Below is the reported pretest and posttest means, standard deviations,and paired t-test statistics for each semester.

**Bodie Scale:** The Active-Empathetic Listening Scale was developed to assess effective vs ineffective listening from the point of view of customers. The scale has been refined and adapted to a more general social context. We are using this scale as our main score to measure empathy. One who actively listens to others is empathetic. The total score from the 11 questions from this category (a Likert Scale) is a measure of empathy, and the mean score per person is out of 7 from a 1-7 scale. 7 is the highest score of the most empathy.

**Busiol Scale:** A 1-7 scale that determines how well a person does at actively attempting to understand a person in a conversation. This is its own type of empathy. The Bodie Scale is a measure of how a person naturally empathizes with others, but the Busiol measures what a person actively does in order to understand the other person. Bodie is reactionary, Busiol is more of the action taken. 7 is the highest score of how much effort a person puts into active listening.

**Fassaert Scale:** A 1-5 scale that determines the characteristics of a person who is already interested in what another person is saying. A person who has a 5 in the Fassaert Scale does an excellent job of showing physical reciprocation, and keeps the conversation going as well and as deep as possible. The difference between the three empathy scales listed above is very slight and takes a bit of contemplating to understand. But the measures are separate, and we believe that they all do a great job of measuring empathy.

**Experiencing Bias Scale:** 1-7 Scale showing how much a person has experienced bias in their own lives or witnessed in others lives. This number shouldn’t change from beginning to the end of the semester, unless the humanities class opened their eyes to concepts they hadn’t considered before.

**Diversity Scale:**  1-5 Scale that shows how much people get along naturally with people of other cultures. This scale did not change much from the beginning to the end of the semester, whether the class was a humanities class or a control.

\*\*\* For control groups, Busiol data was not separated into Busiol and Fassaert scores. The busiol p-values were generally huge so we didn’t dig into this further.

**Experimental, Fall 2019**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Pre/Post** | **Scale** | **Mean** | **Std. Dev** | **N** |
| Pretest | Bodie | 5.23 | 0.72 | 64 |
|  | Busiol | 5.22 | 0.72 |  |
|  | Fassaert | 4.04 | 0.51 |  |
|  | Experiencing Bias | 4.21 | 0.88 |  |
|  | Diversity | 3.25 | 0.29 |  |
| Posttest | Bodie | 5.45 | 0.74 | 64 |
|  | Busiol | 5.47 | 0.59 |  |
|  | Fassaert | 4.25 | 0.59 |  |
|  | Experiencing Bias | 4.13 | 0.82 |  |
|  | Diversity | 3.36 | 0.27 |  |

T-Test Results:

Bodie Scale:

* Difference of the Means: 0.230
* P-Value: 0.085
* T-Statistic: 1.75
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. However, we are very close. We haven’t separated the different Humanities classes at this point, but looking at the other semesters results, we can *almost* conclude that a person’s Bodie score has gone up after the Humanities class

Busiol Scale:

* Difference of the Means: 0.247
* P-Value: 0.019
* T-Statistic: 2.406
* Conclusion: Reject the null hypothesis. There is a significant difference between the means, and we can say that a person’s Busiol score has gone up after taking the Humanities Class.

Fassaert Scale:

* Difference of the Means: 0.206
* P-Value: 0.036
* T-Statistic: 2.14
* Conclusion: Reject the null hypothesis. There is a significant difference between the means, and we can say that a person’s Fassaert score has gone up after taking the Humanities Class.

**Psychology (Control), Fall 2019**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Pre/Post** | **Scale** | **Mean** | **Std. Dev** | **N** |
| Pretest | Bodie | 5.09 | 0.82 | 22 |
|  | Busiol | 4.98 | 0.55 |  |
|  | Fassaert |  |  |  |
|  | Experiencing Bias | 4.42 | 0.76 |  |
|  | Diversity | 3.42 | 0.40 |  |
| Posttest | Bodie | 5.17 | 0.82 | 22 |
|  | Busiol | 5.77 | 0.52 |  |
|  | Fassaert |  |  |  |
|  | Experiencing Bias | 4.04 | 0.93 |  |
|  | Diversity | 3.19 | 0.34 |  |

Bodie Scale:

* Difference of the Means: 0.08
* P-Value: 0.746
* T-Statistic: -0.327
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.7457, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

Busiol Scale:

* Difference of the Means: 0.035
* P-Value: 0.747
* T-Statistic: 0.328
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.7466, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

