Childhood Executive Function Predicts Later Autistic Features and Adaptive Behavior in Young Autistic People: a 12-Year Prospective Study

Lorcan Kenny 1 · Serena J. Cribb 2 · Elizabeth Pellicano 1,3

© The Author(s) 2018

Abstract

Longitudinal studies of autistic people show that the behavioral features of autism generally endure into adulthood. Yet the prognostic indicators remain far from certain, especially for cognitively able individuals. Here, we test the predictive power of specific cognitive skills, namely theory of mind and executive function, measured in childhood, on young people’s autistic features and adaptive behavior 12 years later. Twenty-eight young autistic people (2 female) were seen twice within the space of 12 years. At Time 1 (M = 5 years; 7 months, SD = 11 months), participants were assessed on components of executive function (planning, inhibition and cognitive flexibility) and theory of mind (false-belief understanding). At Time 2, 12 years later (M = 17 years 10 months, SD = 1 year; 2 months), we measured participants’ autistic features and adaptive behavior. Only Time 1 executive function skills predicted significant variance in autistic adolescents’ autistic features, over and above variance attributable to early age, intellectual ability and theory of mind skills. Furthermore, early EF skills, in addition to early verbal ability and nonverbal ability, predicted significant variance in young people’s adaptive behavior at the 12-year follow-up. These long-term longitudinal findings clearly demonstrate that executive function measured in early childhood has prognostic significance in a sample of young autistic people approaching emerging adulthood and underscore their importance as a key target for early intervention and support.

Keywords Autism · Theory of mind · Executive function · Development · Outcomes · Longitudinal

Autistic 1 children grow up to become autistic adults. Yet the predominant focus of research on childhood (Interagency Autism Coordinating Committee 2017; Pellicano et al. 2014) has meant that we know remarkably little about adult

---

1 In this article, we use ‘identify-first’ language (i.e. ‘autistic person’) rather than person-first language (i.e. ‘person with autism’), because it is the preferred term of autistic activists (e.g., Sinclair 1999) and many autistic people and their families (Kenny et al. 2016) and is less associated with stigma (Gernsbacher 2017).

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10802-018-0493-8) contains supplementary material, which is available to authorized users.

---

© Elizabeth Pellicano
liz.pellicano@mq.edu.au

Lorcan Kenny
l.kenny@ucl.ac.uk

Serena J. Cribb
serena.cribb@uwa.edu.au

1 Centre for Research in Autism and Education (CRAE), UCL Institute of Education, University College London, 55-59 Gordon Square, London WC1H 0NU, UK

2 School of Psychology, University of Western Australia, 35 Stirling Highway, Nedlands 6009, Australia

3 Department of Educational Studies, Macquarie University, 29 Wally’s Walk, Sydney, NSW 2109, Australia

Published online: 13 November 2018
autistic lives, and the factors that promote their life chances, especially during the transition to adulthood. The few studies that have followed autistic children into adulthood have highlighted the striking variability in their long-term outcomes, even among the most cognitively able. Although a minority achieve ‘very good’ (Howlin et al. 2004) or ‘very positive’ (Anderson et al. 2014; see also Fein et al. 2013) outcomes, autistic adults are far less likely than their non-autistic counterparts to have a job, live independently or to have extensive social networks (see Henninger and Taylor 2014, and Howlin and Magiati 2017, for review). Many also have difficulties with their physical, emotional and material wellbeing, which can require ongoing support (Billstedt et al. 2011; Bishop-Fitzpatrick et al. 2016; Renty and Roeyers 2006). Understanding the potential source(s) of this variability in long-term outcomes is critical to identifying where best to target intervention efforts.

Existing longitudinal studies have focused on identifying prognostic indicators in areas thought to be critical for achieving independence in adulthood – including intellectual functioning, autistic features and adaptive functioning (see Howlin and Magiati 2017, for review). These studies have repeatedly shown that childhood language ability, especially the acquisition of language before age 5 (Eaves and Ho 2008; Gotham et al. 2012; Lotter 1974; Magiati et al. 2014; Pickles et al. 2014; Szatmari et al. 2003; Venter et al. 1992) and general intellectual ability (Bal et al. 2015; Gillberg and Steffenburg 1987; Howlin et al. 2013) are significant predictors of autistic features and adaptive behavior in autistic people. This is not necessarily the case, however, for cognitively-able individuals, whose adult outcomes are notoriously variable and less than straightforward to predict on the basis of early language and intellectual functioning alone (e.g., Anderson et al. 2014; Howlin et al. 2004). In the current study, we sought to go beyond language ability and IQ to examine the predictive power of specific cognitive skills, namely childhood theory of mind and executive function, in a group of autistic participants considered to be cognitively-able in childhood followed over a 12-years period.

