Citation for published version (APA):

**The association between theory of mind, executive function and the symptoms of autism spectrum disorder**

Catherine R.G. Jones¹, Emily Simonoff², Gillian Baird³, Andrew Pickles⁴, Anita J.S. Marsden⁵ Jenifer Tregay⁶ Francesca Happé⁷, Tony Charman⁸

¹School of Psychology, Cardiff University, Cardiff, UK
²Department of Child & Adolescent Psychiatry and NIHR Biomedical Research Centre for Mental Health, King’s College London, Institute of Psychiatry Psychology and Neuroscience, London, UK
³Guy’s & St Thomas’ NHS Foundation Trust, London, UK
⁴Biostatistics Department and Biomedical Research Centre for Mental Health, King’s College London, Institute of Psychiatry Psychology and Neuroscience, London, UK
⁵Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK
⁶Oxford University Hospitals NHS Foundation Trust, Oxford, UK
⁷SGDP Research Centre King’s College London, Institute of Psychiatry Psychology and Neuroscience, London, UK
⁸Department of Psychology, King’s College London, Institute of Psychiatry Psychology and Neuroscience, London, UK

Address for correspondence and reprints: Catherine R.G. Jones, School of Psychology, Cardiff University, Park Place, Cardiff, CF10 3AT, UK. Email: jonescr10@cardiff.ac.uk
Grant sponsor Medical Research Council; Grant number: G0400065
Scientific abstract

It has been strongly argued that atypical cognitive processes in autism spectrum disorder (ASD) contribute to the expression of behavioural symptoms. Comprehensive investigation of these claims has been limited by small and unrepresentative sample sizes and the absence of wide-ranging task batteries. The current study investigated the cognitive abilities of 100 adolescents with ASD (mean age = 15 years 6 months), using 10 tasks to measure the domains of theory of mind (ToM) and executive function (EF). We used structural equation modelling as a statistically robust way of exploring the associations between cognition and parent-reported measures of social communication and restricted and repetitive behaviours (RRBs). We found that ToM ability was associated with both social communication symptoms and RRBs. EF was a correlate of ToM but had no direct association with parent-reported symptom expression. Our data suggest that in adolescence ToM ability, but not EF, is directly related to autistic symptom expression.

Lay Summary

The behaviours that are common to autism spectrum disorder (ASD) have been linked to differences in thinking ability. We assessed autistic adolescents and found that social communication difficulties and the presence of restricted and repetitive behaviours related to difficulties in understanding other peoples’ minds (theory of mind). In contrast, these behaviours were not associated with the general thinking abilities involved in planning and executing tasks (executive function).

Keywords: adolescents, executive functioning, restricted/repetitive behaviours, social cognition & theory of mind
Introduction

The causes of the behavioural symptoms of autism spectrum disorder (ASD), which include social communication difficulties and restricted and repetitive behaviours (RRBs) (American Psychiatric Association 2013), remain elusive. Recognition that these behavioural symptoms are continuously distributed in the general population and can occur in isolation has meant that focus is on identifying multiple causes, which likely vary across the autistic population (e.g. Constantino 2011; Happé et al. 2006). An important consideration in unpicking the complex constellation of biological and environmental influences on autistic symptoms is the role of cognition. It has been proposed that multiple cognitive difficulties are relevant to ASD and that they may have distinct interactions with different behavioural symptoms (Happé and Ronald 2008; Happé et al. 2006). Characterising cognitive-behaviour associations in ASD is an important step towards enhanced understanding and better targeted interventions.

The two cognitive accounts of ASD that have received most attention are theory of mind (ToM; e.g. Frith et al. 1991) and executive function (EF; e.g. Pennington et al. 1997; Russell 1997). ToM, or mentalising, is the ability to infer the mental states of other people and to use this information to predict behaviour. Difficulties in passing classic measures of ToM, the most common of which are false belief tasks, are well documented in ASD (see Frith 2012). False belief tasks ask direct questions about the mental states of others, although they also require competency in a range of other cognitive skills, such as inhibiting reality and response selection (e.g. Baillargeon et al. 2010). The ToM account, at least as a complete explanation, has been challenged by evidence that some individuals with ASD can pass false belief tasks (see Boucher 2012) as well as more advanced ToM assessments (e.g. Scheeren et al. 2013). However, verbal ability is a consistent correlate of ToM (Ronald et al. 2006; Scheeren et al. 2013) and one interpretation is that verbally able autistic individuals ‘hack out’ mentalising explanations when given time and structure but this ability does not
withstand the complexities of everyday life, where mentalising has to be intuitive, fast and reflexive (Happé 1995). Indeed, real life mentalising requires attention to relevant social detail, which may not be forthcoming in ASD (Chevallier et al. 2012). More recent studies measuring implicit mentalising (e.g. spontaneous looking patterns that reflect intuitive tracking of another person’s belief state) are clear in documenting difficulties in adults with ASD, despite competency on classic explicit mentalising tasks (e.g. direct questioning about another person’s belief state) (e.g. Senju et al. 2009). Some tests of mentalising require emotion recognition to infer mental states (e.g. Baron-Cohen et al. 2001), and this can also be impaired in ASD.