Diversity Scale:

* Difference of the Means: -0.054
* P-Value: 0.556
* T-Statistic: -0.60
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.556, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

Experiencing Bias Scale:

* Difference of the Means: -0.177
* P-Value: 0.3016
* T-Statistic: -1.06
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.3016, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

**Kinesiology (Control), Fall 2019**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Pre/Post** | **Scale** | **Mean** | **Std. Dev** | **N** |
| Pretest | Bodie | 5.36 | 0.72 | 49 |
|  | Busiol | 4.69 | 0.51 |  |
|  | Fassaert |  |  |  |
|  | Experiencing Bias | 3.89 | 0.63 |  |
|  | Diversity | 3.20 | 0.42 |  |
| Posttest | Bodie | 5.35 | 0.788 | 49 |
|  | Busiol | 4.69 | 0.68 |  |
|  | Fassaert |  |  |  |
|  | Experiencing Bias | 4.00 | 0.65 |  |
|  | Diversity | 3.22 | 0.41 |  |

Bodie Scale:

* Difference of the Means: -0.0008
* P-Value: 0.9938
* T-Statistic: -0.0078
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.9938, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

Experiencing Bias Scale:

* Difference of the Means: 0.1159
* P-Value: 0.302
* T-Statistic: 1.04
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.302, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

Busiol Scale:

* Difference of the Means: -0.004
* P-Value: 0.9672
* T-Statistic: -0.04
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.9672, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

Diversity Scale:

* Difference of the Means: -0.177
* P-Value: 0.7659
* T-Statistic: 0.299
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.7659, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

**Experimental Spring, 2020**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Pre/Post** | **Scale** | **Mean** | **Std. Dev** | **N** |
| Pretest | Bodie | 5.27 | 0.82 | 58 |
|  | Busiol | 5.32 | 0.71 |  |
|  | Fassaert | 4.16 | 0.53 |  |
|  | Experiencing Bias | 4.28 | 0.776 |  |
|  | Diversity | 3.30 | 0.28 |  |
| Posttest | Bodie | 5.65 | 0.876 | 58 |
|  | Busiol | 5.65 | 0.77 |  |
|  | Fassaert | 4.38 | 0.61 |  |
|  | Experiencing Bias | 4.20 | 0.85 |  |
|  | Diversity | 3.38 | 0.35 |  |

Bodie Scale:

* Difference of the Means: 0.388
* P-Value: 0.003
* T-Statistic: 3.12
* Conclusion: Reject Null Hypothesis. The difference of .388 from pre- to post- prove that the humanities class raises the Bodie empathy scores of the students.

Experiencing Bias Scale:

* Difference of the Means: -0.069
* P-Value: 0.610
* T-Statistic: -0.51
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.610, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

Busiol Scale:

* Difference of the Means: 0.332
* P-Value: 0.01
* T-Statistic: 2.65
* Conclusion: Reject Null Hypothesis. The difference of .331 from pre- to post- prove that the humanities class raises the Busiol empathy scores of the students.

Fassaert Scale:

* Difference of the Means: 0.219
* P-Value: 0.0218
* T-Statistic: 2.36
* Conclusion: Reject Null Hypothesis. The difference of .219 from pre- to post- prove that the humanities class raises the Fassaert empathy scores of the students.

Diversity Scale:

* Difference of the Means: 0.079
* P-Value: 0.1076
* T-Statistic: 1.635
* Conclusion: Fail to report a significant difference from pre- to post- over their experience with diversity.

**Control Spring, 2020**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Pre/Post** | **Scale** | **Mean** | **Std. Dev** | **N** |
| Pretest | Bodie | 5.42 | 0.67 | 41 |
|  | Busiol | 4.76 | 0.59 |  |
|  | Fassaert |  |  |  |
|  | Experiencing Bias | 4.11 | 0.95 |  |
|  | Diversity | 3.16 | 0.28 |  |
| Posttest | Bodie | 5.37 | 0.94 | 41 |
|  | Busiol | 4.65 | 0.61 |  |
|  | Fassaert |  |  |  |
|  | Experiencing Bias | 4.20 | 1.03 |  |
|  | Diversity | 3.22 | 0.30 |  |

Bodie Scale:

* Difference of the Means: -0.040
* P-Value: 0.6832
* T-Statistic: -0.41
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.6832, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

Experiencing Bias Scale:

* Difference of the Means: 0.09
* P-Value: 0.4599
* T-Statistic: 0.75
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.4599, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

Busiol Scale:

* Difference of the Means: -0.11
* P-Value: 0.2818
* T-Statistic: -1.09
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.2818, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

Fassaert Scale: N/A

Diversity Scale:

* Difference of the Means: 0.069
* P-Value: 0.1134
* T-Statistic: 1.62
* Conclusion: Fail to report a significant difference from pre- to post- over their experience with diversity.

