Supplementary Materials for Manuscript 'Exploring the Neurocognitive Correlates of Challenging Behaviours in Young People with Autism Spectrum Disorder'

Supplementary Materials 1

Neurocognitive Tasks

Testing took place in a quiet testing area and tasks were presented in one of four carefully selected orders. The battery was completed over two days of testing, with a median gap of 21 days (range 1-259 days) between sessions. Seventeen participants required a final day of testing to complete the battery.

Theory of Mind Tasks

The Strange Stories task (Happé 1994) was used as a general measure of mental state understanding. Participants were read a series of stories, which were also available in front of them and accompanied by an appropriate illustration. At the end of each story, participants were asked a question about the text. Correct answers demonstrated an understanding of the characters thoughts, feelings and intentions. The outcome variable was the average score across the four theory of mind stories (score range 0-2).

The Frith–Happé animations (Abell et al. 2000) consist of a series of silent videos of two-dimensional animations, requiring the participant to understand intentionality behind the moving shapes. Four animations depicted theory of mind interactions and two goal- directed interactions. The outcome variable was the average intentionality score, based on degree of mental state attribution for the four theory of mind trials (score range 0-5).

False Belief Composite Score. A composite score was generated based on performance on two false belief tasks. The first was the 'combined false belief task', which is a combination of first- and second-order false belief tasks based on previous tasks measuring false belief understanding (Hughes et al., 2000; Sullivan, Zaitchik & Tager-Flusberg, 1994). The second task was the 'second order false belief task', which had greater verbal demands than the combined task. A total score of performance on the combined and second order false belief tasks was used, with points awarded for correctly passing and justifying each false belief question (score range 0-8).

The Reading the Mind in the Eyes task (Baron-Cohen et al. 2001). The eyes test requires the participants to understand mental/emotional state "concepts" and match them to expression of eyes from black and white photos. Participants were shown black and white photographs of just the eye region of the face of 28 people. Participants were asked to pick which of four inner state words best described what the person in the photo is thinking or feeling. A point was awarded for each correct trial (score range 0-28).

The Penny Hiding task (Baron-Cohen 1992) was used as a naturalistic and non- verbal measure, specifically indexing the participant's ability to deceive the experimenter. The participant was given six trials of hiding the penny. Responses are coded for the type of deception errors made, with a total score calculated. It was possible to display more than one error on a trial. Given the distribution of the scores this variable was re-coded as ordinal (score range 0/1= '1', 2/3= '2', 4/5= '3', ≥6= '4').

Executive Functioning Tasks

The Card Sort task was used as a measure of cognitive flexibility and response reversal adapted from a child-friendly version of the Wisconsin Card Sorting Task (Tregay et al. 2009). Participants had to correctly sort cards to one of three alternative sets across three trials, with the correct set varying in each trial. The key variable was the number of incorrect responses made across all three trials.

The adapted Trail Making task was included as a measure of attentional switching and response reversal (Reitan and Wolfson 1985). Participants were asked to 'join the dots' in numerical order, then, in a second trial, in alphabetical order, followed by a third trial switching between numbers and letters. The difference between the time taken on the first trial and the third trial comprised a measure of switching ability.

The Opposite Worlds task was taken from the Test of Everyday Attention for Children (Manly et al. 2001) and was included as a measure of interference inhibition. The task included a "same world" trial, where participants read out a series of the numbers 1 and 2; and the "opposite world" trial, where participants had to say the opposite to the number they were reading. Two same world trials and two opposite world trials were presented. The time taken to complete each world was recorded in seconds. The outcome variable was the subtraction of the mean same worlds completion time from the mean opposite worlds completion time.

The Score! Task was also taken from the Test of Everyday Attention for Children (Manly et al. 2001) and was included as a measure of sustained attention. Participants have to keep a count of the number of 'scoring' sounds they hear on a tape across 10 trials. A trial was coded as correct if the correct amount of sounds was identified at the end of the trial. Given ceiling effects in the scores, the variable was re-coded as ordinal (0 incorrect trials = '1', 1/4 incorrect trials= '2', 5/10 incorrect trials= '3').

