THEORY OF MIND, LANGUAGE, COGNITION, AND SYMPTOMS OF AUTISM IN YOUTH WITH AUTISM SPECTRUM DISORDER

PATRYCJA KRZEMINSKA CZAPINSKI

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN

PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN PSYCHOLOGY

YORK UNIVRSITY

TORONTO, ONTARIO

APRIL 2020

Abstract

The purpose of the current study was to explore the development of theory of mind (ToM) as measured by the Reading the Mind in the Eyes Test-Children (RMET-C) in youth with autism, and to determine evidence for a relationship between performance on RMET-C and cognitive skills, language skills, as well as impairments associated with autism as measured by the ADOS. The data was derived from a genetics study on autism and in total, there were 91 participants diagnosed with autism spectrum disorder (ASD) who were divided into five groups based on age: 6-8, 8-10, 10-12, 12-16, and 16-20 years. First, compared to the published data for typically developing youth for mean performance on the RMET-C, participants with ASD in the 8-10, 10-12, and 12-16 age groups performed significantly lower; however, unlike participants in the typically developing groups, they continued to improve their performance beyond 12 years of age. Based on the DSM-IV classification of autism, individuals with Asperger's performed significantly better than PDD-NOS and ASD PDD groups on the RMET-C before adjusting for verbal and nonverbal IQ. Second, for youth with ASD, when language skills (expressive and receptive) were statistically taken into account, verbal IQ ceased to be predictive of performance on the RMET-C. Language skills were significantly correlated with RMET-C performance across four age groups (all except 6-8), whereas verbal IQ was significantly correlated in two of the age groups, 8-10 and 10-12. Non-verbal IQ was significantly correlated with RMET-C performance in three age groups, 8-10, 10-12, and 12-16. Third, participants with ASD were less efficient at identifying positive feelings and slightly less efficient at identifying positive mental states than negative feelings and negative mental states, although there were differences in frequencies of responses to specific items on the RMET-C.

The question with the highest frequency of correct response was from the negative valence category ("upset") and one with the least frequent correct response was from positive valence category ("friendly"). Fourth, based on ADOS-II classification for Module 3, there was no association between the increased symptoms of autism on the Social Affect nor on the RRB scale and performance on the RMET-C. Based on the ADOS I classification of the same module, better scores on the Communication domain were related to increased performance on the RMET-C. On ADOS II for Module 4, better performance on the Communication domain and SBRT scale was associated with higher performance on the RMET-C. There was no significant correlation between Reciprocal Social Interaction and performance on the RMET-C. Correlations between individual items on the ADOS (Modules 3 and 4) and performance on the RMET-C were also explored. Lastly, anxiety scores on Module 3 and 4 of the ADOS were not significantly correlated to performance on the RMET-C. The present findings are discussed with respect to existing empirical literature and theoretical paradigms (Eye Avoidance, Affective, Metarepresentation, Language, Alexithymia and Evolutionary theories). Implications for practice, policy, and future research are also presented.

Dedication

To my parents, Lucyna Krzemińska and Stanisław Krzemiński for inspiring my love of sciences and discovery, and for being my first teachers of critical thinking.

To my husband, Swavek Czapiński, for sharing my passion for the quest for knowledge, for your strength and humour, and for filling our home with books.

To my children, Paulina, Konrad, and Natalia whose growing curiosity and wonder about the world are both inspiring and humbling. I am proud of you.

Thank you all for your love and support.

Acknowledgements

I am deeply honoured for the mentorship and guidance of my supervisor, Dr. Debra Pepler, whose dedication and knowledge has shaped my development as a researcher and a clinician. Dr. Pepler's unwavering commitment to her students, research and communities is exceptional and I am most grateful for having had the opportunity to learn from such an inspiring role model.

I am grateful to Dr. Susan Bryson who introduced me to the field of autism and with whom I began the journey of research prior to graduate school. I feel most privileged to have had the opportunity to learn from someone who is so committed to her students and research.

My appreciation extends to Dr. Stephen Scherer and all co-investigators on the genetics study for supporting my research interests by providing me with access to the data.

I am thankful to my committee members, Dr. Mary Desrocher and Dr. Irene Drmic for their enthusiasm about my research and insightful suggestions. My gratitude extends also to Dr. Suzanne MacDonald, Dr. Terry Bennett, and Dr. Kate Tilleczek for their interest in my study, expertise and thoughtful questions.

I am grateful for my supportive friends who were always ready to hear about my research, cheered me on and provided reassurance, especially when I needed it most.

Last but not least, my heartfelt appreciation extends to my wonderful family, both immediate and extended for their continuous interest and encouragement during the various stages of my graduate studies. I feel fortunate to have such a caring and supportive family.

Table of Contents

List of Tables

List of Figures

Theory of Mind, Language, Cognition, and Symptoms of Autism in Youth with Autism Spectrum Disorder

The general aim of the current study was to explore the development of the theory of mind (ToM) in children and adolescents with autism and to provide evidence for a relationship between ToM abilities, cognitive skills and language skills. Additionally, I explored the link between ToM and impairments associated with autism, including, social, communication and behaviour difficulties. This was accomplished by first examining the accuracy and pattern of responses on a ToM measure across different age intervals and gender in individuals with Autism Spectrum Disorder (ASD), both within the group and compared to published data for typically developing individuals. Second, I explored the associations between performance on the ToM measure and standardized language and cognitive assessment measures. Additionally, I conducted qualitative analyses involving the response patterns related to valence of emotions on the ToM task. Lastly, I compared the degree of impairment in the areas of social, communicative and behaviour symptoms of autism in individuals with ASD to participants' with ASD performance on the ToM task and standardized assessment measures of cognition and language. This study contributes to both theoretical and practical understanding of the links between social cognition, language and characteristics of autism.

The current study begins with a description of the autism disorder and the psychological theories of autism, including the deficient "theory of mind" hypothesis. The introduction follows with a review of research on ToM in typical development and then the role of language on the development of ToM in typically developing children. There is also a review of research of language skills in individuals with autism. The focus then turns to studies involving the widely used ToM task, namely, the Reading the Mind in the Eyes Test-Children (RMET-C) in both typically developing children and adolescents and those with a diagnosis of autism. Preceding this review is a presentation of literature on nonverbal communication as it relates to the development of ToM. Research related to other possible correlates of the development of ToM is explored as it relates to symptoms of autism. Specifically, the review focuses on studies exploring correlations between symptoms of ASD on the Autism Diagnostic Observational Scale (ADOS) and performance on the ToM tasks.

Autism Spectrum Disorder

Autism Spectrum Disorder (ASD) is a life-long neurodevelopmental disorder that emerges in early childhood (APA, 2013). In the Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (DSM-IV) (APA, 1994), autism was part of the spectrum of developmental disorders, referred to as autism spectrum disorders (ASDs; also known as the pervasive developmental disorders; PDDs, including Aspergers syndrome, PDD-not otherwise specified, and Childhood Disintegrative Disorder). Since the development of DSM-5 (APA, 2013), the group of PDDs has been reorganized into a single category of Autism Spectrum Disorders (ASD s) thus eliminating PDD-NOS, Aspergers syndrome, and Childhood Disintegrative Disorder classifications. The term ASD and autism was used interchangeably in this paper, unless otherwise specified.

ASD is a disorder defined by impairments in two areas: social interaction and communication, and by the presence of repetitive or restrictive activities and interests (APA, 2013). As part of the ASD diagnosis, deficits associated with these two areas of impairments are specified by three levels of support that reflect the spectrum of difficulties (from mild to severe) in individuals with autism (APA, 2013). The prevalence of autism is estimated to be 1/68 (CDC, 2014), with intellectual deficits reported in 30-50% of cases (Chakrabarti & Fombonne, 2005; CDC, 2014). There is wide heterogeneity in the presentation of symptoms, abilities, and skills among children and youth with autism. The ratio of males to females with ASD is 4:1 but the cause of this discrepancy remains unknown (Zwaigenbaum et al., 2012). ASD is a heterogeneous disorder in terms of presentation with equally variable aetiologies, including genetic, epigenetic and environmental variables that influence one another in a multitude of ways (Anagnostou et al., 2014). Autism co-occurs with other genetic and neurological conditions (e.g., Fragile X – 2.1%, Down Syndrome – 3.7%, and epilepsy – 18.2%) (Kielinen, Rantala, Timonen, Linna, & Moilanen, 2004) as well as psychiatric disorders both internalizing (e.g., depression) and externalizing (e.g., ADHD) and tic disorders (Mazzone, Ruta, & Reale, 2012).

