FACES Statistical Consultation - Final Report
Client: Dr. Jamie Pearson
Consultants: James Carr, Joseph Oliveira, Shantel Ward

Introduction

Fostering Advocacy, Communication, Empowerment, and Support (FACES) is a 4 to 6 week parent training intervention program aimed at educating African American parents of children with Autism Spectrum Disorder (ASD), such that these parents feel confident in implementing ASD interventions, accessing relevant information, and discussing options with medical professionals. The central goal is to foster empowerment and advocacy by improving parents' knowledge of ASD and improving their use of communication and behavior management strategies. The program is at a preliminary stage of inquiry, and aims to investigate a spanning array of research questions regarding the central pillars of empowerment as it relates to the outcomes of the FACES program. Therefore, data on FACES participants' ability to advocate for their children, understand implementation of social communication strategies, manage characteristic ASD behavior, communicate with professionals, and pertinent feelings of confidence were collected for analysis.

Purpose

The purpose of this statistical consultation is to examine if the FACES intervention had an effect on improving advocacy, communication, empowerment, and support for African American families with autistic children by comparing the survey results of the intervention group and the control group. The secondary aim of this consultation is to provide guidance on the participant size needed to answer the research questions. This will give Dr. Pearson a target size of participants to acquire in the next phases of the FACES program. Our goal is for this analysis to assist the FACES program with training as many families as possible and helping to improve the quality of life for those families.

Research Questions

The research questions were formally implemented into a study led by Dr. Pearson and formulated in her dissertation. As consultants our objectives are to analyze participant survey data collected from the study participants, build understanding of FACES's impact on empowering African American families of children with ASD, and subsequently report our statistical analysis to be used for Dr. Pearson's Institute of Education Sciences (IES) grant proposal submission as well as support for future FACES publications. Specifically, our consultations goals are to examine the following questions[1]:

1. Does the FACES program increase empowerment in African American parents of children with ASD?
2. Does the FACES program increase parents' knowledge of ASD?
3. Two part question related to confidence:

[1] Pearson, J. N. (2017) *Fostering Advocacy, Communication, Empowerment and Support (FACES) for African American Families of Children with Autism: A Pilot Study* [Unpublished Doctoral Dissertation], University of Illinois

   a. Does the FACES program strengthen parents' understanding of and confidence in their ability to implement social communication strategies?
   b. Does the FACES program strengthen parents' understanding of and confidence in managing behavior?
4. Does the FACES program increase parent-professional communication as reported by parents?
5. In what ways and to what extent does the FACES program strengthen parents' reported perceptions of their ability to advocate effectively for services for their children with ASD?

The immediate challenge presented by investigating these research questions is the inherent subjectivity in measuring dispositions of confidence, sense of being supported by professionals, and ability to advocate. As a result, substantiated statistical methodologies will be leveraged as a benchmark of performance and measurement in these areas. Thus a methodology that is coherent with standard practice and previous research in Autism Support programs is required to inform our statistical methodology. Difficult as this may be, an effort to align our methodology to previous work ensures that results can be reliably communicated by using common language and numerical summaries.

While it is known that African American children with ASD are less likely to be identified, there are also no known parent-advocacy programs designed specifically to meet the needs of African American families. Additionally, according to Dr. Pearson, little is known about the experiences of African American parents affected by having children with ASD. In an effort to resolve the lack of public support and educational outreach to African American parents, FACES was implemented as "parent-training interventions that promote advocacy and empowerment" (Dr. Pearson), thereby stimulating community efforts to recognize ASD earlier and to provide early childhood social support.

Background

This is the second iteration of this study by Dr. Pearson. In her 2017 dissertation , she discusses her results from the pilot study, which were successful at showing the strengthening of empowerment for the families in her study. We will be drawing upon this previous work in performing statistical analysis, and, therefore,we expect a similar outcome. In particular, the results from the FES (Family Empowerment Scale) survey should align with historic results since it most directly measures empowerment.

Data

The data set we analyzed is a collection of survey responses administered by Dr. Pearson and co-investigators in conducting the FACES intervention and training program. The data was sourced from digital pre and post surveys that were completed by the participants. The survey captured responses pertaining to six different categories of measurements including the novel implementation of the FACES program. Participant demographic information was collected during the early stages of the study.

Participant selection criteria specified both parents or guardian and child identified as African-American, the child was between the age of 3 and 21 with a primary diagnosis of autism spectrum disorder (ASD), and parents or guardians were both willing to participate in a four week program and at least 20 years old at the start of the study. With the use of flyers, both digital and hard copy, participants were recruited into the program. Additional families were recruited via word of mouth formally implemented as snowball sampling. Once a recruitment pool was created, there was a phone screening process to confirm inclusion criteria was met - specifically that ASD diagnoses were confirmed via the Social Communication Questionnaire (SCQ).

The recruitment process resulted in sixteen families that were then randomly selected for experimental and control (waitlist) groups. The experimental group contained ten families while the control group contained six families. Both groups took a survey before and after the intervention. The treatment we tested is the FACES intervention program, which is a multi-part training course that takes place over four weeks. The experimental group received these training courses and the control group did not. The survey responses tested are derived from six survey categories of multiple questions with all answers ranging from 1-5:

1. Fostering Advocacy, Communication, Empowerment and Support Scale (FACES) - seven questions
2. Autism Knowledge Scale (AKS) - twenty-one questions
3. Family Empowerment Scale (FES) - thirty-four questions
4. School Communication Scale (SCS) - seven questions
5. Family Professional Partnership Scale (FPPS) - eighteen questions
6. Special Education Advocacy Scale (SEAS) - ten questions

The response tested is the paired/unpaired differences between groups for each of the six response survey categories. Several combinations of differences in pre/post surveys and experimental/control groups were tested. These combinations will be described in more details in the Statistical Methods section.

Data Transformations

Our response variables are discrete Likert scale responses, ranging from 1 to 5, for a series of questions from 6 survey categories. The responses for each survey category were summed up to create total scores for data analysis.