Researchers have long sought to identify specific aspects of cognition, which might be more proximal to the child’s behavior than general sources of individual and developmental differences (Frith et al. 1991) and therefore might be better placed to explain, at least in part, the variation in autistic individuals’ behavioral outcomes. One such candidate is atypicalities in ‘theory of mind’ (ToM) or difficulties in the ability to infer the mental states of others, which was once proposed to explain certain behavioral features of autism, especially difficulties with social interaction and communication (Baron-Cohen 2000; see Tager-Flusberg 2007, for review). An alternative, rival candidate is atypicalities in ‘executive function’ (EF) or difficulties with those skills necessary for flexible, goal-oriented behavior (including planning, cognitive flexibility, inhibition), which were linked in particular to the presence of restricted and repetitive behaviors in autism and problems managing everyday routines (Damasio and Maurer 1978; Ozonoff et al. 1991; Turner 1997).

Cross-sectional studies have demonstrated some evidence for the explanatory power of these cognitive skills in autistic people. For example, individual differences in ToM have been shown to predict children’s social interaction and communication difficulties (Jones et al. 2018; Tager-Flusberg 2003) and can reliably discriminate between levels of autism severity (Hoogenhout and Malcolm-Smith 2016), even when adjusted for the effects of verbal ability and age. Similarly, variation in EF has been consistently shown to relate to autistic features, including both restrictive, repetitive behaviors (Kenworthy et al. 2009; Lopez et al. 2005; South et al. 2007; Turner 1997) and social communication difficulties (McEvoy et al. 1993; though see Jones et al. 2018, and Tager-Flusberg 2003), as well as everyday adaptive behavior (Gilotty et al. 2002; Pugliese et al. 2016).

Despite the extensive literature on ToM and EF in autism (see reviews by Tager-Flusberg 2007, and Demetriou et al. 2017, respectively), two significant oversights remain. The first is that there is only a handful of longitudinal studies examining the predictive relationship between specific cognitive functions and behaviors – which means we know very little about whether early skills in these cognitive domains act as prognostic indicators. Notwithstanding, these few studies have shown that autistic features, particularly children’s social communication difficulties, are predicted by both early ToM (Tager-Flusberg 2003) and early EF performance (Griffith et al. 1999) over a one-year period. Similarly, EF, particularly cognitive flexibility, has been shown to be “a significant prognostic marker” for everyday adaptive behavior in adulthood, as measured by the Vineland Adaptive Behavior Scales (Sparrow et al. 1984) – over 3-years (Berger et al. 2003) and much longer periods (Pugliese et al. 2016), including up to 27 years after intake (Szatmari et al. 1989).

These studies demonstrate the potential prognostic significance of specific cognitive skills, particularly EF, on autistic features and adaptive behavior over long-term periods. Yet both theoretical and longitudinal empirical work in typical (Hughes and Ensr 2007) and autistic (Pellicano 2007; 2010; Russell 1996, Russell 1997) children have demonstrated that ToM and EF themselves are linked – and in one particular direction, such that children’s emerging EF plays a critical role in shaping the development of ToM (but not vice versa). This point brings us to the second oversight – that identifying the unique contribution of early ToM and EF skills to autistic individuals’ behavioral outcomes requires testing both sets of skills during the initial phase of longitudinal studies. To our knowledge, only one study has measured cognitive abilities in more than one domain at intake, and examined their links with behavior. Pellicano (2013) assessed 37
cognitively-able autistic children on tests tapping verbal and nonverbal ability, ToM, EF, and central coherence (local processing) at intake and autistic features 3 years later. Early executive skills – but not ToM – were a significant predictor of autistic features 3 years later, over and above variation in age, verbal ability, nonverbal ability and early ToM skills, suggesting that individual differences in early EF play a key role in shaping autistic children’s later behaviors – at least over a short-term period.

While Pellicano’s (2013) study provides initial evidence for the prognostic significance of early EF skills, it is unclear whether this result extends over a longer period, as children move on up into adulthood, and to everyday life skills – an outcome that may be more meaningful to young autistic people and their parents than autism severity (see Cribb et al., submitted, for discussion). The aims of the current study were therefore twofold. First, we investigated the predictive validity of these childhood cognitive skills (EF and ToM) on behavioral outcomes over a longer time period (12 years). Second, we examined the long-term impact of early specific cognitive skills on two standard outcome measures, autistic features and everyday adaptive behavior.

To address these aims, we saw a group of participants twice within the space of 12 years, once in childhood (Time 1) during which we measured both general cognitive ability and more specific aspects of cognition (ToM and EF\(^2\)) and, again, as participants approached emerging adulthood, 12 years later (Time 2), at which point we focused on standard behavioral outcomes. Based on existing findings and amassing evidence that executive skills are an essential ingredient in typical children’s social, academic and day-to-day lives (Crone et al. 2017; Hughes and Leekam 2004; Riggs et al. 2006; Zelazo et al. 2017), we predicted that individual differences in childhood EF skills would uniquely predict young people’s autistic features and adaptive behavior measured 12 years later.