Perceptual Processing

Auditory Processing

In each dinosaur pairing, the participant was presented with one 'standard' stimulus, which did not change across the particular task, and a probe stimulus that varied. A detection threshold was established using a two-down/one-up (after 2 correct trials the perceptible difference between the two stimuli reduces; after 1 incorrect trial the perceptible difference between the two stimuli is increased) adaptive staircase procedure, where the task was made easier/harder dependent on ongoing participant performance. This was used to determine the threshold at which the participant was correct on 75% of trials. The task was terminated after 6 reversals (changes in direction in the two-down/one-up procedure) or after 40 trials, and the final threshold score was the mean threshold value from the 4th reversal. Across both tasks, a higher threshold indicated a greater amount of information required to detect differences in the two stimuli.

Visual Processing

Three tasks were presented (motion coherence, form-from-motion, and biological motion), and each task was preceded by a five trial practice, where feedback and discussion of their

decision ensured that all participants understood the task. Similar to auditory tasks, a detection threshold was established using a two-down/one- up adaptive staircase procedure, where the task was made easier/harder depending on ongoing performance. A task was terminated after seven reversals of the staircase. The threshold score was calculated as the average signal-to-noise ratio (signal/signal + noise) of the seven reversals. Across all three tasks, a higher threshold indicated a greater amount of information required to detect differences in the two stimuli.

Motion coherence task

This task established a threshold for the ability to detect coherent motion. Both panels contained randomly positioned white dots. Dots moved with translational motion and were either signal elements that moved coherently (in the same direction) or random noise. The participant had to select the panel that contained the dots that "moved the same way".

Form-from-motion task

This task establishes a threshold for the ability to use motion cues to detect form. In one panel a rectangle was positioned vertically and in the other it was positioned horizontally; the location of the rectangles within the panels was assigned randomly. The participant was shown an example of the target shape and asked "Where is the shape?".

Biological motion task

This task establishes a threshold for the ability to detect biological motion. One display panel depicted a centrally positioned walker. The other panel presented a spatially identical but temporally scrambled version of the walker point light display, with the trajectories of the dots played temporally out of phase with each other (e.g. instead of the dots representing a foot and knee moving forward together, they now might move in the opposite direction). The participant had to point to the panel that contained the "man walking".

Supplementary Materials 2

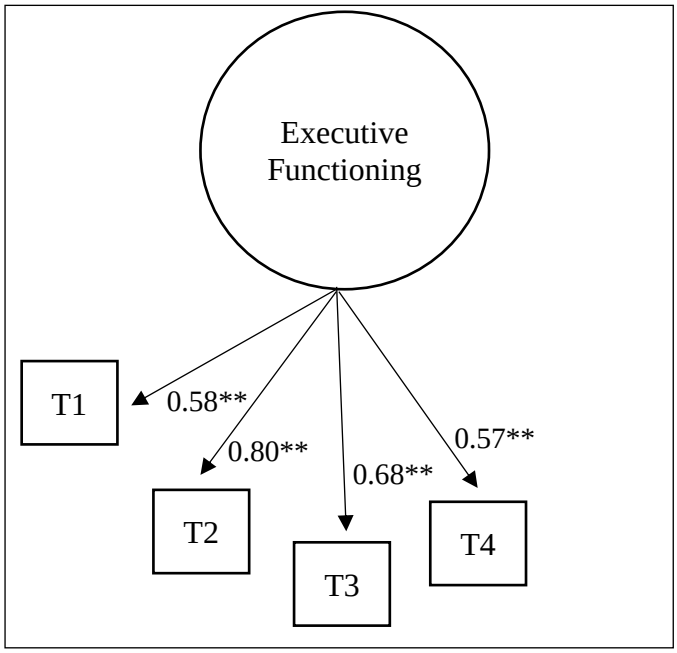

Figure 1. Structural Model for Executive Functioning Latent Variable

T1: Opposite Worlds, T2: Trail Making, T3: Score!, T4: Card Sort. **p<0.01 (relative χ^2=0.12, RMSEA=0.00, CFI=1.00, TLI=1.00).