Each observation was represented by a distinct combination of participant (*p*), time (*t*), and group (*g*). For each observation, we aggregated the responses for each of the six question categories by summing the individual responses for questions that make up that survey category. The following six features were created:

let p be the participant number
let t be the Time factor with 2 *levels* : *Pre and Post*
let g be the Group factor with 2 *levels* : *Experimental and Control*
let q be the number of questions on the survey for a given survey category

$$FACES\ Total_{p,t,g} = \sum_{q=1}^{7} FACES_{p,t,g,q}$$

$$AKS\ Total_{p,t,g} = \sum_{q=1}^{21} AKS_{p,t,g,q}$$

$$FES\ Total_{p,t,g} = \sum_{q=1}^{34} FES_{p,t,g,q}$$

$$SCS\ Total_{p,t,g} = \sum_{q=1}^{7} SCS_{p,t,g,q}$$

$$FPPS\ Total_{p,t,g} = \sum_{q=1}^{18} FPPS_{p,t,g,q}$$

$$SEA\ Total_{p,t,g} = \sum_{q=1}^{10} SEA_{p,t,g,q}$$

Missing Data

Upon review of the raw data, missing response values were identified. Missing data was summarized using software to determine the best course of action. For each distinct grouping of *time*, *group*, and *survey category*, a *participant's* missing response data was tallied. If 2 or fewer missing values were observed within the aforementioned distinct group, then the missing value(s) was(were) imputed with the mean of the non-missing responses from the corresponding distinct grouping. If more than 2 missing values were observed within the distinct grouping, then *all* the responses from that participant within the distinct grouping were dropped.

Participant #	Group	Time	Survey	Missing Observations	Observations Removed
6	Experimental	Pre	FPPS	18	18
6	Experimental	Post	FPPS	18	18
6	Experimental	Pre	SCS	7	7
6	Experimental	Post	SCS	7	7
6	Experimental	Pre	SEAS	9	10
6	Experimental	Post	SEAS	9	10
8	Experimental	Post	AKS	1	0
8	Experimental	Post	FES	1	0
8	Experimental	Pre	FPPS	18	18
8	Experimental	Post	FPPS	18	18
8	Experimental	Pre	SCS	7	7
8	Experimental	Pre	SEAS	5	10
14	Control	Post	FACES	1	0
18	Control	Pre	FES	2	0
18	Control	Pre	FPPS	18	18
18	Control	Pre	SCS	7	7
19	Control	Post	FACES	1	0
19	Control	Post	FPPS	18	18
19	Control	Post	SCS	7	7
			Total	172	173

Table 1: Missing Data

Of the 3,104 observations there are 172 missing data points, or about 6% of our observations were missing. 6 observations were imputed (<0.2%), and the remaining 166 observations resulted in 173 observations being removed.

Data Limitations

One of the main challenges associated with this study, and one that the next iteration of the FACES study aims to resolve, is the constraint of sample size. Our methodology and assumptions factor in this constraint. Additionally, power test simulations were conducted to purview the effects of various sample sizes.

Methodology

- Software: R version 4.0.0 for data clean up, manipulation, statistical analysis, and data visualizations.

- Power Calculations: Non-parametric power simulation approach was taken to view the effective power in relation to various sample sizes of the experimental and control groups. Using R, we conducted a bootstrap sampling technique on the same samples used in our analysis. Our scope of interest was the effective difference between the pre intervention scores versus the post intervention scores. Post intervention simulation samples were compared to pre intervention samples via the Wilcoxon Signed-Rank test, and the proportion of the ten thousand simulations run that contained a p-value less than 0.05 determined our predicted power.

- Wilcoxon Signed-Rank Test: We used the Wilcoxon Signed-Rank test instead of the Paired t-test because the Paired t-test requires the data to be normally distributed and we cannot assume normality. We conducted a myriad of Wilcoxon Signed-Rank tests using the six Total features that were derived from the survey responses.

    - The total of Pre intervention responses for the Control group was compared to the total of Pre intervention responses for the Experimental group. This unpaired test was conducted for each of the six Total features described in the Data Transformation section.

      H_0 : *There is no difference between the pre intervention responses for the experimental group and the pre intervention responses for the control group*
      H_A : *Pre intervention responses for experimental group* $\neq$ *pre intervention responses for control group*

    - The total of Post intervention responses for the Control group was compared to the total of Post intervention responses for the Experimental group. This unpaired test was conducted for each of the six Total features described in the Data Transformation section.

      H_0 : *There is no difference between the post intervention responses for the experimental group and the post intervention responses for the control group*
      H_A : *Post intervention responses for experimental group* $\neq$ *post intervention responses for control group*

- For the Experimental group, the total of Pre intervention responses were compared to the total of Post intervention responses. This paired test was conducted for each of the six Total features described in the Data Transformation section.

  H_0 : *For the experimental group, there is no difference between the pre intervention responses and the post intervention responses*
  H_A : *Pre intervention responses for experimental group* $<$ *post intervention responses for experimental group*

- For the Control group, the total of Pre intervention responses were compared to the total of Post intervention responses. This paired test was conducted for each of the six Total features described in the Data Transformation section.

  H_0 : *For the control group, there is no difference between the pre intervention responses and the post intervention responses*
  H_A : *Pre intervention responses for controlgroup* $<$ *post intervention responses for control group*

- Outlier Detection - The differences between post intervention scores and pre intervention scores for all participants were standardized and graphed in a qq plot. Potential outliers are identified by points that are far away from the line. The Wilcoxon Signed-Rank Test was run with and without participants identified as potential outliers and the results were compared to determine if potential outliers should be removed from the analysis.

Results

With the limited number of participants included in this study, we would caution against using the findings presented here for statistical inference.

Experimental Group vs Control Group: Pre Intervention Wilcoxon Test

The results of the Wilcoxon Signed-Rank Test to determine if the pre intervention responses for the experimental group were any different than the pre intervention responses for the control group are displayed in Table 2. The results here emphasize that the experimental and control groups had similar responses for all surveys except the FPPS survey. The FPPS survey test yielded the lowest p-value (p=0.05). It was encouraging to see that the two groups had similar pre intervention responses for the majority of the surveys. The test results gave us more confidence to use the control group as a comparison for the post intervention responses.

Table 2: Pre Group Results

Time	Survey	Group A	Group B	W	p
Pre	AKS	Experimental	Control	20.0	0.290
Pre	FACES	Experimental	Control	30.0	1.000
Pre	FES	Experimental	Control	33.0	0.786
Pre	FPPS	Experimental	Control	28.0	0.050
Pre	SCS	Experimental	Control	21.5	0.395
Pre	SEAS	Experimental	Control	22.5	0.897

W = Wilcoxon Test Statistic, p = p-value

Experimental Group vs Control Group: Post Intervention Wilcoxon Test

The results of the Wilcoxon Signed-Rank Test to determine if the post intervention responses for the experimental group were any different than the post intervention responses for the control group are displayed in Table 3. Based on the results here we cannot detect a significant difference in post intervention responses for the experimental group and the control group for any of the surveys. A larger sample size is recommended and could help with detecting a difference in post responses for the experimental group and the control group. Power and sample size analysis results are discussed later in the results section.