Intuitively, an association would be expected between mentalising difficulties and a range of atypical social communicative behaviours characteristic of ASD. Indeed, correlations between ToM and social communication have been identified (e.g. Ames and White 2011; Joseph and Tager-Flusberg 2004; Lerner et al. 2011; Nagar Shimoni et al. 2012), although not consistently (e.g. Cantio et al. 2016; Pellicano et al. 2006; Scheeren et al. 2013; Travis et al. 2001; Wilson et al. 2014). In contrast, the ToM account is less able to explain RRBs (Brunsdon and Happé 2014) and this relationship is less studied. However, valid theoretical links between RRBs and ToM can be drawn. For example, RRBs could conceivably alleviate anxiety that is elicited in challenging social situations. Further, limited insight into how the self is being perceived could serve to facilitate the development and maintenance of RRBs. Previous research has largely reported no significant associations between RRBs and ToM (Cantio et al. 2016; Pellicano et al. 2006; White et al. 2009; Wilson et al. 2014). However, Joseph and Tager-Flusberg (2004) reported a correlation between RRBs and ToM that was significant when nonverbal mental age, although not language ability, was controlled. In addition, Nagar Shimoni et al. (2012) found a correlation between
mentalising ability and the observed presence of stereotypic and limited interests, although there were no significant correlations with parent reported RRBs.

Whereas the ToM hypothesis is specific to social cognition, the EF hypothesis proposes domain-general cognitive difficulties in ASD (Pennington et al. 1997; Russell 1997). EFs encompass a range of interacting cognitive processes, subserved by the frontal lobes, which are particularly relevant to successful engagement in complex, novel and goal-oriented behaviours. Evidence of difficulties across the breadth of EFs, including planning, inhibition, cognitive flexibility, generativity and working memory, have all been reported in ASD (see Hill 2004). It has been proposed that EFs, particularly monitoring actions and acting with volition, are prerequisites for self-awareness and therefore for mentalising (Russell 1997). Impaired EFs are subsequently hypothesised to limit the ability of individuals with ASD to reflect on own and others’ mental states (Pennington et al. 1997; Russell 1997). This position has been bolstered by evidence that early EF ability predicts later proficiency in ToM in children with ASD (Pellicano 2010). A more prosaic explanation is that adequate EFs are necessary for coping with theory of mind tests, which are inherently cognitively demanding and require inhibition of reality/true beliefs (see Moses 2001). However, this account fares less well when considering difficulties with implicit mentalising, where executive demands are limited. Another interpretation is that EFs facilitate social interaction, thereby supporting the development of mentalising ability through exposure to relevant social exchange (e.g. Hughes 1998). Regardless, accounts of EF and social communication in ASD generally conceive of executive difficulties as having a cascading impact on mentalising ability, which is a more specific indicator of ASD. Evidence of significant association between poor EF and impairment in social communication exists (e.g. Dichter et al. 2009; Kenworthy et al. 2009; McEvoy et al. 1993) but null findings are more common (e.g. Cantio

Impairment in EFs has also been proposed to be functionally associated with RRBs. For example, difficulty in generating new ideas could lead to rigid routines and difficulty in managing when routine is disrupted (Turner 1997). This hypothesis has been supported by evidence of correlations between EF and RRBs in children with ASD (e.g. D'Cruz et al. 2013; Kenworthy et al. 2009; Lopez et al. 2005; Miller et al. 2015; Mosconi et al. 2009; Mostert-Kerckhoffs et al. 2015; Reed et al. 2013; South et al. 2007; Turner 1997; Yerys et al. 2009), although the finding is not universal (Cantio et al. 2016; Dichter et al. 2009; Faja and Dawson 2014; Joseph and Tager-Flusberg 2004; Liss et al. 2001; Pellicano et al. 2006; Wilson et al. 2014).

A cluster of studies have explored cognition in ASD across domains and using multiple measures (Brunsdon et al. 2015; Cantio et al. 2016; Kimhi et al. 2014; Lai et al. 2012; Lam 2013; Losh et al. 2009; Narzisi et al. 2013; Pellicano et al. 2006; Wilson et al. 2014; Yang et al. 2009). Direct attempts to correlate symptom severity with performance in specific cognitive domains were only examined in three of the studies and did not yield significant associations (Cantio et al. 2016; Pellicano et al. 2006; Wilson et al. 2014).

Thus, although there are compelling theoretical accounts of close association between cognition and behaviour in ASD, indeed these theories and suppositions drive the fertile investigation into the cognitive profile in ASD, the evidence base is surprisingly limited and contradictory. Here we report data from 100 adolescents (aged 14-16 years) with ASD who completed a battery of ToM and EF measures. Measurement of social communication and RRB symptoms was obtained via parent report. Critically, our sample spanned the range of testable participants (full scale IQ range 50-119), enabling a representative sample that was not confined to intellectually able participants. The sample size and multiple assessments
allowed us to explore the theorized associations between cognition and behaviour using structural equation modelling (SEM). Given the limitations and varied findings of previous research, our approach was agnostic, with our initial model predicting that both cognitive factors would be associated with both behavioural factors. The identification of the cognitive impairments that are associated with core autistic behaviours will provide a test of the ‘real life validity’ of the putative cognitive phenotypes of ASD.

**Method**

**Participants**

One hundred adolescents (91 male) with a consensus clinical ICD-10 (World Health Organisation, 1993) diagnosis of ASD were tested. Participants were from the Special Needs and Autism Project (SNAP; see Baird et al. 2006; Charman et al. 2011) and were on average 15 years 6 months (SD= 6 months; range 14 years 8 months – 16 years 9 months) at the time of testing. Participants had to have been able to successfully engage with the Wechsler Intelligence Scale for Children (WISC-III UK; Wechsler 1992) during the previous phase of SNAP, when they were between 9 and 10 years old. Further details on diagnostic procedure and sample characteristics can be found in Charman et al. (2011). The study was approved by the South East Research Ethics Committee (05/MRE01/67).

**Materials and procedure**

Participation involved completing a large battery of tasks, only some of which are reported here (see Charman et al. 2011). 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 third day of testing to complete the battery.

