%. Taking all the evidence into account, it is likely that the neurofeedback program had a positive impact on attendance rates in the Winter term of 2020.

7. Conclusion

In this study, we investigated the effects of a Neurofeedback program recently implemented by ScenicView Academy on their student body. In order to do that, we split our sample of 71 students into a treatment group (participants) and a control group (non-participants) and compared their performance on two outcome variables: attendance and scores on an Outcome Questionnaire test.

Through a mixture of qualitative (graphs) and quantitative (mean and median tests and linear regression) methods, and with the help of matching techniques, we found that the Neurofeedback program did not have an impact on OQ scores, but that there is qualified evidence that it did have an impact on attendance. In other words, it is likely that the neurofeedback intervention increased attendance rates.

Given the relatively new implementation of the Neurofeedback program at ScenicView Academy, it is not surprising that our study did not turn up definitive evidence of its impact on student performance. Nevertheless, our findings should be seen as a first step in that direction. We expect our continuing collaboration with ScenicView Academy to generate more and higher quality data, allowing us to widen the scope of our studies and produce a new batch of results.

References:

Marzbani, H., H.R. Marateb, and M. Mansourian (2016). "Neurofeedback: A Comprehensive Review on System Design, Methodology and Clinical Applications." *Basic and clinical neuroscience* 7(2), p. 143-158.

Sitaram, R. et al. (2017). "Closed-Loop Brain Training: The Science of Neurofeedback. *Nature Reviews Neuroscience* 18, p. 86-100.