Table 3: Post Group Results

Time	Survey	Group A	Group B	W	p
Post	AKS	Experimental	Control	36.0	0.547
Post	FACES	Experimental	Control	42.0	0.211
Post	FES	Experimental	Control	43.0	0.174
Post	FPPS	Experimental	Control	27.5	0.060
Post	SCS	Experimental	Control	17.5	0.864
Post	SEAS	Experimental	Control	32.5	0.299

W = Wilcoxon Test Statistic, p = p-value

Pre Intervention vs Post Intervention: Experimental Group Wilcoxon Test

The results of the Wilcoxon Signed-Rank Test to determine if the post intervention responses were greater than the pre intervention responses for the experimental group are displayed in Table 4. The test showed evidence that the post intervention responses are significantly greater than the pre intervention responses in the FES (p=0.002), SCS (p=0.011), and SEAS (p=0.020) surveys. However, we cannot conclude that the post intervention responses are significantly greater than the pre intervention responses in the FACES (p=0.179), FPPS (p=0.175), and AKS (p=0.092) surveys. We did not expect the FPPS and AKS surveys to show significant differences between the pre intervention and post intervention responses because these two measurements are time sensitive and will likely take more than 4 weeks to have a significant impact on the participants. We did expect to have significant increases in the post intervention responses for the FACES survey. We hope that a large sample size will remedy this in future studies.

Table 4: Experimental Group Results

Group	Survey	Pre	Post	W	p
Experimental	AKS	Pre	Post	14	0.092
Experimental	FACES	Pre	Post	18	0.179
Experimental	FES	Pre	Post	1	0.002
Experimental	FPPS	Pre	Post	8	0.175
Experimental	SCS	Pre	Post	0	0.011
Experimental	SEAS	Pre	Post	3	0.020

W = Wilcoxon Test Statistic, p = p-value

Pre Intervention vs Post Intervention: Control Group Wilcoxon Test

The results of the Wilcoxon Signed-Rank Test to determine if the post intervention responses were greater than the pre intervention responses for the control group are displayed in Table 5. The test showed evidence that the post intervention responses are not significantly greater than the pre intervention responses in all the surveys. Since the control group did not receive the 4 week FACES intervention, we expected that there would not be significant differences between the pre intervention and post intervention responses.

Table 5: Control Group Results

Group	Survey	Pre	Post	W	p
Control	AKS	Pre	Post	11	0.584
Control	FACES	Pre	Post	10	0.791
Control	FES	Pre	Post	2	0.089
Control	FPPS	Pre	Post	2	0.188
Control	SCS	Pre	Post	1	0.099
Control	SEAS	Pre	Post	7	0.500

W = Wilcoxon Test Statistic, p = p-value

Further Exploration of Research Questions

1. Does the FACES program increase empowerment in African American parents of children with ASD?

   The Family Empowerment Scale (FES) was used to address this question. The Wilcoxon Signed-Rank Test findings indicated that the FACES program increased FES scores for the experimental group. FES Total Scores at the participant level are displayed in Figure 1. This chart shows that participant # 8 was the only participant to experience a decrease in FES scores. As a result of the FACES program, we see that participant # 5 experienced the largest increase in FES scores but still falls below the median for pre and post FES scores.

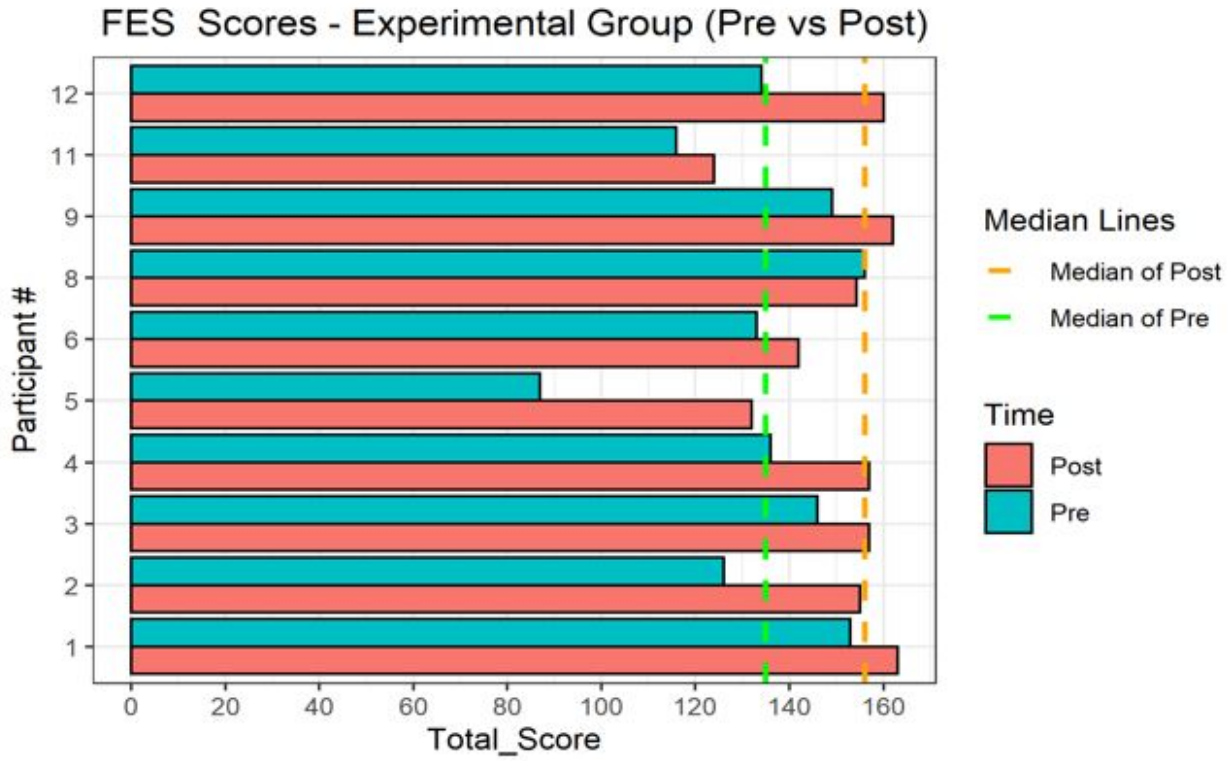

Figure 1: Bar chart for FES Total Scores

2. Does the FACES program increase parents' knowledge of ASD?

   The Autism Knowledge Scale (AKS) was used to address this question. The Wilcoxon Signed-Rank Test findings indicated no evidence that the FACES program increased AKS scores for the experimental group. AKS Total Scores at the participant level are displayed in Figure 2. This chart shows that 4 of the 10 participants experienced a decrease in AKS scores. The pre and post AKS scores for participant # 4 falls far below the median for pre and post AKS scores.

