Journal of Development and Psychopathology

Developmental trends of hot and cool executive function in school aged children with and without autism spectrum disorder: links with theory of mind

Evangelia-Chrysanthi Kouklari, Stella Tsermentseli & Claire P. Monks,

Department of Psychology, Social Work, and Counselling, University of Greenwich, London, UK

Acknowledgements

The authors are grateful for the Vice Chancellor Scholarship Funding from the University of Greenwich. The authors would like to acknowledge their gratitude to the participating schools and children.

Conflicts of Interest: The authors declare that they have no conflict of interest**.**

Ethical approval: “All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.”

Informed consent: “Informed consent was obtained from all individual participants included in the study.”

Abstract

The development of executive function (EF) in autism spectrum disorder (ASD) has been investigated using only “cool”-cognitive- EF tasks while there is limited knowledge regarding the development of “hot”-affective- EF. Although cool EF development and its links to theory of mind (ToM) have been widely examined, understanding of the influence of hot EF to ToM mechanisms is minimal. The present study introduced a longitudinal design to examine the developmental changes in cool and hot EF of children with ASD (n=45) and matched (to age and IQ) controls (n=37) as well as the impact of EF on ToM development over a school year. For children with ASD, although selective cool (working memory, inhibition) and hot (affective decision making) EF domains presented age-related improvements, they never reached the performance level of the control group. Early cool working memory predicted later ToM in both groups but early hot delay discounting predicted later ToM only in the ASD group. No evidence was found for the reverse pattern (early ToM predicting later EF). These findings suggest that improvements in some EF aspects are evident in school age in ASD and highlight the crucial role both cool and hot EF play in ToM development.

Autism spectrum disorder (ASD) is a multifaceted neurodevelopmental disorder that significantly impairs children's verbal and nonverbal communication, social interactions, and behaviours (Diagnostic and Statistical Manual of Mental Disorder— DSM-5, American Psychiatric Association, 2013). Executive dysfunction has received increased attention, as deficits in aspects of executive function (EF) have been demonstrated consistently in ASD samples across the lifespan (see Hill's review (2004) and Demetriou et al.'s (2017) meta-analysis). EF refers to a set of future-oriented and goal-directed cognitive skills, closely associated to the prefrontal cortex, that are crucial for problem solving and social behaviour, as well as the ability to organise oneself (Best & Miller, 2010). EF problems in ASD typically manifest as deficits in domains such as working memory (Alloway et al., 2009; Geurts et al., 2014; Kercood et al., 2014), inhibition (Christ et al., 2007; Happé et al., 2006a; Xiao et al., 2012), planning (Chen et al., 2016; Kimhi et al., 2014; Verté et al., 2005) or cognitive flexibility (Gioia et al., 2002; South et al., 2007; Yeung et al., 2016). Although research has shown that impairments in specific components of EF (mostly cognitive flexibility and planning) have been conceptualized as being key aspects in discriminating ASD from other neurodevelopmental disorders (e.g. ADHD) (for a review see Craig et al., 2016), the additive effect comorbid ADHD has been found to have on EF deficits such as working memory (Andersen et al., 2013; Yerys et al., 2009) and inhibition (Sinzig et al., 2008) in ASD, should not be overlooked either.

EF development is generally considered to influence the development of children's social cognition (Devine et al., 2016) and is intimately tied to Theory of Mind (ToM), another cognitive domain known to develop atypically in ASD (Lantz, 2002). ToM is the ability to infer mental/ emotional states in order to predict and explain behaviour (Goldman, 2012). ToM is a multifaceted cognitive skill that develops gradually; its development commences in infancy and continues to improve throughout middle childhood and adolescence. The understanding of false belief (understanding that one's belief/ representation about the world can contrast with reality), which is a critical aspect of ToM, measured by first-order false belief tasks, typically emerges at the age of 3-4 years (Schug et al., 2016). As children grow up, they present age-related performance gains and become capable of solving more complex, high-order ToM tasks (e.g. emotion understanding) across middle and later childhood (Devine & Hughes, 2013; Dumontheil et al., 2010). Evidence from ASD suggests that some children with ASD fail to fully develop the prerequisites of ToM and that the ToM trajectory could follow either a delayed (Steele et al., 2003) or deviant (Peterson et al., 2005; Serra et al., 2002) pathway in ASD.

Evidence from both typical development and ASD has consistently shown that EF is strongly associated to ToM across childhood (Bock et al., 2015; Carlson et al., 2013; Im-Bolter et al., 2016). Several developmental psychologists have argued that there is in fact a more fundamental link between EF and ToM with functioning in one domain being a necessity for the emergence of the other. Perner (1998) and Perner & Lang (1999) proposed that the acquisition of ToM is a prerequisite of children's ability to control their behaviour (EF) while alternatively, Russell (1996) suggested that the ability to control one's actions (EF) must be attained first as it then crucially influences the development of ToM. Several longitudinal studies on the EF-ToM relationship in early childhood in typical development indicated that children's performance on EF measures predicted later performance on ToM false belief tasks (independent of age, verbal ability, and earlier ToM scores) but not vice versa (Carlson et al., 2004; Flynn, 2007; Hughes, 1998; Hughes & Ensor, 2007). These findings highlight that EF is a prerequisite for the development of ToM in early life as it scaffolds the emergence of ToM mechanisms (Apperly, Samson, & Humphreys, 2009; Sabbagh et al., 2006). However, as this theoretical position does not account for potential EF-ToM associations beyond the fifth year, less is known about the extension of this developmental relationship in school aged children with and without ASD. Relevant longitudinal research in school age is very limited (e.g. Austin et al., 2014) and has demonstrated weak evidence for the account that early ToM predicts later EF, but stronger support of the early EF influencing later ToM in school age. This could suggest that EF might play a substantial role in children's developmental outcomes, particularly in relation to social cognition.

The developmental nature of the EF-ToM relationship in ASD has been vastly theoretically debated due to this coexistence of deficits in both domains. Building on the initial propositions regarding the executive dysfunction influencing autism symptomatology (Pennington & Ozonoff, 1996; Russell, 1997), there is growing evidence showing that the ASD functional/ social outcomes, such as ToM, may be due to differences in emerging EF abilities (Demetriou et al., 2017; Leung et al., 2016; Pellicano 2012; Russo et al., 2007). Performance on measures of EF and ToM (mainly false belief understanding tasks) have indeed been found to be correlated in ASD (Ozonoff et al., 1991; Kimhi et al., 2014; Kouklari et al., 2017; Pellicano, 2007). Evidence regarding the link between EF and ToM in ASD mainly derives from the preschool period and full consideration of this developmental issue has not yet been possible due to the limited number of studies analysing this issue within a longitudinal design. Pellicano (2010) in a longitudinal study of two time points with children aged 4-7 years, (followed after 3 years), indicated that cool EF abilities (planning, cognitive flexibility, and inhibition) at the first time point were a significant predictor of the changes in children's ToM skills at the second time point (over and above the variance of age, verbal, and non-verbal ability) in ASD. Moreover, Tager-Flusberg and Joseph (2005) found that children's (5-14 years) performance on early EF tasks predicted later ToM abilities after 1 year independent of initial ToM scores and verbal ability. These studies show that in ASD, as in typical development, EF skills play a critical role in shaping the development of ToM. The important contribution of EF in typical development, together with these promising findings from studies in ASD provide good reason to suspect that the development of EF might critically influence children's developmental trajectories of sociocognitive profiles, particularly ToM skills (with poor EF being a risk factor for poor developmental outcomes) in ASD.

It should be noted that the determination of the precise nature of the EF developmental pathway and its influence on ToM in ASD first requires research to shed more light on the actual nature of EF itself (Pellicano, 2012). EF is by nature a complex construct to characterise, as most attempts to determine its nature reflect that "EF is by no means a unitary concept" (Elliott, 2003). EF has traditionally been viewed through a purely cognitive lens. However, the theoretical distinction between cool and hot EF aspects, suggested by Zelazo and Müller (2002), could aid in better understanding the effect of EF development on ToM and social cognition in general. According to this distinction, cool EF includes processes evoked in motivationally neutral, non-affective situations such as inhibition, working memory, and planning (Zelazo & Carlson, 2012) that are usually tapped by decontextualised EF tasks, such as Go/No-Go, Tower of London, Digit Span tasks (Anderson et al., 2008). In contrast to cool EF, hot EF are elicited under motivationally significant, affective conditions such as delay discounting (*the tendency to choose more immediate, smaller rewards*) and affective decision making (*mental processing occurring on the selection of one or more possible options under risk where one employs both rational and emotional processes*). Hot EF is therefore mainly measured by tasks with meaningful rewards and losses for the individual (i.e. the Gambling task and/or Delay discounting tasks).

EF emerges early in life, evidences critical changes in school age and matures by adolescence (Best & Miller, 2010). EF development has been mainly assessed by tasks assessing cool EF despite recent evidence supporting separate domains of cool and hot EF (Kim et al., 2014; Willoughby et al., 2011). Generally little is known about the developmental course of "hot"-affective EF processes and whether cool and hot EF present similar developmental changes.

Hot EF is suggested as presenting rapid developments in the early years of life in typical development followed by age-related improvements across middle childhood and adolescence (Prencipe et al., 2011; Zelazo & Müller, 2002). Only a few cross-sectional studies to date have investigated the development of hot EF (in conjunction with cool EF) mainly in typical development and have yielded mixed results. No significant differences have been found in the development of hot and cool EF in the preschool period (3-5 years) with both domains being correlated and exhibiting performance gains after the third year of life (Hongwanishkul et al., 2005). Further evidence from middle childhood and adolescence (Hooper et al., 2004; Prencipe et al., 2011) suggested that the weakly correlated hot and cool EF develop independently, and hot EF is likely to follow a differentiated developmental trajectory beyond the 5 years of age. The potential differences in the developmental trends of cool and hot EF is an open topic of debate, and such examination in ASD could make it plausible for separate EF domains to be found specifically affected or have specific developmental relations to other outcomes such as ToM.

The limited number of longitudinal EF studies in ASD have focused only on cool aspects to date and due to mixed results it is not clear whether the development of EF in ASD follows the same pathway as that of typical development. Early studies (Griffith, Pennington, Wehner, & Rogers, 1999; Ozonoff & McEvoy, 1994) indicated a lack of age-related improvements and poorer performance for children and adolescents with ASD relative to matched controls in measures of selective cool EF (i.e. planning and cognitive flexibility) across time. More recently though, Pellicano (2010), found that EF planning of pre-schoolers with ASD (mean age 5.5 years) improved significantly over a 3 year period, surprisingly at an even faster rate than the controls. Identifying the developmental pathways of both cool and hot EF could provide a solid ground to overcome the limitations of current theories of EF development and lead to a better understanding of the heterogeneity in neurocognitive impairments in ASD.