**Cognitive tasks**
Tasks are summarised in Table 1, and comprehensively described in the online supplementary materials. Task selection aimed at measuring ToM and EFs as broadly as possible, enabling a wide source of variance. ToM measures included: (i) False belief, as the gold standard measure of mentalising, (ii) Strange stories, as a general measure of mental state understanding, requiring understanding of the intent of one character to manipulate the mental state of another, (iii) Frith-Happé animations, as a general measure of mental state understanding, requiring attribution of intentions based only on movement patterns, (iv) Reading the mind in the eyes task – Children’s version (RMET), as a perception based measure of recognising psychological states (including emotions). EF tasks included: (i) Opposite worlds, as a measure of inhibition of a verbal response, (ii) Card sort, as a measure of cognitive set-shifting/flexibility, (iii) Category fluency, as a verbal measure of generativity, (iv) Design fluency, as a non-verbal measure of generativity, (v) Backwards digit span, as a measure of working memory, (vi) Planning drawing, as a measure of visuo-spatial planning.

Insert Table 1 about here

IQ and language

Verbal, performance and full-scale IQ was measured using the Wechsler Abbreviated Scale of Intelligence-UK (WASI; Wechsler 1999). A measure of language was obtained using the electronic version of the Test for the Reception of Grammar (TROG-E; Bishop 2005). This task assesses receptive grammar by requiring participants to choose pictures that correspond to sentences of increasing grammatical complexity. Both tasks used standard scores.

Parent-report measures of ASD symptoms
Social communication symptoms were measured using the social awareness, social cognition, social communication and social motivation raw subscores, from the Social Responsiveness Scale (SRS; e.g. Constantino and Gruber 2005). The SRS rates behaviours from 1 (not true) to 4 (almost always true), with a mean calculated for each subscale. RRBs were measured with the Repetitive Behavior Scale-Revised (RBS-R; Bodfish et al. 2000), using five empirically derived behaviour subscales: stereotypy, self-injurious, compulsive, ritualistic/sameness and restricted (Lam and Aman 2007). The RBS-R rates behaviours on a 0 (behaviour does not occur) – 3 (behaviour occurs and is a severe problem) scale, with the mean score for each subscale calculated.

**Analysis**

Data preparation and descriptive and correlational analyses were carried out in Stata 12 (StataCorp 2011) and SPSS 20.0 (IBMCorp 2011), while SEM was conducted in MPlus 7 (Muthén and Muthén 1998-2012). Variables were assessed for skewness and Box-Cox transformed, where appropriate. Transformed variables were Opposite Worlds, Card sort, and the RBS-R subscales stereotypy, self-injurious, and compulsive behaviours. All cognitive variables for SEM were treated so that a higher score indicated worse performance, this meant the ToM variables and the category fluency, design fluency, digit span, and planning drawing variables were all reverse scored. Collection of the complete dataset was not possible for a variety of reasons including time restrictions, participant engagement and ability, and parent availability for questionnaire completion. We dealt with missing data by using multiple imputation (see Schafer 1999) within MPlus and imputed 50 datasets for each analysis. Multiple imputation that handles missing data by creating replicates of an original dataset and replacing the missing data in each with imputed values. Analysis is then carried out on each dataset and averaged to create a single output (see Sterne et al. 2009). Descriptive and correlational data presented are based on the true dataset.
SEM enabled theoretical models of the interrelationships between multiple measures to be tested and compared. The structural component of SEM assesses the relations between latent variables and it is therefore essential that these latent variables are psychometrically sound (see Byrne 2011). Preliminary confirmatory factor analyses (CFA) were used to create two measurement models, which established the latent variables of ToM, EF (cognitive CFA), and social communication and RRBs (behavioural CFA) (Step 1). The second phase (Step 2-5) used an incremental approach to explore the structural relationships between cognition and behaviour by imposing a regression structure on the confirmed latent variables. In Step 2 we focused on the direct and basic regressions between ToM and behaviour and EF and behaviour in two separate models. Step 3 progressed to analysing the relationship between cognition and behaviour in a combined model. This meant cognition-behaviour associations were explored in a context in which both types of cognition were controlled. Step 4 repeated the structure of the combined model but additionally regressed each latent variable onto a measure of receptive language (TROG-E). This meant that patterns of association could be explored in a context that controlled for the effects of receptive language ability on task performance. Step 5 replicated Step 4 but regressed each latent variable onto a measure of IQ rather than language. Classic model generating frameworks (Jöreskog 1993) systematically drop non-significant paths, starting with the most non-significant, to identify the most parsimonious model. However, we considered it important to maintain paths that could potentially confound cognitive-behaviour associations, even if non-significant. The model estimator was maximum likelihood. Model fit was assessed using the comparative fit index (CFI), and root mean square error of approximation (RMSEA). A CFI of $\geq .95$ and a RMSEA of $\leq .08$ were considered suggestive of a reasonable fitting model (see Byrne 2011). In cases where model fit did not improve, the removal of consecutive paths was assessed using chi-square; a significant worsening of model fit was indicated by a drop of $\geq 3.84$. 
Results

Descriptive statistics are illustrated in Table 2. Participant performance was heterogeneous and used the full range of the scales. Correlations between tasks and behaviours are shown in Tables 3-4 at the item level, with correlations between individual cognitive tasks presented in the supplementary materials (Supplementary Table 1). For all correlations, if the data were Box-Cox transformed for the modelling then this transformation was used.

