Data Analytics CA Pair Project - VR Jungian Sandplay

Charlie Duff, Adam Heaton, Ademide Adenuga

December, 2022

Contents

Abstract

Aim and Rationale

The aim and rationale of this project is to formulate a single hypothesis based on a set of data we were given on an experiment conducted to attempt to determine the effectiveness of use of VR technology in Jungian Sandbox therapy, comparing the effectiveness of the approaches based on the data we were provided, and cleaning up any data that requires cleanup.

Participants and setting

Participants in the study are psychotherapy patients undergoing treatment for PTSD with Jungian Sandplay alongside a therapist, using either traditional Cognitive Behavioural Therapy or in a virtual reality environment.

Experiment design

Each participant spent 50 minutes per week for 12 weeks with a therapist, under three testing groups:

Control: Undergoing traditional CBT with a therapist.

Static: Using the VR app with non-animated model content.

Animated: Using the VR app with animated content.

Results gathering

Participants and their therapists were asked to rate their level of PTSD between 0-10 before and after their treatment, with 0 being no PTSD and 10 being the highest level of PTSD.

Major findings

We found that the average PTSD rating for each testing group decreased after treatment. Post-trial results were more widely dispersed than pre-trial results, leading us to conclude that the treatment was not as effective for some participants as others. Self-reported results are more correlated than observer-reported results, leading us to conclude that the treatment seemed more beneficial for participants than observers.

Findings/Implications

An outlier was discovered in the female control section of our data. This outlier's data could have affected our mean calculations for each set of data. Both post trials reports failed tests for normality showing a possibility of outliers being present in our data or clusters of specific values being found in our data.

Introduction

Topic and context

Jungian Sandplay is a psychotherapeutic method where participants play with sand and miniature objects to create an image that can reflect internal experiences and difficulties. Post-Traumatic Stress Disorder (PTSD) is a psychological condition caused by witnessing or experiencing an event or situation. Jungian Sandplay has been shown to decrease levels of PTSD in individuals.

Rationale

The Rationale behind this experiment was to determine if using Virtual Reality (VR) technology in a Jungian Sandbox setting could be beneficial in reducing the PTSD levels in patients.

Hypothesis

Our null hypothesis for this experiment is that post-trial values for OR results should correlate with post-trial CPSS results. This would infer that therapists observe as much of a benefit from the trial as the participant.

Our alternate hypothesis is that OR values will not correlate with the CPSS values.

Following tests done to check normality done to our data, we found that the post CPSS and OR values of the data were not normally distributed. This casts doubt on the quality of data collected for the post CPSS and OR values as outliers or skewed clusters could be present in our data causing us to reject our null hypothesis. Our dot plot shows a more positive correlation between pre and post-trial CPSS than pre and post-trial OR.

Method

Participants

Participants of this experiment were psychotherapy patients undergoing treatment for PTSD, aged between 18 - 25 years old with no exact ages of the participants recorded. 150 participants were recorded in this study with an equal amount of male and female participants recorded (Male = 75 / Female = 75).

Participants were broken into 3 groups using random sampling. These groups were then assigned one of three testing groups (Control, Static, Animated).

Design

Each participant had their level of PTSD rated between 0-10 both before and after their 12 weeks of treatment, assessed by themselves (Child PTSD Symptom Scale Seld Report - SR) and by their therapist (observer reported – OR). These results were collated in a CSV file tracking the participants' gender, test group, pre and post-trial self-reported results, and pre and post-trial observer reported results.

Materials

The materials used for this experiment were the results of the pre and post-trials, both independently observed, and the results stemming from records by the therapists. Additionally, RStudio was used to collate all the data and information needed to be documented.

Procedure

Patients within this experiment underwent 12 weeks of Jungian sand play treatment. Results of this treatment were recorded at the start and end of this experiment. Results we compiled into four sections. The CPSS (Self report) and OR (Observed report) were recorded both pre and post this experiment.

We imported the dataset into R studio and began running a series of tests to determine the validity of the data and to generate a hypothesis based on the data set we had been provided. Using imputation, we were able to resolve the missing data in our "pre_trial_ccpp" section by using the mean of that column to get a value to insert into that missing data. We calculated the mean, median, and standard deviation of the datasets along with graphing our datasets to better understand our data to help us formulate a hypothesis. We used the Shapiro wilk test to determine the normality of our dataset. We also calculated the 95% confidence interval for each of our datasets to help generate a more accurate representation of our true mean.

Results

Table 1: Patient Data

X	gender	test_group	pre_trial_cpss	post_trial_cpss	pre_trial_or	post_trial_or
1	Male	Static	4.54	5.77	4.48	5.95
2	Male	Static	6.36	5.29	6.08	5.24
3	Male	Static	5.17	6.86	5.08	6.90
4	Male	Static	4.56	5.59	4.09	5.36
5	Male	Static	3.84	5.43	4.33	5.26
6	Male	Static	7.09	6.92	7.58	6.69
7	Male	Static	4.91	5.46	4.89	5.50
8	Male	Static	7.24	5.65	7.37	5.72
9	Male	Static	6.24	7.54	6.46	7.73
10	Male	Static	8.52	5.51	8.51	5.70
11	Male	Static	5.83	6.10	6.16	6.31
12	Male	Static	7.78	5.18	7.28	5.00
13	Male	Static	5.76	5.11	5.35	5.10
14	Male	Static	6.29	6.19	6.82	6.37
15	Male	Static	7.31	5.18	7.16	5.17
16	Male	Static	6.61	4.19	6.54	4.15
17	Male	Static	5.73	4.43	5.75	4.37
18	Male	Static	6.93	6.46	6.79	6.45
19	Male	Static	7.41	7.91	7.46	7.68
20	Male	Static	7.00	6.13	7.49	6.35
21	Male	Static	7.88	5.47	7.90	5.32
22	Male	Static	6.25	4.97	6.00	5.18
23	Male	Static	5.88	5.12	6.04	4.93
24	Male	Static	5.79	7.24	6.30	7.18
25	Male	Static	7.56	6.47	7.60	6.64
26	Female	Static	6.28	5.72	5.98	5.95
27	Female	Static	5.50	4.65	5.50	4.50
28	Female	Static	5.81	5.07	6.07	5.17
29	Female	Static	5.06	4.03	4.53	3.96
30	Female	Static	6.17	5.58	6.54	5.69
31	Female	Static	6.06	5.41	5.87	5.58
32	Female	Static	6.56	6.11	6.96	6.30