Finally, it should be noted that research into the ToM-EF association in school aged participants has mainly employed cool EF tasks, despite hot EF and ToM considered as being mediated by the same medial regions of the prefrontal cortex (Chan et al., 2008; McDonald, 2013; Sabbagh et al., 2009). For instance, social interactions that involve ToM abilities may require the control of behaviour or thought under emotionally significant situations (hot EF) (Zelazo & Müller, 2002). Thus, hot EF could be more central to the emergence of ToM. There is no study in ASD to date either that has investigated whether hot EF influences ToM across time and how they may interact within the cognitive profile of school aged children with ASD. The present study may hopefully enhance our understanding of the higher-order cognitive deficits that underpin social interaction problems in ASD.

Current Objectives

Research on the development of hot EF both in typical development and in ASD lags behind that of cool EF and the understanding of the potential link between hot EF and ToM is quite limited. Longitudinal studies investigating the links between EF and ToM in the school age period (> 5 years) are very limited (Devine & Hughes, 2014). Extending the developmental EF-ToM research in school aged children is critical in order to examine whether developmental patterns found in early life persist across the course of children’s development (McAlister & Peterson, 2013). Moreover, based on the evidence presented above (Hooper et al., 2004; Prencipe et al., 2011), distinct EF domains (hot and cool) are expected to follow different developmental trends and become more specialised across middle childhood and adolescence. The present study therefore attempted to address these gaps in the literature.

The first aim of this study was to compare the developmental changes in cool and hot EF and ToM abilities between school aged children with and without ASD after a one year interval. Taken together, previous studies show there is no clear developmental framework of cool EF in ASD; with some studies reporting age-related improvements (Pellicano, 2010) and others not (Griffith et al., 1999; Ozonoff & McEvoy, 1994). Moreover, no research to date has investigated the development of hot EF in ASD. Due to these mixed findings and the minimal longitudinal evidence regarding the development of hot EF in ASD, the present study was exploratory and specific predictions could not be made. We sought to determine whether there are similarities or deviance/delay relative to controls in the hot and cool EF developmental pathways followed in ASD.

The second aim was to shed more light on the longitudinal association between cool and hot EF and ToM in school aged children with and without ASD. Based on previous research, specific predictions can be made only about the cool EF-ToM link, due to the current lack of relevant longitudinal hot EF research. Specifically, there is stronger evidence for the cool EF to predict the emergence of ToM rather than the opposite in early childhood and the preschool period in typical development (Marcovitch et al., 2014) and ASD (Pellicano, 2010, 2012). We hypothesised that early cool EF would predict later ToM also in school aged children with and without ASD. Taking into consideration the theoretical notion that ToM may be more strongly related to hot EF than cool EF (Zelazo et al., 2005) we attempted to examine whether later ToM performance could be also predicted by early hot EF performance after controlling for potential covariates and cool EF and whether the association of hot EF variables to ToM was stronger in either controls or ASD, by including hot EF X ASD diagnosis interaction terms.

Method

Participants

Forty five (45) children with a clinical diagnosis of ASD (38 males) (M=9.07 years, SD=1.42) and thirty seven (37) controls (M=9.03 years, SD=1.17) (35 males) aged 7-11 years old were recruited to participate in the present study. They were followed up one year after the initial assessment (two time points). At the second time point all 82 children were followed up (0% attrition). All ASD participants were high functioning (IQ >70), held a clinical diagnosis by a qualified clinician using DSM-IV (American Psychiatric Association, 1994) or DSM-V (American Psychiatric Association, 2013) criteria and qualified for a "broad ASD" on the Autism Diagnostic Interview/Autism Diagnostic Interview-Revised (ADI/ADI-R; Le Couteur et al., 1989; Lord, Rutter, & Le Couteur, 1994) and/or the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) , in accordance to National Institute for Health and Clinical Excellence (NICE, 2011) guidelines. They were also in receipt of a Statement of Special Educational Needs (SEN), a legal document that details the child's needs and services that the local authority has a duty to provide, which specified ASD as their primary need. All clinical records were inspected and any individual lacking detailed information about the official source of diagnosis was excluded from the study. Additional exclusion criteria for the ASD group included the presence of a diagnosed psychiatric illness, comorbid conditions (i.e. ADHD, seizures or colour blindness) and Full Scale Intelligence Quotient (FSIQ) below 70 as determined by the abbreviated version of the Wechsler Intelligence scales (two subtests: vocabulary and matrix reasoning; Wechsler, 1999). Typically developing participants were required to have no diagnosis, and no family history of ASD or other mental health disorders, dyslexia or learning disability. Participants were matched for chronological age ($t\ (80) = -.13$, $p = .89$) and FSIQ ($t\ (80) = 1.73$, $p = .09$). Most of the participants were white from mixed socioeconomic backgrounds and spoke English as their first language. Specific data on the socioeconomic status or educational attainment levels of their parents were not recorded. Ethical approval for the study was obtained and all participants' parents/carers gave written informed consent (consistent with the Declaration of Helsinki) in compliance to the University Research Ethics Committee. Table 1 shows descriptive characteristics (means and standard deviations) of participants of both groups.

[Table 1 should be placed here]

Measures

Cool Executive Function

Inhibition. The 'R' and 'P' version of the Go/No-Go paradigm (Mueller & Piper, 2014) was used in the present study to assess participants' response inhibition. An image of either the letter P or letter R appeared in the centre of the screen (for 1500 milliseconds) on a black background. Participants were then instructed to press the button only when the letter P was shown (Go trials) and to avoid pressing it for the letter R (No-Go trials). On the second block of trials, the pattern was reversed and the participants were asked to press the button when the letter R appeared (Go trials) and to avoid pressing it when P was presented (No-Go trials) this time. There was no feedback provided after a correct or incorrect response. Before each block, participants first completed 10 practice trials followed by the actual 320 test trials. In order to measure participants' response inhibition, the proportion of incorrect No-Go trials was recorded. Lower scores indicated better performance.

Planning. Participants' planning ability was measured by the Tower of London (ToL) task (Shallice, 1982). As a practice, participants were presented with three 3-move problems, followed by the 12 actual trials of the original problem set (two 2-move tasks; two 3-move tasks; four 4-move tasks; and four 5-move tasks). Following the procedure of Monks et al.

(2005) and Poland et al. (2015), successful performance required participants to solve each problem moving only one bead each time and in the number of moves required; participants were given two minutes to complete each problem. The task was stopped when the participant completed all problems or failed two consecutively. In terms of the scoring, we measured the number of problems each participant completed successfully. One point was given to participants if they completed the problem successfully and 0 points if they failed to complete the problem. Scores ranged from 0 to 12.

Working Memory. The digit span forward and backwards subtests from the Wechsler Intelligence Scale for Children-3rd edition were employed to measure participants' verbal working memory (WISC-III; Wechsler, 1991). Participants were asked to recall the sequence presented by the researcher (at a rate of one number per second) in the exact same order. In the backwards digit recall task, the series of numbers should be repeated in reverse order. If participants responded successfully to all trials (4) within a block, the researcher proceeded to the next block. Each block included 2 trials at each span length. In terms of scoring, participants were awarded 1 point for each correct trial and the task was terminated when the participant failed both trials at any given span length. The sum of the points awarded for both the forward and backward subtest created a working memory score which was then converted into a standardised score.

Hot Executive Function

Affective Decision Making. A modified computerised version of the IOWA gambling task (IGT; Bechara et al., 1994) was employed to measure participants' affective decision making. Participants were presented with four decks of cards (A, B, C, and D) and were told they should select a card from any of the four decks each time. Decks A and B were equivalent in terms of overall net loss, whereas decks C and D were equivalent in terms of overall net winning. For each card selection, the wins and losses were set in a way that in every block of 20 cards from Decks A or B there was a potential total gain of £1,000, interrupted by potential losses up to £1,250. Losses were less frequent but of a larger magnitude in deck B whereas in Deck A losses were more frequent but in smaller amounts. For Decks C and D, the wins for each block were £500 totally while the potential net losses £250. In Deck D losses were less frequent and of higher magnitude relative to those in Deck C. Thus, Decks A and B were equally “disadvantageous” relative to Decks C and D that were equally “advantageous”. We measured whether participants made mainly more advantageous or disadvantageous decisions. Based on the approach used by Verdejo-Garcia et al. (2006), scores were calculated by subtracting the number of disadvantageous choices (decks A and B) from the number of advantageous choices (decks C and D) divided then by the total number of trials.

Delay Discounting. In line with previous research studying hot EF (Hongwanishkul et al., 2005; Prencipe et al., 2011), the Delay Discounting task was used in the present study in a computerised version to measure the extent to which participants do discount future rewards (Richards et al., 1999). This task originally included the forced-choice between different amounts of money after different delays or with different chances. However, as the task was being given to school aged participants it was decided to modify it and completely remove the probability questions. Participants were told that they had to choose (hypothetically) between an immediate amount of money or £10 available after a delay. The test consisted of about 70 such questions (i.e. (a) Would you rather have £10 for sure in 30 days or (b) £2 for sure right now?). The amount of immediate money was adjusted across trials until an amount was reached that was determined by previous choices as being equivalent to a delayed £10 reward; until the participant was indifferent between the two choices (random adjusting procedure; for more details see Richards et al., 1999). For every participant, this indifference point (the amount of immediate money judged to be equivalent to £10) signified the subjective value of the delayed large reward (Richards et al., 1999). Delay discounting was determined by five delays (0, 10, 30, 180, and 365 days later). In terms of scoring we followed the procedure described in Myerson et al. (2001), where the indifference points were used to estimate delay discounting. Thus, indifference points were established within participants and were plotted against time (delay). Indifference points and delays were normalised, by expressing indifference points as proportions of the amount of the maximum delayed reward (£10) and the delays as proportions of the maximum delay (365 days). These normalised values were used as the x (delay) and y (indifference points) axes in order to plot the discounting function. Four separate trapezoids were then created by drawing vertical lines from each data point on the x axis. The formula $(x_2-x_1) \cdot [(y_1 + y_2)/2]$ was used to calculate the area of each trapezoid. The areas under these discounting curves (AUC) were calculated by summing the resulting trapezoids.

Theory of Mind

False belief. In order to measure participants' false belief understanding, the Sandbox Task (Begeer et al., 2012) was used. Participants were told that this task was about a father and a daughter (Sanne) planting flower bulbs in a sandbox. The researcher showed the participants the picture of a sandbox and told them the father decided to bury the flower bulb at the location of the cross. When the father went away to bring a watering can, Sanne decided to move the flower bulb and bury it in a different location. Before asking the false belief question, the researcher asked whether participants had a good look at the pictures and remembered where each character (Sanne and her dad) placed the flowerbulb, in order to ensure that they have the requisite attentional or memory capacities necessary to demonstrate their theory of mind knowledge (control question). The researcher then asked participants one false belief question: "When Sanne's dad comes back with the watering can where will he give water to the flower bulb? You have to draw a cross". In terms of scoring, the difference between the original hiding location of the flower bulb (0 mm) and the location where participants indicated dad would look for it was measured (in millimetres). If participants indicated a location towards the direction of Sanne's hiding location of the flower bulb (63 mm) they received a positive bias score. If participants indicated a location in the opposite direction of the flower bulb, to the right of the original hiding location, they received a negative bias score. Lower scores indicated better performance. This paradigm was employed because, as the object is buried and reburied in the sandbox, a continuum is created between locations in contrast to the categorical approach of the classic false belief task that has been found to sometimes omit the subtle variance in false belief reasoning at different ages (Bernstein et al., 2011).