Method

Participants

Participants who took part in the original study (Time 1: Pellicano et al. 2006; \(n = 45\)) were invited to take part in the 12-years follow-up (Time 2). Twenty-eight participants (26 boys, 2 girls) participated at two time points, once during childhood (\(M = 5\) years; \(SD = 7\) months) and again as they approached emerging adulthood (\(M = 17\) years; 7 months, \(SD = 1\) year; 2 months) (\(M\) duration = 12 years; 2 months; \(SD = 6\) months). Of the 28 young people, 23 had a childhood diagnosis of autism, four of Pervasive Developmental Disorder – Not Otherwise Specified (PDD-NOS) and one of Asperger’s syndrome according to Diagnostic and Statistical Manual of Mental Disorders (4th edition) criteria (DSM-IV; American Psychiatric Association 1994) and all had met full or partial criteria for an autism diagnosis according to the Autism Diagnostic Interview, Revised (ADI-R; Lord et al. 1994). All children were considered to be cognitively able during childhood, that is, they obtained a verbal and nonverbal IQ score of 80 or above on standardized measures. See Table 1 for scores on measures.

At Time 2, 17 of the 45 participants who were seen at Time 1 (3 girls, 38% of original sample) did not take part because either they were untraceable (\(n = 11\)) or did not wish to participate (\(n = 6\)). There were no significant differences between those children who participated at Time 2 (\(n = 28\)) and those who did not (\(n = 17\)) in terms of Time 1 age (\(p = 0.99\)), verbal ability, (\(p = 0.41\)), non-verbal ability (\(p = 0.55\)) or ADI-R algorithm score (\(p = 0.32\)). Only parent-report data were available for four of the 28 participants who took part at Time 2 either because the family had moved interstate and could not be seen face-to-face (\(n = 3\)) or the young person declined to participate (\(n = 1\)).

The inclusion criteria in the original study meant that all participants had English as a first language, none had any additional medical or developmental condition and none were in receipt of medication (see Pellicano et al. 2006). According to parent report at Time 2, young people had since received additional diagnoses of attention deficit hyperactivity disorder (ADHD) (\(n = 8\)), developmental coordination disorder (\(n = 2\)), dyslexia (\(n = 2\)) and mild intellectual disability (\(n = 2\)). For some participants, parents also reported co-occurring medical conditions (e.g., epilepsy, hypotonia, haemochromatosis) (\(n = 4\)) and mental health difficulties (e.g., anxiety, depression, obsessive compulsive disorder) (\(n = 3\)). Seven participants were in receipt of psychoactive medication at Time 2. The majority (\(n = 24\); 86%) were of White ethnic background, 11% were from Asian backgrounds (\(n = 3\)) and 3% were from Mixed backgrounds (\(n = 10\)). Young people’s psychosocial outcomes are described more fully in Pellicano et al. submitted and Cribb et al. submitted.

Measures

Cognitive Measures: Time 1

General Cognitive Ability The Peabody Picture Vocabulary Assessment, third edition (PPVT-III; Dunn and Dunn 1997) was used to assess receptive language ability and
four subtests (Matching, Associated Pairs, Forward Memory, Attention Sustained) of the Leiter International Performance Scale—Revised (Leiter-R; Roid and Miller 1997) were used to index nonverbal ability. Standard scores \((M = 100; SD = 15)\) were derived for Verbal IQ and Performance IQ (see Table 1) but raw scores were used in analyses since such scores have not been adjusted for age and therefore reflect children’s ability.

**Theory of Mind** To assess children’s ToM, they completed three tasks: (1) three first-order unexpected contents tasks, including ‘own’ and ‘other’ false-belief questions (six questions total) (based on Perner et al. 1987), (2) six first-order unexpected transfer tasks, each with a critical false-belief question (based on Baron-Cohen et al. 1985) and (3) two second-order unexpected transfer tasks, each with a critical false-belief test question (based on Perner and Wimmer 1985). Children were given one point for each correctly answered false-belief test question. Higher scores are indicative of better ToM ability. Reliability estimates were high (Cronbach’s alpha = 0.86).

### Table 1: Descriptive statistics for developmental variables and cognitive measures at intake (Time 1) and behavioral measures at the 12-years follow-up (Time 2)