Impairment in social interaction is considered the most salient feature of autism, with profound deficiencies in verbal and nonverbal communication. Kanner (1943) was the first to describe the “extreme autistic aloneness” of the syndrome, which underscores the assumption that children with autism are happiest when left alone. During encounters with other people, individuals with autism exhibit deficient nonverbal communication skills, such as eye gaze, imitation, and joint attention (Fodstad, Matson, Hess, & Neal, 2009). For example, children with autism have difficulties sharing and directing attention, and following another person's point or gaze. In the area of verbal communication, lack of speech (Bryson et al., 1998) has been identified as one of the main concerns in people with autism. For those individuals who do develop verbal skills, their speech often lacks communicative intent, is not functional, and lacks fluency (Johnson & Myers, 2007). Their speech is often repetitive in nature, focusses strictly on an individual's topic of interest, sounds scripted, and is characterized by echolalia or repeating another person's speech.

Individuals with autism are also described as having an inflexible and repetitive repertoire of behaviours, such as hand flapping (Chawarska & Volkmar, 2005), or spinning objects and lining up toys, as well as, narrow and focussed interests such as learning all transit routes. It is challenging for people with ASD to adapt to changes in their routine and any variation in their environment from a regularly followed pattern causes significant distress. Finally, individuals with autism tend to be overly sensitive to various sensory stimuli (DSM-5, 2013), such as sound, texture, or odour.

Psychological Theories of Autism and Empirical Support

Over the past decades, different accounts of autism have focused on socio-emotional processes, as reflected by Hobsons (1992) "impaired social relatedness" theory, as well as cognitive processes, as reflected by the "impaired executive function" theory (e.g., Ozonoff, Pennington, & Rogers, 1991) or the "deficient theory of mind" hypothesis (Baron-Cohen, 1990).

Hobson (1993) proposed that the basic dysfunction in autism lies in affective systems and those underlie the social and higher order cognitive deficits in autism. Although Kanner (1943) was the first one to state that children with autism have an inborn deficiency of affective contact, Hobson developed this perspective further and postulated that the primary deficit in autism is a lack of personal relatedness. Specifically, according to Hobson, children with autism lack the innate ability to both perceive and respond to the affective expressions of other people. This, in turn, renders the children incapable of receiving the social experiences necessary to develop cognitive structures underlying social understanding. Impairments in language and cognition are viewed as secondary to these social deficits.

Research on affect in children with autism has focused on the areas of perception and expression. Hobson (1986) found that relative to matched non-autistic mentally handicapped and normal children, the children with autism were impaired in choosing facial expressions that corresponded with gestures, vocalizations, and situations/contexts, thus revealing these childrens deficits in understanding facial expressions of emotion. Similarly, Assumpcaio, Sprovieri, Kuczynski, and Farinah (1999) found that compared to normal controls, children and adults with autism were deficient in identifying facial expressions of basic emotions such as happiness, sadness, anger, and surprise. In the area of expression of emotion, one of the most consistent findings is that children with autism display less positive affect during free play and joint attention situations in both familiar and unfamiliar contexts (Snow, Hertzig, & Shapiro, 1987; Kasari, Sigman, Mundy, & Yirmiya, 1990; Bieberich & Morgan, 1998; Joseph & Tager-Flusberg, 1997; Czapinski & Bryson, 2002). In one study, the expression of positive affect in autism was less likely to be directed at a partner and more likely to occur during self-absorbed activities (Snow, Hertzig, & Shapiro, 1987). In terms of the quality of affect, Yirmiya, Kasari, Sigman, and Mundy (1989) found that relative to children who were intellectually impaired and typically developing, children with autism were more likely to exhibit neutral or ambiguous expressions and display blends of emotion (e.g., anger with joy at the same time).

The expression of emotion has been noted by Trevarthen (1979) as playing a fundamental role in infancy in communicating with others, engaging in social relationships, and understanding that others are like the self. If the ability to express emotion adequately is impaired, as it has been reported in children with autism, then this in turn may result in deficient communicative pathways, such as difficulties in nonverbal and verbal communication between the child and the caregiver and subsequent deficiencies in higher order thinking, such as the development of ToM.

In a theoretical paradigm implicating higher brain areas, researchers have argued that individuals with autism have an impairment in cognitive processes. For example, executive function (EF) is a higher order cognitive construct used to describe goal directed behaviors, such as planning, inhibition of pre-potent responses, flexibility, organized search, self-monitoring, and use of working memory (Hill, 2004). Ozonoff et al. (1991) argued that individuals with autism have deficits in executive functions, including planning, organization, and thinking flexibly, and this has been demonstrated in reduced performances of individuals with ASD on the standardized assessment measures such as the Wisconsin Card Sorting Task and the Tower of Hanoi task. In addition to autism, other disorders with potential deficiencies of frontal lobe connectivity and thus executive dysfunction, include attention deficit hyperactivity disorder (ADHD), obsessive compulsive disorder, and Tourette syndrome (Hill, 2004); these disorders also co-occur in individuals with autism. Craig et al. (2016) have posited that difficulties with executive functioning impact the development of ToM. Although a relationship between executive functioning and ToM has been reported in individuals with ASD (Joseph & Tager-Flusberg, 2004; Pellicano, 2007) the causal direction remains uncertain. It may be that difficulties with executive functioning limit the development of ToM or that specific deficits in the development of ToM interfere with the development of EF.

In the theoretical area of deficient cognitive processes, Baron-Cohen (1990) argued that children with autism lack "theory of mind", that is, they fail to recognize that others have minds and mental states, such as knowing and desiring, which are not directly observable. To understand and predict other people's behavior during social interaction, individuals need to understand the mental states of others (Philpott, Rinehart, Gray, Howlin, & Cornish, 2013). If children with autism are unaware that other people's actions are governed by what they feel, know, desire, and believe, then they lack "metarepresentation" or "theory of mind."

Theory of Mind – Studies in Typical Development

Premack and Woodruff (1978) were the first ones to study understanding of goal directed behavior in chimpanzees and coined the term "theory of mind". One of the most prevalent topics in cognitive development over the past decades has been the study of the ToM in children. The development of the ToM has been investigated using different tasks and concepts; however, the central focus has been on children's understanding of false belief because the mental state of "belief" is internal and separate from real world situations (Wellman, Cross, & Watson, 2001). A widely referenced example of a false belief task described by Wimmer and Perner (1983) involves a child, Maxi, who placed his chocolate in a kitchen cupboard and left the room to play. While he was away, his mother moved the chocolate from the cupboard to the drawer so Maxi was holding a false belief about the location of his chocolate. The children are asked where Maxi will look for his chocolate when he returns to the kitchen, whether it will be in the cupboard or in the drawer. Many 4- and 5-year old children pass this task concluding that Maxi will look in the cupboard whereas 3-year old children typically have difficulty with this task and state that Maxi will look in the drawer where his mother moved the chocolate (Wellman, Cross,

& Watson, 2001). In order to answer the question correctly, the children need to understand the "belief" of the character, rather than the physical reality involving the object.

The understanding of false belief provides a major milestone in children's social understanding that beliefs are mental representations and not direct reflections of reality (Miller, 2009). Although it is understood that the first-order reasoning task such as the false belief task described above develops sometime between the ages of 3 and 5, there is less consensus in research about the timing of development of the second-order false belief task in subsequent stages of childhood.

While first order reasoning deals with one target for assigning belief and one proposition, the second order reasoning tasks involve two targets and two propositions that are integrated (Miller, 2009). Tasks used to measure second-order reasoning are similar to a standard false belief tasks except that in a second-order ToM situation the individual is capable of thinking about a person's beliefs, thoughts or intentions about another person's beliefs, thoughts, or intentions. For example, in a situation described above involving Maxi, another character could be introduced in the scenario so that the second order question would pertain to what Maxi thinks about this new character thinking about Maxi or his mother. The original second-order false belief scenario by Perner and Wimmer (1985) involved two characters (John and Mary) who were individually informed about an object's (van's) unexpected transfer to a new location. John and Mary knew where the van was but there was a mistake in John's second-order belief about Mary's belief since he did not know that Mary was also informed about the change of location of the van. Typically, the age of mastery of second order false belief tasks is 1 to 2 years after the development of first order tasks (5-6 years) and this variation in age is dependent on methodology and samples involved in the studies (Miller, 2009).