Insert Table 2 about here

SEM

Step 1: Initial CFA measurement models

The cognitive CFA model fit was only moderate ($\chi^2(34)=67.02$, p<.001; CFI=.861; RMSEA=.099) but with all variables significantly loading onto their latent factor (all p<.001). The correlation between factors was high (r=.87 (p<.001)). The behavioural CFA model showed good model fit ($\chi^2(26)=40.01$, p=.04; CFI=.966; RMSEA=.076), all variables significantly loaded onto their latent factor (all p<.01), and the correlation between factors was high (.77, p<.001).

Step 2: Separate SEM of ToM and behavioural symptoms and EF and behavioural symptoms

The SEM of ToM and the behavioural latent factors (see Figure 1a) showed good model fit ($\chi^2(62)=80.51$, p=.06; CFI=.963; RMSEA=.055). Paths between social communication and ToM (β=.43, p<.001) and RRB and ToM (β=.40, p=.001) were significant. The SEM of EF and the behavioural latent factors (see Figure 1b) also showed good model fit ($\chi^2(87)=105.38$, p=.09; CFI=.964; RMSEA = .046). Paths between social communication and EF (β=.26, p=.03) and RRB and EF (β=.29, p=.02) were significant.

Insert Figure 1 about here
Step 3: Combined SEM of ToM, EF and behavioural symptoms

The initial SEM combining paths between both cognitive latent factors and behavioural symptoms showed reasonable model fit ($\chi^2(146)=213.86$, $p<.001$; CFI=.902; RMSEA=.068). Paths from ToM to social communication ($\beta=.60$; $p=.12$) and RRBs ($\beta=.52$; $p=.20$), and from EF to social communication ($\beta=-.22$; $p=.57$) and RRBs ($\beta=-.16$; $p=.69$) were not significant. The correlations between cognitive ($r=.86$, $p<.001$) and behavioural ($r=.73$, $p<.001$) latent factors were both highly significant.

Non-significant paths were systematically removed, starting with the regression of RRBs on EF, which increased model fit ($\chi^2(147)=214.07$, $p<.001$; CFI=.903; RMSEA=.068). Subsequent removal of the regression of social communication on EF also improved fit ($\chi^2(148)=214.38$, $p<.001$; CFI=.904; RMSEA=.067) with all remaining paths significant (see Figure 2). Thus, the best fitting model indicates a direct association between ToM, but not EF, and behavioural symptoms.

Insert Figure 2 about here

Step 4: Combined SEM of ToM, EF and behavioural symptoms, controlling for receptive language

The initial SEM including all paths between cognitive and behavioural latent factors as well as regressing all factors onto the TROG-E showed reasonable model fit ($\chi^2(161)=232.12$, $p<.001$; CFI=.908; RMSEA=.066). Initial paths from ToM to social communication ($\beta=.66$; $p=.08$) and RRBs ($\beta=.57$; $p=.15$), and from EF to social communication ($\beta=-.10$; $p=.80$) and RRBs ($\beta=-.05$; $p=.90$) were not significant. Additionally, the paths between TROG-E and social communication ($\beta=.20$; $p=.48$) and TROG-E and RRBs ($\beta=.18$; $p=.40$) were also non-significant. These latter pathways represent associations that may confound the primary relationships of interest between cognition and behaviour. Therefore, they were maintained in
the model to control for these effects. The first pathway removed was between RRB and EF, which increased model fit ($\chi^2(162)=232.22, p<.001$; CFI=.910; RMSEA=.066). Subsequent removal of the path between social communication and EF improved model fit ($\chi^2(163)=232.37, p<.001$; CFI=.911; RMSEA=.065) and produced a final model solution with all paths significant apart from those between TROG-E and social communication and TROG-E and RRBs (see Figure 3). In summary, the model indicates an association between ToM and social communication and RRBs when controlling for receptive language ability.

Insert Figure 3 about here

**Step 5: Combined SEM of ToM, EF and behavioural symptoms, controlling for IQ**

We were additionally interested in exploring the pattern of cognition-behaviour associations while controlling for full-scale IQ. The paths between EF and IQ ($\beta=-.86; p<.001$) and ToM and IQ ($\beta=-.93; p<.001$) were extremely high in the initial model ($\chi^2(161)=251.62, p<.001$; CFI=.888; RMSEA=.075). The strong effect of full-scale IQ on cognitive-behaviour associations can also be seen in Tables 3 and 4. While maintaining all paths with IQ, systematically removing the non-significant paths between RRB and EF ($\chi^2(162)=251.43, p<.001$), social communication and EF ($\chi^2(163)=250.68, p<.001$), RRB and ToM ($\chi^2(164)=251.82, p<.001$) and social communication and ToM ($\chi^2(165)=252.84, p<.001$; CFI =.891; RMSEA =.073) indicated a model that did not improve in fit incrementally, albeit with no decreases in model fit of statistical significance ($\chi^2$ difference $\geq3.84$). Exploratory analysis with performance IQ and verbal IQ found a similar pattern for performance IQ, while there were issues with model convergence for verbal IQ. In summary, controlling for full-scale IQ produced an unstable model fitting process and no cognition-behaviour paths were significant.
**Discussion**

ToM and EFs are cognitive domains argued to be central to the behavioural presentation of ASD. However, thorough exploration of cognitive-behaviour associations using multiple measures has been surprisingly limited. We addressed this by investigating ToM and EF capabilities in 100 adolescents with ASD, alongside parent report measures of ASD symptomatology. Using SEM, which meant we could account for both ToM and EF in one model, we established that mentalising difficulties were associated with more severe social communication symptoms and RRBs. In contrast, the model did not support a direct relationship between EF and behavioural symptoms.