<table>
<thead>
<tr>
<th></th>
<th>Time 1 ((n = 28))</th>
<th>Time 2 ((n = 28))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chronological age</strong></td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>(in months)</td>
<td>67.21 (11.45)</td>
<td>214.11 (14.26)</td>
</tr>
<tr>
<td></td>
<td>49–88</td>
<td>192–242</td>
</tr>
<tr>
<td><strong>Verbal ability</strong></td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>(\text{Max} = 6)</td>
<td>97.04 (10.71)</td>
<td>91.29 (20.02)</td>
</tr>
<tr>
<td></td>
<td>80–122</td>
<td>51–120</td>
</tr>
<tr>
<td><strong>Nonverbal ability</strong></td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>(\text{Max} = 1)</td>
<td>113.93 (13.73)</td>
<td>99.17 (19.36)</td>
</tr>
<tr>
<td></td>
<td>83–141</td>
<td>59–133</td>
</tr>
<tr>
<td><strong>Theory of mind (ToM)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First-order unexpected-contents (\text{Max} = 6)</td>
<td>2.14 (2.17)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0–6</td>
<td></td>
</tr>
<tr>
<td>First-order unexpected location (\text{Max} = 6)</td>
<td>1.78 (1.83)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0–6</td>
<td></td>
</tr>
<tr>
<td>Second-order unexpected location (\text{Max} = 2)</td>
<td>0.07 (0.26)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0–1</td>
<td></td>
</tr>
<tr>
<td><strong>Executive function (EF)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mazes total score (\text{Max} = 26)</td>
<td>13.75 (5.34)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4–22</td>
<td></td>
</tr>
<tr>
<td>Luria’s hand-game conflict score (\text{Max} = 10)</td>
<td>7.21 (1.83)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4–10</td>
<td></td>
</tr>
<tr>
<td>Tower of London (\text{No. trials solved in min. no. moves} (\text{Max} = 16)</td>
<td>6.39 (2.92)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2–13</td>
<td></td>
</tr>
<tr>
<td>Set-shifting (proportion of errors following first sort to criterion) (\text{Max} = 1)</td>
<td>0.31 (0.07)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.18–0.45</td>
<td></td>
</tr>
<tr>
<td><strong>Behavioral outcome measures</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADOS-2 Calibrated Severity Scores(c)</td>
<td>6.92 (2.18)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2–10</td>
<td></td>
</tr>
<tr>
<td>Vineland-2f Adaptive Behavior Composite score(d)</td>
<td>71.68 (9.88)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>53–100</td>
<td></td>
</tr>
</tbody>
</table>

\(a\) Verbal ability was measured using the Peabody Picture Vocabulary Test—Third edition (Dunn and Dunn 1997) at Time 1 \((n = 28)\) and the Verbal Comprehension Index from the Wechsler Abbreviated Scales of Intelligence—2nd edition (WASI-2; Wechsler 2011) at Time 2 \((n = 24)\), standard scores reported here

\(b\) Nonverbal ability was measured using the Leiter International Performance Scale—Revised (Leiter-R; Roid and Miller 1997) at Time 1 \((n = 28)\) and the Perceptual Reasoning Index from the WASI-2 (Wechsler 2011) at Time 2 \((n = 24)\)

\(c\) ADOS-2: Autism Diagnostic Observation Schedule—2nd edition (Lord et al. 2012) at Time 3 \((n = 24)\), calibrated severity scores reported here (ranging from 1 to 10; Hus and Lord 2014)

\(d\) Vineland-2: Vineland Adaptive Behavior Scales—2nd edition (Sparrow et al. 2005) at Time 3 \((n = 28)\), standard scores reported here
Executive Function At Time 1, children completed four tasks tapping executive function (see Pellicano et al. 2006). The Mazes task from the Wechsler Preschool and Primary Scales of Intelligence–Revised (Wechsler 1989) tested children’s planning ability through a set of increasingly complex mazes. To succeed on each trial, children needed to plan their route ahead, to reach the opening of the maze while making minimal errors. Standard scoring procedures were applied (see Wechsler 1989). High scores reflect good planning ability (maximum score = 26).

The Tower of London task (Shallice 1982; see also Hughes 1998) assessed children’s higher order planning ability. Children were presented with three colored beads (red, white, black) arranged in a particular configuration (start state) on a wooden tower structure consisting of three vertical pegs of increasing size. They were then shown a picture of the beads in a different configuration (goal state) and asked to move the beads one at a time and only onto the peg board, all within the minimum number of moves. After three practice trials, children were given problem sets (four trials each) of increasing difficulty, including one-, two-, three-, and four-move sets. Testing ceased if children failed all four trials within a problem set. The number of moves taken and rule violations were recorded. Children were given one point for each trial if they reached the goal state within the minimum number of moves and without violating any rules. The dependent variable of interest was the total number of trials completed in the minimum number of moves. High scores indicate good planning ability (maximum = 16).

Luria’s hand-game was used to assess children’s motor inhibition (see Hughes 1996; Luria 1966). The task began with an ‘imitation control’ condition in which, on each trial, the experimenter showed the child a hand movement (either “made a fist” or “pointed a finger”) and asked the child to copy the movement. Next, in the critical, ‘conflict condition’, children were asked to play the ‘opposites’ game: “Now, if I point a finger, I want you to show a fist, and if I show a fist I want you to point a finger, so we’re not making the same shapes. What do you do if I show a fist?... And if I point a finger?” (Hughes 1996, p. 231). There were five trials for each hand movement, presented in a randomized order. Children received one point if they immediately and accurately made the correct hand movement on each trial (maximum = 10). Higher scores in the conflict condition reflect better inhibition.