Theory of Mind and language in typical development. A relationship between language and theory of mind has been well documented, that is, language plays a vital role not only during conversation but also during internal, verbal representation (Astington & Baird, 2005). The extent of the association between language and ToM is complex. Recently, Derksen, Hunsche, Giroux, Connolly, and Bernstein (2018) reported that in addition to the development of attention and executive function, language is a precursor to the development of the theory of mind. For example, children talk more about their own thoughts as well as mental states of other people shortly before the development of theory of mind (Derksen et al., 2018). Hayward (2012) administered some theory of mind measures in an attempt to clarify which tasks are most appropriate for middle childhood and preadolescence. Although the author hypothesized that age would be a predictor of performance across the developmental period that she studied, age did not predict significant amount of variance in performance across all ToM tasks; however, a significant proportion of variance in a ToM Eyes task was accounted by language ability.

Language Skills in Individuals with Autism Spectrum Disorder

In a study evaluating parental detection of the early signs of autism, De Giacomo and Fombonne (1998) found that the most common causes of parental concern were language abnormalities followed by atypical social-emotional responses. On average, children with autism, develop their first words at 38 months, whereas children with Aspergers produce them at 15 months (Howlin, 2003). Although some children develop delayed speech, there is a group of children with autism who fail to develop phrase speech by school years and remain minimally verbal (McGonigle-Chalmers, Alderson-Day, Fleming, Joanna, & Monsen, 2013).

In individuals with autism who do develop speech, the most consistently reported language deficit is in the pragmatics of language (Kelley, Paul, Fein, & Naigles, 2006). Pragmatics is the use of language for the purpose of communication during social interactions and it involves the use of both verbal and nonverbal cues (head shaking, eye contact) (Eigsti, Bennetto, & Dadlani, 2007). In this domain, individuals with autism have been reported to be deficient in maintaining conversation and even though their linguistic skills advance structurally, their content of speech does not change (Tager-Flusberg & Anderson, 1991). Children with autism were also impaired in their discourse abilities (a concept related to pragmatics and characterized by more extensive periods of speech (Eigsti et al., 2007). Capps, Kehres, and Sigman (1998) compared children with autism to children with developmental delay matched on language ability during a semi-structured conversation. Compared to the control group, children with autism more frequently failed to respond to questions or comments and less frequently produced personal experience narratives or offered new information.

Studies have shown that individuals with autism are impaired in structural aspects of language, which is generally referred to as syntax, or an ability to organize words into phrases. Volden and Lord (1991) compared children with high and low functioning autism to typically developing (TD) controls matched on receptive language on their use of neologisms and idiosyncratic words and phrases. Compared to the TD controls, individuals with autism used more neologisms and idiosyncratic speech. Low-functioning children with ASD also had lower mean length of utterance (MLU: the average number of morphemes per utterance) as compared to younger TD individuals. Additionally, Eigsti et al. (2007) found that a sample of children with autism produced syntactically less complex spontaneous language (e.g., shorter MLUs) compared to TD and developmentally delayed children.

Delayed and deficient acquisition of language skills in young children with ASD may influence their development of theory of mind. Fisher, Happé, and Dunn (2005) determined that children with ASD are more dependent on language to understand social tasks than typically developing children. In the light of these findings, the development of social understanding for individuals with autism may rely more strongly on language compared to typically developing individuals.

Reading the Mind in the Eyes Task (RMET) – Autism Spectrum Disorder

The RMET measures the ability to infer mental states from facial cues in the eye region of the face in absence of social context (Baron-Cohen, Wheelwright, Spong, Scahill, & Lawson, 2001). Although RMET was designed to detect social cognitive deficits in individuals with autism and Asperger's syndrome, it has been used with typically developing individuals as well. The original RMET (Baron-Cohen, Jolliffe, Mortimore, Robertson, & 1997) was developed for adults and was later adapted for use with children by using child friendly vocabulary (Baron-Cohen et al., 2001). The Reading the Mind in the Eyes Task-Child version (RMET-C) includes 28 black and white photos that are shown for three seconds each, depicting eye region of the face. Children are asked which of four choice words best describes what the person is thinking or feeling in the photo. To date, only a few studies have used the Child version of the RMET (RMET-C) (Baron-Cohen et al. 2001) in youth with autism and those studies are outlined below.

In the first paper on RMET in children, Baron-Cohen et al. (2001) compared the performance of typically developing children divided into three age groups: 6 to 8 (n = 20), 8 to 10 (n = 14) and 10 to 12 (n = 19) to a group of children with Asperger's syndrome age 8 to 14 (n = 15). Children with Asperger's were of at least average intelligence using the Wechsler scales and typically developing children were assumed to have average intelligence based on school performance, although standardized intelligence testing did not take place in this group. Based on the results, the typically developing children in the 8-10 and 10-12 groups performed significantly higher than both the typically developing children in the 6 to 8 group and the children in the Asperger's syndrome group. Consequently, Baron-Cohen et al. reported that children with Asperger's syndrome have difficulty inferring mental states, as compared to typically developing children and are thus deficient in "folk psychology", or the ability to infer social causality as oppose to "folk physics" or the ability to infer physical causality. In the latter, participants with Asperger's performed superior to typically developing participants. Based on these findings, the ability to infer mental states from images of the eyes solidifies during middle childhood for typically developing children.

Consistent results were also presented by Kaland, Callesen, Møller-Nielsen, Lykke Mortensen, and Smith (2008) who compared the performance of 21 children and adolescents with a diagnosis of Asperger syndrome (AS, age 10-20 years) and 20 typically developing controls (age 9-20 years) on three advanced theory of mind tasks: The Eyes Task (or RMET), the Strange Stories, and the Stories from Everyday Life. Compared to typically developing individuals, children and adolescents with AS demonstrated lower performance on all three tasks. One of the limitations of this study is that there was a significant difference between the

Full Scale IQ of the two groups; therefore, differences in cognitive capacity may have played a role in the results. Although the studies by Kaland et al. and Baron-Cohen et al. revealed significant differences between children with ASD and typically developing participants, it is unclear based on those studies whether this difference was due to other variables such as cognitive abilities or language skills since these domains were not controlled in those studies.

A study by Brent, Rios, Happe, and Charman (2004) addressed this issue by administering both IQ and language measures to participants in their study. The authors compared performance on the child version of the RMET of ASD children (mean age=8.3 years) to that of typically developing children (mean age=8.8 years). The two groups did not differ on age or IQ but children with ASD had lower performance on the Clinical Evaluation of Language Fundamentals–Revised CELF-R, a measure of language including expressive and receptive abilities. The two groups were not matched on standardized score on the CELF and only language age was included. Despite equivalent IQ scores, the ASD participants performed significantly lower than typically developing controls on the RMET-C. Performance on the Eyes task was significantly correlated with "language age" on the CELF-R ($r = 0.60$, $p < 0.01$) in children with ASD only. In summary, in the presence of intact cognitive abilities, children with ASD performed lower than typically developing controls; however, this difference could be attributed to the lower scores on the language task achieved by participants with ASD. Moreover, only the Full Scale IQ measure was provided so the extent to which Verbal and Nonverbal cognitive abilities may have had an impact on their performance on the RMET-C remains unknown.

Peterson, Slaughter, and Brownell (2015) used both the original and a revised version that constituted of only first 12 items of the child version of the RMET and compared 34 children (aged 5-12) with ASD and 41 typically developing (TD) controls. The two groups were matched for age and according to the authors on verbal intelligence using the Peabody Picture Vocabulary Test–Revised (PPVT-R), a standardized, normative receptive vocabulary test based on pointing responses to pictures. The typically developing children performed better than children with ASD on both the standard and the revised version of the RMET-C, and VIQ was reported not to be related to their performance. One of the limitations of this study is that the authors did not use a standardized measure of intelligence since VIQ was represented by PPVT, a measure of receptive vocabulary rather than verbal cognitive potential.

In contrast to aforementioned studies, a study by Philpott, Rinehart, Gray, Howlin, and Cornish (2013) revealed a different pattern of results. The participants included 12 typically developing children (mean age: 12.0 years, range: 9.9–14.8 years) and 12 high-functioning children with ASD (mean age: 11.0 years, range: 9.1–13.6 years). The children in the two groups were matched on verbal and nonverbal IQ using the Wechsler scales and were also administered the RMET-C as well as another ToM test. Results indicated that children in the ASD group were not impaired on the RMET-C nor the additional ToM task relative to typically developing children, although the authors acknowledged small sample size as a potential limitation. Additionally, the authors did not administer a standardized language measure. Philpott et al. (2013) indicated, however, that the results on the RMET-C were comparable to results of Baron-Cohen et al (2001) once intellectual ability and age were taken into an account.