Emotion recognition

The Reading the Mind in the Eyes test (children's version; Baron-Cohen et al., 2001) was used to assess mental state/ emotion recognition. This task was chosen as it is a widely used ToM test that measures the ability to decode the feelings and thoughts of others from the eyes. The test can also be considered an emotion recognition test (Vellante et al., 2013). It consists of photographs of the eye regions of 28 faces. Participants were asked to make a choice between four words presented at the bottom of the page on which each picture appeared, choosing the one that best described what the person of the photograph was feeling or thinking. Successful performance required participants to select the correct mental and emotional state. Participants were asked to choose one of the terms even if they said that any term was quite right, thus conforming to a forced-choice procedure. One point was given to each correctly reported response.

Procedure

The current study followed a longitudinal design with two time points, approximately one year apart. At both assessment points children undertook the tests individually across two sessions (each lasted 40 minutes) with a female researcher in a quiet space at their school. All tasks were addressed in a fixed order across the two assessment sessions: session 1: IQ test, Sandbox task, Digit Span, ToL, and Eyes Test; session 2: Go/No-Go, IGT, delay discounting test.

Results

Statistical analyses were performed using SPSS-23®. All variables were checked for normality and homogeneity assumptions of parametric tests. No extreme outliers were found. Analyses was conducted at two levels. The developmental changes in cool & hot EF and ToM of children with ASD relative to neurotypical controls across the 12 months were firstly examined. Secondly the predictive relation between hot & cool EF and ToM across the two time points was investigated. No violations of multivariate assumptions for these variables were found. All tests were two-tailed and statistical significance was set at $p < .05$.

Comparison of the developmental changes in EF and ToM between the two groups

Table 2 presents descriptive statistics for hot & cool EF and ToM at each time point. The developmental changes in children's EF and ToM abilities across time points were examined by carrying out a mixed ANOVA. Time was set as the within-subject factor (T1, T2) and Group as the between-subject factor (ASD or control). Post hoc tests were not performed for Group because there were fewer than three groups. Within group comparisons were assessed by paired sample t-tests. [Table 2 should be placed here]

Working memory (digit span). Significant main effects of time F [(1, 80) = 50.6, p = .001, $ŋ_p^2 = .39$] and group [F (1, 80) = 13.52, p < .001, $ŋ_p^2 = .15$] were found. The interaction between time and group was also significant, [F (1, 80) = 11.73, $p = .001$, $\eta_p^2 = .13$]. Figure 1 presents mean working memory score from T1 to T2 for each group. Planned comparisons demonstrated that the ASD group showed poorer performance in working memory than neurotypicals both at Time 1, F (1, 80) = 17.09, $p < .001$, $\eta_p^2 = .18$, and at Time 2, F (1, 80) = 6.24, $p = .02$, $\eta_p^2 = .07$. Further analyses showed that children’s performance improved significantly over time in both groups; ASD: t (45) = -7.96, $p < .001$; controls: t (37) = -2.45, $p = .019$. The significant interaction lays in the pattern of improvements over time. While both groups demonstrated developmental changes after 12 months, improvements for the ASD group were steeper (see figure 1). The ASD group demonstrated a poorer performance on digit span scores throughout this developmental period.

Planning (Tower of London). The main effect of time [F (1, 80) = 86.65, $p < .001$, $\eta_p^2 = .52$], was found significant. Neither the main effect of group, [F (1, 80) = 1.58, $p = .21$, $\eta_p^2 = .2$], nor the interaction between time and group, [F (1, 80) = 1.67, $p = .2$, $\eta_p^2 = .02$], were found significant. Figure 1 presents mean planning score from T1 to T2 for each group. Planned comparisons demonstrated that the ASD group showed equal performance in planning to neurotypicals both at Time 1, F (1, 80) = .21, $p = .65$, $\eta_p^2 = .003$, and at Time 2, F (1, 80) = 3.62, $p = .061$, $\eta_p^2 = .043$. Further analyses showed that children’s performance improved significantly over time in both groups; ASD: t (45) = -7.10, $p < .001$; controls: t (37) = -6.14, p < .001. Figure 1 reveals the similar developmental improvements after 12 months.

Inhibition (go/no-go). The main effect of time on inhibition, [F (1, 80) = 14.61, $p <.001$, $\eta_p^2 = .15$], and group [F (1, 80) = 9.68, $p = .003$, $\eta_p^2 = .11$], were found significant. No significant interaction between time and group, [F (1, 80) = 2.48, $p = .12$, $\eta_p^2 = .03$] was found. Figure 1 presents mean inhibition score from T1 to T2 for each group. Planned comparisons demonstrated that the ASD group showed poorer performance in inhibition than neurotypicals both at Time 1, F (1, 80) = 7.32, $p = .008$, $\eta_p^2 = .08$, and at Time 2, F (1, 80) = 12.39, $p = .001$, $\eta_p^2 = .13$. Further analyses showed that children's performance improved significantly over time in both groups; ASD: t (45) = 4.3, $p < .001$; controls: t (37) = 2.52, $p = .015$; Both groups demonstrated developmental performance gains after 12 months but the ASD group demonstrated a poorer performance on the Go/No-Go task throughout this developmental period.

[Put Figure 1 here]

Affective decision making (IOWA). The main effect of time [F (1, 80) = 17.33, $p < .001$, $\eta_p^2 = .18$] and group [F (1, 80) = 4.76, $p = .03$, $\eta_p^2 = .06$] were found to be significant. The interaction between time and group, [F (1, 80) = .46, $p = .49$, $\eta_p^2 = .01$] was not significant. Figure 2 presents mean affective decision making from T1 to T2 for each group. Planned comparisons demonstrated that the ASD group showed poorer performance in affective decision making than neurotypicals both at Time 1, F (1, 80) = 4.31, $p = .041$, $\eta_p^2 = .05$, and at Time 2, F (1, 80) = 4.03, $p = .048$, $\eta_p^2 = .05$. Further analyses showed that children's performance improved significantly over time in both groups; ASD: t (45) = -2.91, $p = .006$; controls: t (37) = -4.03, $p = .001$. Figure 2 shows that both groups demonstrated developmental performance improvements after 12 months, but with ASD children showing a poorer performance throughout this developmental period.

Delay Discounting (delay discounting task). The effect of time [F (1, 80) = .09, $p = .76$, $\eta_p^2 = .001$], group [F (1, 80) = 3.34, $p = .071$, $\eta_p^2 = .04$], and the interaction between time and group, [F (1, 80) = 1.91, $p = .17$, $\eta_p^2 = .02$], were not found significant. Figure 2 presents mean delay discounting score from T1 to T2 for each group. Planned comparisons demonstrated that the two groups showed equal performance in delay discounting both at Time 1, F (1, 80) = 5.15, $p = .06$, $\eta_p^2 = .06$, and at Time 2, F (1, 80) = .46, $p = .49$, $\eta_p^2 = .006$. Further analyses showed that children’s performance did not improve significantly over time in either group; ASD: t (45) = 1.41, p = .17; controls: t (37) = -.65, p = .52. Neither group presented developmental changes after 12 months.

[Put Figure 2 here]

False belief (Sandbox). The main effect of time [F (1, 80) = 4.29, p = .04, η_p^2 = .05] was found significant. The effect of group [F (1, 80) = 2.12, p = .15, η_p^2 = .03] was not significant while interaction between time and group, [F (1, 80) = 8.54, p = .005, η_p^2 = .09] was found significant. Figure 3 presents mean false belief score from T1 to T2 for each group. Planned comparisons demonstrated that the ASD group showed poorer performance in false belief than neurotypicals only at Time 1, F (1, 80) = 6.18, p = .015, η_p^2 = .07, but equal scores at Time 2, F (1, 80) = .06, p = .8, η_p^2 = .001. Further analyses showed that children’s performance improved significantly over time only in the ASD group; ASD: t (45) = 3.32, p = .002; controls: t (37) = -.69, p = .49.

Mental state/ emotion recognition (Reading the Mind in the Eyes). The main effects of time [F (1, 80) = 92.04, p < .001, η_p^2 = .54], group [F (1, 80) = 12.13, p = .001, η_p^2 = .13], and the interaction between time and group, [F (1, 80) = 4.84, p = .03, η_p^2 = .06] were all found significant. Figure 3 presents mean mental state/ emotion recognition score from T1 to T2 for each group. Planned comparisons demonstrated that the ASD group showed poorer performance in mental state/ emotion recognition than neurotypicals both at Time 1, F (1, 80) = 16.43, p < .001, η_p^2 = .17, and at Time 2, F (1, 80) = 5.52, p = .021, η_p^2 = .07. Further analyses showed that children’s performance improved significantly over time in both groups; ASD: t (45) = -8.62, p < .001; controls: t (37) = -5.11, p < .001. Figure 3 shows that both groups demonstrated an improved performance on mental state/emotion recognition after 12 months, but the ASD group showed a steeper developmental change (source of interaction). The ASD group’s poorer performance insisted throughout this developmental period. [Put Figure 3 here]

Longitudinal relations between EF and ToM in children with and without ASD

Pearson's correlations were conducted to investigate the relation between hot & cool EF and ToM across the two time points (see table 3) in children with and without ASD. Table 3 shows that cool EF (digit span, ToL, Go-No Go scores) and both ToM tasks were significantly correlated at both time points. In terms of hot EF, only delay discounting was significantly related to ToM Eyes Test whereas performance on IOWA (affective decision making) was not significantly related to ToM at any of the time points. [put table 3 here]

Following the correlational analysis, the predictive association between EF and ToM in children with and without ASD was examined by running two series of hierarchical regression models. The first series of regressions investigated whether early EF predicted later ToM (at T2). Block 1 of predictors included ASD diagnosis, concurrent age, concurrent Full Scale Intelligence Quotient (FSIQ) and early ToM (control variables). Block 2 of predictors introduced the individual cool EF skills in order to examine the predictive role of EF on ToM over and above control variables. Block 3 included the hot EF (only delay discounting; IOWA scores were not included in the regression models as they did not correlate with any ToM task at either time point) in order to assess whether hot EF can predict ToM over and above control variables and cool EF. Finally Block 4 of predictors included the hot delay discounting X ASD diagnosis interaction term, computed from the cross-product of effect-coded ASD (-1=ASD, +1=Control) and the centred hot delay discounting scores (Aiken & West, 1991), in order to examine whether the association of hot EF to ToM was stronger in either controls or ASD. The second series of regressions investigated whether early ToM predicts later EF (at T2), after controlling for concurrent age and FSIQ, and early EF.