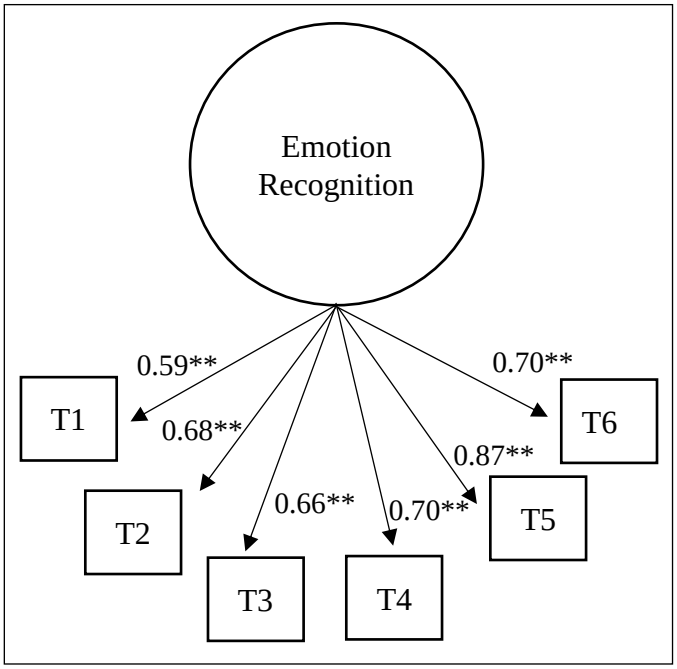

Figure 2. Structural Model for Emotion Recognition Latent Variable

T1: Happiness recognition, T2: Sadness recognition, T3: Fear recognition, T4: Anger recognition, T5: Surprise recognition, T6: Disgust recognition. **p<0.01 (relative χ^2=2.17, RMSEA=0.11, CFI=0.97, TLI=0.95).

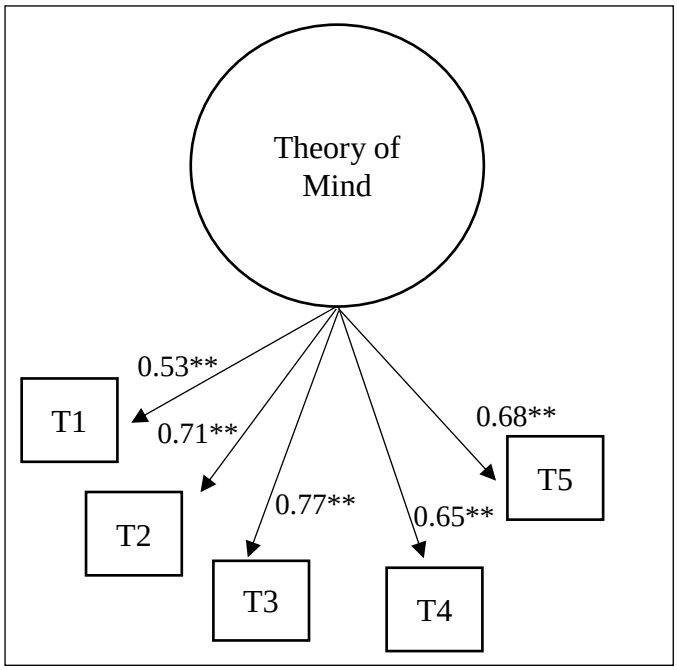

Figure 3. Structural Model for Theory of Mind Latent Variable

T1: Penny Hiding, T2: Strange Stories, T3: Frith–Happé Animations, T4: False Belief, T5: Reading the Mind in the Eyes. $^{**}p<0.01$ (relative $\chi^2=1.75$, RMSEA=0.09, CFI=0.96, TLI=0.92).