Cognitive flexibility was assessed using the teddy-bear set-shifting task (Hughes 1998), a developmentally sensitive card-sorting task, conceptually similar to the traditional Wisconsin Card Sort Task (Heaton 1981), which tested children’s ability to switch flexibly between cognitive sets in response to feedback. Children were shown three card decks, which differed in terms of color (green vs. pink, blue vs. red, or yellow vs. purple), picture shown (hearts vs. diamonds, squares vs. moons, or stars vs. happy faces), and size of picture (small vs. large). Children were first presented with one deck of cards and asked to work out which cards teddy liked best. On each trial, children were shown a card from one deck and asked, “Is this one of teddy’s favorites?” If the card was one of teddy’s favorite cards, children posted it into a post-box. If it was not one of teddy’s favorite cards, children placed it face down on the table. On each trial, children were provided with immediate verbal feedback. Children learned to sort the cards according to one of three rules (color, shape, size). When the child had successfully sorted six cards consecutively or were given a maximum of 20 trials, the sorting rule (e.g., color, shape, size) changed. Importantly, children were never explicitly told that the rule had changed. This was implicit in the fact that the child was presented with a new teddy bear and new deck of cards. The dependent variable was the proportion of errors committed following the first sort to criterion. A low score (i.e., minimal errors) indicates good cognitive flexibility.

Behavioral Measures: Time 2

General Cognitive Ability At the 12-years follow-up, we administered the Wechsler Abbreviated Scales of Intelligence – 2nd edition (WASI-2; Wechsler 2011) to assess IQ, which incorporates a Verbal Comprehension Index (comprised of Vocabulary and Similarities subtests) and the Perceptual Reasoning Index (comprised of Block Design and Matrices subtests). Standard scores are reported in Table 1. We note a discrepancy (specifically, a reduction) in participants’ IQ scores between Time 1 (indexed by the PPVT-III and the Leiter-R) and Time 2 (indexed by the WASI-2). Given that these different instruments place distinct demands on participants’ language and EF skills, we cannot be sure whether the reduction in scores is real or whether it is attributable to the change in instrument use.

Autistic Features The Autism Diagnostic Observation Schedule – second edition (ADOS-2; Lord et al. 2012) was used to measure severity of autistic features. The ADOS-2 is a 40-min, standardized observational scale, administered by a trained examiner, designed to provide opportunities or ‘presses’ for the evaluation of social, communicative and repetitive behaviors. Four of the 28 participants did not take part in the ADOS-2 at Time 2; three had moved out of region and could not be seen face-to-face and one declined to participate. The majority of participants completed Module 4 (n = 22), while one participant completed Module 2 and another Module 3. Total algorithm scores were converted to ADOS calibrated severity scores (CSS) according to Gotham et al. (2009) and Hus and Lord (2014). Higher scores indicate greater severity of autistic behaviors (maximum = 10).

Adaptive Behavior Parents completed the Vineland Adaptive Behavior Scales – second edition (Vineland-2; Sparrow et al.
– a standardized measure that is designed to assess a variety of typical developmental milestones in three domains: socialization, communication and daily living skills. Scores are derived for each domain, as well as an overall adaptive behavior composite (ABC) score. Standardized scores (M = 100; SD = 15) are reported in Table 1. Higher scores reflect better adaptive skills.

**General Procedure**

At each time-point, participants completed all measures on either one or two separate occasions, either at their home or the University. Ethical approval for this follow-up study was granted by the Human Research Ethics Office at the University of Western Australia (RA/4/1/6992). The parents of all participants, and at Time 2 also the participants themselves, gave written informed consent prior to taking part.

**Results**

Descriptive statistics for developmental variables and performance on cognitive and behavioral measures administered at each time point are shown in Table 1. To begin, we conducted correlations between and within behavioral and cognitive variables (see Supplementary Table 1). As expected, scores on tasks tapping each cognitive domain were significantly inter-related, indicating good convergent validity for each construct. Robust composite scores were therefore created by averaging the standardized scores of individual measures for ToM and EF, in order both to minimize the number of predictor variables entered into analyses and to facilitate comparison with Pellicano (2013). Next, we report the results of correlational analyses between our composite ToM and EF variables and behavioral variables. Finally, we conducted hierarchical regression analyses to test whether individual differences in ToM and EF measured in early childhood predicted autistic features, as indexed by ADOS-2 severity scores, and adaptive functioning, as measured by Vineland-2 scores, 12 years later.

**Correlational Analyses**

Table 2 shows raw and partial correlations between Time 1 developmental and cognitive composite variables and Time 2 behavioral variables. Individual differences in Time 1 ToM and EF were significantly related to Time 2 ADOS-2 severity scores – and these relationships remained significant once Time 1 age, verbal ability and non-verbal ability were accounted for in partial correlations. Young people with better early ToM and EF showed fewer autistic features at the 12-years follow-up. There were no significant relationships between Time 1 age, verbal ability and nonverbal ability and young people’s Time 2 ADOS-2 severity scores.

Cross-time correlational analyses also showed that there were significant relationships between Time 1 verbal ability and EF and Time 2 Vineland-2 scores. Once variance attributable to Time 1 age, verbal ability and non-verbal ability was partialled out of the relationship, the EF-Vineland-2 relationship became marginally non-significant, r(19) = 0.41, p = 0.06, likely a result of the small sample size. Young people with better verbal ability and EF ability early in development showed better adaptive functioning 12 years later. No other correlations reached significance.