Rueda, Fernandez-Berrocal, and Baron-Cohen (2015) conducted a detailed study that included analysis of valence by comparing 38 participants (9-17) with Asperger Syndrome to a group of typically developing children and adolescents matched for age and gender. There were no differences in the overall IQ between the two groups measured by the Wechsler scales and IQ was reported to be in the average range for both groups. Children in the AS group performed significantly lower than the typically developing individuals on the RMET-C. Additionally, analyses for each dimension of emotional valence (positive, negative, and neutral) on the RMET task revealed that children with ASD identified fewer positive mental states in the eyes than typically developing individuals. There were no differences in the frequency of identifying negative or neutral emotions in the eyes. One of the limitations of this study is that the authors did not report separate Verbal and Nonverbal IQ results and no language measure was administered so it is not possible to ascertain whether the difference between the groups was due to language abilities.

Holt et al. (2014) included an assessment of verbal skills as part of their experiment and conducted an fMRI study during administration of RMET (revised version) involving adolescents with autism and their siblings. Specifically, 50 participants with Autism Spectrum Condition (ASC) and 40 unaffected siblings took part in the study as well as 40 matched typically developing individuals. The results revealed that when Verbal IQ was controlled for, males with ASC were significantly more impaired compared to typically developing participants and a trend level of significance was noted in the reduced performance of female youth with ASC as compared to controls. There was no difference in performance between siblings and same-sex participants in the control groups. In terms of brain activation, there were significant differences between the ASC and control groups during the administration of the RMET task. Specifically, typically developing male youth had greater activation in the left inferior prefrontal gyrus, orbitofrontal cortex, temporopolar and middle temporal gyrus than male youth with ASC. Similarly, typically developing female youth exhibited greater activation compared to females with ASC in the left orbitofrontal and temporopolar areas, and also in the bilateral inferior and anterior prefrontal cortices than female youth with ASC. According to the authors, the involvement of frontotemporal regions during the RMET serves as potential neuroendophenotypic marker for ASC. These impacted brain regions have been associated with social cognition and processing of language (Holt et al., 2014). One of the limitations of the study was that a standardized measure of language was not included.

Vogindroukas, Chelas, and Petridis (2014) investigated performance on the Greek version of the RMET-C (child) in three different groups, including typically developing (TD) children (mean age 9.84 years, $n = 23$), children with high functioning autism (HFA) (mean age = 10.06 years, $n = 27$) and typically developing (TD) adults (mean age 38.2 years, $n = 53$). Participants with HFA were assessed using the Greek version of the Wechsler Intelligence Scale III and had a mean IQ score of 93.6 ($SD = 7.7$). The authors revealed that TD adults scored slightly higher than TD children, and children with HFA scored lower than their TD peers on the RMET-C. Children with HFA, however, were able to recognize many of the pictures shown in the test and statistical differences were found in 5 out of 28 items. Age and diagnosis had an effect on the overall score, that is, in 18 out of 28 test items or variables, age was a statistically significant covariate, and more than half of the variables were positively correlated with participants' diagnosis. Children with HFA had a mean score of 13.6 out of 28 questions correct. In terms of qualitative analysis, the test item with the highest number of correct responses was a “negative feeling” for both TD adults and children with HFA, and a “negative mental state” for TD children. In general, all participants recognized more negative feelings than positive ones. One of the limitations of the study is that the overall IQ was provided only for participants with HFA and the verbal abilities for the three groups were not represented by standardized measures.

Bennett, Szatmari, Bryson, Duku, Vaccarella, and Tuff (2013) investigated ToM as a possible mediator between language and adaptive functioning in 39 children with ASD (IQ>70 using the Leiter Scale of Intelligence). The participants were followed prospectively every two years from 4-6 years to 12-14 years. Their language skills were assessed using the Test of Language Development-2 (at age 6-8 years), the “Eyes Test” or RMET at 10-12 years and the Vineland Adaptive Behaviour Scales at age 12-14 years. Based on the results, there was a strong association ($\beta = 0.65$, $p < 0.01$) between language ability and the RMET for participants who had higher IQs. The results of this study are consistent with longitudinal studies by Pellicano (2010), as well as Tager-Flusberg and Joseph (2005) who both reported an association between language ability and subsequent ToM skills measured by the false-belief tasks. One of the limitations of the study by Bennett et al. (2013) is that language and ToM were not measured across the three age intervals; therefore, acquisition of language and social cognitive skills could not be compared over the different developmental periods.

In summary, based on studies outlined above, children and adolescents with ASD generally have more difficulty on the RMET task than typically developing children, even when cognitive abilities are controlled. The role of language abilities in relation to the performance on RMET and the development of ToM in general in individuals with ASD requires further exploration. Additionally, most studies cited above place participants in one age category so the impact of developmental maturation and presence of possible watershed periods on the development of language and ToM in individuals with autism may be missed. Including accuracy and pattern of response on the RMET-C across different age intervals may assist in deciphering the development of ToM at different developmental stages. As a result, the first objective of the present study was to establish which developmental periods are associated with the development of ToM in children and adolescents with autism. The second objective was to determine which components of language abilities (e.g., expressive, receptive language skills) are associated with the development of the ToM in individuals with autism. Lastly, since two studies outlined above reported findings pertaining to individuals with ASD choosing fewer positive emotions and mental states (Rueda et al., 2015) and more negative emotions (Vogindroukas et al., 2014) on the RMET-C task, the third objective of the present study was to extend these findings and determine the frequency of valance of emotions across all items on the RMET-C in individuals with ASD.

Nonverbal Communication and the Development of the Theory of Mind

As has been suggested in some of the studies above, impaired language skills contribute to the development of the deficient ToM in individuals with ASD. As a result, the precursors to the development of language, such as nonverbal communication, including expression and perception of emotion may impact the pathway towards the development of ToM and social cognition, either prior to or during the development of language. Preverbal skills in young children refer to the use of and response to eye contact and gestures during social interactions, and thus comprise the foundations of language development (Mundy & Crawson, 1997). For example, joint attention, which involves the use of behaviours (e.g., showing or pointing) to facilitate the synchronization of shared engagement between people and objects or events (Mundy, Sigman, Ungerer, & Tracy, 1986) is seen as a developmental milestone that usually appears between the ages of 8 and 15 months in typically developing children (Bakeman & Adamson, 1984; Jones, Carr, & Feeley, 2006). It is possible that, as Hobson (1993) has argued, children with autism are born with an inability to adequately express or perceive emotion and this in turn may lead to socio-communicative deficits associated with autism, such as impaired nonverbal and verbal communication, difficulties with the regulation of emotion, stereotypical behaviours, and impaired understanding of the ToM.

Other than deficits in verbal communication and executive dysfunction, it has not been determined whether other impairments associated with autism may be related to the development of a deficient ToM. For example, Li, Liu, Yang, Cao, and He (2012) reported that children with high functioning autism may have impairments in executive function and ToM, and the impairment of executive function and ToM may be correlated with more severe symptoms of autism. In order to decipher what social, communicative (including nonverbal) and behaviour impairments associated with autism may lead to development of the deficient ToM, studies that examined a connection between performance on a task involving assessment of symptoms of autism on the Autism Diagnostic Observational Scale (ADOS) and the ToM tasks are reviewed next.

Theory of Mind and the Autism Diagnostic Observational Scale

The Autism Diagnostic Observation Schedule (ADOS) has been considered a “gold standard" diagnostic tool for assessment ASD. The assessment using the ADOS is dynamic and the tasks on this measure are designed to elicit communicative (verbal and nonverbal), social, and behavioral characteristics consistent with ASD. The ADOS consists of five modules and the selection of specific modules is primarily based on the examinee's level of expressive language either ascertained by formal standardized measures of language or observation by examiner (Lord, Luyster, Gotham, & Guthrie, 2012). To date, a limited number of studies have focused on performance on the ADOS by individuals with ASD as it relates to their performance on ToM tasks. In the following section, I review the limited existing literature on the symptoms of autism on the ADOS and the ToM.