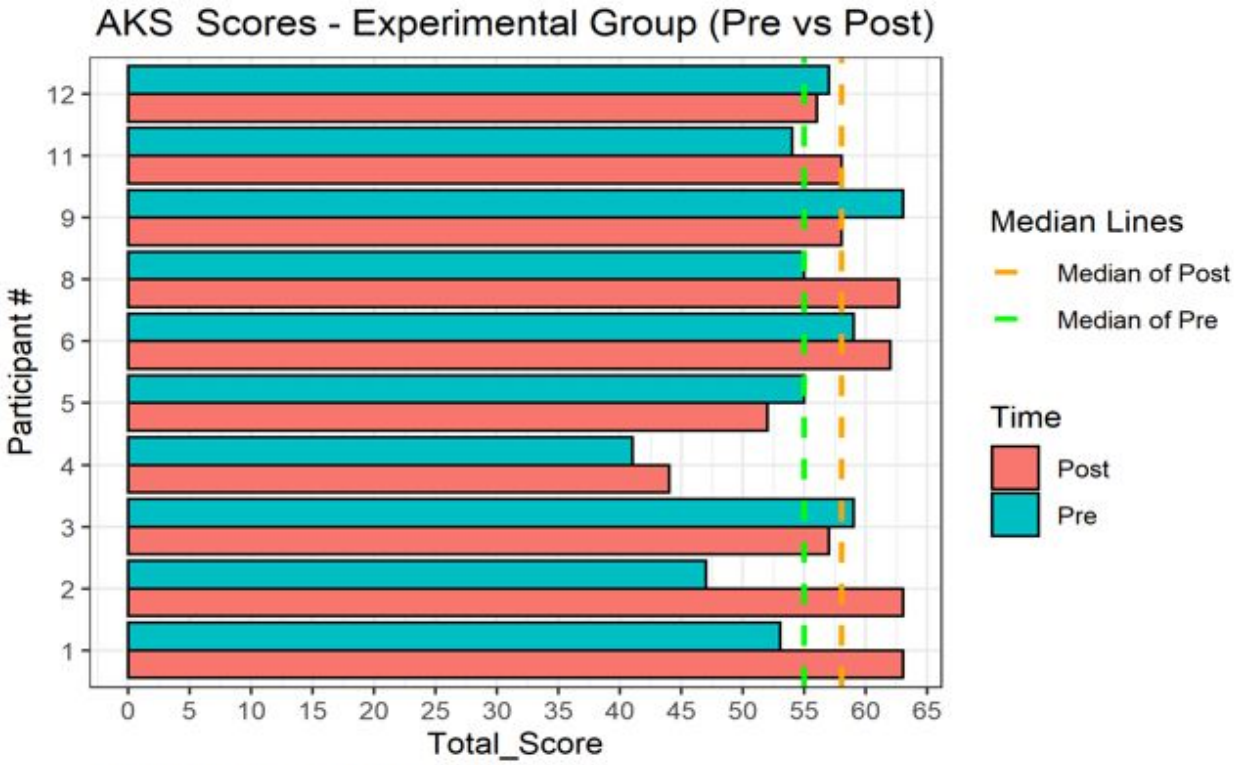

Figure 2: Bar chart for AKS Total Scores

3. Two part question related to confidence:
    a. Does the FACES program strengthen parents' understanding of and confidence in their ability to implement social communication strategies?
    b. Does the FACES program strengthen parents' understanding of and confidence in managing behavior?

The Fostering Advocacy, Communication, Empowerment and Support Scale (FACES) was used to address these questions. The Wilcoxon Signed-Rank Test findings indicated no evidence that the FACES program increased FACES scores for the experimental group. FACES Total Scores at the participant level are displayed in Figure 3. This chart shows that participant # 4 and 8 experienced a sizable decrease in FACES scores.

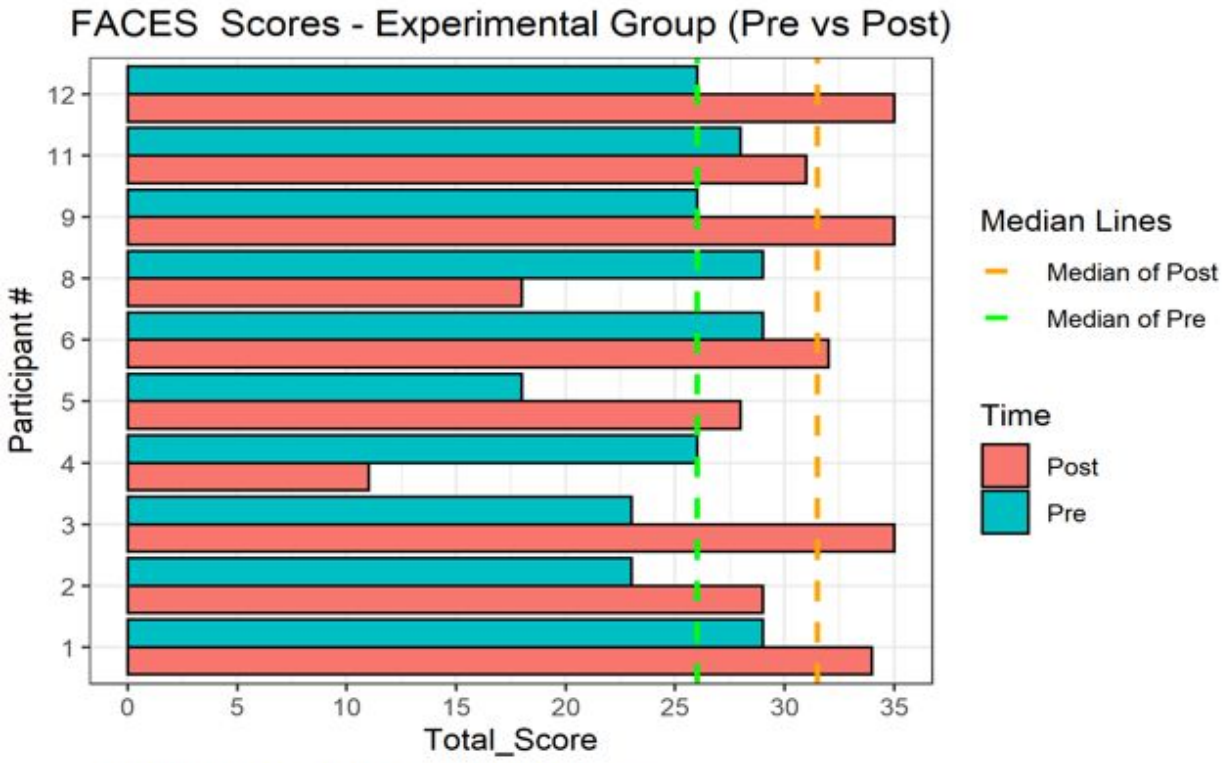

Figure 3: Bar chart for FACES Total Scores

4. Does the FACES program increase parent-professional communication as reported by parents?

   The Family Professional Partnership Scale (FPPS) & the School Communication Scale (SCS) were used to address this question. The Wilcoxon Signed-Rank Test findings indicated that the FACES program increased SCS scores for the experimental group, but indicated no evidence that the FACES program increased FPPS scores for the experimental group. FPPS Total Scores at the participant level are displayed in Figure 4. This chart shows that participant # 3, 4 and 11 experienced a small decrease in FPPS scores. SCS Total Scores at the participant level are displayed in Figure 5. This chart shows that all participants post SCS scores were greater than or equal to their pre SCS scores.