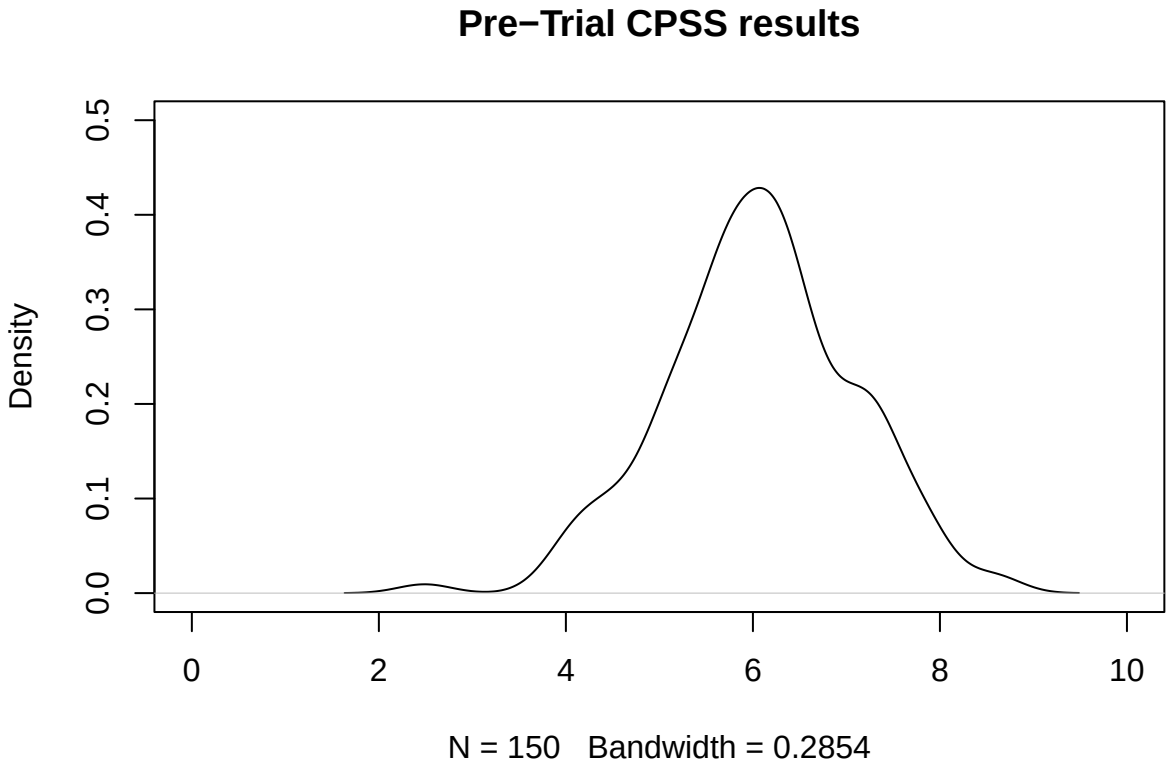
X	gender	test_group	pre_trial_cpss	post_trial_cpss	pre_trial_or	post_trial_or
33	Female	Static	5.82	5.09	5.82	4.86
34	Female	Static	6.24	5.67	6.44	5.75
35	Female	Static	5.79	5.05	5.74	5.01
36	Female	Static	5.50	4.65	5.30	4.58
37	Female	Static	5.66	4.87	5.59	5.01
38	Female	Static	4.49	3.24	4.09	3.03
39	Female	Static	6.04	5.39	6.41	5.24
40	Female	Static	6.90	6.59	6.58	6.37
41	Female	Static	6.04	5.39	5.93	5.26
42	Female	Static	5.28	4.34	5.57	4.30
43	Female	Static	6.09	5.46	5.79	5.49
44	Female	Static	4.10	2.70	4.26	2.50
45	Female	Static	6.19	5.61	6.38	5.67
46	Female	Static	6.92	6.61	7.03	6.70
47	Female	Static	5.37	4.47	5.05	4.28
48	Female	Static	6.13	5.52	5.78	5.35
49	Female	Static	4.86	3.76	4.86	3.94
50	Female	Static	5.74	4.97	5.83	4.95
51	Male	Control	5.60	6.53	5.62	6.66
52	Male	Control	4.37	4.86	4.59	4.88
53	Male	Control	6.86	6.24	6.96	6.27
54	Male	Control	4.97	5.52	4.95	5.75
55	Male	Control	7.40	5.21	7.06	5.45
56	Male	Control	6.66	6.87	7.18	7.08
57	Male	Control	6.17	5.10	6.17	5.13
58	Male	Control	7.24	5.82	7.46	5.61
59	Male	Control	8.00	5.61	8.34	5.49
60	Male	Control	6.30	6.20	6.42	6.25
61	Male	Control	7.23	4.41	6.89	4.46
62	Male	Control	5.73	3.62	5.25	3.69
63	Male	Control	7.90	6.32	7.88	6.23
64	Male	Control	4.93	5.98	5.09	6.14
65	Male	Control	5.09	4.42	4.97	4.59
66	Male	Control	4.03	5.19	4.02	5.02
67	Male	Control	5.37	6.05	5.21	6.12
68	Male	Control	7.85	4.87	7.70	4.80
69	Male	Control	6.65	5.76	7.11	5.69
70	Male	Control	5.56	5.93	6.07	5.71
71	Male	Control	6.83	7.04	7.29	7.25
72	Male	Control	6.21	4.75	6.44	4.58
73	Male	Control	5.75	6.19	6.07	6.17
74	Male	Control	6.06	6.64	5.76	6.67
75	Male	Control	6.75	5.23	7.25	5.35
76	Female	Control	5.14	4.14	5.28	4.04
77	Female	Control	6.45	5.96	6.11	5.95
78	Female	Control	5.03	3.99	5.24	3.90
79	Female	Control	6.33	5.80	6.03	5.68
80	Female	Control	4.66	3.48	5.04	3.51
81	Female	Control	6.20	5.61	6.30	5.57
82	Female	Control	5.68	4.90	5.24	5.01
83	Female	Control	5.02	3.98	5.12	3.94
84	Female	Control	5.74	4.98	5.33	5.20

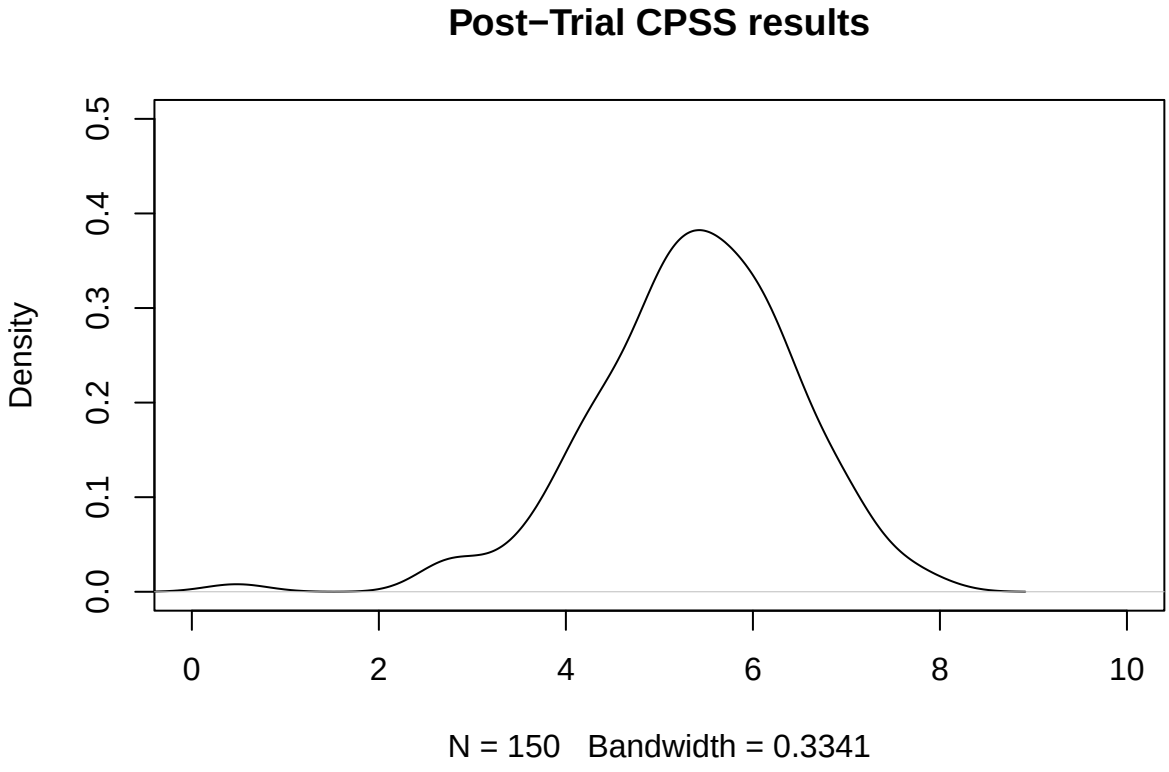
X	gender	test_group	pre_trial_cpss	post_trial_cpss	pre_trial_or	post_trial_or
85	Female	Control	7.27	7.10	7.79	7.16
86	Female	Control	6.73	6.35	6.67	6.35
87	Female	Control	5.57	4.73	5.25	4.94
88	Female	Control	6.49	6.01	6.48	5.82
89	Female	Control	5.50	4.65	6.05	4.45
90	Female	Control	7.10	6.86	7.02	6.88
91	Female	Control	6.16	5.56	6.47	5.65
92	Female	Control	5.17	4.19	4.86	4.05
93	Female	Control	5.89	5.19	6.16	5.02
94	Female	Control	6.47	5.99	6.46	5.87
95	Female	Control	2.49	0.48	2.97	0.25
96	Female	Control	4.18	2.82	4.04	2.71
97	Female	Control	5.30	4.36	5.53	4.16
98	Female	Control	7.64	7.61	7.50	7.54
99	Female	Control	5.93	5.24	6.30	5.11
100	Female	Control	6.06	6.17	7.00	6.30
101	Male	Animated	7.34	4.99	6.85	4.83
102	Male	Animated	6.32	6.18	6.48	6.35
103	Male	Animated	7.62	5.49	7.82	5.27
104	Male	Animated	5.11	6.36	4.75	6.14
105	Male	Animated	7.29	4.64	7.71	4.86
106	Male	Animated	6.42	4.28	6.79	4.34
107	Male	Animated	6.50	4.29	7.03	4.08
108	Male	Animated	5.29	3.56	5.56	3.46
109	Male	Animated	6.42	6.05	6.84	6.05
110	Male	Animated	5.52	6.38	5.13	6.19
111	Male	Animated	5.59	5.11	5.83	5.35
112	Male	Animated	7.21	3.79	6.90	3.72
113	Male	Animated	5.61	6.79	5.73	6.83
114	Male	Animated	4.63	6.61	4.23	6.73
115	Male	Animated	5.87	6.56	5.60	6.57
116	Male	Animated	4.28	4.87	4.44	4.93
117	Male	Animated	6.10	4.88	5.94	4.98
118	Male	Animated	4.65	4.56	4.99	4.45
119	Male	Animated	8.63	6.13	8.74	5.99
120	Mal	Animated	7.18	6.45	7.22	6.69
121	Male	Animated	7.21	5.49	7.78	5.66
122	Male	Animated	7.83	3.90	8.37	3.65
123	Male	Animated	6.84	6.96	6.33	6.79
124	Male	Animated	5.77	5.82	6.02	5.90
125	Male	Animated	7.41	5.08	7.25	5.17
126	Female	Animated	5.40	4.50	4.99	4.49
127	Female	Animated	5.83	5.10	6.40	5.11
128	Female	Animated	5.87	5.15	5.87	5.22
129	Female	Animated	4.22	2.87	3.98	2.88
130	Female	Animated	5.74	4.98	5.45	5.21
131	Female	Animated	6.24	5.67	5.67	5.82
132	Female	Animated	6.63	6.21	6.33	6.43
133	Female	Animated	5.38	4.48	5.71	4.69
134	Female	Animated	5.96	5.29	6.23	5.34
135	Female	Animated	6.59	6.16	6.35	6.23
136	Female	Animated	6.34	5.80	6.83	5.72