Livingston, Colvert, the Social Relationships Study Team, Bolton, and Happe (2018) studied 136 adolescents with ASD (aged 10-15 years). The ToM task involved administration of a computerized Frith-Happe Animations task involving false belief (Abell, Happe, & Frith, 2000). The authors used Module 3 of the ADOS and included "social and non-social symptoms" associated with features of autism. Participants who were assigned to a group of "High Compensators" had "good" ADOS scores and thus low number of social difficulties associated with autism despite poor ToM performance. They also included a group of "Low Compensators" who exhibited similarly poor ToM performance but also reduced performance on the ADOS. Based on the results, higher IQ (Verbal), better EF, and greater self-reported anxiety were accounted for in the process of compensating for ToM difficulties in "high compensating" individuals with ASD with fewer symptoms of ASD. When comparing individuals with good versus poor ADOS performance, it is of note that no differences were found when ToM performance was well developed across these groups suggesting that the severity of ADOS scores (and thus symptoms of autism) is independent of ToM.

As described earlier, both language (and Verbal IQ) as well as EF have been strongly associated with ToM. It can be inferred from the above study that despite strong associations in these areas, some individuals with ASD continue to demonstrate reduced ToM capacities, but have fewer symptoms associated with ASD. In effect, ToM difficulties may not always be predictive of ASD symptoms. At the same time, the authors have suggested that anxiety may mask symptoms of autism on the ADOS although the mechanisms by which this could happen was not explained/explored. Additionally, the increased anxiety scores and lower scores on the ToM task may point to an inverse association between these two areas, with the possibility that anxiety may also play a role in the development of the ToM. In effect, high anxiety may be associated with reduced scores on the ToM but help with other symptoms of autism. One of the limitations of the study by Livingston et al. (2018) is absence of an exploration of a possible relationship between observed anxiety (e.g., on the ADOS) and its correspondence to the symptoms of autism and performance on the ToM tasks. As well, individual items on the ADOS pertaining to social and nonsocial functioning were not included in the analysis but rather presented as composite scores.

Consistent with the above finding that anxiety interferes with the development of intact ToM, Kamp-Becker, Ghahreman, Smidt, and Remschmidt (2009) analyzed whether early symptoms associated with autism as measured by the ADI-R could be predictive of symptoms of autism evaluated through the ADOS, as well as, adaptive, IQ and ToM measures. Participants included 104 individuals with ASD diagnosis (aged 6-24) with a Full Scale IQ greater than 70 and 36 typically developing individuals in the same age range. The theory of mind skills were evaluated by three tests, including, face recognition (Benton, Sivan, Hamsher, Varney, & Spreen, 1983), facial emotion recognition (www.candit.com: Facial Emotion Matching), and social attributions test (social attribution test, Klin and Jones 2006). One value was provided for theory of mind skills by adding the number of correct responses on the three tests. The authors reported that the social interaction and communication domains were closely related and emerged as a single factor they called "social communication". The key finding, according to the researchers was emergence of ''anxiety and compulsion'' factor in early development that was found to be associated with "current" social communication functioning. The authors did not find a correlation between the main social communication factors of ADOS and ADI and only modest associations between early ADI and current ADOS scores were found in relation to ''anxiety and compulsions''.

In terms of the ToM, consistent with the study by Kamp-Becker et al. (2009), the more symptoms of ''anxious and compulsive behavior'' exhibited in early childhood, the lower were theory of mind abilities. In contrast to the study by Livingston (2018), Kamp-Becker et al. (2009) demonstrated that the symptoms of ''anxious and compulsive behavior'' in early childhood were related to more symptoms of autism (higher score on social communication scale). The authors noted that the IQ did not predict the main symptoms, that is, the performance on IQ measures did not correlate with "earlier" or "current" social communication functioning; however, it is unclear what role language played in these findings.

The above two studies found an association between low ToM and high anxiety but they were contradictory in terms of the number of symptoms present in relation to social functioning.

In the first study, the authors reported low ToM, high anxiety and low number of symptoms associated with social communication on the ADOS. In the second study, high anxiety early in life was related to low ToM and more symptoms associated with social communication difficulties on the ADOS. A state of anxiety may interfere with the development of ToM but it is unclear how anxiety and ToM are related to symptoms of social functioning in individuals with autism.

Romero, Fitzpatrick, Roulier, Duncan, Richardson, and Schmidt (2018) studied the relationship between measures of joint attention and theory of mind (as measured by verbal and nonverbal false belief tasks) in relation to predictors of ASD severity. Severity was assessed with the ADOS severity score, the Restrictive and Repetitive Behaviors (RRB), and the Social Affect subscale (SA), as well as whole body movements (coordination and variability) during conversation with the clinician. There were 45 participants with a diagnosis of ASD and a mean age of 8.5 years. They found that the traditional measures of social cognitive deficits (theory of mind and joint attention) were related to body movements. This subscale of body movements includes the frequency of typical social behaviors such as the use of gestures, making eye contact, and appropriateness of responses during social interactions, in essence different aspects of communication. The ToM task, along with RRB on the ADOS and whole bodily movement measures, accounted for 35% of variance. The ADOS Social Affect subscale, ToM, and joint attention loaded onto a second factor that explained 25% of the variance. These results suggest that ToM was comprised of two factors, namely socioemotional functioning and repetitive behaviours and interests. Although the authors did not directly address anxiety, the category of repetitive interests and behaviours may be linked to anxiety in that individuals who perform compulsive acts may do so in order to reduce their levels of anxiety.

In summary, there appears to be a relationship between ToM and symptoms associated with autism, such as, socioemotional factors, behaviour difficulties, and anxiety. Although the direction of these relationships needs to be investigated further, generally, higher anxiety is associated with lower performance on the ToM tasks and in the majority of studies described above, the increased socioemotional symptoms associated with autism were related to lower performance on the ToM tasks. No study to date has examined a relationship between Reading Mind in the Eyes Test in children (RMET-C) and performance on the ADOS.

In order to decipher which components of socioemotional and behavioural functioning on the ADOS are related to the development of the ToM in individuals with autism, the fourth objective of the present study was to examine performance between composite scores that comprise social, communicative and behavioural domains and performance on the RMET-C, as well as, to examine performance between individual items on the ADOS that comprise these domains and performance on the RMET-C. Lastly, the fifth objective was to explore the relationship between observed anxiety scores on the ADOS and performance on the RMET-C.

The Present Study (Objectives and Hypotheses)

The general purpose of the current study was to explore the performance of individuals with autism on one of the most widely used ToM tasks, the RMET-C, across different developmental stages and to determine a possible correlation between performance on the RMET-C task and cognitive skills, language skills, as well as social, communicative, and behaviour symptoms associated with autism. The study had the following five objectives, outlined in detail below with the resulting hypotheses.

The first objective of the present study was to establish which developmental periods are associated with the development of ToM abilities in individuals with autism, specifically, to determine variability in performance on the RMET-C across different developmental stages (age intervals). This objective was accomplished by categorizing participants with ASD into different age categories and analyzing their performance across these, as well as, comparing these age group scores to published data for typically developing individuals. I hypothesized that there would be an increase in response accuracy on the RMET-C across ascending age intervals up to age 12, based on the trend in typically-developing populations (Baron-Cohen, 2001). Based on studies by Kaland et al. (2007) and Baron-Cohen et al. (2001), I expected that participants with ASD would perform less well than the typically developing youth. I hypothesized that there would be gender differences in the performance of participants with ASD, consistent with Holt et al. (2014) who reported significantly lower performance on the RMET-C in male participants with ASD and a trend toward lower performance in female participants with ASD. An exploratory analysis of performance on the RMET-C for participants who received diagnosis based on the criteria from the DSM-IV was also provided since these criteria were used in studies prior to the development of the DSM-5. Participants with Asperger's were hypothesized to have the highest performance on the RMET-C because of absence of significant delay in language (APA, 1994).

The second objective of the present study was to determine which components of cognitive abilities, as well as expressive and receptive language skills are associated with the development of theory of mind in individuals with autism. This was accomplished by comparing the performance of individuals with ASD on standardized measures of intelligence and language skills to their performance on the RMET-C. I hypothesized that consistent with the studies by Brent et al. (2004) and Bennett et al. (2013) who found strong association between language skills and performance on the RMET that there would be a strong correlation between expressive and receptive language skills and performance on the RMET-C. Consistent with the studies by Brent et al. (2004) and Rueda et al. (2015), I expected that for individuals' with ASD, performance on the RMET-C would be independent of their intellectual abilities.