Due to the large number of predictors (n=13) and the relatively small sample size (see appendix for full results of regression analysis and reports of collinearity diagnostics; VIFs and Tolerance values), it was decided to set the cut off for significance at p < .01 rather than p < .05, for the following regressions, to counteract the possibility of Type I error (Robinson et al., 2016). Although this threshold is not as stringent as a standard Bonferroni correction, the latter has been widely criticized for being problematically conservative, particularly for exploratory research such as the following regression analysis (Perneger, 1998). Thus we assumed that p < .01 was satisfactory for this analysis; stringent but not overly so.

EF predicting later ToM mental state/ emotion recognition (EyesTest at T2):

The first block of predictors (ASD diagnosis, concurrent age & FSIQ, and early Eyes Test scores (T1) contributed significantly to the variance of the later Eyes Test scores, F (4, 77) = 22.11, p < .001, explaining 53.5% of the variance. Neither for cool EF entered in block 2 [F (5, 72) = 0.97, p =.44] nor hot EF in block 3 [F (2, 70) = .66, p = .52], significant additional variance was explained. Finally for the Hot delay discounting X ASD interaction terms entered in block 4, the total variance explained rose to 66.2% [F (2, 68) = 9.07, p < .001]. Later ToM Eyes Test scores (T2) were significantly predicted by the early (T1) (p = .002) and later (T2) (p = .002) hot delay discounting X ASD diagnosis interactions terms. Thus, early and later delay discounting predicted later ToM mental state/ emotion recognition only in school aged children with ASD.

EF predicting later ToM false belief (Sandbox scores at T2)

The first block of predictors (ASD diagnosis, concurrent age & FSIQ, and early false belief scores (T1) contributed significantly to the variance of the ToM false belief ability, F (4, 77) = 6.52, p < .001, explaining 25.3% of the variance. For cool EF entered in block 2 the total variance explained rose to 39.8%, representing a significant increase of 14.5% [F (5, 72) = 3.48, p =.007] additional variance explained. Neither for hot EF entered in block 3 [F (2, 70) = 0.31, p =.74] nor for the hot delay discounting X ASD interaction entered in block 4, [F (2,

68) = 2.07, *p* =.13] significant additional variance was explained. Later ToM false belief scores (T2) were significantly predicted by early working memory (Digit span T1) (*p* = .0004) and later working memory (Digit span T2) (*p* = .0002) overall in school aged children with and without ASD.

Early ToM predicting later hot and cool EF skills

None of the regression models with early ToM skills predicting each one of the individual later hot and cool EF skills were significant. Early ToM skills did not predict later EF in school aged children with and without ASD.

Discussion

The present study examined the developmental changes of cool and hot EF and their associations to ToM across one year in school aged children with and without ASD. This longitudinal analysis demonstrated that for children with ASD, selective aspects of EF (working memory, inhibition, and affective decision making) presented significant age-related gains after one year but their impairments, present from the initial assessment, remained throughout development without reaching the levels of neurotypicals. For "cool" planning, ASD participants showed equal performance and the same developmental gains relative to controls while for "hot" delay discounting there were no deficits or developmental changes found in ASD or typical development. With regards to ToM abilities, the mental state/ emotion recognition presented age-related improvements but demonstrated a pattern deviant to controls as children's deficits remained present across time in ASD. For false belief, our results suggested it followed a delayed development in ASD. Second, results of the longitudinal association between cool and hot EF and ToM in school aged children with and without ASD revealed that selective early aspects of EF (working memory, and delay discounting) predicted later ToM abilities which supports the well documented theoretical account of early EF predicting later ToM. No evidence was found to support the argument that ToM abilities predict later EF. Cool EF working memory predicted later ToM false belief overall in children with and without ASD, while hot delay discounting predicted later ToM mental state/ emotion recognition over and above cool EF and control variables only in the ASD participants. This is the first study to date to investigate the developmental changes of both cool and hot EF across time and report that early hot delay discounting predicts later ToM above and beyond cool EF in school age in ASD. These findings highlight the multidimensional nature of EF and how its influence on other developmental landmarks such as ToM may increase our understanding of the higher-order cognitive deficits that underpin social interaction problems in ASD.

Development of EF and ToM across time in school age

Our results demonstrated that children with ASD showed age-related improvements in all cool EF aspects suggesting that during school age specific aspects of EF (working memory, inhibition, planning) present developmental gains in ASD. These results support previous evidence of performance gains in EF during childhood both in typical development (Carlson et al., 2013; Gur et al., 2012) and in ASD (Pellicano, 2010). These findings contradict two of the three previous longitudinal EF studies in ASD that reported no developmental improvements in EF, either in young children (Griffith et al., 1999) or adolescents (Ozonoff & McEvoy, 1994). Both of these studies indicated very few EF changes across time and suggested that probably there is a ceiling on the development of such cognitive abilities in ASD. In line with the third longitudinal EF study though (Pellicano, 2010), our findings paint a more positive picture of children's with ASD EF developmental trends, indicating perhaps the likelihood of a window of plasticity in ASD as well. Notably, the contradicting studies (Griffith et al., 1999; Ozonoff & McEvoy, 1994) included ASD participants much less able than in our study that could account for their lack of significant developmental changes. However, it should be noted that despite the reported developmental changes in cool working memory and inhibition, children with ASD presented impairments in these aspects relative to matched neurotypicals which remained present across development and never reached the performance level of the control group. This evidence in a way supports Happé et al.’s (2006b) proposition that there may be a particular profile of “coexisting cognitive atypicalities” in ASD that pertain across development. One could argue that this could suggest that children with ASD might eventually reach a performance ceiling in some EF aspects as Ozonoff and McEvoy (1994) implied in their study. Our data failed to provide more evidence about this issue as the present longitudinal design included only two time points and the sample did not include adolescents or young adults which could shed more light on the maturity peaks of EF in ASD (if they ever develop up to the same level as controls). The present data indicate thus that despite the significant age-related improvements in working memory and inhibition, the performance of the ASD group never reached that of controls which in a way implies *deviant* development across the age range of the present sample. The suggested deviant development should not allow though for the present data to be overlooked. More specifically, it is worth considering the present data from the maturation processes perspective of Luna et al. (2007) that proposed that if deficits in EF persist across development, it could imply that impairments in the underlying EF brain mechanisms are not related to the brain developmental/ maturation processes. This, in conjunction with the emerging developmental improvements we found here, could suggest that the developmental processes may be intact for ASD participants across this specific age range (middle childhood). The fact that planning ability presented an intact profile and similar developmental improvements to the control group (no deviance) as also shown in a previous study (Happé et al., 2006a) could actually add more support to this notion that, at least in school age, developmental/ maturation processes of selective cool EF are intact in ASD. The developmental pattern of cool EF gains in school age could be explained in relation to the prefrontal cortex (the underlying brain region of EF) undergoing substantial maturation during this period (Otero & Barker, 2014). School age is a crucial period of rapid developmental improvements and increased cognitive demands where children have to process and understand both their own sense of self and their sense of others as well as learning to interact effectively with the world around them (Siegel, 2013) which could justify these advances.

School age is an important developmental period demanding not only cool EF improvements, but also hot EF gains, as the social contexts children are faced with involve advanced emotional and motivational processes. To our knowledge, this is the first study to examine the developmental changes of hot EF in ASD across time and showed that only affective decision making presented significant age-related gains in school aged children with ASD. Our results are in line with previous research (Hooper et al., 2004; Prencipe et al., 2011) that reported age-related performance gains in affective decision making (Iowa Gambling Task) across childhood and adolescence as well as with developmental theories proposing that the development of hot EF would be protracted to the extended development of the underlying brain region that is the ventromedial prefrontal cortex, across school age (Segalowitz & Davies, 2004). As with the cool EF developmental framework discussed above, despite the emerging developmental improvements, children with ASD presented deficits in affective decision making relative to the control group that did not become less marked with age (*deviant* to control group development across middle childhood). The emerging developmental gains in this hot EF aspect, despite the persisting deficits, highlight the importance of this finding since it suggests that the developmental/maturation processes of the brain structures underpinning selective hot systems in ASD continue across school age. The cognitive maturation processes of the areas of ventromedial prefrontal cortex regulating affective decision making seem to progress across school age in ASD. It should be noted at this point that our assumption that the persisting performance deficits of the ASD group (despite the developmental improvements found) suggest a deviant rather than a delayed pattern should be treated with caution. We have made sure to clarify at all relevant points in discussion above that these suggested as deviant developmental patterns apply only to the specific age range employed in the present study. It is likely that group differences may eventually lessen with further age if performance was to be assessed at an older age for example (late adolescence or young adulthood).

Contrary to the significant results of cool EF and hot affective decision making, hot delay discounting demonstrated an intact profile (in ASD) and non-significant developmental changes neither in ASD nor typical development. These results contradict previous cross sectional evidence having indicated age-related improvements in the delay discounting task (Scheres et al., 2006) across childhood and adolescence. Scheres et al. (2006) employed a temporal and probabilistic discounting task that had a differentiated design with shorter time delays, smaller immediate monetary rewards as well as levels of probability for the delayed reward. Besides this, the monetary awards they offered were real, contrary to ours being hypothetical due to the impractical cost and ethical issues raised within the school contexts. Such measures discrepancies could have perhaps made their older participants more motivated to wait for the larger rewards during the task relative to the younger ones resulting in the reported age-related gains. As the differentiated designs of delay discounting tasks are multifaceted and with levels of difficulty/complexity, we could assume that our delay discounting task dimensions were not developmentally sensitive enough to capture subtle age-related differences across school age in either group. Steinberg et al. (2009) in fact suggest that differentiated task designs do not likely follow the same developmental pathway. Besides task discrepancies though, the lack of significant developmental changes in hot delay discounting relative to the other hot aspect (affective decision making) is quite surprising and raises questions about their underlying brain structures and developmental course. One should expect the delay discounting trajectory to progress across school age as children are in greater need of their impulsivity control (tapped by delay discounting) within the more demanding social and educational settings. As findings from early childhood (Zelazo & Carlson, 2012) and adolescence (Scheres et al., 2014) have both demonstrated age-related gains, it makes us consider the possibility that delay discounting may present a developmental pause during middle childhood (6/7-11) years before continuing to progression in adolescence. This assumption needs to be cautiously interpreted as the present design did not include a third or more time points across school age and adolescence that could clarify this issue. Finally, as several measures thought to tap hot EF have been criticized for lacking enough "heat" or not being ecologically valid (Welsh & Peterson 2014), one potential explanation behind the lack of significant developmental changes here could be the delay discounting task not being so hot for this specific age range. One could argue that differing quantities of imaginary money is quite an abstract construct that failed to enhance young children's motivation or increase their sensitivity to money loss.