It is important to consider the auxiliary demands inherent to EF and ToM tasks that may confound results, although previous studies have tended not to control for non-specific task demands or general ability. However, we replicated our findings in a model that also controlled for receptive language, thus accounting for varying ability in understanding task demands. In contrast, our attempts to control for full-scale IQ indicated that it was too highly correlated with the cognitive tasks to provide sensitivity for investigating cognitive-behaviour associations. This may reflect recognised issues with the generalised nature of IQ assessments when attempting to parcel out specific cognitive or perceptual confounds (see Dennis et al. 2009).

*ToM and ASD symptoms*

We found that difficulties with ToM were related to the degree of autistic symptoms. Notably, the strength of the standardised coefficients between social communication and ToM and between RRBs and ToM were similar. The current study uses cross-sectional data and was not designed to examine bidirectional effects between cognition and behaviour. Therefore, although the theories under discussion describe the impact of cognition on behaviour, the current results cannot directly speak to causality. However, the significant
association between ToM and social communication symptoms fits with the argument that impairments in understanding other minds might underlie complex and varied impairments in social interaction (e.g. Frith et al. 1991). The majority of studies that have found no significant association have limited their measurement of ToM to false belief. Our broad approach included more challenging tasks that are better able to measure individual differences, as well as the RMET (Baron-Cohen et al. 2001), which engages socio-perceptual processing (although see Oakley et al. 2016 for a critique of this task). Our data suggest that the association between ToM and autistic behaviours may be best identified by using an inclusive approach that goes beyond narrow measures of false belief.

The significant association between ToM and RRBs is perhaps surprising, although Ronald et al (2006) found a significant relationship between parent-reported RRBs and ToM ability in a community sample of 9-year-old twin pairs, which persisted when verbal ability was controlled. However, the most common finding across the small sample of studies that have directly assessed this association in ASD is that ToM does not correlate with RRBs (Cantio et al. 2016; Joseph and Tager-Flusberg 2004; Nagar Shimoni et al. 2012; Pellicano et al. 2006; White et al. 2009; Wilson et al. 2014). With the exception of Cantio et al. (2016), these studies measured RRBs using clinical assessment through observation or parent-interview. Both the ADOS and ADI-R have been criticised for undersampling RRBs (e.g. Esbensen et al. 2009) and the observational format of the ADOS favours certain RRBs (e.g. motor stereotypies) over others (e.g. restricted interests). In contrast, we used a targeted questionnaire designed to gather information about the breadth of RRBs observed in ASD.

Despite the advantages of our measure, the reliance on parent-report measures of behaviour is a limitation. For example, parent-report of child behaviour can be influenced by parent depression (e.g. Randazzo et al. 2003). An optimal study design would accommodate
multiple sources of behaviour measurement, including direct observation and a variety of informants (e.g. parent, teacher, self).

When considering the theoretical link between ToM and RRBs, a bewildering social world due to impoverished mentalising abilities could lead to RRBs that lessen anxiety and reduce confusion. RRBs have been associated with anxiety in ASD and interpreted as forming a ‘buffer’ to alleviate anxiety (Lidstone et al. 2014). An interplay has also been observed between RRBs, anxiety and intolerance of uncertainty in ASD (Joyce et al. 2017; Wigham et al. 2015). Arguably those with poor mentalising could be more prone to experience social events as uncertain and unpredictable, leading to elevated anxiety.

It is also possible that a limited understanding of or interest in how the self is perceived could reduce motivation to suppress or modify RRBs. Related to this, a recent meta-analysis has demonstrated the regions of the superior temporal gyrus and medial prefrontal cortex are involved in both classic ToM and self-awareness (van Veluw and Chance 2014), while theoretical links have been drawn between ToM and social motivation (Chevallier et al. 2012). Within the potentially complex relationship between social understanding, social motivation and RRBs, difficulties in engaging with and understanding the social world could lead to the development of idiosyncratic and unusually intense interests, and certainly to situations where RRBs ‘win out’ over more conventional and social pursuits. Indeed, young autistic people have described how their RRBs are used as a way to be alone and to avoid people (Joyce et al. 2017).

There may also be value in taking a developmental perspective, which recognises that RRBs are an intrinsic part of typical development (Evans et al. 1997). A mentalising deficit, including precursors such as atypical joint attention (Charman et al. 2000), could disrupt the experience-dependent brain and behaviour development that leads to the typical trajectory of
RRBs. This reflects evidence from animal models in which restricted environments produce elevated repetitive behaviours (Lewis and Kim 2009).

In summary, the current study suggests that the theoretical position that ToM is not relevant to RRBs needs reconsidering. We suggest that there may be multiple ways in which ToM might associate with RRBs and that taking a fine-grained approach, which enables investigation of RRB subtypes, may prove illuminating.

EF and ASD symptoms

Our simple model of EF and behaviour showed a significant association between EF and both RRBs and social communication. However, our aim was to model EF and ToM simultaneously, thus controlling for the association between these variables as well as the effects of concurrent cognitive-behaviour associations. Using this technique, which is novel to the field, we failed to find a direct association between EFs and autistic behaviours. Therefore, we have concluded that EFs have no unique association with autistic behaviours but have an indirect effect through their association with ToM. This strong association between EF and ToM is an established finding (e.g. Pellicano 2007).