The third objective of the present study was to determine whether individuals with ASD are more likely to choose negative or positive emotions or mental states on the RMET-C. This was established by classifying responses on the RMET-C into positive and negative mental states, as well as positive and negative emotions, then tabulating the percentage of correct responses for each category. Consistent with the study by Vogindroukas et al. (2014) who revealed that children with High Functioning Autism (HFA) identified more negative than positive feelings, and slightly more negative mental sates than feelings, I hypothesized that participants in the current study would identify more negative emotions and mental states than positive emotions and mental states.

The fourth objective was to identify which components of socioemotional and behavioural functioning associated with symptoms of autism are related to the development of the ToM. This was accomplished by examining both a) participants' performance on individual items on the ADOS that comprise social, communicative and behaviour domains, as well as b) performance on those domains and results on the RMET-C. No study to date has examined the correlation between performance on the ADOS and the RMET-C task; however, based on the limited literature on ToM and ADOS, I hypothesized that increased socio-communicative symptoms on the ADOS would be related to reduced performance on the RMET-C. This hypothesis was consistent with the study by Kamp-Becker et al. (2009) who reported that socio-communicative deficits are associated with lower ToM and with the study by Li et al. (2012) who noted that the severity of autism may be associated with lower performance on the ToM task. I also hypothesized that performance on the RMET-C would be correlated with Social Affect and Restrictive and Repetitive Behaviors (RRB) performance, based on Romero et al.'s (2018) findings that ToM along with RRB and body movements account for the same amount of variance on RMET-C.

The fifth objective was to explore the relationship between anxiety and performance on the RMET-C, which has not been studied before. This was accomplished by comparing the anxiety score on the ADOS to performance on the RMET-C for individuals with ASD. I hypothesized that higher anxiety would be predictive of lower scores on the RMET-C consistent with the findings by Kamp-Becker et al. (2009) and Livingston et al. (2018).

Method

Data for the current study were derived from participants with ASD who took part in the "Molecular and Genomic Analysis of Autism Spectrum Disorders and Related Conditions" (Dr. Stephen Scherer, Principal Investigator), a genetics study designed to provide information on genetic variants in individuals with ASD, which also collected cognitive and socioemotional functioning data. The inclusion criterion for the current study was completion of the RMET-C. Each participant entering the study had a diagnosed of Autism Spectrum Disorder and as part of the research protocol, this diagnosis was confirmed using the Autism Diagnostic Interview-

Revised (ADI-R) or the Autism Diagnostic Observational Scale-Generic (ADOS-G). Participants were recruited through the Hospital for Sick Children (HSC) (Toronto), Holland Bloorview Kids Rehabilitation Hospital (Toronto), Chedoke-McMaster Hospital (Hamilton), and Lawson Health Research Institute (London). Information was also collected on medical, developmental, and family history, and physical measures of the participants. As part of the Genetics study, the participants were administered various psychological and medical assessment measures; however, for the purpose of the current study, the measures of interest included the Reading the Mind in the Eyes Task for Children (RMET-C), the ADOS-G (Modules 3 and 4), standardized language measure and intellectual measures.

Participants

Participants with complete RMET-C data who were at least 6 years of age and completed intellectual measures in verbal and nonverbal domains were selected for this study. The number of participants who had data for the RMET-C as well as verbal and nonverbal intellectual measures was 91 (70 males and 21 females) and of these, 90 participants had data for the Oral and Written Language Scales-Second Edition (OWLS-II), 59 participants had data for ADOS II-Module 3, and 24 participants had data for ADOS II-Module 4. All 91 participants met criteria for Autism Spectrum Disorder (ASD) based on the DSM-5 classification and of these there were 83 participants who also had a diagnosis based on the DSM-IV criteria: Asperger syndrome (n = 30), Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS) (n =11), Autism (n = 16), and Autism Spectrum Disorder/PDD (n =26). The remaining 8 participants had diagnosis consistent only with the DSM-5 classification, the Autism Spectrum Disorder (ASD) so they were not included as part of the statistical analysis related to the diagnostic criteria.

The 91 participants ranged in age from 6.5 to 20.7 years with a mean age of 11.8 years and a SD of 34. To examine performance based on developmental maturation, the participants were divided into five groups representing different age intervals. The first three age groups, 6-8, 8-10, and 10-12 were determined based on the original study by Baron-Cohen et al. (2001) who presented the first normative data for typically developing children in those age groups. Given the heterogeneity of age intervals presented in the studies outlining the findings of RMET-C during the adolescent period as described in the introduction, the subsequent two age groups, 12-16 and 16-20, were determined based on the developmental stages associated with middle adolescence and late adolescence to adulthood, respectively. Table 1 presents the characteristics of participants in the present study.

Table 1
Characteristics of participants

Age Groups	Gender	*N*	*M* (years)	*S.D.*
6 to 8	Male	6		
	Female	3		
	Total	9	7.3	0.6
8 to10	Male	24		
	Female	6		
	Total	30	9.2	0.5
10 to12	Male	13		
	Female	4		
	Total	17	11.2	0.5
12 to16	Male	18		
	Female	3		
	Total	21	14.2	1.2
16 to 20	Male	9		
	Female	5		
	Total	14	17.6	1.5
Total (all age groups)	Male	70		
	Female	21		
	Overall Total	91	11.8	3.4

All participants completed the Verbal Comprehension Index of either the WPPSI4, WISC4, or WASI2. There were no significant differences in performance on the Verbal Comprehension Index between the different age groups, (ANOVA) $F(4, 86) = 0.51$, $p = 0.73$. Additionally, all participants completed the Fluid Reasoning Index of the WPPSI4 or the Perceptual Reasoning Index of the WISC4 or WASI2, and similarly, no significant differences in performances on these Nonverbal Indices of the Wechsler scales between the different age groups were found, (ANOVA) $F(4, 86) = 1.02$, $p = 0.40$). Tables 2 and 3 display the psychometric data for all ASD participants.

Table 2

Psychometric data for the ASD participants: Verbal IQ-Standard Scores

Age Groups	*N*	Mean	*SD*	Minimum	Maximum
6 to 8	9	103.56	24.13	81	147
8 to 10	30	102.63	18.79	70	144
10 to 12	17	101.06	26.83	55	146
12 to 16	21	94.57	19.06	63	134
16 to 20	14	102.07	25.19	49	142
Total	91	100.48	21.84	49	147

Table 3

Psychometric data for the ASD participants: Noverbal IQ-Standard Scores

Age Groups	*N*	Mean	*SD*	Minimum	Maximum
6 to 8	9	102.56	9.61	83	116
8 to 10	30	101.53	19.42	64	142
10 to 12	17	95.88	19.42	61	134
12 to 16	21	92.62	23.08	55	132
16 to 20	14	103.50	20.40	49	129
Total	91	98.82	19.82	49	142

Measures

The Reading the Mind in the Eyes Test Child version (RMET-C) was developed by Baron-Cohen et al., (2001) and consists of 28 photographs of the eye region of the face for different individuals. The child version of the RMET was derived from the original 36-item adult RMET by changing more complex mental state words to a child friendly vocabulary (Baron-Cohen et al., 2001). As part of the administration, the examiner asks the child to look carefully at pictures of people's eyes and then choose the word that best describes what the person in the picture is thinking or feeling. There are four response choices of corresponding mental states or feelings and only one out of the four is correct. For example, for the practice item (Figure 1), the examiner asks: "Do you think he is feeling jealous, scared, relaxed or hate?"

The examiner is required to point to the words while reading them and provide encouraging feedback. The correct number of responses is calculated out of 28 questions. The RMET-C can be downloaded through the Autism Research Website (http://www.autismresearchcentre.com/arc_tests). There are no norms available for this test; however, a number of published studies have included psychometric information on typically developing populations (e.g. Baron-Cohen et al., 2001).