Regarding the developmental course of ToM in ASD, discussion will not go in great depth as our main focus in the present study was its longitudinal association to EF across school age. ToM mental state/ emotion understanding ability and false belief understanding ability both made substantial progress in ASD across the 1-year period and expand relevant findings from longitudinal studies in preschool period (Pellicano, 2010; Steele et al., 2003) to school age as well. The ASD performance in mental state/ emotion recognition (Eyes task) never reached up the level of the control group despite the age-related gains which implies that, as with cool EF described above, the developmental trend of this ToM ability for this specific age range could be argued as *deviant* of the typical development one. Once again, such an assumption is not warranted as group differences may eventually not be significant at an older age (after middle childhood). The small but significant changes however highlight that ToM gains, should they occur, may be present beyond preschool period in ASD. In contrast to this, false belief understanding (Sandbox task) in ASD presented a differentiated developmental course as impairments were present only at the initial assessment point. Age-related gains emerged only in the ASD group, with false belief performance reaching up the controls' level at the second time point. This could imply that the false belief ability in ASD presents a *delayed* developmental pattern in our study. Failure to report developmental gains in the control group is due to ceiling effect (Sandbox task probably not sensitive enough to developmental trends of typical development). It seems that the ASD heterogeneity and the unique ToM profiles of distinct tasks cannot allow for neither the delayed nor the deviant development hypothesis to fully explain the ToM deficits in ASD. Our data support Baron-Cohen's (1991) proposition that ToM development in ASD fits a hypothesis of both deviance and delay.

Longitudinal associations between hot and cool EF and ToM across time in school age

Our results showed that EF and ToM are developmentally linked across school age. In line with the vast majority of previous studies (see for a review Devine & Hughes, 2014), early EF predicted later ToM rather than the reverse pattern (early ToM predicting later EF). Therefore these findings add more to the theoretical account suggesting that emerging EF in childhood is a potent, although not exclusive platform for the development of ToM both in typical development (Flynn, 2007; Hughes, 1998) and ASD (Pellicano, 2010). The emergence account of ToM posits that early EF skills predict later ToM; thus children would first need to obtain sufficient EF skills and then understand and process ToM false beliefs or mental states **(**Russell, 1996**).** Indeed, after controlling for concurrent age, FSIQ, and prior ToM, we found that early working memory predicted later ToM false belief while early delay discounting predicted later

ToM mental state/emotion recognition in ASD. These findings thus do not support Perner’s (1998) proposition that the acquisition of ToM is a prerequisite of children’s, according to which, longitudinal predictions from earlier EF are not expected for a ToM task.

In line with our hypothesis and previous cross sectional and longitudinal studies that presented associations between cool EF and ToM in typical development and ASD (Carlson & Moses, 2001; Pellicano, 2007; Kimhi et al., 2014), we found that later ToM false belief was predicted by early cool working memory. Children between 3 and 5 years of age present dramatic and rapid improvements in EF and ToM (Anderson, 2008) but the present findings suggest that developmental changes in ToM mechanisms across school age (beyond 5 years) may require EF to facilitate the emergence of more sophisticated ToM abilities. Advanced needs for cognitive executive control during school age are more than expected as children have to maintain and manipulate new, complex knowledge while socially interacting with their environments (Del Giudice, 2014). This evidence supports the working memory hypothesis according to which the working memory development is an important factor influencing children's developing understanding of false belief (Davis & Pratt, 1995) in early childhood, as replicated by other studies too (Gordon & Olson, 1998; Keenan et al., 1998). It could be then argued that also across school age children with or without ASD need a heavier (improved) cognitive load on working memory towards a successful ToM development. Working memory and inhibition are generally considered central to the EF-false belief relation (Carlson et al., 2002; Devine & Hughes, 2014), a notion for which we provided limited support. Contradicting prior studies in early childhood (Carlson et al., 2013; Carlson et al., 2004), inhibition did not predict later ToM false belief in school age in our study. One could thus argue that inhibition may be more central to the emergence of false belief at the early years of childhood (Tillman et al., 2015) and as ToM abilities progress across childhood, other EF may be more central to the development of ToM.

Another important finding of the present study was the significant longitudinal predictive association found between hot delay discounting and ToM mental state/ emotion recognition, over and above cool EF and control variables in ASD. Current findings corroborate to an extent that there may be a developmental relation of the underlying brain mechanisms of selective hot processes and ToM present in ASD across time. Delay discounting could be linked with the emergence of ToM as, in order for children with ASD to understand the mental states of others, “hot” motivational or emotional processes need to be evoked (Zelazo et al., 2005) across school age. This emerging association between two seemingly unrelated constructs suggests that the ability of school aged children with ASD to disengage from the present while considering more long-term goals/ temporal perspectives (delay discounting) may provide a platform for the development of one’s emotion understanding ability. Stolarski (2011) has also suggested that emotional functioning is linked with the development of temporal perspectives (i.e. delay discounting). As this longitudinal association between delay discounting and mental state/ emotion recognition was found only in the ASD group, it could imply there is a specificity in the relation between ToM and this hot executive process in ASD. However as delay discounting did not predict later false belief understanding (the other ToM task), this assumption has to be examined cautiously. The developmental association between these delay related motivational processes and ToM mechanisms being stronger in the ASD group needs to be tackled by future imaging studies investigating the structure of the underlying brain regions. For example, previous fMRI research in clinical population showed an association between ADHD and activity in the ventral striatum (brain region responsible for preference for small sooner rewards over large later rewards; McClure et al., 2004) during reward anticipation in delay tasks (Scheres et al., 2007; Ströhle et al., 2008). Relevant research could perhaps clarify if that could also be the case for ASD.

The present findings need to be interpreted cautiously in the light of limitations and be also corroborated with results from larger longitudinal studies with more than two time points across school age. First, as the present study followed the approach of the convenience sampling, the sample of the ASD participants was quite small (e.g. the between-groups IQ statistical difference of $p = 0.09$ indicates a medium effect size) and may not reflect the broader ASD population. Furthermore the participants of the present study were aged between 7-11 years old only. It thus remains to be examined whether these results can be generalised to younger children, adolescents or adults across the spectrum. The fact that we did not include a validated screening measure to corroborate the clinical diagnostic reports was another important limitation. Finally, it should be noted that clear conclusions about the longitudinal association between hot EF and the broader ToM mechanism cannot be drawn as early hot delay discounting predicted only one of the two later ToM measures addressed here. Moreover, these two ToM measures tap only some of the various ToM skills and cannot be considered as the only crucial ToM measures. Future longitudinal research thus could investigate the impact of hot EF to several other ToM tasks such as the strange stories (Happé, 1994), second-order ToM (Perner & Wimmer, 1985) or the Faux Pas test (Baron-Cohen et al., 1999).

In conclusion, the present study demonstrated that for children with ASD, selective cool and hot EF skills, and ToM abilities presented significant developmental changes across time in school age. These data highlight the need to shed more light on the underlying brain structures as the reported impairments in EF are likely not related to the maturation processes. Furthermore, our data provided more to the theoretical account that cool EF influences the development of ToM and not vice versa in ASD and typical development, while expanding these longitudinal associations of ToM to hot EF as well suggesting that specific hot EF skills (delay discounting) also provide a platform for the emergence of ToM across school age in ASD. Although research into hot EF in childhood has recently received increased attention, knowledge of its developmental trajectory still lags behind that of cool EF both in typical development and ASD, mainly due to the limited number of tasks that actually tap hot EF skills. Future research should direct more attention towards the development of more relevant tasks to measure these hot EF skills. Finally, findings of specific EF predicting later ToM contributed support to an emergence account (Russell, 1996, 1997) in typical development and ASD. Studying the developmental trends of hot and cool EF and their longitudinal associations to ToM may aid in gaining a greater understanding of the link between cognition and behaviour in typical development and of the development of higher-order cognitive impairments being a risk factor for poor developmental/ social outcomes in children with ASD. Our findings highlight the need to address both hot and cool EF in clinical practice as they could contribute more towards future diagnosis or intervention projects.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, London: Sage.

Alloway, T. P, Gathercole, S. E, Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child development, 80*(2), 606-621. doi: 10.1111/j.1467-8624.2009.01282.x.

Andersen, P. N., Hovik, K. T., Skogli, E. W., Egeland, J., & Oie, M. (2013). Symptoms of ADHD in children with high-functioning autism are related to impaired verbal working memory and verbal delayed recall. *PloS One, 8*(5), e64842. doi: org/10.1371/journal.pone.0064842

Anderson, P. J. (2008). Towards a developmental model of executive function. In V. Anderson, R. Jacobs & P. J. Anderson (Eds.), *Executive Functions and the Frontal Lobes: A Lifespan Perspective* (pp. 3–23). New York, USA: Psychology Press.

Anderson, V., Anderson, P. J., Jacobs, R., & Spencer-Smith, M. (2008). Development and assessment of executive function: From preschool to adolescence. In V. Anderson, R. Jacobs & P. J. Anderson (Eds.), *Executive functions and the frontal lobes: A lifespan perspective* (pp. 123–155). New York, USA: Psychology Press.

American Psychiatric Association (APA). (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington DC: Author.

American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Apperly, I. A., Samson, D., & Humphreys, G. W. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental psychology*, *45*(1), 190-201. doi: org/10.1037/a0014098.

Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: a 1-year longitudinal perspective. *Frontiers in psychology*, *5*, 655. doi: 10.3389/fpsyg.2014.00655.

Baron-Cohen, S. (1991). The development of a theory of mind in autism: deviance and delay? *The Psychiatric Clinics of North America*, *14*(1), 33-51.

Baron-Cohen, S., O'Riordan, M., Jones, R., Stone, V. E. & Plaisted, K. (1999). A new test of social sensitivity: Detection of faux pas in normal children and children with Asperger syndrome. *Journal of Autism and Developmental Disorders*, *29*, 407-418.

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

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to prefrontal cortex. *Cognition*, *50*, 7–15. doi: org/10.1016/0010-0277(94)90018-3.

Begeer, S., Bernstein, D. M., van Wijhe, J., Scheeren, A. M., & Koot, H. M. (2012). A continuous false belief task reveals egocentric biases in children and adolescents with Autism Spectrum Disorders. *Autism, 16*(4), 357-366. doi: org/10.1177/1362361311434545.

Bernstein, D. M, Thornton, W. L., & Sommerville, J. A. (2011). Theory of mind through the ages: Older and middle-aged adults exhibit more errors than do younger adults on a continuous false-belief task. *Experimental Aging Research*, *37*, 481–502. doi: org/10.1080/0361073X.2011.619466.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child development*, *81*(6), 1641-1660. doi: 10.1111/j.1467-8624.2010.01499.x.

Bock, A. M., Gallaway, K. C., & Hund, A. M. (2015). Specifying links between executive functioning and theory of mind during middle childhood: Cognitive flexibility predicts social understanding. *Journal of Cognition and Development*, *16*(3), 509-521. doi: org/10.1080/15248372.2014.888350.