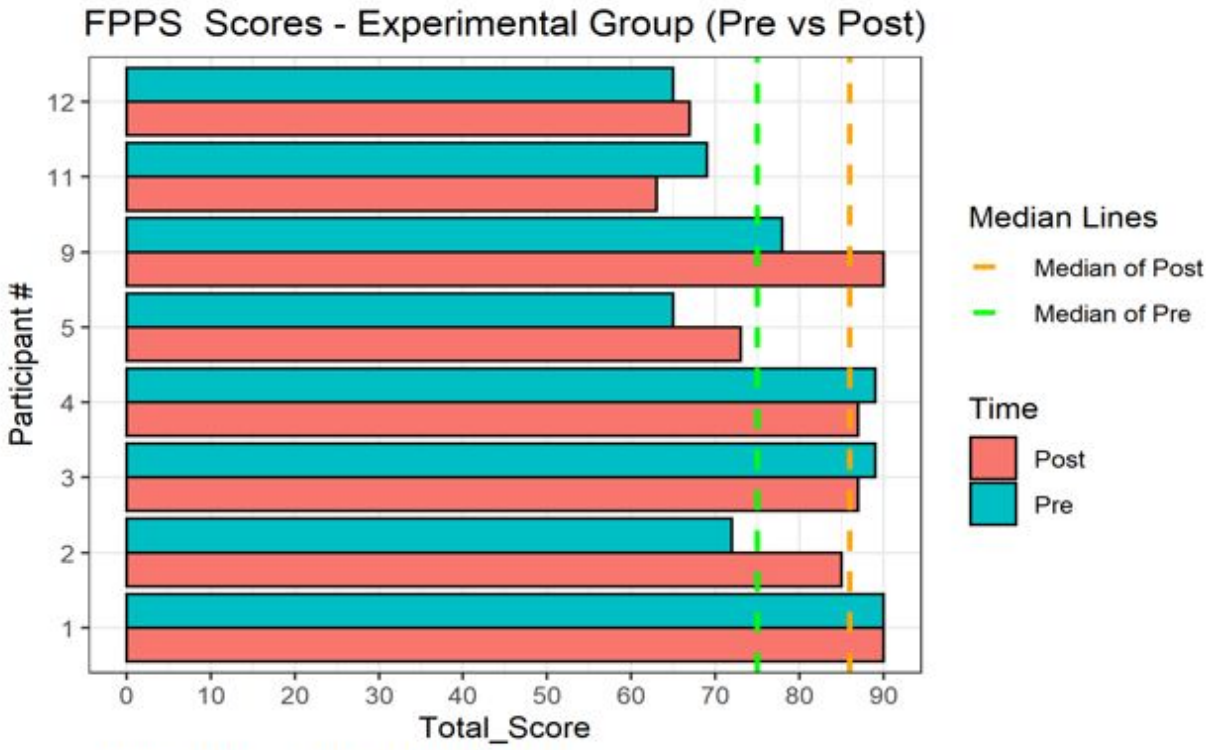

Figure 4: Bar chart for FPPS Total Scores

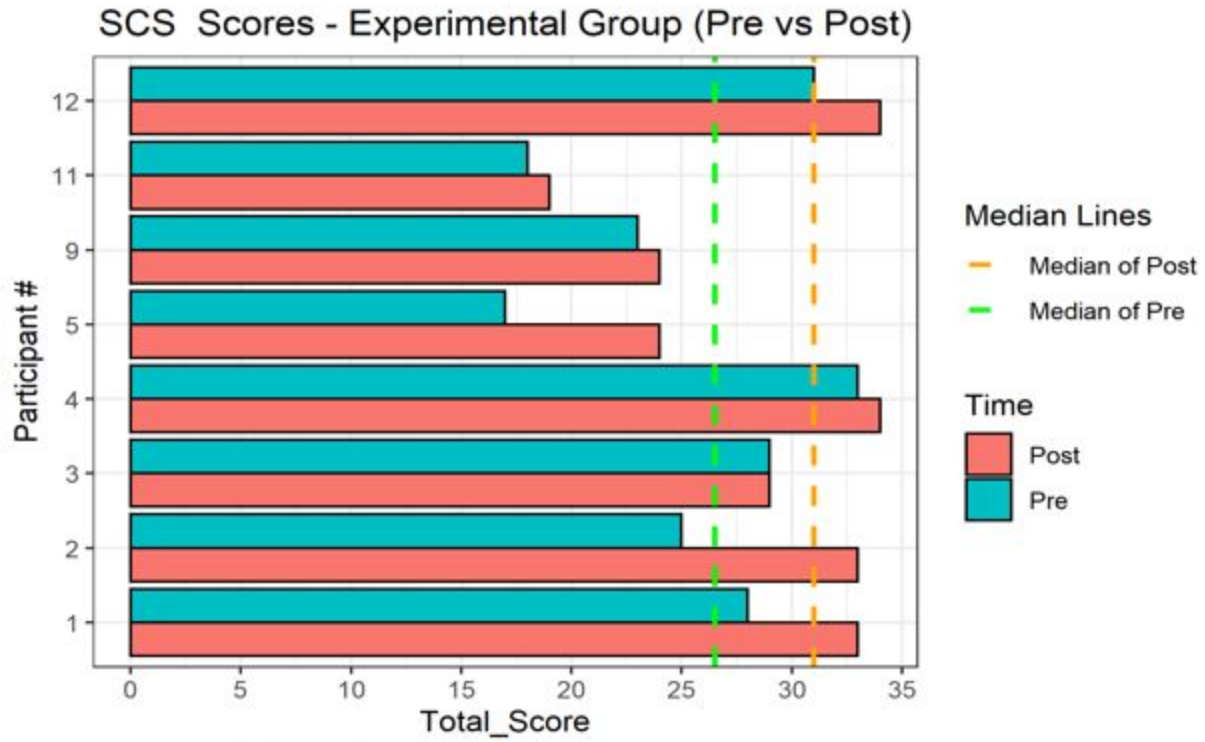

Figure 5: Bar chart for SCS Total Scores

5. In what ways and to what extent does the FACES program strengthen parents' reported perceptions of their ability to advocate effectively for services for their children with ASD?