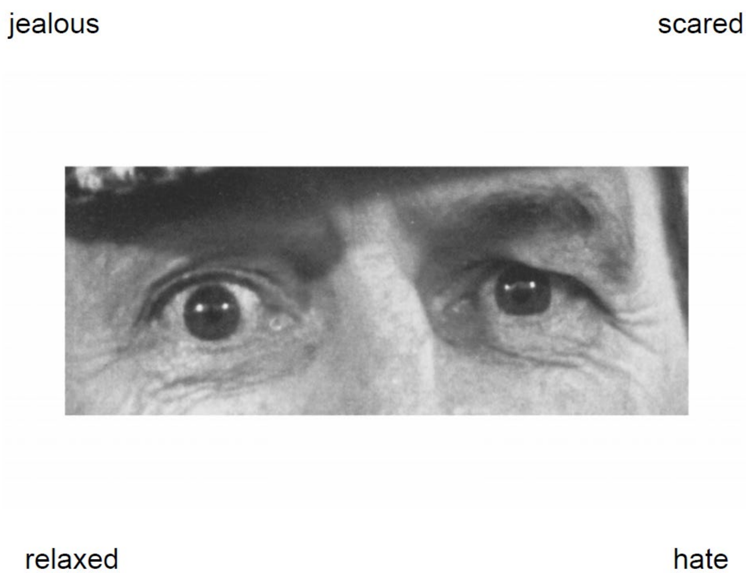

Figure 1. Sample item from RMET-C (Baron-Cohen, 2001).

The ADOS-G (Lord, Rutter, DiLavoire, & Risi, 1999) is a semi-structured assessment protocol that consists of various activities that facilitate observation of social interaction, communication, play, and behaviours consistent with symptoms of ASDs. The Autism Diagnostic Observation Schedule, Second Edition (ADOS-2; Lord et al., 2012) is an updated version of the original ADOS and was designed to improved precision of diagnostic algorithms of the original version. The ADOS-2 consists of 5 modules, each requiring 35-40 minutes to administer. The individual being evaluated is given one module, selection of which is primarily dependent on examinee's expressive language level established by formal assessment or examiner's observation, and secondarily on examinee's chronological age and interest in relation to items administered. The items on the ADOS comprise individual and domain scores, as well as data for diagnostic criteria. Module 3 has been updated to include two (from previous three) domains, including Social Affect (SA) and Restrictive and Repetitive Behaviours (RRB) for which a score is provided (McCrimmon & Rostad, 2014). SA, consists of the formerly known Communication and Reciprocal Social Interaction domains. For Modules 1 through 3, logistic regressions revealed that both, the SA and RRB domains made significant independent contributions to the prediction of diagnosis and the overall total score produced the highest predictive value, supporting the use of this unified score in diagnostic decisions (McCrimmon and Rostad, 2014).

In the current study, Modules 3 and 4 were used. McCrimmon and Rostad (2014) reviewed the updated ADOS-2 measure and described Module 3 as best suited for verbally fluent children and adolescents for whom action figure toy play is age appropriate (approximately 16 years or younger). This module consists of 14 activities that inform 29 coded items pertaining to SA and RRB. Module 4 is suited for verbally fluent older adolescents and adults. It consists of 10 to 15 activities that inform 32 coded items. Module 4 was not revised with the ADOS-2 and its coded items are structured for separate Communication, Reciprocal Social Interaction, and RRB domains. Each child's response is scored on a 4-point Likert scale with a score of 0 indicating no evidence of atypical behavior, 1 mild difficulties, 2 definite concerns, and 3 profoundly atypical behavior. Other codes may be used to depict specific language difficulties or behaviours, or items that cannot be scored (McCrimmon and Rostad, 2014). The ADOS is widely considered to be the “gold standard” observational assessment tool for the diagnosis of ASD. Both Module 3 and Module 4 include an item assessing anxiety on a 3-point scale (0 = no anxiety, 1 = mild, and 2 = pronounced).

The language skills of participants were assessed using the Oral and Written Language Scales (OWLS, ages 3 to 21 years), specifically the Listening Comprehension and Oral Expression Indices as well as the Oral Language Composite that encompasses these two aforementioned indices (Carrow-Woolfolk, 2011). The Listening Comprehension scale measures oral language reception, which is the understanding of spoken language. The examiner orally presents increasingly difficult words, phrases, and sentences and the examinee is asked to respond by pointing to or stating which of four pictures is correct. The Oral Expression scale measures the use of spoken language. The examiner presents a verbal prompt along with a picture and the examinee is asked to respond orally to the prompt with increasingly difficult language.

To assess cognitive ability, one of the Wechsler Scale of Intelligence measures was chosen depending on the age of the participant. Specifically, Wechsler Preschool and Primary Scale of Intelligence (WIPPSI-IV, ages 2:6 to 7:3) (Wechsler, 2012), or Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV, ages 6:0 to 16:11) (Wechsler, 2003), or Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV, ages 16:00 to 90:11) (Wechsler, 2008) or the abbreviated version, the Wechsler Abbreviated Scale-1st edition (WASI, ages 6:00 to 90:00) (Wechsler, 1999).

Procedure

The youth were assessed at one of the health care centres outlined above by a psychometrist who worked under a supervision of a psychologist. The consent was obtained from parents/guardians of children who were under the age of 16 and consent from participants over the age of 16. Additionally, assent was obtained from youth who were able to provide it. The consent included, but was not limited to, the explanation of the purpose of the study, possible harms and benefits (if any), confidentiality, and answering questions. Participants received a summary report from psychological assessment as well as remuneration for parking at Sick Kids when they participated in the study.

Results

An alpha level of 0.05 was used for statistical significance unless otherwise specified. The Sidak multiple comparisons test was used for the purpose of post hoc analyses where indicated. The data was analyzed using the SPSS versions 25 and 26.

Objective 1. Performance of ASD Participants on the Reading the Mind in the Eyes Task –Children (RMET-C)

The effect of age and gender. Performance on the RMET-C task was examined in terms of mean accuracy, that is, the percentage of correct responses of the target words for each participant was determined and represented across the five age groups and gender (see Table 4). An ANOVA was conducted to determine whether there was a difference in performance across the two fixed factors, age group and gender. The results revealed that the mean accuracy on the RMET-C task was not significantly different between the five age groups (ANOVA), $F(4, 86) = 1.51$, $p = .21$). However, after controlling for the effect of Verbal and Nonverbal IQ using the

ANCOVA, there was a significant effect of age group on mean accuracy on the RMET-C task, $F(4, 84) = 2.54$, $p = .04$. The first covariate, Verbal IQ was a significant predictor of the performance on the RMET-C task, $F(1, 84) = 8.79$, $p = .004$, as was the second covariate, the Nonverbal IQ, $F(1,84) = 4.92$, $p = .029$. After adjusting for both covariates, the pairwise comparisons revealed no differences between the age groups. The mean accuracy on the RMET-C task was not significantly different based on gender, (ANOVA) $F(1, 89) = .02$, $p = 0.88$, and remained insignificant after controlling for the effects of Verbal and Nonverbal IQ, $F(1,87) = 1.07$, $p = .303$.

Table 4

Mean accuracy on the RMET-C task across five age groups

Age Groups	Mean	*SD*
6 to 8	13.89	4.01
8 to 10	15.80	4.59
10 to 12	15.71	4.03
12 to 16	16.19	3.50
16 to 20	18.14	5.11

Comparison to means for typically developing youth. There is no single source of standardized normative data for the RMET-C task and as a result, only published investigations involving typically developing participants can be used as a comparison. The first study to reference normative data for typically developing children was by Baron-Cohen et al. (1997) and since that time several studies have outlined performances for typically developing individuals, as well as for clinical populations across variable age groups. One of the challenges of using the data from different investigators for comparative analysis is the heterogeneity of age of the participants across the different studies. For the purpose of the present study, the most suitable comparison for the 6-8, 8-10, and 10-12 group was the normative data from Baron-Cohen et al. (2001) since identical age intervals were used. For the 12-16 age group, the study by Muller and Gmunder (2014) provides results for typically developing participants with similar age and this is also the largest sample size published to date with that age group (participants aged 13-15 years). Lastly, for the 16-20 age group, a comparison to Moor et al.'s (2012) 19-23-year-old participants was made. There is a paucity of studies for the 16-18 age range. Although studies exist that encompass wider age intervals, such as, 12-19 years (Overgaauw, van Duijvenvoorde, Moor, & Crone 2015), these would not be appropriate comparisons for the age intervals in the present study due to the broad age ranges.

One sample t-tests were conducted to compare the number of correct responses on the RMET-C task for participants with ASD in the current study to typically developing youth from the existing studies outlined above by using the mean and standard deviation from the published studies (see Table 5). Based on the results, there was no significant difference between the observed and the expected mean for typically developing youth in the youngest age group (6-8 years), $p = .72$, nor in the oldest group (16-20 years), $p = .27$. There were, however, significant differences in the three middle age groups: 8-10 years, $p = .02$, 10-12 years, $p < .001$, and 12-16 years, $p = .05$. Specifically, the ASD participants in the present study scored significantly lower in those three groups than the same age typically developing participants reported by Baron-Cohen et al. (2001) and Muller and Gmunder (2014). On the other hand, participants in the lowest age range (6-8 years) performed closely to the participants in the study of Baron-Cohen et al. (2001) and participants in the highest range (16-20 years) had similar performance to participants in the study of Moor et al. (2012).