**Experimental Fall, 2020**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Pre/Post** | **Scale** | **Mean** | **Std. Dev** | **N** |
| Pretest | Bodie | 5.52 | 0.66 | 71 |
|  | Busiol | 5.48 | 0.68 |  |
|  | Fassaert | 4.19 | 0.56 |  |
|  | Experiencing Bias | 4.17 | 0.69 |  |
|  | Diversity | 3.33 | 0.34 |  |
| Posttest | Bodie | 5.85 | 0.74 | 71 |
|  | Busiol | 5.63 | 0.97 |  |
|  | Fassaert | 4.39 | 0.55 |  |
|  | Experiencing Bias | 4.24 | 0.68 |  |
|  | Diversity | 3.39 | 0.38 |  |

Bodie Scale:

* Difference of the Means: 0.328
* P-Value: 0.004
* T-Statistic: 3.018
* Conclusion: Reject Null Hypothesis. The difference of .328 from pre- to post- prove that the humanities class raises the Bodie empathy scores of the students.

Experiencing Bias Scale:

* Difference of the Means: 0.066
* P-Value: 0.5391
* T-Statistic: 0.617
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.5391, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

Busiol Scale:

* Difference of the Means: 0.15
* P-Value: 0.2269
* T-Statistic: 1.219
* Conclusion: \*\*\* Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.2269, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0. (surprising result)

Fassaert Scale:

* Difference of the Means: 0.193
* P-Value: 0.024
* T-Statistic: 2.31
* Conclusion: Reject Null Hypothesis. The difference of .193 from pre- to post- prove that the humanities class raises the Fassaert empathy scores of the students.

Diversity Scale:

* Difference of the Means: 0.062
* P-Value: 0.2335
* T-Statistic: 1.20
* Conclusion: Fail to report a significant difference from pre- to post- over their experience with diversity.

**Control Fall, 2020**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Pre/Post** | **Scale** | **Mean** | **Std. Dev** | **N** |
| Pretest | Bodie | 5.13 | 0.89 | 43 |
|  | Busiol | 4.72 | 0.47 |  |
|  | Fassaert |  |  |  |
|  | Experiencing Bias | 3.95 | 0.81 |  |
|  | Diversity | 3.20 | 0.27 |  |
| Posttest | Bodie | 5.28 | 0.83 | 43 |
|  | Busiol | 4.73 | 0.54 |  |
|  | Fassaert |  |  |  |
|  | Experiencing Bias | 3.91 | 0.77 |  |
|  | Diversity | 3.27 | 0.29 |  |

Bodie Scale:

* Difference of the Means: 0.156
* P-Value: 0.2487
* T-Statistic: 1.17
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.2487, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

Experiencing Bias Scale:

* Difference of the Means: 0.04
* P-Value: 0.7558
* T-Statistic: 0.31
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. With a p-value of 0.7558, we are way far off from rejecting the null hypothesis. We cannot say that the true difference in means is not equal to 0.

Busiol Scale:

* Difference of the Means: 0.142
* P-Value: 0.142
* T-Statistic: 1.50
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. We cannot say that the true difference in means is not equal to 0.

Fassaert Scale: N/A

Diversity Scale:

* Difference of the Means: 0.065
* P-Value: 0.1747
* T-Statistic: 1.38
* Conclusion: Fail to report a significant difference from pre- to post- over their experience with diversity.

