Strengths and Difficulties Questionnaire Analysis:
A Self-Report Behavioral Screening Questionnaire for Children Ages 4-17
Prepared by: Emilie Campos
December 16, 2017

1. Abstract

The Strengths and Difficulties Questionnaire is a self-report behavioral screening questionnaire for children and adolescents ages 4 through 17. The questionnaire and scores on the subscales -- emotional symptoms, conduct problems, inattention-hyperactivity, peer problems, and prosocial behavior – were calculated at each of 4 different time points (assessments) corresponding to the timeline of intervention: baseline (prior to the intervention), exit (immediately following completion of the intervention), follow-up 1 (3 months after completion of the intervention), and follow-up 2 (9 months after completion of the intervention). All 5 of the subscales range from 0-10 (in theory). Higher scores indicate healthier behavior on the Prosocial Behavior Subscale. For the other 4 subscales, lower scores indicate less problematic behavior. The intervention is designed to target the behaviors measured in these subscales with the hope that a decrease in problematic behaviors and an increase in healthy behaviors will be observed from baseline to exit and will be sustained during the 9 months following completion of the intervention. We will create a SAS macro to perform paired t-tests from baseline to the last completed assessment. Next, we will simulate the data in R in order to find similar results with simulated data.

2. Methods

Rather than working with the entire data set, we subset the data by a random number assigned to each student. My random number was 3. This resulted in 876 of the 40,000 assessments. Since there were 4 assessment time points per child, this is 219 children. I then converted this longitudinal data set to a wide data set in order to create the tables in section 4. In order to find if there was a statistically significant improvement in scores, we needed to find the "last available" assessment time point for each child and then compare their baseline score to their last assessment.

3. Results

Without going into extreme detail, we will highlight some of the results from the following tables and figures, keeping in mind that all of the children completed at least the baseline assessment.

Pointing to Table 1, we find that as time progressed, the proportion of missed assessments increased, with only half of the children completing the second follow-up. Proceeding to Table 2, we find that 43% of the children attended all 4 assessments and only 11% dropped out after only the baseline assessment. On the topic of drop out, Table 3 shows that 11% dropped out after the baseline assessment, 17% after the exit assessment, 18% after follow-up 1, and the remaining 54% remained in the study through the assessment.

Next, with Table 4, we find that there were more male children than female children in the study at every time point. Looking at the trend, it seems like all of the subscales expect pro-social behavior and peer problems are decreasing. The pro-social behavior subscale seems to be increasing.

When performing a t-test on each of the subscales for their baseline score minus their last available score we find that there is a significant difference in scores at the .05 significance level for 4 of the 5 subscales (Table 5). Also, the differences for emotional symptoms, conduct problems, inattention hyperactivity, are positive, with entirely positive 95% confidence intervals, suggesting these scores fell indicating improvement. The pro-social behavior scores have a negative mean difference and an entirely negative confidence interval suggesting these scores rose with the treatment and since higher scores indicate healthier behavior, this also suggests the intervention led to less problematic behavior.

Figures 1 and 2 were created using simulated data in R. By simulating 500 data sets using the summary statistics found in the observed data set, we were able to find the proportion of t-tests for the Emotional Symptoms subscale that were statistically significant at the .05 significance level. From this procedure, we see that as the sample size increases and the proportion of missing values stays the same, the proportion of statistically significant t-tests increases to 1. When the proportion of missing completely at random assessments increases, we find that the proportion of statistically significant differences decreases rapidly.

4. Tables and Figures

Assessment Time Point	Attendance	Frequency	Percentage*
Baseline	Missed	0	0.0
	Attended	219	100.0
Exit	Missed	39	17.8
	Attended	180	82.2
Follow-up 1	Missed	74	33.8
	Attended	145	66.2
Follow-up 2	Missed	103	47.0
	Attended	116	53.0

Table 1: Assessment Attendance
Percentages based on total number of children - 219

No of Assess	Frequency	Percentage*
1	25	11.4
2	41	18.7
3	59	26.9
4	94	42.9

Table 2: Number of Assessments
Percentages based on total number of children – 219

Baseline	Exit	Follow-up 1	Follow-up 2	Frequency	Percentage*
1	0	0	0	25	11.4
1	0	1	0	3	1.4
1	0	1	1	11	5.0
1	1	0	0	38	17.4
1	1	0	1	11	5.0
1	1	1	0	37	16.9
1	1	1	1	94	42.9

Table 3: Combinations of Assessments
Percentages based on total number of children – 219