We measured EF as a composite of a variety of executive skills, which aligned with our parsimonious approach and avoided issues with identifying distinct EFs, which are rarely isolated in any one executive task (Van Eylen et al. 2015). However, hypotheses have been made about specific executive abilities and specific RRBs, for example, poor generativity restricting the range of behaviours (Turner 1997), and there is some evidence that different EFs have differential association with RRBs (e.g. Kenworthy et al. 2009; Lopez et al. 2005). Further, RRBs are often considered as two distinct subtypes, repetitive sensory and motor behaviours and insistence on sameness (see Barrett et al. 2015), with some evidence of distinct relationships with EF (Mosconi et al. 2009). A limitation of our parsimonious
approach is that it did not allow for this type of nuanced investigation, which could prove informative in future work. It is also worth considering Brunsdon et al.’s (2015) finding that the number of cognitive tasks on which participants performed atypically correlated positively with autistic symptoms. Further research that explores the cumulative effect of difficulties with both ToM and EF could be beneficial.

Although our data do not suggest a direct association between EFs and behaviour, the Triple I hypothesis (White 2013) argues EF deficits are driven by difficulties in ‘Inferring Implicit Information’. This theory suggests that difficulties relevant to mentalising drive the pattern of impairment across executive tasks. In support, children with ASD fare better on structured executive tasks compared to open-ended tasks, where the correct behaviour has to be implicitly inferred (Van Eylen et al. 2015). The Triple I hypothesis would not predict a strong association between EF and ASD symptoms but would predict our observed correlation between EF and ToM. Indeed, two of our EF tasks, planning drawing and card sort, had particularly high demands in terms of requiring inference, and open-endedness is inherent in generativity tasks.

However, our previous research, also using the SNAP sample, found evidence that poorer executive skills related to higher levels of anxiety (Hollocks et al. 2014), which has subsequently been replicated (Lawson et al. 2015; Wallace et al. 2016). It may be that the pathway between EFs and RRBs is indirect and mediated by anxiety. For example, poorer cognitive control could lead to hyperattentiveness to negative information and subsequent anxiety (see Hollocks et al., 2014); attempts to manage the anxiety could then lead to RRBs (see Spiker et al. 2012). As far as we are aware, this mediation hypothesis has yet to be tested.

EFs also distinguish themselves from ToM by being prevalent across other developmental disorders, particularly attention deficit hyperactivity disorder (ADHD) (see
Craig et al. 2016). Our population was drawn from SNAP, 28% of whom met criteria for co-morbid ADHD (Simonoff et al. 2008). There is evidence of shared genetic overlap between autistic and ADHD traits, which is particularly strong for RRBs (Polderman et al. 2013; Ronald et al. 2014; Taylor et al. 2015). Within this context, it could be argued that the strength of association between RRBs and EFs is being tempered by complex comorbidity. This is consistent with the research domains criteria (RDoC) initiative that proposes behaviour and cognition should be considered within a dimensional framework, unrestricted by diagnostic classification (Insel et al. 2010). Related to this, our population was confined to those with an ASD diagnosis, and therefore levels of RRBs and social communicative difficulties were high; whether the pattern of results would replicate in a population with a broader range of symptom presentation remains to be established.

An important consideration when interpreting the current findings is that ASD is a developmental disorder and the current study took a snapshot of cognitive-behaviour associations in adolescence. Pellicano (2013) found that executive skills of 4-7 year old children with ASD predicted both social communication abilities and RRBs three years later, with no predictive relationship established for ToM. Therefore, executive difficulties might become less directly relevant to autistic symptoms as development progresses. Both cognitive task performance and behavioural symptoms could be moderated by the development of compensatory strategies, be they externally imposed or internally generated (e.g. avoidance of triggers; use of communication or organisational tools; social rule learning), which would dilute the degree of association. Alongside the limitation of a cross-sectional design, we were also restricted to exploring ASD at the group level. Phenotypic heterogeneity is a recognised characteristic of the disorder (e.g. Georgiades et al. 2013) and our approach may have masked distinct subgroups.

**Concluding comments**
In a modelling approach that considered co-occurring cognitive and behaviour associations, we present evidence that ToM ability, but not EF ability, is directly associated with both social communication and RRBs in adolescents with ASD. This finding suggests that training in ToM may impact positively on autistic behaviours. A recent Cochrane review of ToM interventions in ASD concluded that although ToM can be taught, evidence of generalisation beyond task performance was limited (Fletcher-Watson et al. 2014). However, the authors called for more longitudinal research and improved outcome measures to better evaluate the effectiveness of ToM interventions.

The multiple deficit account would have predicted that ToM associated uniquely with social communication symptoms and that EFs associated uniquely with RRB symptoms (Brunsdon and Happé 2014). However, there are many reasons why a single cognitive deficit is an unlikely explanation, particularly the low correlation between core behavioural features in both the general (Ronald et al. 2006) and autistic (Dworzynski et al. 2009) population. These core behaviours also have relatively independent heritability (e.g. Ronald et al., 2006; see Happé & Ronald (2008) for a fuller discussion). Although further research with large sample sizes and multiple measures are required, the data presented here question whether a simple cognition-behaviour relationship for EFs and RRBs exists. Our findings also suggest that the relationship between ToM and RRBs needs to be re-examined both theoretically and experimentally.