**Regression Analyses**

**Autistic Features** To examine the early predictors of young people’s later autistic features, as indexed by their Time 2 ADOS-2 severity scores, participants’ Time 1 developmental variables (chronological age, verbal ability and nonverbal ability) were entered simultaneously into the first step of the model. These variables accounted for a negligible amount (1.4%) of variation in young people’s autistic features. The model was not significant, F(3,20) = 0.09, p = 0.96, R² = 0.014. Next, Time 1 composite ToM and EF variables were entered stepwise to test the additional – and potentially unique – contribution of these variables to later autistic features. Time 1 EF scores significantly improved model fit (B = −1.114, p = 0.003) and explained a further 38% of the variance in participants’ autistic features, F(1,19) = 11.69, p = 0.003, ΔR² = 0.38 (Time 1 ToM scores did not: B = −0.344, p = 0.12). The negative beta value suggests that better EF skills at Time 1 were significantly predictive of fewer autistic features 12 years later (Table 3). The final model was significant, F(4, 19) = 3.03, p = 0.04.

### Table 2 Pearson correlations between Time 1 developmental and cognitive variables and Time 2 behavioral variables (n = 24)

<table>
<thead>
<tr>
<th>Time 1</th>
<th>Autistic featuresa</th>
<th>Adaptive behaviorb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>−0.09</td>
<td>0.03</td>
</tr>
<tr>
<td>Verbal ability</td>
<td>−0.11</td>
<td>0.37*</td>
</tr>
<tr>
<td>Nonverbal ability</td>
<td>−0.09</td>
<td>0.02</td>
</tr>
<tr>
<td>Theory of mind</td>
<td>−0.52** (−0.53**)</td>
<td>0.20 (0.18)</td>
</tr>
<tr>
<td>Executive function</td>
<td>−0.50** (−0.62**)</td>
<td>.34* (0.41*)</td>
</tr>
</tbody>
</table>

Partial correlations (df = 19) adjusting for early/Time 1 chronological age, verbal ability, and nonverbal ability are shown in parentheses

*aSignificant at the 0.05 level (2-tailed)

**Significant at the 0.01 level (2-tailed)

a as indexed by calibrated severity scores on the Autism Diagnostic Observation Schedules – 2nd edition (ADOS-2; Lord et al. 2012)

b as indexed by the Vineland-2 Adaptive Behavior Composite score (Sparrow et al. 2005)

c p = 0.06
Adaptive Behavior

We used a similar model to examine the predictive validity of early cognition on young people’s later adaptive behavior, as indexed by Vineland-2 scores measured at the 12-years follow-up. When participants’ Time 1 age, verbal ability and nonverbal ability were entered simultaneously as predictors into the first step of the model, these variables accounted for 27% of the variance in young people’s Time 2 Vineland-2 scores, $F(3,24) = 2.95, p = 0.05, R^2 = 0.27$. When Time 1 ToM and EF were entered stepwise into the second block, only Time 1 EF composite scores made an independent contribution to the prediction of young people’s Time 2 Vineland-2 scores (B = 0.665, $p = 0.047$), explaining an additional 12% of the variance, $F(1,23) = 4.42, p = 0.047, \Delta R^2 = 0.118$ (Time 1 ToM composite scores were not significant: B = 0.027, $p = 0.89$). Table 3 shows that the final model was significant, $F(4,23) = 3.64, p = 0.019$. Better verbal ability, nonverbal ability and EF early in development were predictive of higher Vineland-2 scores 12 years later.

Discussion

This study examined the predictive power of specific cognitive skills measured in childhood for a group of cognitively-able autistic children followed over a 12-year period. Early EF skills were significantly predictive of later behavioral outcomes, including autistic features and adaptive behavior, over and above variance attributable to age, verbal ability, nonverbal ability and ToM skills at intake. The current findings support and extend existing longitudinal work demonstrating significant links between early EF skills and later behavior in cognitively-able autistic people (Berger et al. 2003; Pellicano 2013; Pugliese et al. 2016; Szatmari et al. 1989), suggesting that measures of specific aspects of cognition (namely, EF) might well be better predictors of behavioral outcomes than general intellectual functioning, at least for autistic individuals considered to be cognitively able in childhood.

Critically, we showed that the long-term links between early EF and later outcomes were independent of the influence of early ToM skills – skills that are also held to be linked theoretically to autistic children’s behavior (see Tager-Flusberg 2007, for review). This finding is particularly noteworthy since no study has examined the predictive power of specific cognitive skills over a long-term period, measuring such skills in more than one domain. The only study that had done so was an earlier follow-up of the current sample, in which EF skills measured in childhood were uniquely predictive of autistic children’s social communication and repetitive behaviors 3 years later (Pellicano 2013). In both the 3-years and 12-years follow-up studies, it was not that early individual differences in ToM failed entirely to predict children’s later behavior (see Table 2 and Supplementary Table 1). Rather, it was that such differences did not contribute unique variance to behavioral outcomes beyond that already accounted for by early verbal and nonverbal ability – and, critically, children’s EF.