Carlson, S. M., Claxton, L. J., & Moses, L. J. (2013). The Relation between executive function and theory of mind is more than skin deep. *Journal of Cognition and Development*, *16*, 186–197. doi: org/10.1080/15248372.2013.824883.

Carlson, S., Mandell, D. J., & Williams, J. L. (2004). Executive function and theory of mind: Stability and prediction from ages 2 to 3. *Developmental Psychology*, *40*, 1105–1122. doi: org/10.1037/0012-1649.40.6.1105.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*, 1032–1053. doi: 10.1111/1467-8624.00333.

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development*, *11*(2), 73-92. doi: 10.1002/icd.298.

Chan, R. C., Shum, D., Toulopoulou, T., & Chen, E. Y. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of clinical neuropsychology*, *23*(2), 201-216. doi: org/10.1016/j.acn.2007.08.010.

Chen, S. F., Chien, Y. L., Wu, C. T., Shang, C. Y., Wu, Y. Y., & Gau, S. S. (2016). Deficits in executive functions among youths with autism spectrum disorders: an age-stratified analysis. *Psychological medicine*, *46*(8), 1625-1638. doi: org/10.1017/S0033291715002238.

Christ, S. E, Holt, D. D., White, D. A., & Green, L. (2007). Inhibitory control in children with autism spectrum disorder. *Journal of autism and developmental disorders*, *37*(6), 1155-1165. doi: org/10.1007/s10803-006-0259-y.

Craig, F., Margari, F., Legrottaglie, A. R., Palumbi, R., de Giambattista, C., & Margari, L. (2016). A review of executive function deficits in autism spectrum disorder and attention-deficit/hyperactivity disorder. *Neuropsychiatric Disease and Treatment*, *12*, 1191–1202. doi: 10.2147/NDT.S104620.

Davis, H. L., & Pratt, C. (1995). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology, 47*(1), 25-31.

Del Giudice, M. (2014). Middle childhood: An evolutionary-developmental synthesis. *Child Development Perspectives, 8*(4), 193-200. doi: 10.1111/cdep.12084.

Demetriou, E. A., Lampit, A., Quintana, D. S., Naismith, S. L., Song, Y. J. C., Pye, J. E., Hickie, I., Guastella, A. J. (2017). Autism spectrum disorders: a meta-analysis of executive function. *Molecular Psychiatry.* doi: 10.1038/mp.2017.75.

Devine, R. T. & Hughes, C. (2013). Silent films and strange stories: Theory of mind, gender and social experiences in middle childhood. *Child Development, 84,* 989 - 1003. doi: 10.1111/cdev.12017.

Devine, R. T., & Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: A meta-analysis. *Child Development, 85,* 1777–94. doi: 10.1111/cdev.12237.

Devine, R. T., White, N., Ensor, R., & Hughes, C. (2016). Theory of Mind in Middle Childhood: Longitudinal Associations with Executive Function and Social Competence. *Developmental Psychology, 52* (5), 758-771. doi: org/10.1037/dev0000105.

Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science,* 331-338. doi: 10.1111/j.1467-7687.2009.00888.x.

Elliott, R. (2003). Executive functions and their disorders Imaging in clinical neuroscience. *British medical bulletin, 65*(1), 49-59. doi: org/10.1093/bmb/65.1.49.

Flynn, E. (2007). The role of inhibitory control in false belief understanding. *Infant and Child Development, 16*(1), 53-69. doi: 10.1002/icd.500.

Geurts, H. M., de Vries, M., & van den Bergh, S. F. (2014). Executive functioning theory and autism. In: S. Goldstein & J. Naglier (Eds.), *Handbook of executive functioning* (pp.121-141). New York: Springer.

Gioia, G. A., Isquith, P. K., Kenworthy, L., & Barton, R. M. (2002). Profiles of everyday executive function in acquired and developmental disorders. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence, 8*(2), 121–37. doi: org/10.1076/chin.8.2.121.8727.

Goldman, A. I. (2012). Theory of mind. In: E. Margolis, R. Samuels & S. Stich (Eds.), *The Oxford handbook of philosophy of cognitive science* (pp. 402-424). Oxford: Oxford University Press.

Gordon, A. C., & Olson, D. R. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *Journal of Experimental Child Psychology*, *68*(1), 70-83. doi: org/10.1006/jecp.1997.2423.

Griffith, E. M., Pennington, B. F., Wehner, E. A., & Rogers, S. J. (1999). Executive functions in young children with autism. *Child development*, *70*(4), 817-832. doi: 10.1111/1467-8624.00059.

Gur, R. C., Richard, J., Calkins, M. E., Chiavacci, R., Hansen, J. A, Bilker, W. B., … Gur, R. E. (2012). Age group and sex differences in performance on a computerized neurocognitive battery in children age 8-21. *Neuropsychology*, *26*, 251–65. doi: org/10.1037/a0026712.

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of autism and Developmental disorders*, *24*(2), 129-154. doi: org/10.1007/BF02172093.

Happé, F., Booth, R., Charlton, R., & Hughes, C. (2006a). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: examining profiles across domains and ages. *Brain and cognition, 61*(1), 25-39. doi: org/10.1016/j.bandc.2006.03.004.

Happé, F., Ronald, A., & Plomin, R. (2006b). Time to give up on a single explanation of autism. *Nature Neuroscience*, *9*, 1218–1220. doi: 10.1038/nn1770.

Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental review*, *24*(2), 189-233. doi: org/10.1016/j.dr.2004.01.001.

Hongwanishkul, D., Happaney, K. R., Lee, W. S., & Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: Age-related changes and individual differences. *Developmental neuropsychology, 28*(2), 617-644. doi: org/10.1207/s15326942dn2802_4.

Hooper, C. J., Luciana, M., Conklin, H. M., & Yarger, R. S. (2004). Adolescents’ performance on the Iowa Gambling Task: Implications for the development of decision-making and ventromedial prefrontal cortex. *Developmental Psychology*, *40*, 1148–1158. doi: org/10.1037/0012-1649.40.6.1148.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, *16*, 233–253. doi: 10.1111/j.2044-835X.1998.tb00921.x.

Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from ages 2 to 4. *Developmental Psychology*, *43*, 1447–59. doi: org/10.1037/0012-1649.43.6.1447.

Im-Bolter, N., Agostino, A., & Owens-Jaffray, K. (2016). Theory of mind in middle childhood and early adolescence: Different from before? *Journal of Experimental Child Psychology, 149*, 98–115. doi:10.1016/j.jecp.2015.12.006.

Keenan, T., Olson, D. R., & Marini, Z. (1998). Working memory and children's developing understanding of mind. *Australian Journal of Psychology*, *50*(2), 76-82. doi: 10.1080/00049539808257537.

Kercood, S., Grskovic, J. A., Banda, D., & Begeske, J. (2014). Working memory and autism: A review of literature. *Research in Autism Spectrum Disorders*, *8*(10), 1316-1332. doi: org/10.1016/j.rasd.2014.06.011.

Kim, S., Nordling, J. K., Yoon, J. E., Boldt, L. J., & Kochanska, G. (2014). Effortful control in "hot" and "cool" tasks differentially predicts children's behaviour problems and academic performance. *Journal of Abnormal Child Psychology, 41*, 43–56. doi: org/10.1007/s10802-012-9661-4.

Kimhi, Y., Kugelmas, D., Agam Ben Artzi, G., Ben Moshe, I., & Bauminger-Zviely, N. (2014). Theory of mind and executive function in preschoolers with typical development versus intellectually able pre-schoolers with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 44*(9), 2341-2354. doi: org/10.1007/s10803-014-2104-z.

Kouklari, E. C., Thompson, T., Monks, C. P., & Tsermentseli, S. (2017). Hot and cool executive function and its relation to theory of mind in children with and without autism spectrum disorder. *Journal of Cognition and Development.* doi: org/10.1080/15248372.2017.1339708.

Lantz, J. (2002). Theory of mind in autism: Development, implications, and interventions. *The Reporter*, 7(3), 18-25.

Le Couteur, A., Rutter, M., Lord, C., Rios, P., Robertson, S., Holdgrafer, M., & McLennan, J. D. (1989). Autism Diagnostic Interview: A standardized investigator-based instrument. *Journal of Autism and Developmental Disorders*, *19*, 363–387. doi: org/10.1007/BF02212936.

Leung, R. C., Vogan, V. M., Powell, T. L., Anagnostou, E., Taylor, M. J. (2016). The role of executive functions in social impairment in Autism Spectrum Disorder. *Child Neuropsychology, 22*, 336–344. doi:10.1080/09297049.2015.1005066

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2000). Autism Diagnostic Observation Schedule (ADOS). Los Angeles, California: Western Psychological Services.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*, 659–685. doi: org/10.1007/BF02172145.

Luna, B., Doll, S. K., Hegedus, S. J., Minshew, N. J., & Sweeney, J. A. (2007). Maturation of executive function in autism. *Biological psychiatry*, *61*(4), 474-481. doi: org/10.1016/j.biopsych.2006.02.030.

Marcovitch, S., O'Brien, M., Calkins, S. D., Leerkes, E. M., Weaver, J. M., & Levine, D. W. (2014). A longitudinal assessment of the relation between executive function and theory of mind at 3, 4, and 5 years. *Cognitive Development*, *33*, 40–55. doi.org/10.1016/j.cogdev.2014.07.001.

McAlister, A. R., & Peterson, C. C. (2013). Siblings, theory of mind, and executive functioning in children aged 3–6 years: New longitudinal evidence. *Child development*, *84*(4), 1442-1458. doi: 10.1111/cdev.12043.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*(5695), 503-507. doi: 10.1126/science.1100907.

McDonald, S. (2013). Impairments in social cognition following severe traumatic brain injury. *Journal of the International Neuropsychological Society*, *19*, 231–246. doi: org/10.1017/S1355617712001506.

Monks, C. P., Smith, P. K., & Swettenham, J. (2005). Psychological correlates of peer victimisation in preschool: Social cognitive skills, executive function and attachment profiles. *Aggressive Behavior*, *31*(6), 571–588. doi: 10.1002/ab.20099.

Mueller, S. T., & Piper, B. J. (2014). The psychology experiment building language (PEBL) and PEBL test battery. *Journal of neuroscience methods*, *222*, 250-259. doi: org/10.1016/j.jneumeth.2013.10.024.

Myerson, J., Green, L., & Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *Journal of the experimental analysis of behaviour*, *76*(2), 235-243. doi: 10.1901/jeab.2001.76-235.

National Institute for Health and Clinical Excellence (2011). Autism: recognition, referral and diagnosis of children and young people on the autism spectrum. (Clinical guideline 128.)

Otero, T. M., & Barker, L. A. (2014). The frontal lobes and executive functioning. In S. Goldstein & J. A. Naglieri (Eds.), *Handbook of Executive Functioning* (pp. 29–45). New York, USA: Springer.