Acknowledgements

This study was funded by the Medical Research Council (G0400065). We are grateful to the adolescents and families who took part in this study. With thanks to Rebecca Phillips for help with data collection and to Rhonda Booth for providing the illustrations for the combined 1\textsuperscript{st} and 2\textsuperscript{nd} order false belief story.
**Figures**

Figure 1: Separate structural equation models of ToM and behavioural symptoms and EF and behavioural symptoms (Step 2 of SEM analysis)

Animations = Frith-Happé animations; Plan draw = Planning drawing; EF = executive functions; RMET = Reading the mind in the eyes task; SRS = Social Responsiveness Scale; ToM = Theory of mind; RBS = Repetitive Behavior Scale-Revised; RRBs = Restricted and repetitive behaviours

Figure 2: Combined structural equation model of ToM, EF and behavioural symptoms (Step 3 of SEM analysis) Animations = Frith-Happé animations; Plan draw = Planning drawing; EF = executive functions; RMET = Reading the mind in the eyes task; SRS = Social Responsiveness Scale; ToM = Theory of mind; RBS = Repetitive Behavior Scale-Revised; RRBs = Restricted and repetitive behaviours

Figure 3: Combined SEM of ToM, EF and behavioural symptoms, controlling for receptive language (Step 4 of SEM analysis)

Dotted lines represent non-significant paths, including the regression of TROG-E on social communication behaviours ($\beta=.23$, $p = .31$) and the regression of TROG-E on RRBs ($\beta=.19$, $p = .42$). Animations = Frith-Happé animations; Plan draw = Planning drawing; EF = executive functions; RMET = Reading the mind in the eyes task; TROG-E = Test of Reception for Grammar; SRS = Social Responsiveness Scale; ToM = Theory of mind; RBS = Repetitive Behavior Scale-Revised; RRBs = Restricted and repetitive behaviours
Figure 1

Figure 1a

Figure 1b
Figure 2
Figure 3
Table 1: Summary of tasks used.

<table>
<thead>
<tr>
<th>Task</th>
<th>Key variable</th>
<th>Number of trials</th>
<th>Score range</th>
<th>Reference for task procedure*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Theory of Mind</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FB</td>
<td>1&lt;sup&gt;st&lt;/sup&gt; and 2&lt;sup&gt;nd&lt;/sup&gt; order FB score (sum)</td>
<td>2 stories. Three FB questions (1 1&lt;sup&gt;st&lt;/sup&gt; order; 2 2&lt;sup&gt;nd&lt;/sup&gt; order) and 3 justification questions.</td>
<td>0-8</td>
<td>Bowler (1992); Hughes et al. (2000); Sullivan et al. (1994)</td>
</tr>
<tr>
<td>Strange stories</td>
<td>Mentalising score (average)</td>
<td>4 theory of mind stories</td>
<td>0-2</td>
<td>Ricketts et al. (2013)</td>
</tr>
<tr>
<td>Animations</td>
<td>Mentalising (intentionality) score (average)</td>
<td>4 theory of mind animations</td>
<td>0-5</td>
<td>Jones et al. (2011)</td>
</tr>
<tr>
<td>RMET</td>
<td>Total correct</td>
<td>28</td>
<td>0-28</td>
<td>Baron-Cohen et al. (2001)</td>
</tr>
<tr>
<td><strong>Executive functions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opposite Worlds</td>
<td>Inhibition cost score</td>
<td>4</td>
<td>-</td>
<td>Manly et al. (2001)</td>
</tr>
<tr>
<td>Card sort</td>
<td>Number of errors</td>
<td>3</td>
<td>0-60</td>
<td>Tregay et al. (2009)</td>
</tr>
<tr>
<td>Category fluency</td>
<td>Number of correct responses</td>
<td>2</td>
<td>-</td>
<td>Tregay, Gilmour, and Charman (2009)</td>
</tr>
<tr>
<td>Design fluency</td>
<td>Number of correct responses</td>
<td>1</td>
<td>-</td>
<td>Jones-Gotman and Milner (1977)</td>
</tr>
<tr>
<td>Digit span (backwards)</td>
<td>Raw score Variable</td>
<td></td>
<td>0-14</td>
<td>Cohen (1997)</td>
</tr>
<tr>
<td>Planning drawing</td>
<td>Planning score (sum)</td>
<td>3</td>
<td>0-6</td>
<td>Booth et al. (2003)</td>
</tr>
</tbody>
</table>
Further details for each task provided in online supplementary materials. Animations = Frith-Happé animations; CMS = Children’s Memory Scale; FB = False belief; RMET = Reading the mind in the eyes task; TEA-Ch= Test of Everyday Attention for Children.
Table 2: Descriptive statistics. The listed tasks were the indicator variables for each of the four latent factors (Theory of Mind, Executive function, Social communication, Restricted and Repetitive Behaviours) used in the models.

<table>
<thead>
<tr>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal IQ</td>
<td>100</td>
<td>80.81</td>
<td>18.04</td>
</tr>
<tr>
<td>Performance IQ</td>
<td>100</td>
<td>90.37</td>
<td>18.61</td>
</tr>
<tr>
<td>Full scale IQ</td>
<td>100</td>
<td>84.31</td>
<td>18.03</td>
</tr>
<tr>
<td>TROG-E</td>
<td>98</td>
<td>82.89</td>
<td>17.20</td>
</tr>
</tbody>
</table>

**Cognition: Theory of Mind**

<table>
<thead>
<tr>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>False belief</td>
<td>99</td>
<td>4.75</td>
<td>2.42</td>
</tr>
<tr>
<td>Strange stories</td>
<td>88</td>
<td>.85</td>
<td>.52</td>
</tr>
<tr>
<td>Animations</td>
<td>87</td>
<td>2.87</td>
<td>.94</td>
</tr>
<tr>
<td>RMET</td>
<td>94</td>
<td>17.02</td>
<td>4.44</td>
</tr>
</tbody>
</table>

**Cognition: Executive function**

<table>
<thead>
<tr>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opposite worlds</td>
<td>98</td>
<td>8.37</td>
<td>7.49</td>
</tr>
<tr>
<td>Card sort</td>
<td>98</td>
<td>7.24</td>
<td>6.62</td>
</tr>
<tr>
<td>Category fluency</td>
<td>97</td>
<td>35.27</td>
<td>11.31</td>
</tr>
<tr>
<td>Design fluency</td>
<td>94</td>
<td>7.91</td>
<td>4.00</td>
</tr>
<tr>
<td>Digit span</td>
<td>99</td>
<td>4.66</td>
<td>2.46</td>
</tr>
<tr>
<td>Planning drawing</td>
<td>98</td>
<td>3.56</td>
<td>1.70</td>
</tr>
</tbody>
</table>