Time Point	Gender	N	ESS Mean	ESS Std	CPS Mean	CPS Std	IHA Mean	IHA Std	PPS Mean	PPS Std	PSB Mean	PSB Std
BASELINE	Female	92	4.0	2.4	2.2	1.6	4.0	2.5	1.9	1.6	7.0	2.0
	Male	127	3.1	2.5	2.6	2.0	5.1	3.0	2.4	2.0	6.5	1.9
	All Children	219	3.5	2.5	2.4	1.8	4.6	2.8	2.2	1.9	6.7	2.0
EXIT	Female	73	2.5	2.1	1.6	1.6	2.9	2.6	1.5	1.7	8.2	2.6
	Male	107	2.6	2.4	2.1	2.3	3.8	3.0	2.1	1.9	8.1	2.4
	All Children	180	2.6	2.3	1.9	2.0	3.4	2.8	1.8	1.9	8.1	2.5
FOLLOW-UP 1	Female	58	2.3	2.5	1.5	2.0	2.2	2.7	1.4	1.9	8.5	2.5
	Male	87	2.3	2.7	1.8	2.2	3.1	3.2	2.0	2.1	8.3	2.5
	All Children	145	2.3	2.6	1.7	2.1	2.8	3.0	1.7	2.0	8.4	2.5
FOLLOW-UP 2	Female	36	1.8	2.6	1.6	2.2	2.9	3.2	2.0	2.9	9.1	1.9
	Male	80	2.6	2.8	1.9	2.3	3.7	3.4	2.3	2.4	8.0	3.1
	All Children	116	2.3	2.8	1.8	2.3	3.4	3.3	2.2	2.5	8.3	2.8

Table 4: Means and Standard Deviations of Each Subscale by Gender by Time Point

	N	Mean Difference	Std	95% CI		t Value	p-value	Sig
SDQ_ESS	194	1.064	3.1215	0.60	1.49	4.67	<0.0001	*
SDQ_CPS	194	0.5979	2.2656	0.28	0.92	3.68	0.0003	*
SDQ_IHA	194	1.4433	3.1753	0.99	1.89	6.33	<0.0001	*
SDQ_PPS	194	0.2371	2.4506	-0.11	0.58	1.35	0.1794	
SDQ_PSB	194	-1.6392	2.1627	-1.95	-1.33	-10.56	<0.0001	*

Table 5: Results from t-tests performed on Baseline scores – Last available scores
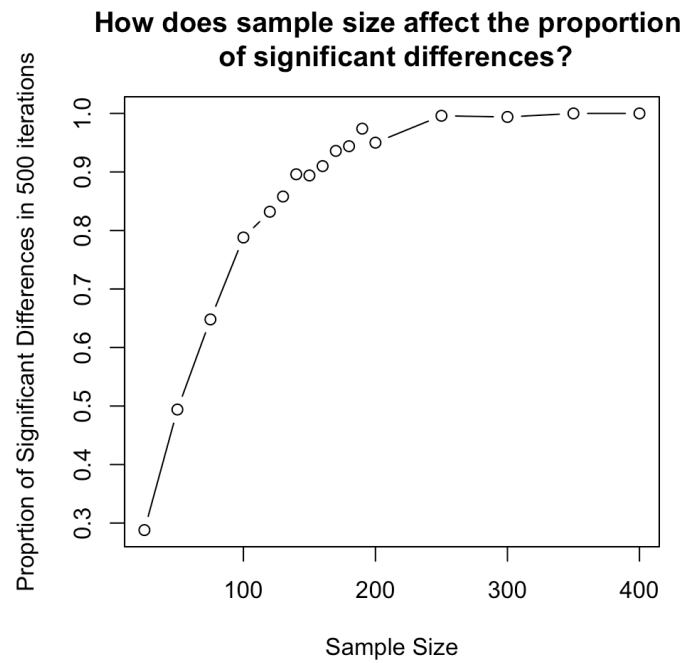
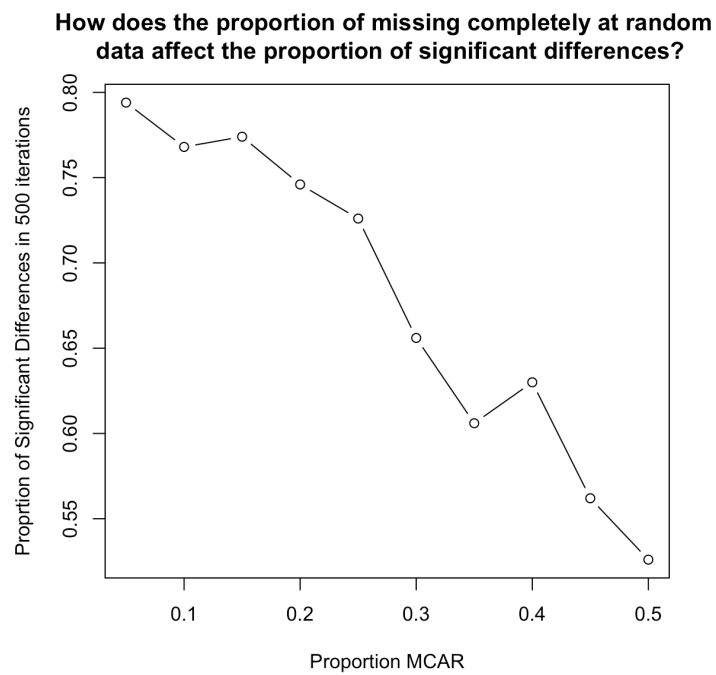Figure 1: Varying Sample Size – t-test on the Emotional Symptoms subscale

Figure 2: Varying Proportion of Missing Completely at Random – t-test
on the Emotional Symptoms subscale