**Experimental Spring, 2021**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Pre/Post** | **Scale** | **Mean** | **Std. Dev** | **N** |
| Pretest | Bodie | 5.45 | 0.79 | 40 |
|  | Busiol | 4.99 | 0.56 |  |
|  | Fassaert | 3.89 | 0.54 |  |
|  | Experiencing Bias |  |  |  |
|  | Diversity |  |  |  |
| Posttest | Bodie | 5.74 | 0.68 | 40 |
|  | Busiol | 5.64 | 0.53 |  |
|  | Fassaert | 4.82 | 0.30 |  |
|  | Experiencing Bias |  |  |  |
|  | Diversity |  |  |  |

Bodie Scale:

* Difference of the Means: 0.2835
* P-Value: 0.1129
* T-Statistic: 1.62
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. However, we are very close. We haven’t separated the different Humanities classes at this point, but looking at the other semesters results, we can *almost* conclude that a person’s Bodie score has gone up after the Humanities class

Experiencing Bias Scale:

Busiol Scale:

* Difference of the Means: 0.65
* P-Value: 4.27 e-06
* T-Statistic: 5.33
* Conclusion: Reject Null Hypothesis. The difference of .65 from pre- to post- prove that the humanities class raises the Bodie empathy scores of the students.

Fassaert Scale:

* Difference of the Means: 0.93
* P-Value: 2.14e-10
* T-Statistic: 8.49
* Conclusion: Reject Null Hypothesis. The difference of .93 from pre- to post- prove that the humanities class raises the Fassaert empathy scores of the students. (0.93 is easily the greatest reported difference between a pre- and post- score of all datasets)

Diversity Scale:

\*\*\*Diversity and Bias scores not calculated in this instance. Neither was Control for Spring 2021. At this point, Kenneth and Kevin have concluded that the diversity and bias scores never jump up significantly, whether its an experimental or a control group. We have also concluded that no control groups have had any jump in empathy, so to continue to evaluate such groups would pointless to us.

**Giant Dataset, Experimental**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Pre/Post** | **Scale** | **Mean** | **Std. Dev** | **N** |
| Pretest | Bodie | 5.37 | 0.75 | 233 |
|  | Busiol | 3.84 | 0.50 |  |
|  | Fassaert | 2.60 | 0.35 |  |
|  | Experiencing Bias |  |  |  |
|  | Diversity |  |  |  |
| Posttest | Bodie | 5.68 | 0.78 | 233 |
|  | Busiol | 4.06 | 0.55 |  |
|  | Fassaert | 2.81 | 0.37 |  |
|  | Experiencing Bias |  |  |  |
|  | Diversity |  |  |  |

Bodie Scale:

* Difference of the Means: 0.308
* P-Value: 3.605e-06
* T-Statistic: 4.75
* Conclusion: Reject null hypothesis! Over 4 years of experimental courses, the average Bodie score has gone up by 0.308 (4.4%)

Experiencing Bias Scale:

Busiol Scale:

* Difference of the Means: 0.225
* P-Value: 8.028e-07
* T-Statistic: 5.07
* Conclusion: Reject Null Hypothesis. The difference of .225 from pre- to post- proves that the humanities class raises the Busiol empathy scores of the students.

Fassaert Scale:

* Difference of the Means: 0.210
* P-Value: 3.526e-10
* T-Statistic: 6.557
* Conclusion: Reject Null Hypothesis. The difference of .210 from pre- to post- prove that the humanities class raises the Fassaert empathy scores of the students.

**Giant Dataset, Control**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Pre/Post** | **Scale** | **Mean** | **Std. Dev** | **N** |
| Pretest | Bodie | 5.32 | 0.77 | 155 |
|  | Busiol | 3.82 | 0.49 |  |
|  | Fassaert | 2.62 | 0.36 |  |
|  | Experiencing Bias |  |  |  |
|  | Diversity |  |  |  |
| Posttest | Bodie | 5.31 | 0.85 | 155 |
|  | Busiol | 3.83 | 0.56 |  |
|  | Fassaert | 2.56 | 0.38 |  |
|  | Experiencing Bias |  |  |  |
|  | Diversity |  |  |  |

Bodie Scale:

* Difference of the Means: 0.02
* P-Value: 0.6957
* T-Statistic: 0.39
* Conclusion: Fail to reject the null hypothesis, at 95% confidence. We not even close. We have once again proven that the control courses have nothing to do with Bodie scale score changes.

Experiencing Bias Scale:

Busiol Scale:

* Difference of the Means: 0.046
* P-Value: 0.3394
* T-Statistic: 0.96
* Conclusion: Fail to reject the null hypothesis. We have once again proven that the control courses have nothing to do with Busiol scale score changes.

Fassaert Scale:

* Difference of the Means: -0.035
* P-Value: 0.3009
* T-Statistic: -1.038
* Conclusion: Fail to reject the null hypothesis. We have once again proven that the control courses have nothing to do with Fassaert scale score changes.