Ozonoff, S., & McEvoy, R. E. (1994). A longitudinal study of executive function and theory of mind development in autism. *Development and psychopathology*, *6*, 415-431. doi: org/10.1017/S0954579400006027.

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry, 32*, 1081–1105. doi: 10.1111/j.1469-7610.1991.tb00351.x.

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental Psychology, 43*, 974–990. doi: org/10.1037/0012-1649.43.4.974.

Pellicano, E. (2010). Individual differences in executive function and central coherence predict developmental changes in theory of mind in autism. *Developmental psychology*, *46*(2), 530-544. doi: org/10.1037/a0018287.

Pellicano, E. (2012). The development of executive function in autism. *Autism Research and Treatment* (Special issue: Autism: cognitive control across the lifespan). doi:10.1155/2012/146132.

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *37*(1), 51-87. doi: 10.1111/j.1469-7610.1996.tb01380.x.

Perneger, T. V. (1998). What’s wrong with Bonferroni adjustments. *British Medical Journal*, *316*, 1236–1238.

Perner, J. (1998). The meta-intentional nature of executive functions and theory of mind. In P. Carruthers & J. Boucher (Eds.), *Language and Thought* (pp. 270–283). Cambridge: Cambridge University Press.

Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in cognitive sciences, 3*(9), 337-344. doi: org/10.1016/S1364-6613(99)01362-5.

Perner, J., & Wimmer, H. (1985). “John thinks that Mary thinks that…” attribution of second-order beliefs by 5-to 10-year-old children. *Journal of experimental child psychology*, *39*(3), 437-471. doi: org/10.1016/0022-0965(85)90051-7.

Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theory of mind development for children with autism, deafness or typical development. *Child Development, 76,* 502–517. doi: 10.1111/j.1467-8624.2005.00859.x.

Prencipe, A., Kesek, A., Cohen, J., Lamm, C., Lewis, M. D., & Zelazo, P. D. (2011). Development of hot and cool executive function during the transition to adolescence. *Journal of experimental child psychology*, *108*, 621-637. doi: org/10.1016/j.jecp.2010.09.008.

Poland, S. E., Monks, C. P., & Tsermentseli, S. (2015). Cool and hot executive function as predictors of aggression in early childhood: Differentiating between the function and form of aggression. *British Journal of Developmental Psychology*, *34*, 181–197. doi: 10.1111/bjdp.12122.

Richards, J. B., Zhang, L., Mitchell, S. H., & Wit, H. (1999). Delay or probability discounting in a model of impulsive behavior: effect of alcohol. *Journal of the experimental analysis of behaviour, 71*(2), 121-143. doi: 10.1901/jeab.1999.71-121.

Robinson, O. C., Dunn, A., Nartova-Bochaver, S., Bochaver, K., Asadi, S., Khosravi, Z., ...Yang, Y. (2016). Figures of Admiration in Emerging Adulthood: A Four-Country Study. *Emerging Adulthood*, *4*(2), 82-91. doi: org/10.1177/2167696815601945.

Russell J. (1996). *Agency: Its Role in Mental Development*. Hove: Lawrence Erlbaum Associates.

Russell, J. (1997). How executive disorders can bring about an adequate theory of mind. In J. Russell (Ed.), *Autism as an executive disorder* (pp. 256–304). Oxford, UK: Oxford University Press.

Russo, N., Flanagan, T., Iarocci, G., Berringer, D., Zelazo, P. D., & Burack, J. A. (2007). Deconstructing executive deficits among persons with autism: Implications for cognitive neuroscience. *Brain and cognition*, *65*(1), 77-86. doi: org/10.1016/j.bandc.2006.04.007.

Sabbagh, M. A., Bowman, L. C., Evraire, L. E., & Ito, J. M. B. (2009). Neurodevelopmental correlates of theory of mind in preschool children. *Child Development*, *80*, 1147–1162. doi: 10.1111/j.1467-8624.2009.01322.x.

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind a comparison of Chinese and US preschoolers. *Psychological science*, *17*(1), 74-81. doi: org/10.1111/j.1467-9280.2005.01667.x.

Scheres, A., Dijkstra, M., Ainslie, E., Balkan, J., Reynolds, B., Sonuga-Barke, E., & Castellanos, F. X. (2006). Temporal and probabilistic discounting of rewards in children and adolescents: Effects of age and ADHD symptoms. *Neuropsychologia*, *44*, 2092 – 2103. doi: org/10.1016/j.neuropsychologia.2005.10.012.

Scheres, A., Milham, M. P., Knutson, B., & Castellanos, F. X. (2007). Ventral striatal hyporesponsiveness during reward anticipation in attention-deficit⁄hyperactivity disorder. *Biological Psychiatry*, *61*, 720–724. doi: org/10.1016/j.biopsych.2006.04.042.

Scheres, A., Tontsch, C., Thoeny, A. L., & Sumiya, M. (2014). Temporal reward discounting in children, adolescents, and emerging adults during an experiential task. *Frontiers in psychology*, *5*, 711. doi: 10.3389/fpsyg.2014.00711.

Schug, J., Takagishi, H., Benech, C., & Okada, H. (2016). The development of theory of mind and positive and negative reciprocity in preschool children. *Frontiers in Psychology*, *7*. doi: 10.3389/fpsyg.2016.00888.

Segalowitz, S., & Davies, P. L. (2004). Charting the maturation of the frontal lobe: An electrophysiological strategy. *Brain and Cognition*, *55*, 116–133. doi: org/10.1016/S0278-2626(03)00283-5.

Serra, M., Loth, F. L., van Geert, P. L. C., Hurkens, E., & Minderaa, R. B. (2002). Theory of mind in children with ‘‘lesser variants’’ of autism: A longitudinal study. *Journal of Child Psychology and Psychiatry*, *43*, 885–900. doi: 10.1111/1469-7610.00104.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London*, B 298, 199–209.

Siegel, D. (2013). *Brainstorm: The power and purpose of the teenage brain*. New York, NY: Jeremy Tarcher/Penguin.

Sinzig, J., Morsch, D., Bruning, N., Schmidt, M. H., & Lehmkuhl, G. (2008). Inhibition, flexibility, working memory and planning in autism spectrum disorders with and without comorbid ADHD-symptoms. *Child and Adolescent Psychiatry and Mental Health*, *2* (4). doi: org/10.1186/1753-2000-2-4.

South, M., Ozonoff, S., & McMahon, W. M. (2007). The relationship between executive functioning, central coherence, and repetitive behaviors in the highfunctioning autism spectrum. *Autism*, *11*, 437–451. doi: org/10.1177/1362361307079606.

Steinberg, L., Graham, S., O’Brien, L., Woolard, J., Cauffman, E., & Banich, M. (2009). Age differences in future orientation and delay discounting. *Child development, 80*, 28-44. doi: 10.1111/j.1467-8624.2008.01244.x.

Steele, S., Joseph, R. M., & Tager-Flusberg, H. (2003). Developmental change in theory of mind abilities in children with autism. *Journal of Autism and Developmental Disorders, 33*, 461–467. doi.org/10.1023/A:1025075115100.

Stolarski, M., Bitner, J., & Zimbardo, P. G. (2011). Time perspective, emotional intelligence and discounting of delayed awards. *Time & Society, 20*(3), 346-363. doi: org/10.1177/0961463X11414296.

Ströhle, A., Stoy, M., Wrase, J., Schwarzer, S., Schlagenhauf, F., Huss, M., Hein, J., Nedderhut, A., Neumann, B., … Heinz, A. (2008). Reward anticipation and outcomes in adult males with attention-deficit ⁄ hyperactivity disorder. *NeuroImage, 39*, 966–972. doi: org/10.1016/j.neuroimage.2007.09.044.

Tager-Flusberg, H., & Joseph, R. M. (2005). Theory of mind, language, and executive functions in autism: A longitudinal perspective. In W. Schneider, R. Schumann-Hengsteler & B. Sodian (Eds.), *Young children's cognitive development. Interrelationships among executive functioning, working memory, verbal ability, and theory of mind* (pp. 239-258). Mahwah, NJ: Erlbaum.

Tillman, C., Brocki, K. C., Sørensen, L., & Lundervold, A. J. (2015). A longitudinal examination of the developmental executive function hierarchy in children with externalizing behavior problems. *Journal of Attention Disorders*, *19*, 496–506. doi: org/10.1177/1087054713488439.

Vellante, M., Baron-Cohen, S., Melis, M., Marrone, M., Petretto, D. R., Masala, C., & Preti, A. (2013). The “Reading the Mind in the Eyes” test: systematic review of psychometric properties and a validation study in Italy. *Cognitive Neuropsychiatry, 18*, 326–354. doi: org/10.1080/13546805.2012.721728.

Verdejo-García, A., Bechara, A., Recknor, E. C., & Perez-Garcia, M. (2006). Decision-Making and the Iowa Gambling Task: Ecological validity in individuals with substance dependence. *Psychologica Belgica*, *46*(1-2), 55–78. doi: org/10.5334/pb-46-1-2-55.

Verté, S., Geurts, H. M., Roeyers, H., Oosterlaan, J., & Sergeant, J. A. (2005). Executive functioning in children with autism and Tourette syndrome. *Developmental Psychopathology, 17*, 415–445. doi: org/10.1017/S0954579405050200.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children*—Third Edition. San Antonio TX: Psychological Corporation.

Wechsler, D. (1999). *The Wechsler abbreviated scale for intelligence*. San Antonio TX: The Psychological Corporation.

Welsh, M., & Peterson, E. (2014). Issues in the conceptualization and assessment of hot executive functions in childhood. *Journal of the International Neuropsychological Society*, *20*, 152–156. doi:10.1017/S1355617713001379.

Willoughby, M., Kupersmidt, J., Voegler-Lee, M., & Bryant, D. (2011). Contributions of hot and cool self-regulation to preschool disruptive behavior and academic achievement. *Developmental neuropsychology*, *36*, 162-180.

Xiao, T., Xiao, Z., Ke, X., Hong, S., Yang, H., Su, Y., ... Liu, Y. (2012). Response inhibition impairment in high functioning autism and attention deficit hyperactivity disorder: evidence from near-infrared spectroscopy data. *PloS one*, *7*(10), e46569. doi: org/10.1371/journal.pone.0046569.

Yerys, B. E., Wallace, G. L., Sokoloff, J. L., Shook, D. A., James, J. D., & Kenworthy, L. (2009). Attention deficit/hyperactivity disorder symptoms moderate cognition and behavior in children with autism spectrum disorders. *Autism Research*, *2*(6), 322-333. doi: 10.1002/aur.103.

Yeung, M. K., Han, Y. M., Sze, S. L., & Chan, A. S. (2016). Abnormal frontal theta oscillations underlie the cognitive flexibility deficits in children with high-functioning autism spectrum disorders. *Neuropsychology*, *30*(3), 281. doi: org/10.1037/neu0000231.