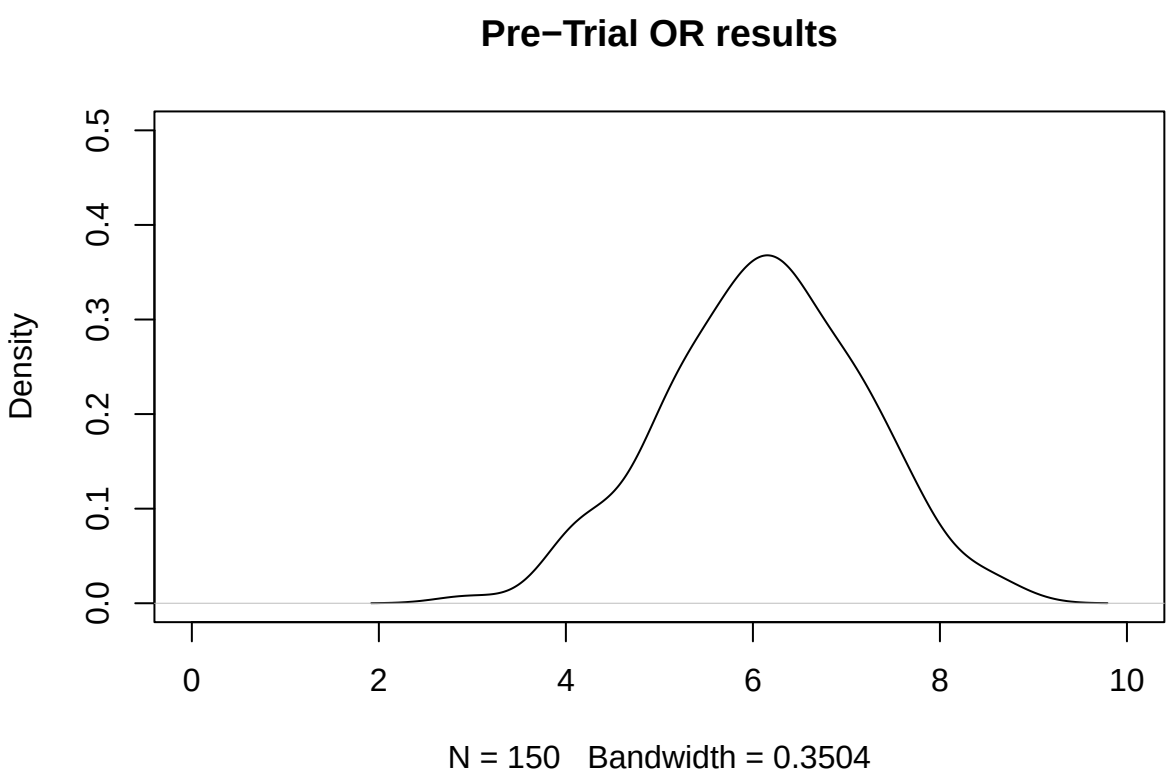
X	gender	test_group	pre_trial_cpss	post_trial_cpss	pre_trial_or	post_trial_or
137	Female	Animated	6.57	6.12	6.38	5.89
138	Female	Animated	7.33	7.17	6.81	7.01
139	Female	Animated	4.05	2.64	3.93	2.86
140	Female	Animated	7.22	7.02	7.20	6.80
141	Female	Animated	5.13	4.13	5.51	3.88
142	Female	Animated	5.22	4.26	5.41	4.24
143	Female	Animated	5.10	4.09	5.45	4.08
144	Female	Animated	6.40	5.90	6.62	5.69
145	Female	Animatd	6.51	6.05	7.05	5.89
146	Female	Animated	6.20	5.62	6.66	5.61
147	Female	Animated	5.98	5.31	6.26	5.23
148	Female	Animated	6.31	5.77	6.37	5.73
149	Female	Animated	5.61	4.80	5.95	4.91
150	Female	Animated	6.25	5.69	6.06	5.83

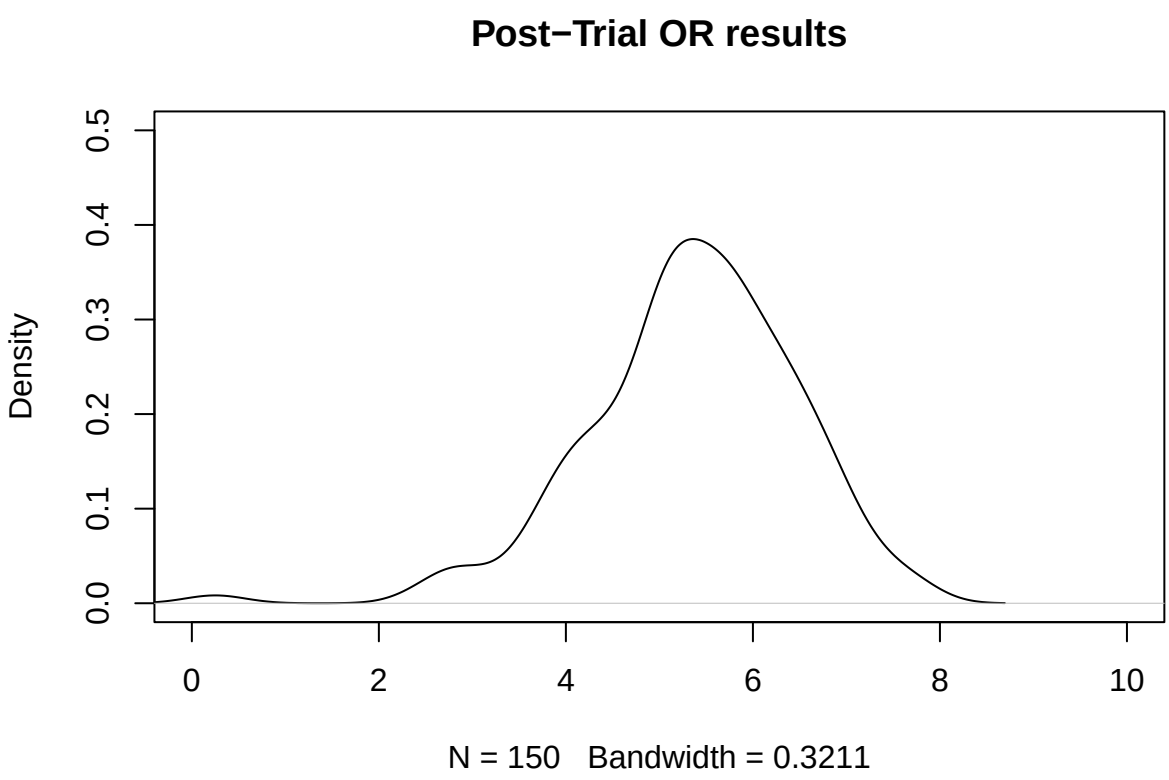
```
##        X              gender           test_group        pre_trial_cpss
##  Min.   :  1.00   Length:150         Length:150         Min.   :2.490
##  1st Qu.: 38.25   Class :character   Class :character   1st Qu.:5.500
##  Median : 75.50   Mode  :character   Mode  :character   Median :6.077
##  Mean   : 75.50                                         Mean   :6.063
##  3rd Qu.:112.75                                         3rd Qu.:6.657
##  Max.   :150.00                                         Max.   :8.630
##  post_trial_cpss  pre_trial_or   post_trial_or
##  Min.   :0.480   Min.   :2.970   Min.   :0.250
##  1st Qu.:4.763   1st Qu.:5.420   1st Qu.:4.838
##  Median :5.460   Median :6.095   Median :5.355
##  Mean   :5.364   Mean   :6.115   Mean   :5.357
##  3rd Qu.:6.117   3rd Qu.:6.848   3rd Qu.:6.140
##  Max.   :7.910   Max.   :8.740   Max.   :7.730
```

