**Supplemental Material**

**Missing Data Analysis**

By design, the Puzzle planning and flexibility items were missing for 47% of observations due to time limitations for one of the treatment trials. Little’s MCAR test indicated that data could be considered missing completely at random for the purpose of data analysis, χ2(13)=18.52, *p*=.14. Missing data were imputed using the r-package MICE so that all observations could be used in the primary analyses (*M*=40, Van Buuren & Groothuis-Oudshoorn, 2011).

**Sensitivity Regressions.**

Three regression models were fit in a hierarchical manner. The first model regressed the outcome on the covariates. As age and verbal intelligence were associated with scores on the EFCT, these variables were used as covariates in a series of regression models. The second model regressed the outcome on the covariates and EFCT Planning and Flexibility Scale scores. The third model regressed the outcome on the covariates and the EFCT Total score. Model 1 was used as a baseline comparison for models 2 and 3. See Supplemental Table 1.

**ANCOVA Sensitivity Analyses.**

A series of one-way ANCOVAs compared the diagnostic groups on the EFCT Total, Planning, and Flexibility scales after covarying for the youth’s sex, age, verbal ability, and parental education level. There were significant differences among the diagnostic groups for EFCT Total (F(2, 128308.18) = 26.69, *p* < .001, η2 = .14, partial η2 = .17), Planning (F(2, 24095.77) = 17.35, p < .001, η2 = .10, partial η2 = .12), and Flexibility, F(2, 47286.65) = 15.23, *p* < .001, η2 = .10, partial η2 = .11.

Supplemental Table 1.

Hierarchical Regressions Predicting EFCT Scores for Construct Validity

|  |  |  |  |
| --- | --- | --- | --- |
| Dependent Variable |  |  | EFCT |
| Measure | Scale |  | Age b | Verbal Intelligence b | Flexibility b | Planning b | Total b |
| BRIEF | Plan/Organize |  | -.94\* | -.26\* |  |  |  |
|  |  | -.38 | -.11 | .96\* | 1.20\* |  |
|  |  | -.39 | -.12 |  |  | 1.06\* |
|  |  | ΔR2 |  |  | .02\* | .02\* | .05\* |
|  | Shift |  | -1.15\* | -.26\* |  |  |  |
|  |  |  | -.35 | -.04 | 1.73\* | 1.43\* |  |
|  |  |  | -.32 | -.04 |  |  | 1.61\* |
|  |  | ΔR2 |  |  | .05\* | .02\* | .11\* |
|  | GEC |  | -1.34\* | -.29\* |  |  |  |
|  |  |  | -.61 | -.09 | 1.43\* | 1.41\* |  |
|  |  |  | -.61 | -.09 |  |  | 1.42\* |
|  |  | ΔR2 |  |  | .04\* | .03\* | .10\* |
| DKEFS | Category Fluency |  | -.01 | .08\* |  |  |  |
|  |  |  | -.06 | .07\* | .05 | -.20 |  |
|  |  |  | -.04 | .07\* |  |  | -.05 |
|  |  | ΔR2 |  |  | .00 | .01 | .00 |
|  | Switch Accuracy |  | .32\* | .11\* |  |  |  |
|  |  |  | .28\* | .10\* | -.08 | -.08 |  |
|  |  |  | .28\* | .10\* |  |  | -.08 |
|  |  | ΔR2 |  |  | .00 | .00 | .01 |
| ADOS | RRB |  | .10 | .00 |  |  |  |
|  |  |  | .06 | .15 | .13 | -.17 |  |
|  |  |  | .11 | .00 |  |  | .01 |
|  |  | ΔR2 |  |  | .04 | .02 | .00 |

\* *p*<.05.

*Note*: ΔR2 is the squared semi partial correlation (e.g., Flexibility ΔR2 controlling for Age, Verbal Intelligence, & Planning).

Autism Diagnostic Observation Schedule (ADOS); Executive Function Challenge Task (EFCT); Behavior Rating Inventory of Executive Function (BRIEF); Global Executive Composite (GEC); Delis-Kaplan Executive Function System (DKEFS); Restricted/repetitive behaviors (RRB)