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives, 6*(4), 354–360. doi: 10.1111/j.1750-8606.2012.00246.x.

Zelazo, P. D., Qu, L., & Müller, U. (2005). Hot and cool aspects of executive function: Relations in early development. In: W. Schneider, R. Schumann-Hengsteler & B. Sodian (Eds), *Young Children's*

Cognitive Development: Interrelationships Among Executive Functioning, Working Memory, Verbal Ability and Theory of Mind (pp. 71-95). Mahwah New Jersey: Lawrence Erlbaum Associates Inc.

Zelazo, P. D., & Müller, U. (2002). Executive function in typical and atypical development. In: U. Goswami (Ed.), *Handbook of Childhood Cognitive Development (*pp.445–470). Oxford: Blackwell Publishing Ltd.-

Table 1. Participants' characteristics

	Group	
	ASD	Control
Variable	(*n*=45)	(*n*=37)
Age (in years)		
M (SD)	9.07 (1.42)	9.03 (1.17)
Range	7-11	7-11
FSIQ total score		
M (SD)	97.05 (12.13)	102.11 (14.3)
Range	70-127	76-135

Table 2. Means and standard deviations for variables across the time points.

	T1				T2			
	Control		ASD		Control		ASD	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Go/No-Go	36.86	20.46	47.77	16.04	34.03	17.06	46.59	15.22
ToL	7.45	1.58	7.26	2.04	9.34	1.45	8.69	1.60
Digit Span	9.71	2.39	7.18	3.03	10.47	2.54	9.34	1.66
IGT	.04	.21	-.05	.19	.08	.16	.02	.16
Delay	.36	1.32	.45	.20	.38	.17	.41	.17
Sandbox	2.76	1.04	3.93	2.81	2.95	1.51	2.85	1.97
Eyes Test	18.20	2.81	15.33	3.47	20.25	2.82	18.6	3.43

Table 3. Correlations between ToM and EF tasks across the two time points.

	False Belief T1	False Belief T2	Eyes test T1	Eyes Test T2
T1 Inhibition	.09	-.05	-.37**	-.18
T1 Planning	-.27*	-.14	.21*	.19
T1 WM	-.14	.10	.24*	.04
T1 IOWA	.014	-.04	.18	.10
T1 Delay	.12	-.03	-.25*	-.24*
T2 Inhibition	.14	-.04	-.38**	-.21
T2 Planning	-.35**	-.16	.31**	.24*
T2 WM	-.16	-.18	.18	.09
T2 IOWA	-.13	-.11	.11	.004
T2 Delay	.12	.02	-.12	-.15

Note. *p < .05, **p < .01.

Figures

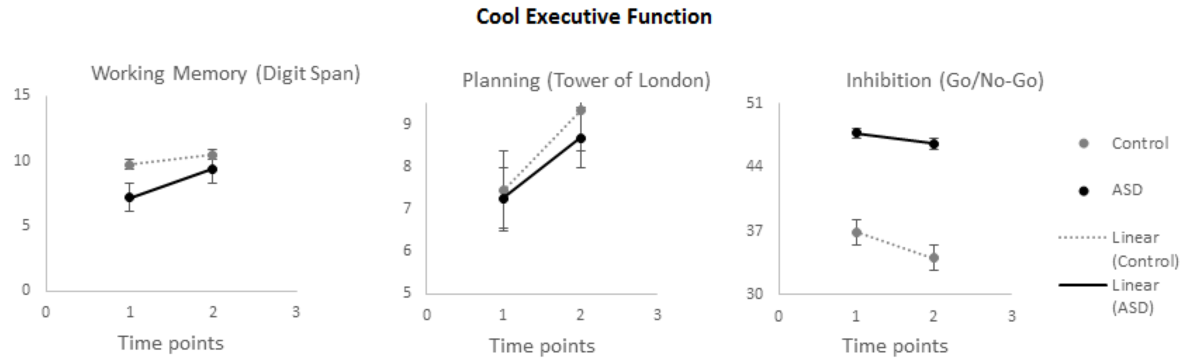

Figure 1. Mean Cool EF scores across 12 months for ASD and control groups.

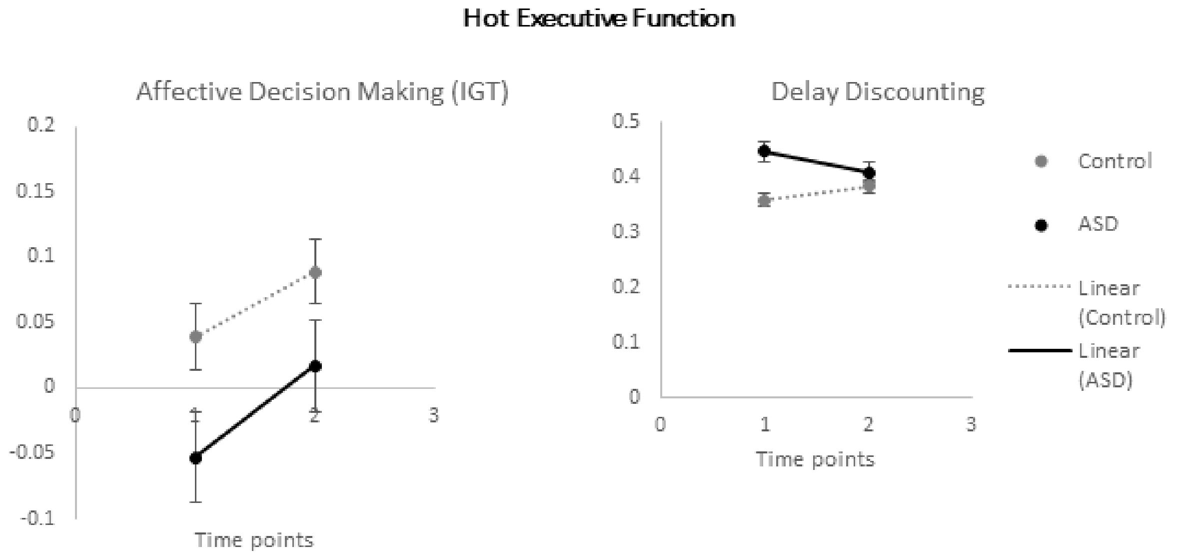

Figure 2. Mean Hot EF scores across 12 months for ASD and control groups.

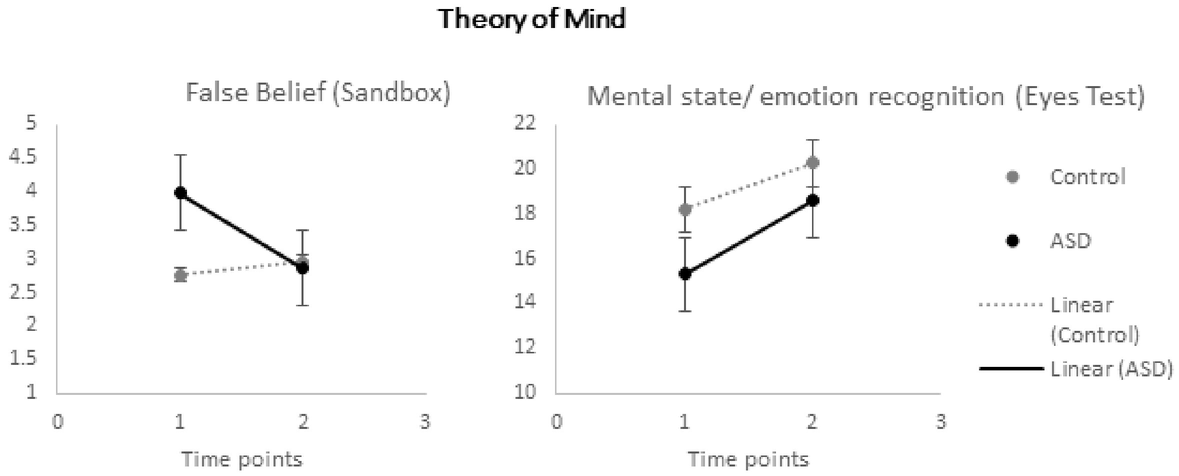

Figure 3. **Mean ToM scores across 12 months for ASD and control groups.**

Appendix

Table A. Hierarchical regression analysis for later ToM false belief (T2) by group and EF variables.

Block		Standardized Coefficients	ΔR^2	*P*	Collinearity Statistics	
		Beta			Tolerance	VIF
1	*Control Variables*		.25	< .001		
	Concurrent Age	-.053		.608	.900	1.111
	ASD diagnosis	.163		.118	.906	1.103
	Concurrent IQ	-.005		.962	.924	1.082
	Early False Belief (T1)	.507		<. 001	.884	1.131
2	*Cool EF*		.15	.007		
	Early Working Memory	.576		< .001	.352	2.839
	Later Working Memory	-.595		< .001	.360	2.781
	Early Planning	-.059		.608	.632	1.583
	Later Planning	.051		.666	.601	1.665
	Averaged Inhibition*	-.029		.772	.866	1.154
3	*Hot EF*		.005	.74		
	Early Delay Discounting	-.057		.577	.825	1.212
	Later Delay Discounting	-.034		.732	.846	1.182
4	*Hot EF X ASD interactions*		.034	.13		
	ASD X Early Del. Disc.	.121		.317	.579	1.728
	ASD X Later Del. Disc.	-.210		.059	.691	1.447

*Averaged Inhibition in block 3 represents the average score of early and later Inhibition variables. It was decided to include this variable because the initial regression model with both inhibition variables suffered from multicollinearity issues due to those variables being highly correlated (*r*= .97).

Table B. Hierarchical regression analysis for later ToM mental state/ emotion recognition (T2) by group and EF variables.

		Standardized Coefficients			Collinearity Statistics	
Block		Beta	ΔR^2	*P*	Tolerance	VIF
1	*Control Variables*		.54	< .001		
	Concurrent Age	.150		.073	.891	1.122
	ASD diagnosis	-.037		.671	.824	1.213
	Concurrent IQ	.004		.962	.843	1.186
	Early Eyes Test (T1)	.708		<. 001	.738	1.356
2	*Cool EF*		.03	.44		
	Early Working Memory	-.230		.083	.352	2.838
	Later Working Memory	.106		.416	.357	2.797
	Early Planning	.054		.579	.638	1.567
	Later Planning	-.019		.850	.606	1.651
	Averaged Inhibition*	.086		.326	.809	1.236
3	*Hot EF*		.008	.52		
	Early Delay Discounting	-.083		.345	.794	1.259
	Later Delay Discounting	-.027		.752	.847	1.181
4	*Hot EF X ASD interactions*		.09	< .001		
	ASD X Early Del. Disc.	-.304		.002	.576	1.736
	ASD X Later Del. Disc.	.281		.002	.677	1.477

*Averaged Inhibition in block 3 represents the average score of early and later Inhibition variables. It was decided to include this variable because the initial regression model with both inhibition variables suffered from multicollinearity issues due to those variables being highly correlated (*r*= .97).