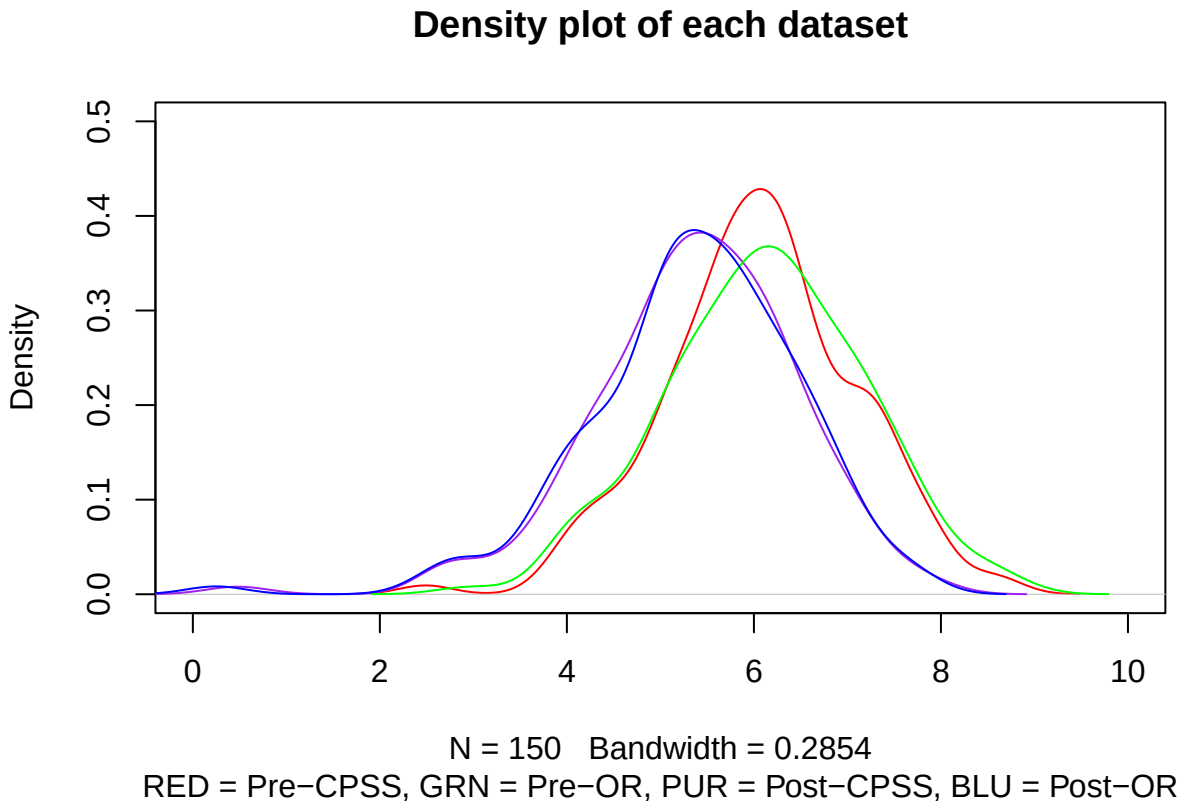
Graphs


Pre−Trial CPSS results
Density
0.0
0.1
0.2
0.3
0.4
0.5
0
2
4
6
8
10
N = 150 Bandwidth = 0.2854

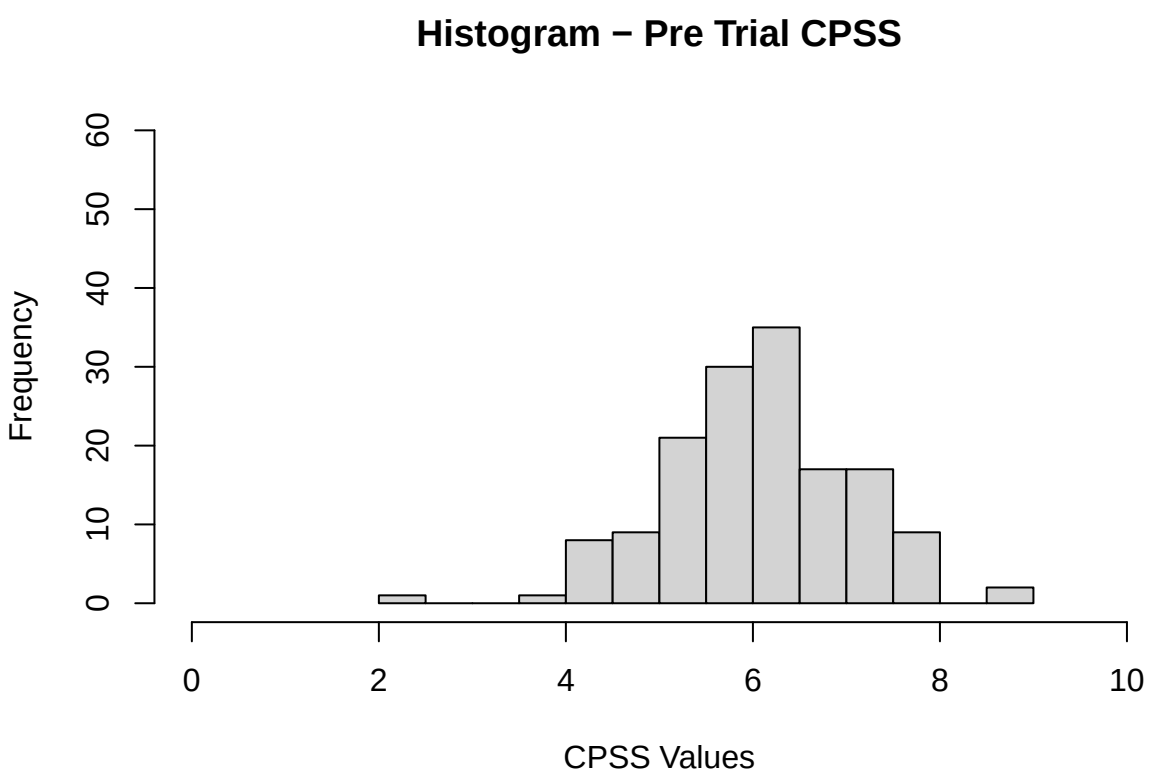

Post−Trial CPSS results
Density
0.0
0.1
0.2
0.3
0.4
0.5
0
2
4
6
8
10
N = 150 Bandwidth = 0.3341

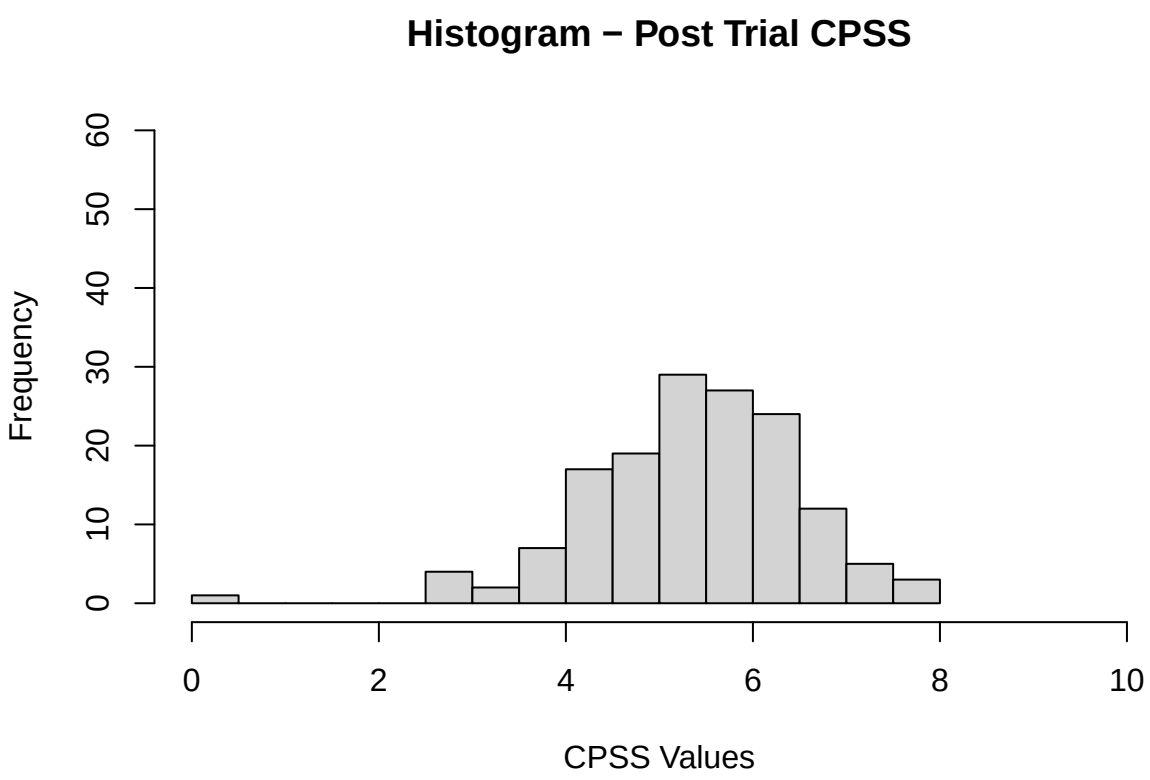

Pre−Trial OR results
Density
0.0
0.1
0.2
0.3
0.4
0.5
0
2
4
6
8
10
N = 150 Bandwidth = 0.3504