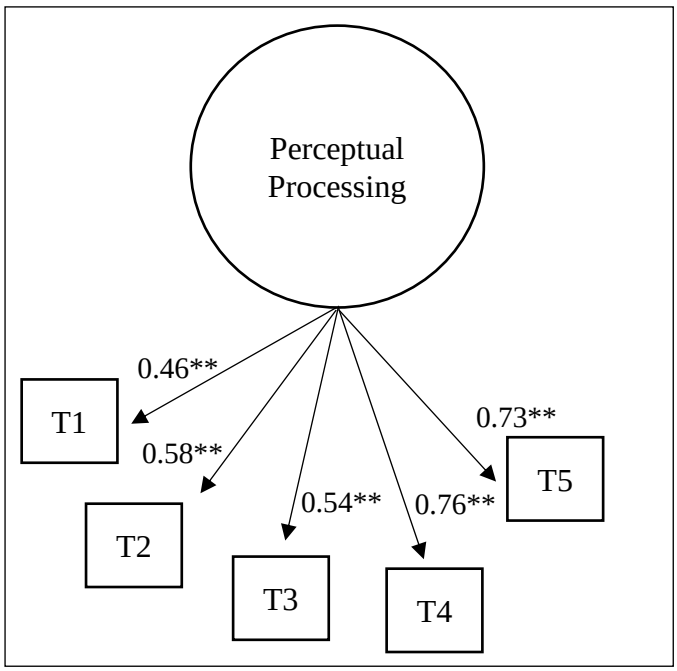

Figure 4. Structural Model for Perceptual Processing Latent Variable

T1: Audio intensity discrimination, T2: Audio duration discrimination, T3: Visual form discrimination, T4: Visual motion discrimination, T5: Visual biological motion discrimination. $^{**}p<0.01$ (relative $\chi^2=0.33$, RMSEA=0.00, CFI=1.00, TLI=1.00).

Supplementary Materials 3

Additional Analysis Using Binary Form of SIB Variable

The analyses were re-run using the binary (as opposed to ordinal) SIB variable. 48% of the sample (48/100) did not report any SIB, leaving 46% reporting some form of SIB, and six participants having missing data for all the SIB items that made up the summed score. A comparable pattern of model fit was found (relative χ^2=1.28, RMSEA=0.05, CFI=0.93, TLI=0.92) and the final pathways matched the final model obtained in Step 1, with a significant association between ToM and SIB (β=0.35, p<0.05) and between PP and externalising behaviours (β=0.29, p<0.01), and significant correlations between SIB and externalising behaviours (r=0.33, p<0.01), and between ToM and PP (r=0.74, p<0.01).

References

Abell F, Happé F and Frith U. (2000) Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development* 15: 1-16.

Baron-Cohen S, Wheelwright S, Spong A, et al. (2001) Are intuitive physics and intuitive psychology independent? A test with children with Asperger Syndrome. *Journal of Developmental and Learning Disorders* 5: 47-78.

Baron-Cohen S. (1992) Out of sight or out of mind? Another look at deception in autism. *Journal of Child Psychology and Psychiatry* 33: 1141-1155.

Happé F. (1994) An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders* 24: 129-154.

Hughes C, Adlam A, Happé F, et al. (2000) Good test-retest reliability for standard and advanced false-belief tasks across a wide range of abilities. *Journal of Child Psychology and Psychiatry* 41: 483-490.

Manly T, Anderson V, Nimmo-Smith I, et al. (2001) The differential assessment of children's attention: The Test of Everyday Attention for Children (TEA-Ch), normative sample and ADHD performance. *The Journal of Child Psychology and Psychiatry* 42: 1065-1081.

Reitan R and Wolfson D. (1985) *The Halstead-Reitan Neuropsychological Test Battery: Theory and Clinical Interpretation.*, Tucson, AZ: Neuropsychology Press.

Sullivan K, Zaitchik D and Tager-Flusberg H. (1994) Preschoolers can attribute second-order beliefs. *Developmental Psychology* 30: 395-402.

Tregay J, Gilmour J and Charman T. (2009) Childhood rituals and executive functions. *British Journal of Developmental Psychology* 27: 283-296.