One explanation for the lack of a unique relationship between early ToM and later behavioral outcomes is that our specific tasks for each cognitive domain may not have been equally sensitive in detecting the potentially subtle cognitive differences in autistic children – although it is worth noting

Table 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>$\beta$</th>
<th>$R^2$ or $\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicting autistic featuresa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1 age</td>
<td>−0.003</td>
<td>0.023</td>
<td>−0.028</td>
<td></td>
</tr>
<tr>
<td>Time 1 verbal ability</td>
<td>0.010</td>
<td>0.015</td>
<td>0.077</td>
<td></td>
</tr>
<tr>
<td>Time 1 nonverbal ability</td>
<td>0.023</td>
<td>0.015</td>
<td>0.423</td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td>0.38*</td>
</tr>
<tr>
<td>Time 1 EF composite</td>
<td>−1.114</td>
<td>0.326</td>
<td>−0.886**</td>
<td></td>
</tr>
<tr>
<td>Predicting adaptive behaviorb</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1 age</td>
<td>−0.239</td>
<td>0.210</td>
<td>−0.274</td>
<td></td>
</tr>
<tr>
<td>T1 verbal ability</td>
<td>0.350</td>
<td>0.144</td>
<td>0.613*</td>
<td></td>
</tr>
<tr>
<td>T1 nonverbal ability</td>
<td>−0.271</td>
<td>0.125</td>
<td>−0.615*</td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td>0.12*</td>
</tr>
<tr>
<td>Time 1 EF composite</td>
<td>0.665</td>
<td>0.316</td>
<td>0.544*</td>
<td></td>
</tr>
</tbody>
</table>

EF: executive function

*aSignificant at the 0.05 level
**Significant at the 0.01 level

a as indexed by calibrated severity scores on the Autism Diagnostic Observation Schedules – 2nd edition (ADOS-2; Lord et al. 2012), n = 24

b as indexed by the Vineland-2 Adaptive Behavior Composite score (Sparrow et al. 2005), n = 28
that reliability estimates for the ToM composite score were high. It is also possible that the narrow focus on ToM as ‘false-belief understanding’ precluded the possibility of revealing significant associations between ToM and behavioral outcomes. Although false-belief tasks were developmentally appropriate for the children when assessed at intake, longitudinal research that adopts a broader approach, which includes both social-cognitive and social-perceptual ToM tasks (Tager-Flusberg 2007), might be beneficial (see Jones et al. 2018 and Kouklari et al. 2018).

Another explanation still relates to the potential primacy of EF over ToM (see Pellicano 2007). EF and ToM skills have long been held to be inextricably linked during development (Moses 2001; Perner 1998; Russell 1997) but the evidence suggests that this relationship exists in one particular causal direction only. Early individual differences in EF have been shown to predict developmental changes in children’s ToM – false-belief understanding – during the toddler (Carlson et al. 2004; Hughes and Ensor 2007) and preschool (Carlson et al. 2004; Hughes 1998) periods in typical children, and even in this sample of autistic children (Pellicano 2010), but there were no significant relationships in the reverse direction; that is, individual differences in early ToM do not predict developmental changes in EF (see also Kouklari et al. 2018, for similar results in school-age children). Furthermore, EF at age 3 has also been shown to significantly predict autistic children’s pre-symbolic and symbolic play skills – precursory ToM skills – at age 6 (Faja et al. 2016). Together, these findings suggest that EF and ToM are related during development, such that early EF plays a critical role in the emergence of ToM in typical and autistic children (Moses 2001; Russell 1997).

It is perhaps not surprising that EF is a stronger predictor of autistic children’s later behavioral outcomes given that, unlike ToM, it is a domain-general function, which facilitates many other cognitive processes (Burgess et al. 2006; Denckla 1996; Diamond 2013; Miller and Cohen 2001;Royall et al. 2002). One key question, however, is how might early EF contribute to later behavior? Given that EF is closely associated with the pre-frontal cortex (Alvarez and Emory 2006), it is possible that early EF difficulties persist during development, which are in turn associated with poor behavioral outcomes. We did not measure specific cognitive skills at the 12-years follow-up because the tasks used would be vastly different from those used in childhood and the sample size (over which we had little control) was insufficient to account for such measurement error. It is unclear therefore whether executive difficulties persisted as young people approached emerging adulthood – at least according to performance on behavioral EF tasks. That said, in-depth qualitative reports from these same young people and their parents during the 12-years follow-up study repeatedly highlighted their executive-related struggles in daily life (see Cribb et al., submitted). Young people spoke of difficulties “making decisions”, switching flexibly from one thing to the next (“like when there’s a change of plan, I just get stressed out and just get frustrated”) and problems with planning and future-oriented thinking: “I’ve never thought of [the future] because I always think of the present and maybe one or two days later”. Parents also described their children’s difficulties with “forward thinking, planning, organization” and “keeping track of time”, all of which they felt to have a negative impact on their ability to do everyday tasks and, ultimately, on their transition to adulthood. These qualitative data point towards persistent everyday executive problems as young people approach emergent adulthood.