**Behaviour: Social communication**

<table>
<thead>
<tr>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRS Social awareness</td>
<td>92</td>
<td>11.83</td>
<td>4.19</td>
</tr>
<tr>
<td>SRS Social cognition</td>
<td>92</td>
<td>17.13</td>
<td>6.55</td>
</tr>
<tr>
<td>SRS Social communication</td>
<td>92</td>
<td>13.58</td>
<td>10.22</td>
</tr>
<tr>
<td>SRS Social motivation</td>
<td>92</td>
<td>15.51</td>
<td>5.67</td>
</tr>
</tbody>
</table>

**Behaviour: Restricted and Repetitive Behaviours**

<table>
<thead>
<tr>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBS-R Stereotypy</td>
<td>82</td>
<td>.42</td>
<td>.48</td>
</tr>
<tr>
<td>RBS-R Self injurious</td>
<td>86</td>
<td>.18</td>
<td>.28</td>
</tr>
<tr>
<td>RBS-R Compulsive</td>
<td>89</td>
<td>.39</td>
<td>.48</td>
</tr>
<tr>
<td>Test Description</td>
<td>Score</td>
<td>Alpha</td>
<td>Beta</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>RBS-R Ritualistic/Sameness</td>
<td>85</td>
<td>.51</td>
<td>.43</td>
</tr>
<tr>
<td>RBS-R Restricted</td>
<td>90</td>
<td>1.08</td>
<td>.86</td>
</tr>
</tbody>
</table>

Animations, Frith-Happé animations; RMET, Reading the mind in the eyes task; SRS, Social Responsiveness Scale, RBS, Repetitive Behavior Scale-Revised
Table 3: Correlations between cognitive tasks and the social communication and restricted and repetitive behaviour latent factors

<table>
<thead>
<tr>
<th></th>
<th>Social communication</th>
<th>RRB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Theory of mind</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False belief</td>
<td>.30** (-.05)</td>
<td>.32** (.01)</td>
</tr>
<tr>
<td>Strange stories</td>
<td>.20 (.01)</td>
<td>.18* (-.03)</td>
</tr>
<tr>
<td>F-H animations</td>
<td>.18 (.02)</td>
<td>.25* (.19)</td>
</tr>
<tr>
<td>RMET</td>
<td>.31** (.17)</td>
<td>.29** (.13)</td>
</tr>
<tr>
<td><strong>Executive functions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opposite worlds</td>
<td>.16 (.02)</td>
<td>.23* (.07)</td>
</tr>
<tr>
<td>Card sort</td>
<td>.23* (-.02)</td>
<td>.23* (-.02)</td>
</tr>
<tr>
<td>Category fluency</td>
<td>.18 (.12)</td>
<td>.22* (.17)</td>
</tr>
<tr>
<td>Design fluency</td>
<td>.04 (-.14)</td>
<td>.05 (-.13)</td>
</tr>
<tr>
<td>Digit span</td>
<td>.21* (.12)</td>
<td>.22* (.12)</td>
</tr>
<tr>
<td>Planning drawing</td>
<td>.14 (.05)</td>
<td>.16 (.06)</td>
</tr>
</tbody>
</table>

F-H animations, Frith-Happé animations; FSIQ, full-scale IQ; RMET, Reading the mind in the eyes task. RRB, restricted and repetitive behaviours *** p < .001, ** p < .01, * p < .05 Correlations in brackets are partialled for FSIQ.
### Table 4: Correlations between social communication and restricted and repetitive behaviour variables and the theory of mind and executive function latent factors

<table>
<thead>
<tr>
<th></th>
<th>Theory of Mind</th>
<th>Executive functions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Social communication</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRS Social awareness</td>
<td>.34** (.19)</td>
<td>.24* (.06)</td>
</tr>
<tr>
<td>SRS Social cognition</td>
<td>.41*** (.25*)</td>
<td>.31** (.11)</td>
</tr>
<tr>
<td>SRS Social communication</td>
<td>.35** (.22)</td>
<td>.20 (-.02)</td>
</tr>
<tr>
<td>SRS Social motivation</td>
<td>.28** (.07)</td>
<td>.20 (-.08)</td>
</tr>
<tr>
<td><strong>Restricted and repetitive behaviours</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBS-R Stereotypy</td>
<td>.49*** (.31**)</td>
<td>.39*** (.12)</td>
</tr>
<tr>
<td>RBS-R Self injurious</td>
<td>.23* (.14)</td>
<td>.19 (.07)</td>
</tr>
<tr>
<td>RBS-R Compulsive</td>
<td>.34** (-.01)</td>
<td>.21* (-.18)</td>
</tr>
<tr>
<td>RBS-R Ritualistic/Sameness</td>
<td>.28* (.19)</td>
<td>.20 (.09)</td>
</tr>
<tr>
<td>RBS-R Restricted</td>
<td>.29** (.22)</td>
<td>.21 (.09)</td>
</tr>
</tbody>
</table>

SRS, Social Responsiveness Scale, RBS, Repetitive Behaviour Scale-Revised *** p < .001, ** p < .01, * p < .05 Correlations in brackets are partialled for FSIQ.
References


Texas, The Psychological Corporation.


StataCorp (2011). Stata Statistical Software: Release 12. . College Station, TX StataCorp LP.