   The Special Education Advocacy Scale (SEAS) was used to address this question. The Wilcoxon Signed-Rank Test findings indicated that the FACES program increased SEAS scores for the experimental group. SEAS Total Scores at the participant level are displayed in Figure 6. This chart shows that participant # 11 was the only participant to experience a decrease in SEAS scores. As a result of the FACES program, we see that participant # 5 experienced the largest increase in SEAS scores but still falls below the median for pre and post SEAS scores.

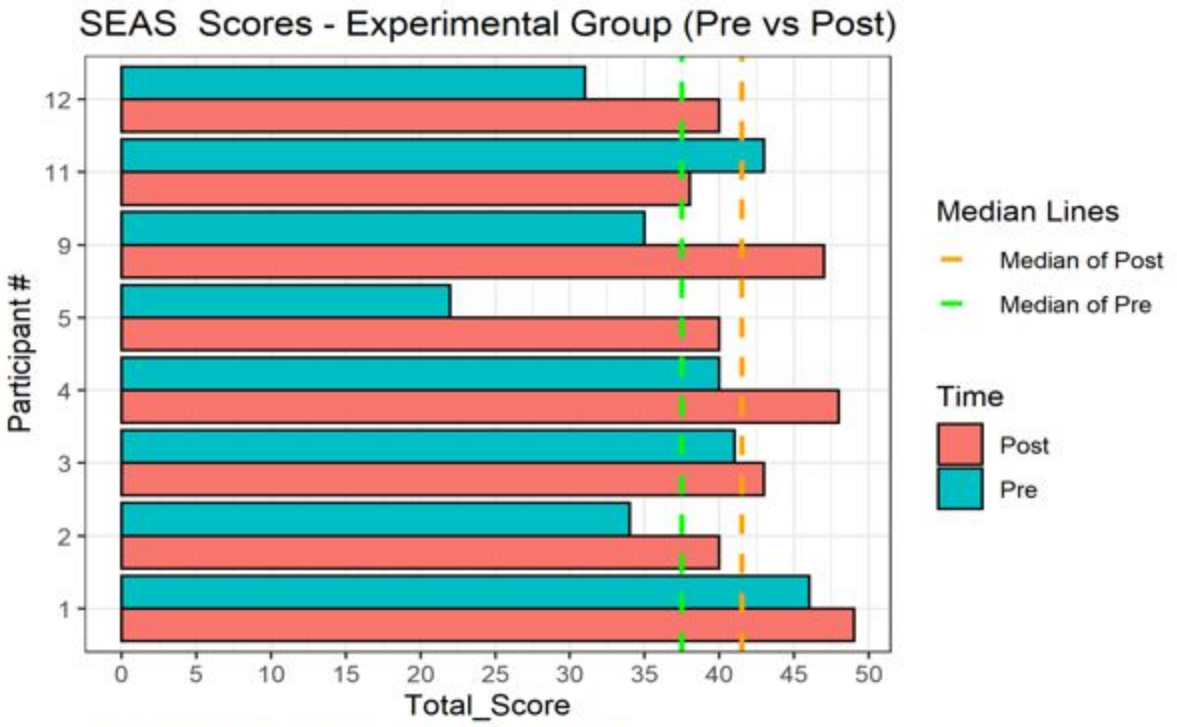

Figure 6: Bar chart for SEAS Total Scores

Experimental Group vs Control Group: Differences Between Post and Pre Total Scores

The difference between pre and post for each participant was calculated (*Post score - Pre score*) for each survey which was then summarized in the boxplots below. The general observed trend is that the experimental group demonstrates a larger positive change in survey scores than the control group. This is the case for all but the FPPS survey and the SCS survey.

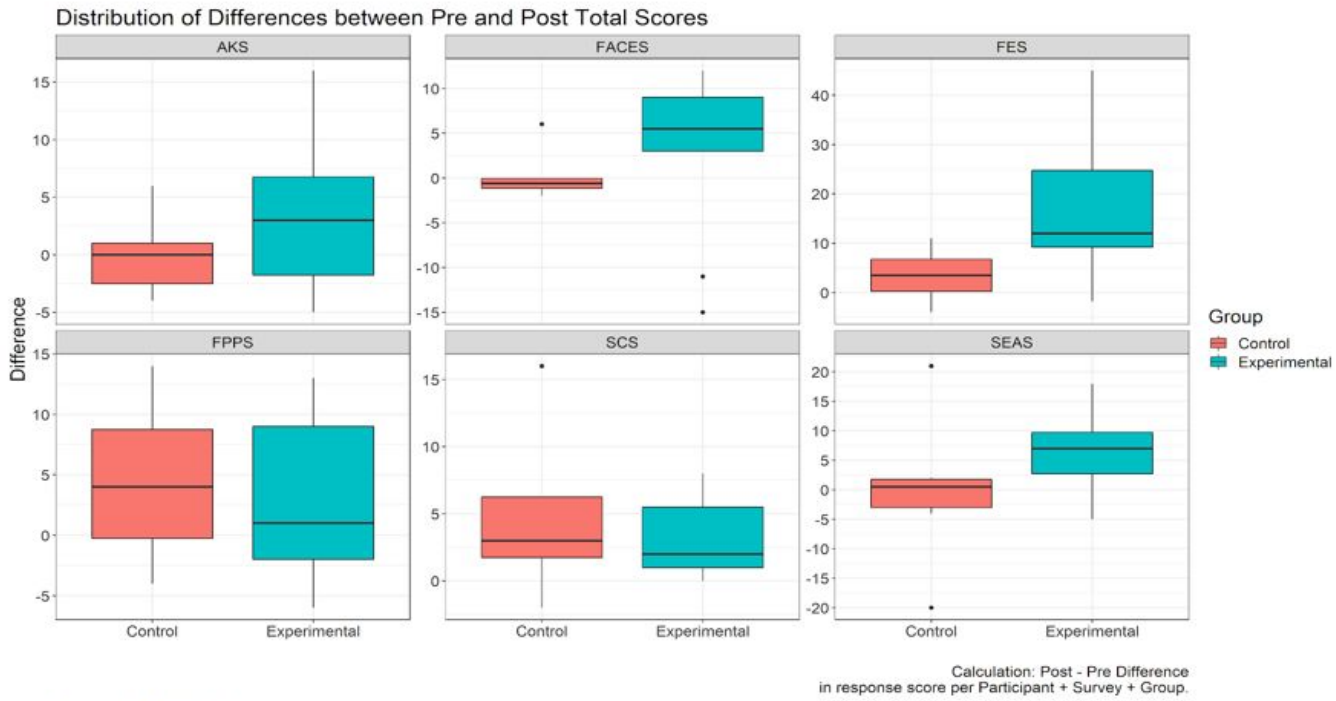

Figure 7: A Boxplot for each survey.