Alternatively, it is possible that EF difficulties, especially during sensitive periods early in development, have a sustained effect on behavior that persists beyond our ability to measure the cognitive atypicality itself. Johnson (2012) posits that poor EF skills early in development are an additional risk factor because children have less capacity to compensate for atypical functioning in other neural systems. In contrast, those with strong EF skills may be better able to adapt and thus are more likely to have better behavioral outcomes (see also Halperin and Schulz 2006). Future, well-powered longitudinal studies that measure cognition and behavior across multiple domains and multiple time-points are needed to disentangle these possible explanations.

Whatever the reason, however, the current results suggest that EF difficulties are likely to place the individual autistic child at greater likelihood for a poor developmental outcome – and therefore highlight EF as a potentially promising candidate for intervention efforts. Despite decades of research on EF in autism, there has been remarkably little attention on EF as a potential target for intervention. One of the first attempts to ‘train’ EF with autistic school-age children failed to show gains in performance on executive tasks following such training (Fisher and Happé 2005). Yet the results of recent work, which has embedded interventions across home and school settings, have been more positive. Kenworthy et al.’s (2014) Unstuck and On Target program targets autistic EF difficulties through a cognitive behavioral program that teaches autistic school-age children self-regulatory EF scripts (e.g., “stuck on a detail”, “big picture”, “flexible thinking”) and how to deploy these scripts across different contexts in order to guide flexible, goal-directed behavior. A randomized controlled trial that tested the effectiveness of this program against a social skills intervention demonstrated that, while 7- to 11-year-old autistic children in both interventions improved, especially with regard to social skills, there were greater gains for those children who received the Unstuck and On Target program, including better EF skills (problem-solving, flexibility, and planning), as well as improved classroom behavior (following rules, making transitions and being flexible) (see also Oswald et al. 2018, for a program in adults that includes executive function as a key focus).
While Kenworthy et al.’s results are extremely encouraging, the findings from the current study suggest that we need to extend EF interventions into the preschool period—a time when the pre-frontal cortex shows a boost in development (Diamond 2013) and may have the greatest chance of influencing a range of important, concurrent skills, including school readiness (Blair and Razza 2007; Pellicano et al. 2017), play (Faja et al. 2016) and theory of mind (Hughes 1998; Pellicano 2010), as well as later behavior. Several intervention studies with typical children that are, once again, embedded in the school curriculum report evidence of the malleability of EF by “exercising” children’s early-emerging EF skills, demonstrating significant positive effects on children’s later EF skills and early academic success (Bierman et al. 2008; Blair and Raver 2014; Diamond et al. 2007; Duncan et al. 2007). Furthermore, those children who show the weakest EF skills (through social disadvantage) appear to benefit the most from these interventions (Blair and Raver 2014; Raver et al. 2012), making these interventions particularly relevant for autistic children who often show poorer EF skills.

In sum, this study showed that EF— but not ToM— was a significant prognostic indicator in a group of young autistic people approaching emerging adulthood. These findings require replication with larger samples, which would also allow one to take a more nuanced approach, identifying whether specific EF components, such as cognitive flexibility, play an especially important prognostic role (cf. Berger et al. 2003; Szatmari et al. 1989), rather than treat EF as a heterogeneous construct as we have done here—and to examine the effects of gender, amongst other potentially confounding factors, including the type, frequency and intensity of interventions (behavioral or pharmacological). Such future studies should also aim to measure both behavioral (e.g., ADOS and Vineland) and specific cognitive skills (e.g., ToM, EF) at multiple points over time in order to specify further the nature of the relationship between such cognitive skills and developmental changes in behavior. Nevertheless, that the current findings support the findings from existing longitudinal studies, including an earlier follow-up of this same sample, warrants confidence in the results and provides compelling evidence of the important contribution of early individual differences in EF in shaping autistic children’s developmental trajectories. Future work must focus on interventions that can bolster the development of early EF skills and help to promote the life chances and opportunities of young autistic people.

Acknowledgements We are extremely grateful to the young people and their parents for continuing to be so generous with their time over the years, to Murray Maybery for providing support throughout the study and to Marc Stears for helpful comments on a previous version of this manuscript. This research was supported in part by funds from the School of Psychology at the University of Western Australia and in part by a 2015 Philip Leverhulme Prize awarded to EP by the UK’s Leverhulme Trust. We have no conflicts of interest of which we are aware.

Compliance with Ethical Standards

Conflict of Interest The authors have no conflicts of interest of which they are aware.

Ethical Approval Ethical approval for this study was granted by the Human Research Ethics Office at the University of Western Australia (RA/4/1/6992).

Informed Consent All participants, parents and young autistic people, gave written informed consent to take part, for their child or for themselves, prior to participation.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References