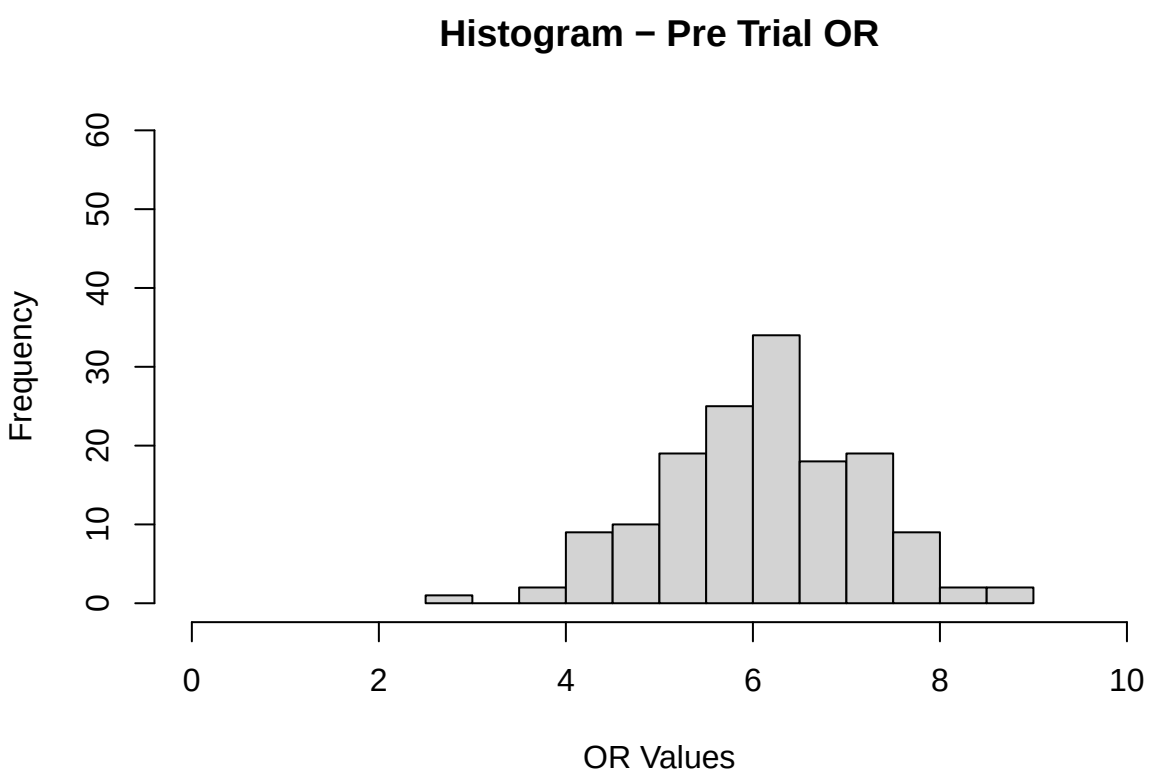

Post−Trial OR results
Density
0.0
0.1
0.2
0.3
0.4
0.5
0
2
4
6
8
10
N = 150 Bandwidth = 0.3211

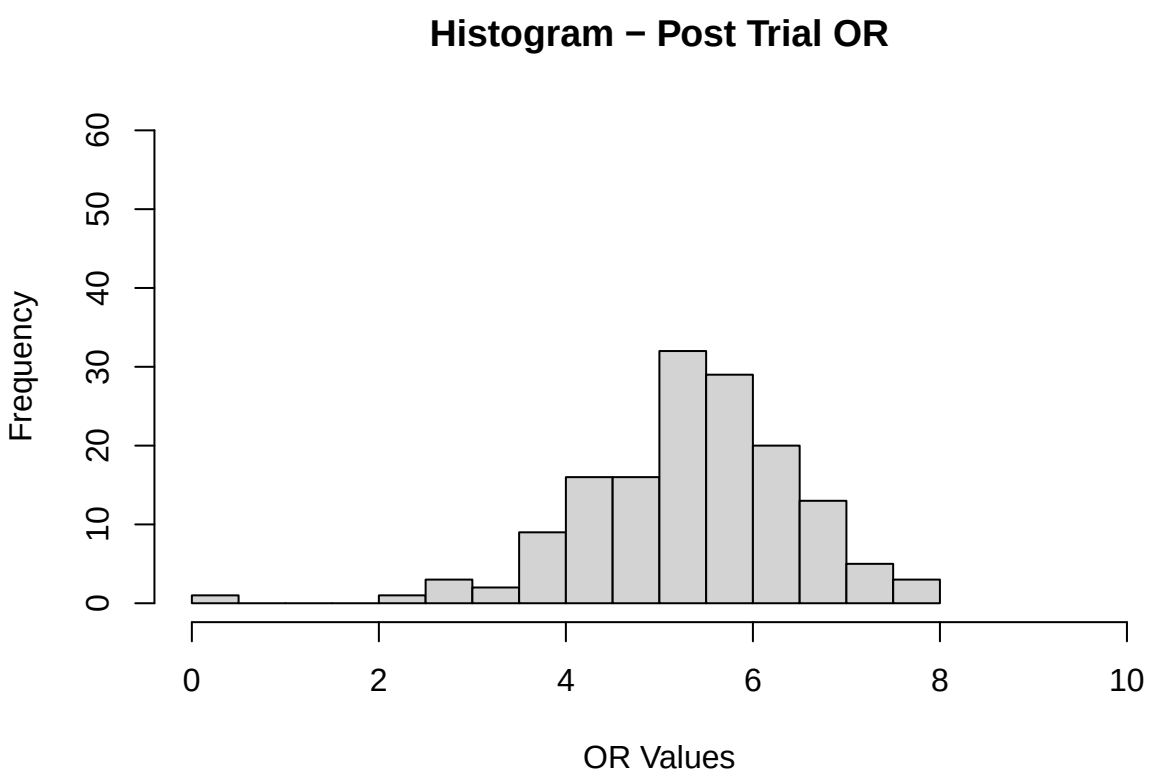


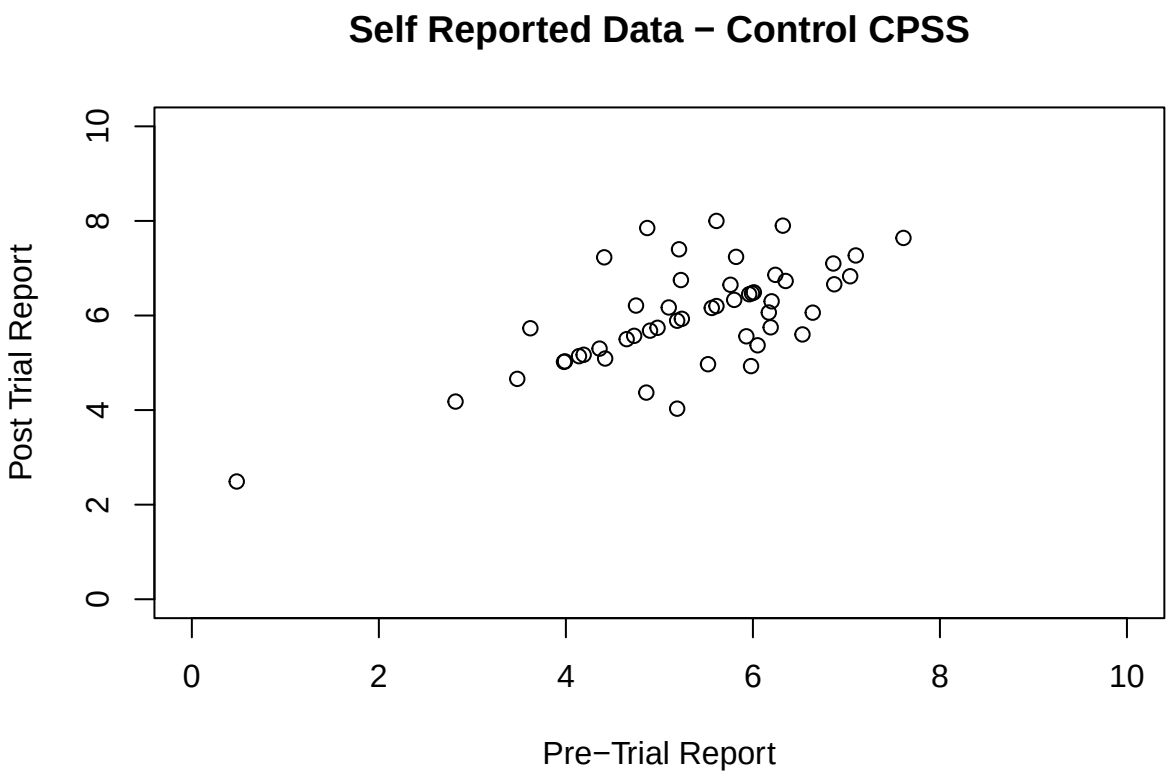
In this graph 'Bandwidth' only applies to the red line.