Outlier Detection

The point circled in Figure 8 was classified as a potential outlier because the point resides far away from the line. The circled point represents participant # 5. We ran the Wilcoxon Signed-Rank Tests with and without participant # 5 and compared the test results. There were some slight differences in the results but those differences were not large enough to change any of the conclusions. For this reason, we decided to leave participant # 5 in the analysis.

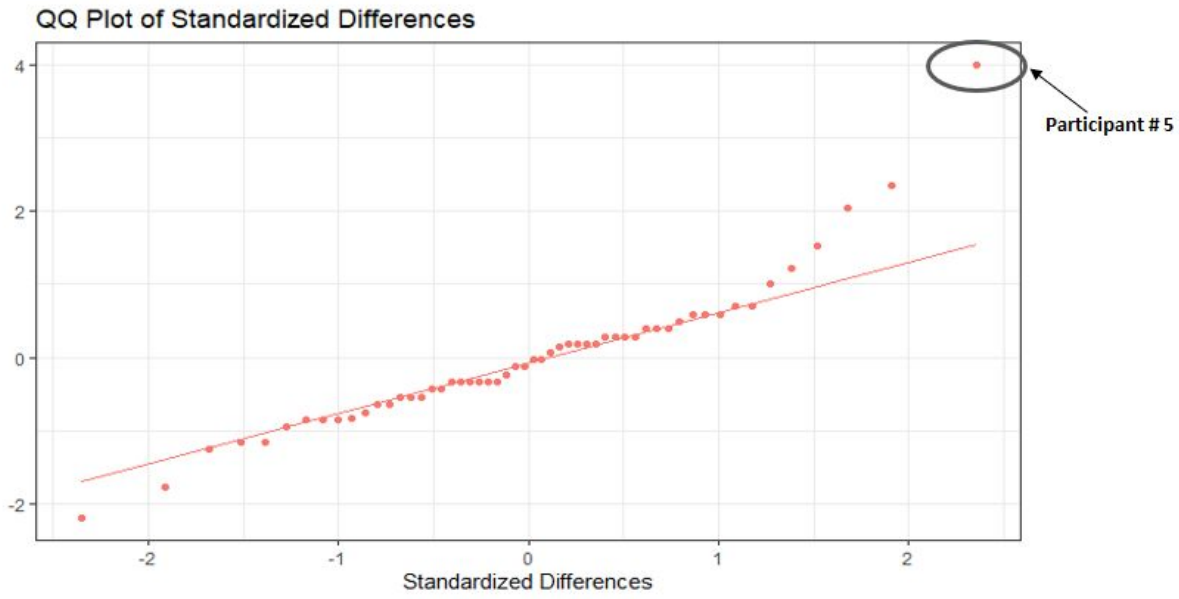

Figure 8: Q-Q (Quantile-Quantile) Plot for Standardized Differences between Pre and Post Scores

Power Analysis and Sample Size Calculations

Power is the probability of failing to reject the null hypothesis when it is false (a Type II Error). We decided to proceed with non-parametric bootstrap simulation. Bootstrap is a statistical methodology that treats the collected data as the "population" one is trying to estimate the metric of. The method entails resampling the collected data, calculating the metric of interest, and repeating a large number of times. The repetitions of this process create a simulated distribution of the metric of interest.

In our simulations, we set out to estimate the p-value of the Wilcoxon Signed-Rank test. We controlled the sample size, and calculated the power level for each survey testing Experimental versus Control group for the Post survey results (unpaired Wilcoxon test) and the Pre versus Post Experimental survey results (paired Wilcoxon test).

For each iteration of the simulation we performed a logical TRUE/FALSE test. If the estimated Wilcoxon p-value was less than 0.05 for a particular comparison, then the results were stored as TRUE, and else it was stored as FALSE. The proportion of simulated results that were TRUE was calculated for various sample sizes.

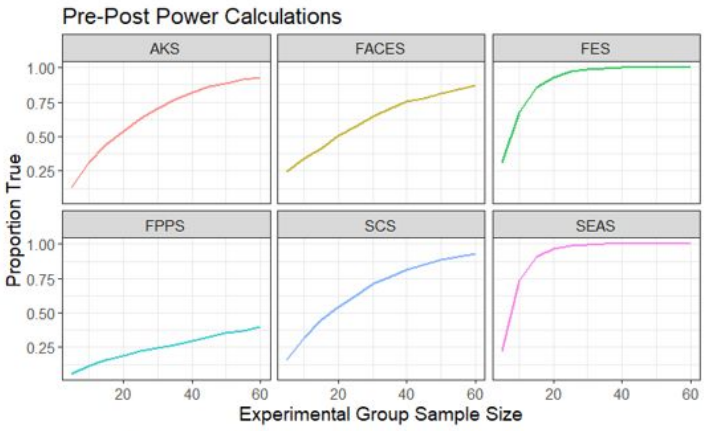

Figure 9: Plot of Power by Sample Size.

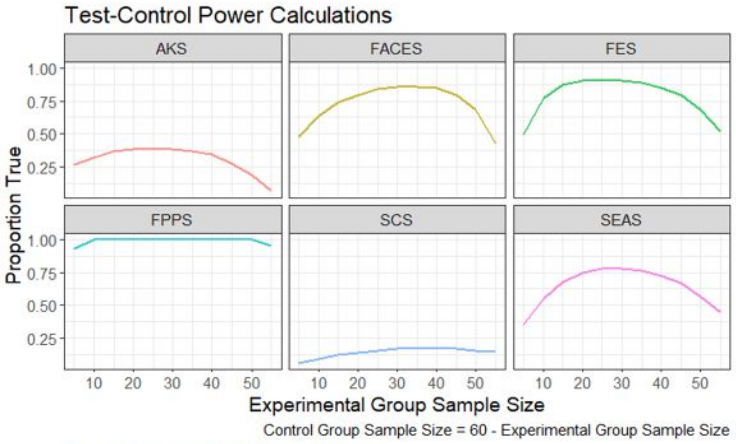

Figure 10: Plot of Power by Sample Size.

In the Pre-Post Experimental Power Calculation (Figure 9), the power can only go up as sample size increases, since this is paired test, whereas in the Test versus Control Calculations (Figure 10), as the number of participants in the experimental group goes up, the number in the control group goes down (and vice versa), which affects the power. It is evident that, with a total number of participants near sixty, having between twenty and forty in each group is sufficient, but there is a noticeable decline in power when one group drops below that threshold of twenty.

