**Supplementary Material**

Assumption Checks

Statistical assumptions for correlational and regression analyses (normality, multivariate normality, linearity, homoscedasticity, multicollinearity and singularity) were checked before analyses were conducted. Visual inspection of scatter plots revealed no violations to linearity or homoscedasticity and no influential outliers were identified. Violations to normality were identified for ADOS-2 Comparison Score (CS) and HSQ frequency (see Figure 1 a-g for probability-probability [P-P] plots for the various scales). The ADOS-2 CS was left untransformed given transformation of standard scores can hinder interpretation of results (Tabachnick & Fidell, 2013).
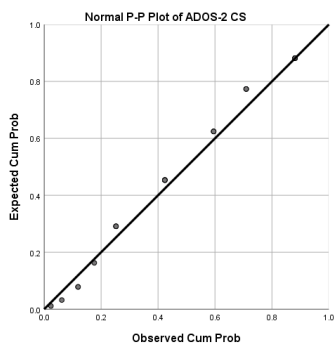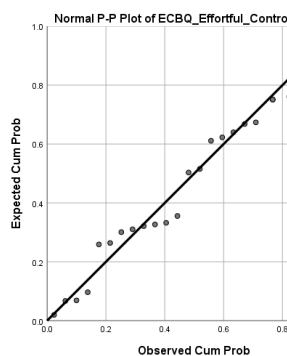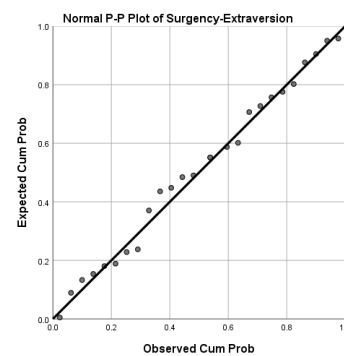
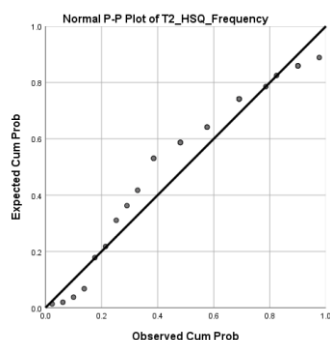a. P-P Plot ADOS2-CS
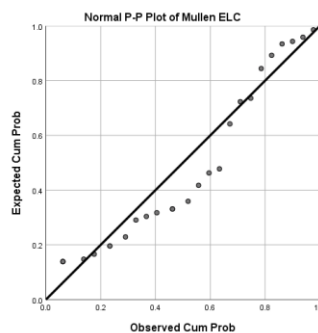
b. P-P Plot Mullen ELC
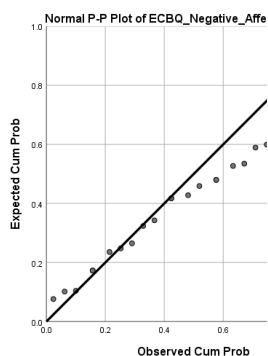
c. P-P Plot Negative Affectivity   d. P-P Plot Effortful Control   e. P-P Plot Extraversion

f. P-P Plot HSQ Frequency
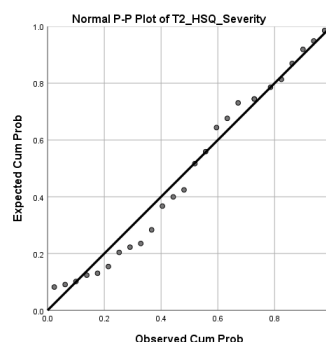
g. P-P Plot HSQ Severity

Figure 1. Normality check (P-P Plots) for all scales

Selection Bias and Confounds

Independent samples *t*-tests were run to examine whether those who were included in the current study ($n = 26$) differed to those who did not participate ($n = 17$) on age, autism severity (ADOS-2 CS), developmental level (Mullen ELC), and temperament (ECBQ). As can be seen in Supplementary

Table 1, no significant differences were identified between the broader cohort and the sub-sample on any of the aforementioned variables.

Supplementary Table 1
T-Tests between Follow-Up Cohort (*n* = 26) and Original Cohort that Did Not Participate in the Follow-Up (*n* = 17) on Age, ADOS-2 CS, Mullen ELC, and ECBQ

	Follow-Up Cohort		Original Cohort			
	M	*SD*	*M*	*SD*	*t*	*p*
Age (months)	39.65	8.20	43.24	9.92	-1.29	.205
ADOS-2 CS	7.27	2.31	6.59	3.02	.837	.408
Mullen ELC	77.46	26.23	76.31	27.39	.136	.893
ECBQ						
Negative affectivity	3.63	.761	3.60	.848	.168	.868
Effortful control	4.13	.657	3.85	.768	1.32	.194
Surgency-extraversion	4.53	.652	4.63	.513	-.572	.571

Note. ADOS-2 CS: Autism Diagnostic Observation Schedule, second edition, Comparison Score. Mullen ELC: Mullen Early Learning Composite. ECBQ: Early Childhood Behavior Questionnaire.

Number of Therapies and Challenging Behavior Outcomes at Follow-up
Given that behavioral therapies and interventions (e.g., Applied Behavioral Analysis, Speech Therapy, Psychology, etc.) may have impacted challenging behavior at follow-up, information regarding the number of therapies children engaged in was analyzed with challenging behavior outcomes (HSQ frequency and HSQ severity). At follow-up all children were engaged in a minimum of two therapies. Information regarding the kinds of therapies children engaged in is provided in the main manuscript (see Table 1). A one-way ANOVA was conducted in order to examine whether the number of interventions a child engaged in influenced the frequency or severity of challenging behavior at follow-up. Children were first categorized into groups based on the amount of therapies they had engaged in over the course of the follow-up: Low (≤2 therapies), Moderate (2-3), Frequent (≥4). Mean scores obtained for HSQ frequency and severity were compared across each group. There were no significant differences between scores for each of the three groups for HSQ Frequency, $F(2, 22) = .437$, $p = .651$, $\omega^2 = .05$, and HSQ Severity, $F(2, 22) = .751$, $p = .484$, $\omega2 < .01$, thus the frequency and severity of challenging behavior at follow-up was found to be unrelated to the amount of therapy children were engaged in.

References
Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate Statistics*. Boston, MA: Pearson/Allyn & Bacon.