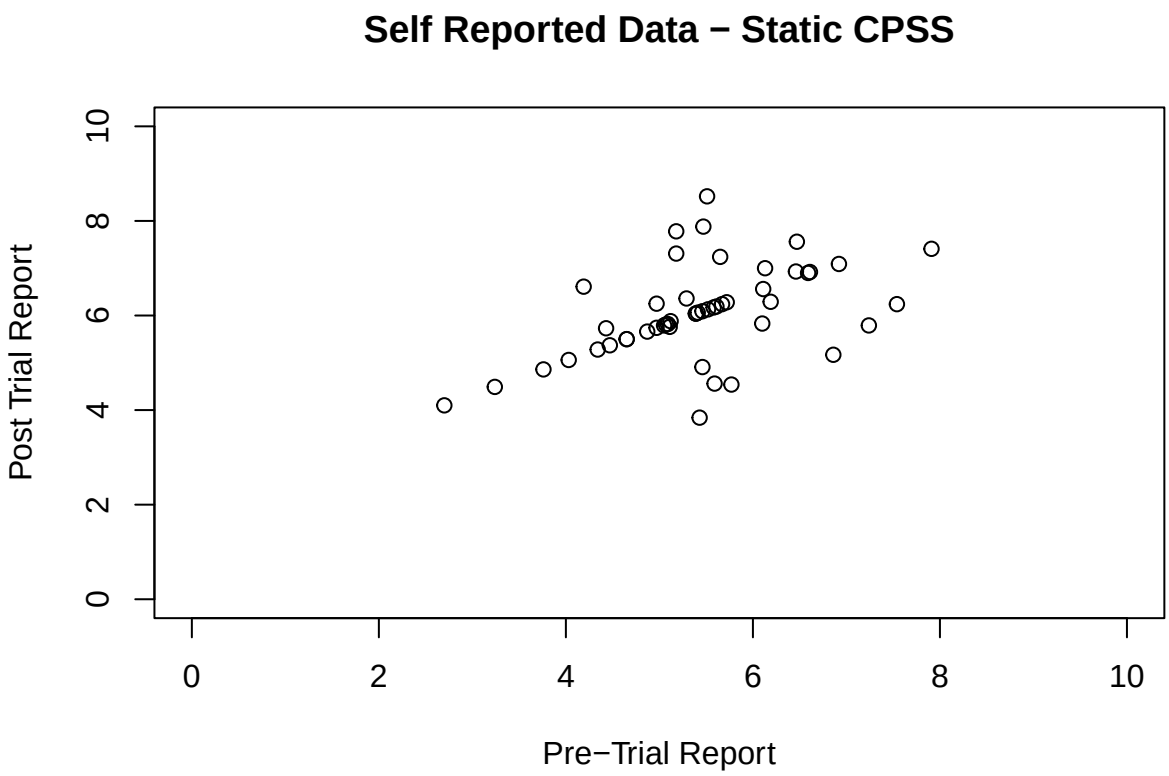

Histogram − Pre Trial CPSS
Frequency
0
10
20
30
40
50
60
0
2
4
6
8
10
CPSS Values

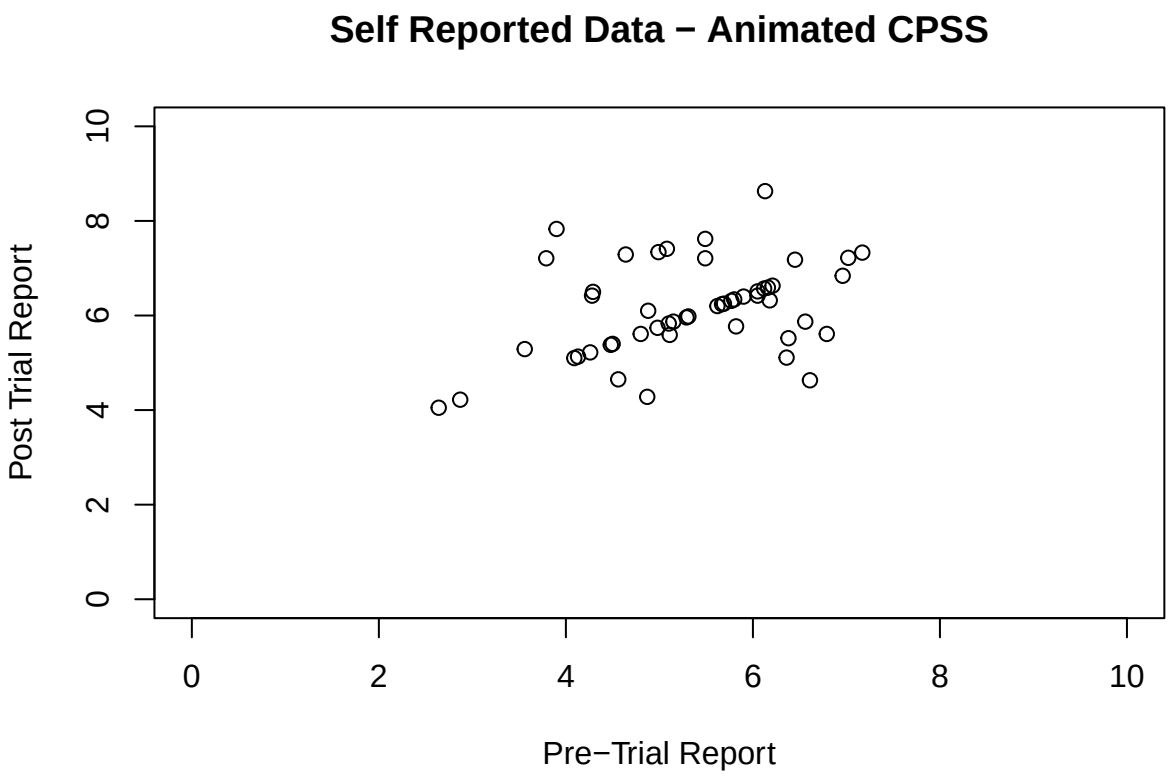

Histogram − Post Trial CPSS
Frequency
0
10
20
30
40
50
60
0
2
4
6
8
10
CPSS Values