From tables 6 and 7, we recommend a ratio of 40 participants in the experimental group, and 20 participants in the control group. This will provide enough power to detect significance in both, while also providing the benefit of the intervention and training to as many families as possible. Cells highlighted in green indicate power levels that are $\geq 80\%$. The extent to which the power can be determined for each survey is limited by the quality and size of the data collected. Meaning that upon the next iteration of the FACES program power tests should be reperformed to ensure the sample size is adequate, and inform future sample size determinations.

Power of test: Experimental versus Control Group Wilcoxon Signed-Rank Test

Sample Size			Power Level*					
Experimental	Control	Total	FACES	AKS	FES	SCS	FPPS	SEAS
30	30	60	85.29%	38.31%	90.70%	17.02%	100.00%	77.84%
35	25	60	85.88%	37.25%	88.82%	17.42%	100.00%	76.20%
40	20	60	84.72%	34.30%	85.01%	18.07%	99.99%	72.50%
45	15	60	79.16%	27.64%	79.00%	17.12%	99.94%	66.72%
50	10	60	68.15%	18.94%	68.50%	16.02%	99.80%	56.70%
55	5	60	42.77%	7.41%	52.27%	15.07%	95.34%	44.79%

Table 6: Table of Power Levels for Experiment vs Control Post Groups

Power of test: Pre versus Post Survey Results Wilcoxon Signed-Rank Test

Sample Size			Power Level*					
Pre	Post	Total	FACES	AKS	FES	SCS	FPPS	SEAS
5	5	5	24.53%	13.09%	30.98%	15.86%	6.09%	22.40%
10	10	10	34.00%	31.12%	67.36%	31.58%	11.71%	73.30%
15	15	15	41.41%	44.36%	85.47%	44.64%	15.96%	90.49%
20	20	20	50.42%	53.87%	93.08%	54.22%	19.18%	96.62%
25	25	25	57.56%	63.08%	97.20%	62.27%	22.59%	98.88%
30	30	30	64.51%	70.63%	98.70%	70.89%	24.29%	99.61%
35	35	35	70.33%	76.73%	99.55%	76.12%	26.86%	99.89%
40	40	40	75.50%	81.68%	99.83%	81.29%	29.75%	99.98%
45	45	45	77.59%	86.10%	99.93%	85.31%	32.98%	99.98%
50	50	50	81.37%	88.20%	99.97%	88.88%	35.69%	100.00%
55	55	55	84.16%	91.23%	99.99%	90.98%	37.08%	100.00%

Table 7: Table of Power Levels for Pre vs Post Experimental Group

*The power level specifies the chance of not making a Type II error.

Usually, the researcher takes the power level as 0.80. In other words, the researcher has an 80% chance of not making a Type II error.

Type II error - Accepts a false null hypothesis

Discussion

Wilcoxon Test Results

The unpaired Wilcoxon test results for Pre Experimental versus Pre Control (Table 2) suggests that the participants selected for the study come from the same survey responses. Worth noting is the FACES Experimental versus Control Table 2 result, p = 1.000. The results from Table 2 provide a sufficient foundation for extrapolation. The unpaired Wilcoxon test results for Post Experimental versus Post Control are illuminating when contrasted against the aforementioned results.

In Table 3 the results for FACES, FES, and SEAS provide stronger evidence that the survey responses are different in the Post survey results from Pre survey results. If a difference between the two groups is in fact the truth, then, based on the design of the experiment, that difference is likely to have been spurred by the FACES intervention program.

Additionally, the results from the Pre FACES survey versus the Post FACES survey for the Experimental group demonstrate that there may be a meaningful difference in the survey results (p = 0.179). Statistically, the evidence does not provide a watertight conclusion, yet a visual spread of the Pre and Post FACES survey results suggests that the lack of statistical rigor may be due to a small sample size. Two outliers in the FACES boxplot (Figure 1) are observed to skew the overall results lower. The reason for their drop in score is not readily surmised from the data nor is it possible to conjecture from the data alone given the sample size. Comparing the results of the Pre FACES survey versus the Post FACES survey for the control group suggest that the FACES treatment had a significantly weaker effect on the control group. This makes sense as the control group did not receive the FACES intervention treatment prior to the Post survey. The evidence that the Pre and Post survey responses are different for the Control group FACES response is therefore remarkable when taken in context of the Pre and Post FACES Experimental group response.

The result for AKS, FPPS, and SCS, provides evidence that the survey responses may be more alike than different in the Post survey results than in the Pre survey results. The SCS and FPPS scales take longer than four weeks for results to manifest in a detectable difference (Dr. Peason). It is worth noting that there is an underwhelming level of significance in the *p*'s of several results that compare the Experimental group to the Control group survey results and even the Pre and Post survey results. That being said, there are observable trends consistent with the principles from which the FACES intervention derived its design from. FACES was constructed on adult education principles aimed at improving the understanding and confidence of African American parents in managing and utilizing social programs available to their children with ASD. The results of the FACES and FES surveys are encouraging, and warrant a next iteration of the FACES program with a larger sample size.

Power

Knowing the result of this test would be used as a recommendation for the distribution of experimental and control participants in a grant request, we wanted to provide robust power calculations. By using a non-parametric bootstrap sampling method, we were able to rely on the

data and create simulated studies to predict how to best allocate participants into experimental and control groups to achieve optimal power while still providing the treatment to the maximum number of people. The control group sample was small, but the simulation allowed us to extrapolate out to estimate how things may look with a larger sample. This is, of course, predicated on the current control group sample being representative of an average control group population, and if the next sample is dissimilar, significance tests may not match our power calculations.

While keeping in mind the current limitation of the data, our estimates and recommendations are intuitive. Having an even ratio or 40-20 split in the test and control groups makes sense and is a format many would use regardless of any data analysis.

Future Extensions

For the future iterations of the FACES program, we would suggest revamping the survey administration and/or response capture process to reduce the number of missing responses. In the data used for our analysis, 11 of the 16 participants (68.75%) answered all questions in the survey. The goal would be to increase the percentage of participants that answer all questions.

Reviewing the data at the participant level shed some light on the fact that some participants experience a decrease in scores from pre to post intervention. For example, participant # 4 scored 26 in the FACES survey pre intervention and dropped to 11 post intervention. An individualized follow up with participants after the post survey could potentially help families improve in areas where they feel deficient. An individualized follow up, might benefit all participants but if there are resource constraints a strategy could be to focus on the families with lowest scores or families that showed reduction in scores post intervention.

Demographic information about each family was captured in the survey but we did not use the demographic data in our analysis. Future iterations of the FACES program, can utilize the demographic data to see if there is correlation between the demographics and survey responses.

In the future, we would like to build a predictive model to see if the pre intervention survey responses are predictive of major changes in post intervention survey responses. This would help predict which type of participants would benefit the most from the FACES program training. This information could be used to develop more sophisticated and targeted recruitment strategies.