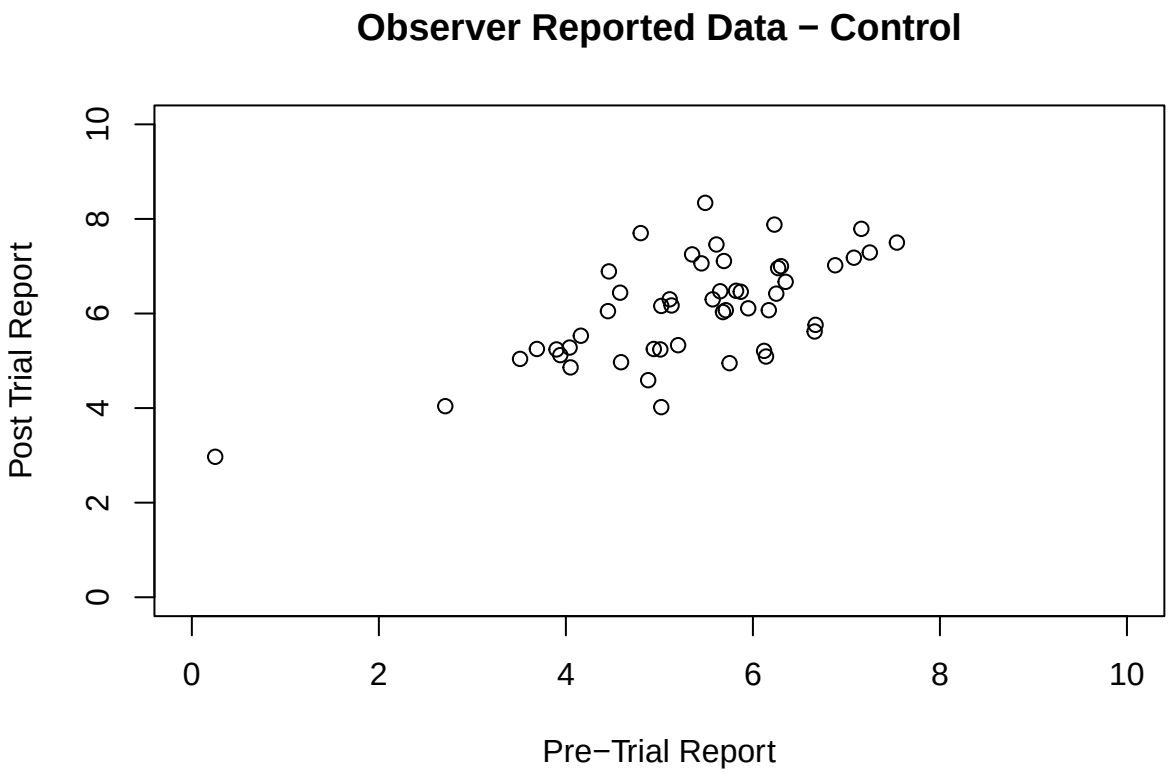

Histogram − Pre Trial OR
Frequency
0
10
20
30
40
50
60
0
2
4
6
8
10
OR Values

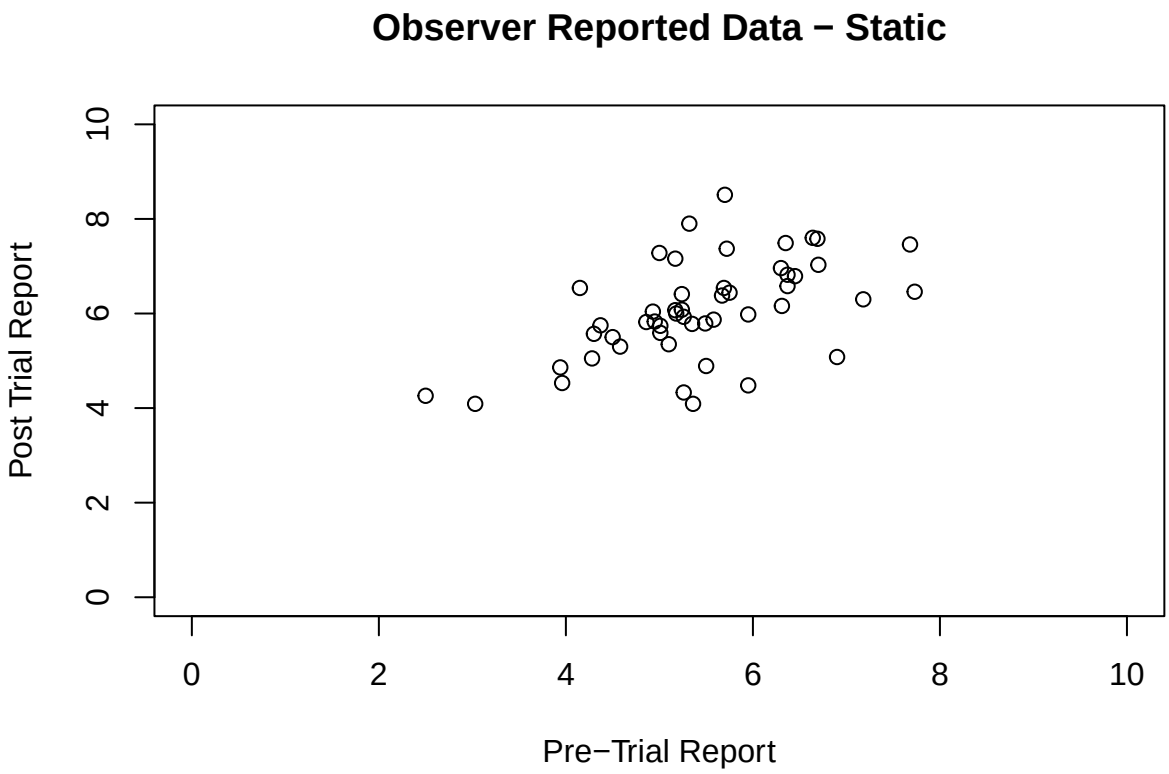

Histogram − Post Trial OR
Frequency
0
10
20
30
40
50
60
0
2
4
6
8
10
OR Values

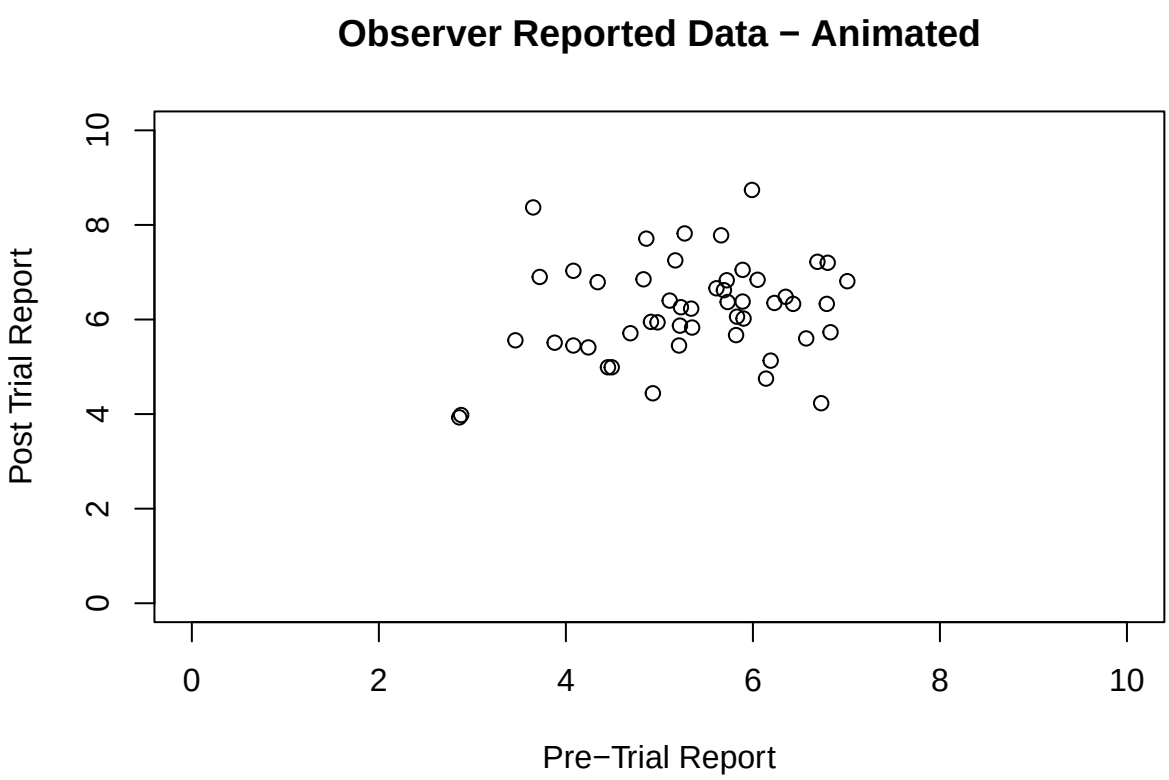

Self Reported Data − Control CPSS
Post Trial Report
Pre−Trial Report
0
2
4
6
8
10


Self Reported Data − Static CPSS
Post Trial Report
Pre−Trial Report
0
2
4
6
8
10


Self Reported Data − Animated CPSS
Post Trial Report
Pre−Trial Report
0
2
4
6
8
10


Observer Reported Data − Control
Post Trial Report
Pre−Trial Report
0
2
4
6
8
10


Observer Reported Data − Static
Post Trial Report
Pre−Trial Report
0
2
4
6
8
10



Descriptive statistics

After calculating the individual means for each of the groups of each data set (static pre_trial_or, control pre-trial_or) we determined that across all the data sets there is a consistent decrease in recorded PTSD levels of the patients post the experiment with a similar decrease in PTSD levels across each group (Control, Static, Animated).

CPSS Means (Static Pre/Post , Control Pre/Post, Animated Pre/Post)

Static Pre trial CPSS Mean = 6.0616

Static Post trial CPSS Mean = 5.4424

Control Pre trial CPSS Mean = 5.9942685

Control Post trial CPSS Mean = 5.3302

Animated Pre trial CPSS Mean = 6.1344

Animated Post trial CPSS Mean = 5.32

OR Means (Static Pre/Post , Control Pre/Post, Animated Pre/Post)

Static Pre trial OR Mean = 6.0682

Static Post trial OR Mean = 5.4338

Control Pre trial OR Mean = 6.0798

Control Post trial OR Mean = 5.322

Animated Pre trial OR Mean = 6.196

Animated Post trial OR Mean = 5.3154

CPSS Standard Deviation (Static Pre/Post , Control Pre/Post, Animated Pre/Post)

Static Pre trial CPSS Standard Deviation = 0.9819875

Static Post trial CPSS Standard Deviation = 1.0074671

Control Pre trial CPSS Standard Deviation = 1.0905146

Control Post trial CPSS Standard Deviation = 1.2355607

Animated Pre trial CPSS Standard Deviation = 0.9796692

Animated Post trial CPSS Standard Deviation = 1.0516886

OR Standard Deviation (Static Pre/Post , Control Pre/Post, Animated Pre/Post)

Static Pre trial OR Standard Deviation = 1.036847

Static Post trial OR Standard Deviation = 1.0484116

Control Pre trial OR Standard Deviation = 1.11438

Control Post trial OR Standard Deviation = 1.2794977

Animated Pre trial OR Standard Deviation = 1.0457162

Animated Post trial OR Standard Deviation = 1.0451759

CPSS Mean (Pre/Post-trial)

Pre trial CPSS Mean = 6.0634228

Post trial CPSS Mean = 5.3642

OR Mean (Pre/Post-trial)

Pre trial OR Mean = 6.1146667

Post trial OR Mean = 5.3570667

CPSS Standard Deviation (Pre/Post-trial)

Pre trial CPSS Standard Deviation = 1.0134715

Post trial CPSS Standard Deviation = 1.0966577

OR Standard Deviation (Pre/Post-trial)

Pre trial OR Standard Deviation = 1.0606125

Post trial OR Standard Deviation = 1.1234167

Inferential statistics

If the participants in this study are an accurate reflection of 18-25 year olds in treatment for PTSD, we can infer that traditional CBT therapy as well the Jungian Sandplay VR app can reduce self-reported levels of PTSD for people with similar attributes in a general population.

Statistical tests

A Shapiro-Wilk normality test was conducted on each of the data sets

Pre-trial CPSS SW Result = 0.5579

Post-trial CPSS SW Result = 0.0029

Pre-trial OR SW Result = 0.9345

Post-trial OR SW Result = 9×10^{-4}

From the output obtained we can assume normality for both pre-trial results as the p-values are greater than 0.05, while both post-trial results fail this normality test as they are less than 0.05.

From the 'Density plot of each dataset' graph we can see a strong correlation between the post-trial results, whereas the pre-trial results are more varied between reports.

Confidence Intervals

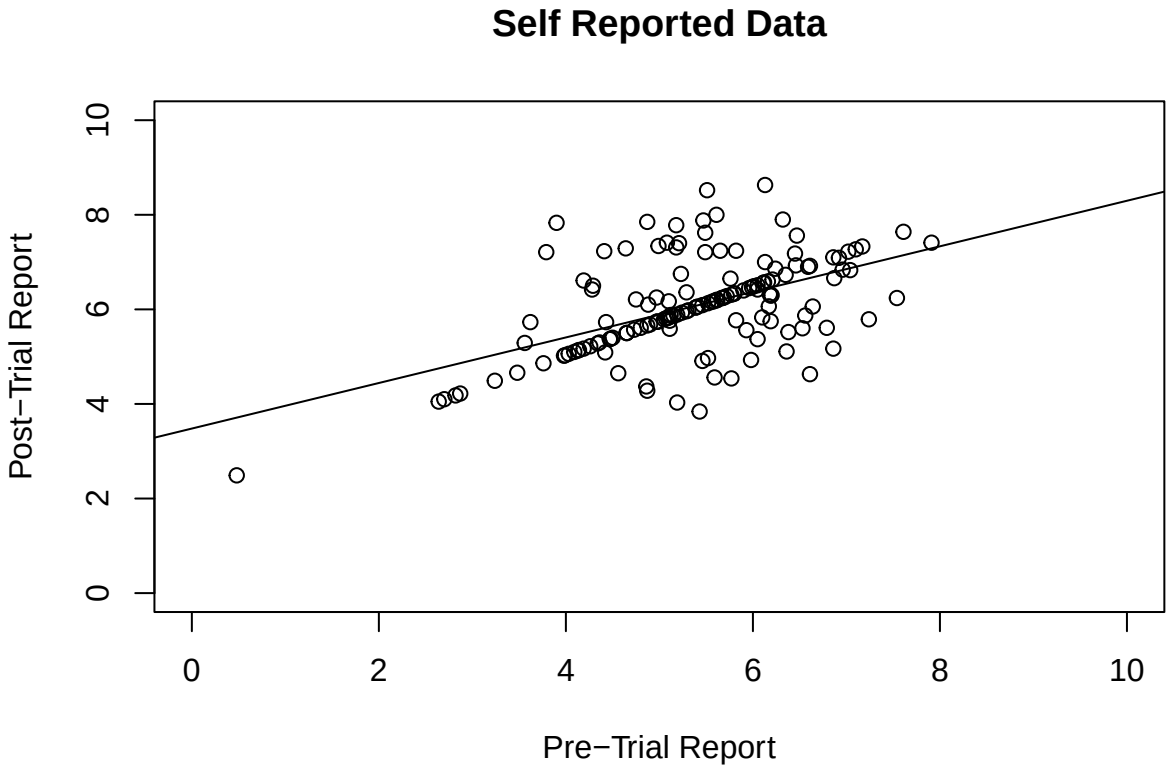
95% CI for pre-trial CPSS = (5.8999085 , 6.2269371)

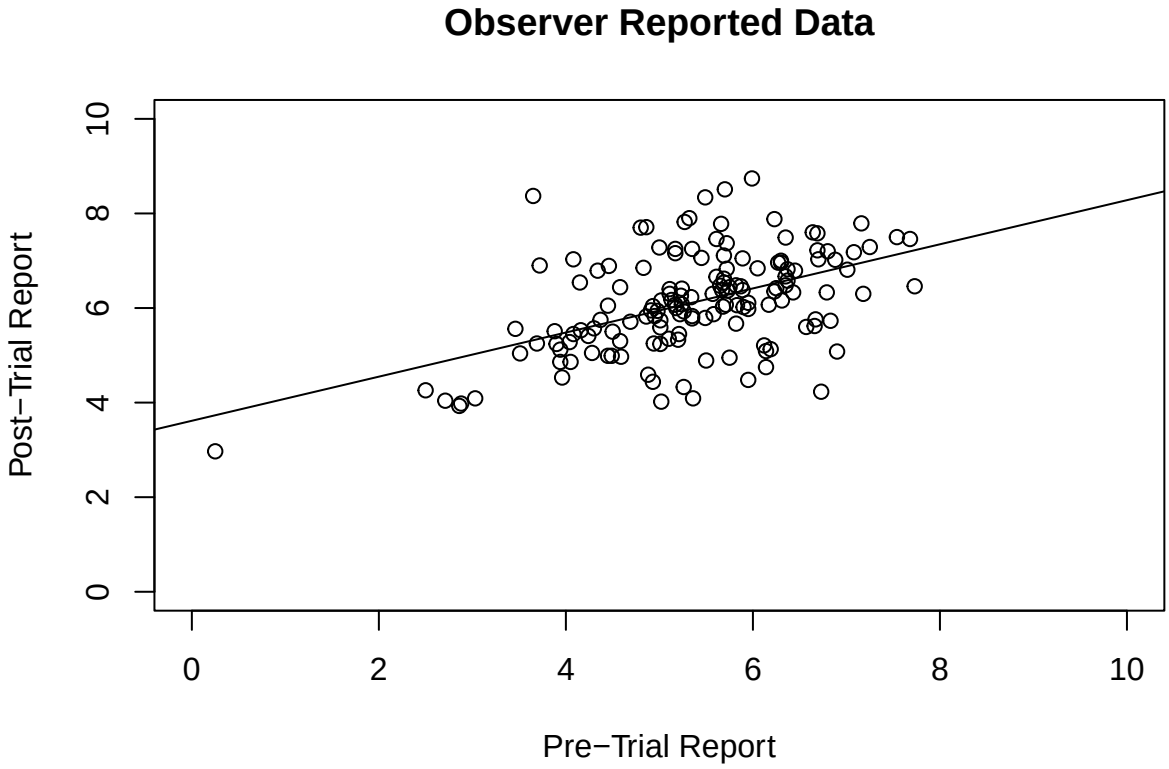
95% CI for post-trial CPSS = (5.1872644 , 5.5411356)

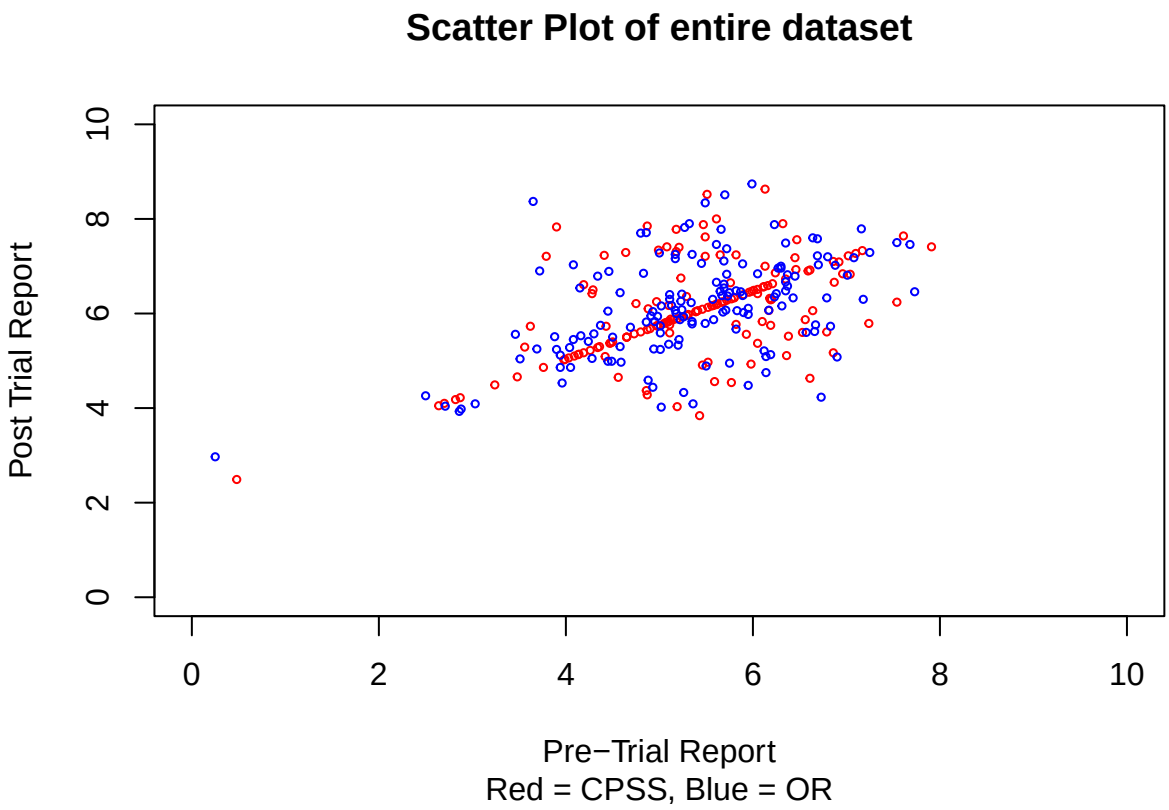
95% CI for pre-trial OR = (5.9435466 , 6.2857867)

95% CI for post-trial OR = (5.1758137 , 5.5383196)

Magnitude and direction of results







By plotting the pre- and post-trial CPSS against one another and the pre- and post-trial OR against one another we can see the CPSS values have a more positive association with one another. The OR values lack this association and appear to be more dispersed and random than the CPSS values.

Discussion

Outline findings and relation to the hypothesis

Our null hypothesis for this experiment is that post-trial values for OR results should correlate with post-trial CPSS results. This would infer that therapists observe as much of a benefit from the trial as the participant.

Our alternate hypothesis is that OR values will not correlate with the CPSS values.

The results of our Shapiro-Wilkes test proved that the Pre-Trial results were normalised, while the post-trial results were not. Our dot plots also visually confirmed that the OR results did not align with the CPSS results. This goes against our null hypothesis that post-trial OR values should correlate with post-trial CPSS results.

Because of these results, we decided that our alternate hypothesis was correct.

Limitations

The initial decision to impute the missing values instead of ignoring them could have led to a different set of results. However, the outcome of the experiment would likely stay the same.

References

Help with getting mean of data while data is missing from column *Stack Overflow

Removing Na's by Column *GeeksforGeeks

How to Handle Missing Data in a Dataset *freeCodeCamp

How to Impute Missing Values in R? *GeeksforGeeks

First 10 entries in a bar plot *StackOverflow

Calculate confidence interval *Cyclismo

How to Create a Scatterplot with a Regression Line in R? *GeeksforGeeks