Table 5
Data for typically developing youth (number of correct responses) on RMET-C in previous studies and comparison to ASD participants in the current study

	Normative Data - Previous Studies			ASD Group-current study			
Studies with norms	Range	Mean	*SD*	Range	Mean	*SD*	p value
Baron-Cohen et al. (2001)	6 to 8	13.4	5.4	6 to 8	13.89	4.01	0.724
Baron-Cohen et al. (2001)	8 to 10	17.9	4.2	8 to 10	15.80	4.59	0.018
Baron-Cohen et al. (2001)	10 to 12	20.6	2.4	10 to 12	15.71	4.03	0.000
Muller and Gmunder (2014)	13 to 15	17.8	3.5	12 to 16	16.19	3.50	0.048
Moor et al. (2012)	19 to 23	19.7	2.00	16 to 20	18.14	5.11	0.275

While the data presented for the five age groups is cross-sectional and results pertain to different participants in these groups, as evident in Table 5, there was generally an increase in the mean response accuracy on the RMET-C across ascending age intervals for participants with autism and the increase in performance continued until early adulthood, that is, for participants with ASD in the present study the mean performance was 13.88, 15.8, 15.7, 16.19, and 18.14 in the 6-8, 8-10, 10-12, 12-16, and 16-20 age groups, respectively. A different trend was observed for typically developing individuals based on the published data for different age groups. Notably, after a steady increase in mean accuracy, there was a decline in middle adolescence (13-15 years) with recovery in mean accuracy in the period of late adolescence to early adulthood (19 years to 23 years), that is the mean performance of typically developing individuals was 13.4, 17.9, 20.6, 17.8, and 19.7 for the 6-8, 8-10, 10-12, 13-15, and 19-23 age groups.

ASD status. All participants in the current study met the diagnostic criteria based on the DSM-5 and there were 83 participants in this group who also received diagnosis based on the DSM-IV classification, including: autism, PDD-NOS, Asperger's Syndrome, and ASD PDD. Although currently the diagnostic criterion for ASD in the DSM-5 encompasses these aforementioned diagnostic categories from the DSM-IV, analysis was conducted to determine the effect of the Clinical Diagnosis based on the DSM-IV classification on mean performance on the RMET-C for these 83 participants (see Table 6). Based on a one-way between-subjects ANOVA, there was a significant effect of Clinical Diagnosis on the mean performance on the RMET-C $F(3, 79) = 6.29$, $p = .001$. Post hoc comparisons using the Sidak test indicated that the mean performance (in percentiles) on the RMET-C for participants with Asperger's diagnosis ($M = 66.89$, $SD = 12.29$) was significantly higher than the remaining groups, including, autism ($M = 52.68$, $SD = 13.01$), PDD-NOS ($M = 50.65$, $SD = 19.61$), and ASD ($M = 53.98$, $SD = 15.58$). After controlling for the effect of Verbal and Nonverbal IQ using an ANCOVA, the effect was approaching significance for Clinical Diagnosis on mean accuracy on the RMET-C task, $F(3, 78) = 2.58$, $p = .06$. The first covariate, Nonverbal IQ was a significant predictor of the performance on the RMET-C task $F(1,78) = 5.21$, $p = .02$; however, the second covariate, the Verbal IQ was not a significant predictor of the performance on the RMET-C $F(1,78) = 2.46$, $p = .12$.

Table 6

Mean accuracy on the RMET-C task across DSM-IV diagnostic classification

Diagnosis	*N*	Mean	*SD*	Min.	Max.	Mean/28
Autism	16	52.68	13.01	28.57	75.00	14.75
PDD-NOS/ Atypical	11	50.65	19.61	17.86	82.14	14.18
Asperger	30	66.89	12.29	39.29	82.14	18.73
ASD/PDD	26	53.98	14.58	32.14	85.71	15.12
Total	83	57.96	15.58	17.86	85.71	16.23

The mean accuracy on the RMET-C for the 8-14 Asperger's participants was 12.6 out of 28 correct in the Baron-Cohen (2001) study whereas it was about three points higher for participants in the current study, 15.8/28 and 15.7/28 for, 8-10 and 10-12 age groups, respectively.

Objective 2. Cognitive and Language measures and performance on RMET-C

Analysis for all youth (6-20 years). Correlations between the Oral Language Composite Standard Score, Verbal IQ Standard Score and Performance IQ Standard Score and mean percentage performance on the RMET-C were examined using the Pearson Correlation Coefficient. There was a strong positive relationship between the Oral Language Composite and performance scores on the RMET-C task, $r = .61$, $p < .001$ ($N = 90$), the Verbal IQ and the RMET-C task, $r = .50$, $p < .001$ ($N = 91$) and Performance IQ and the RMET-C task, $r = .48$, $p < 0.001$ ($n = 91$).

The Oral Language Composite score encompasses Oral Expression and Listening Comprehension Composites. These three domains were significantly correlated with the Mean Percentage Accuracy on the RMET-C task (see Table 7). Since Oral Expression and Listening Comprehension were highly correlated with the Oral Language Composite, a simple linear regression using the Oral Language Composite as a predictor of performance on the RMET-C task was conducted. A significant regression equation was found ($F(1,88) = 52.32$, $p < .001$) and the Oral Language Composite was found to explain 37.3% of variance, with R^2 of 0.37. It was found that the Oral Language Composite significantly predicted performance on the RMET-C task ($\beta = 0.611$, $p < .001$). A simple linear regression using the Verbal IQ as a predictor of performance on the RMET-C task was also conducted. A significant regression equation was found ($F(1,89) = 30.38$, $p < .001$) and the Verbal IQ was found to explain 25.4% of variance, with R^2 of 0.25. Verbal Performance significantly predicted performance on the RMET-C task ($\beta = 0.504$, $p < .001$). Finally, a simple linear regression using Performance IQ as a predictor of performance on the RMET-C task was conducted. A significant regression equation was found ($F(1,89) = 26.98$, $p < 0.001$) and the Performance IQ was found to explain 23.3% of variance. Verbal Performance significantly predicted performance on the RMET-C task ($\beta = 0.48$, $p < 0.001$).

Table 7

Pearson correlations between three domains on OWLS-2 and mean percentage accuracy on the RMET-C

Composites	Pearson r
Listening Comprehension	.588**
Oral Expression	.579**
Oral Language Composite	.611**

** Correlation is significant at the 0.01 level (2-tailed).

Subsequently, multiple regression analysis were performed and two models were created for the purpose of semi-partial (sr) or part correlation analysis. In model 1, two predictors were included, Verbal IQ and Performance IQ. In model 1, using the enter method, I found that Verbal IQ and Performance IQ explain a significant amount of the variance in the performance on the RMET-C task, ($F(2,87) = 17.40$, $p < 0.001$), $R^2 = .29$, $R^2_{Adjusted} = .27$. In model 2, in addition to Verbal IQ and Performance IQ, Oral Language Composite was also included and these three predictors explained more variance, ($F(3,86) = 18.15$, $p < .001$), $R^2 = .39$, $R^2_{Adjusted} = .37$ than predictors in model 1. Specifically, there was a significant R^2 change resulting in 0.10 difference between the R squares of Model 1 and 2 as a result of the third predictor being added.

Furthermore, in model 2 the first two predictors lost significance after the third variable, Oral Language Composite was added and that variable was highly significant. Specifically, the analysis revealed that neither Verbal IQ nor Performance IQ significantly predicted performance on the RMET-C task, ($\beta = .001$, $t(88) = 0.005$, $p > 0.05$), and ($\beta = .158$, $t(88) = 1.326$, $p > 0.05$), respectively. However, Oral Language Composite significantly predicted performance on the RMET-C task ($\beta = .510$, $t(88) = 3.78$, $p < .001$). The part correlation for Verbal IQ was 0 and 0.112 for the Performance IQ whereas the part correlation for the Oral Language Composite was .33. In other words, 10.18 % of variance in performance on the RMET-C was accounted for UNIQUELY by Oral Language Composite independently of the effects of other variables.

In summary, even though Verbal IQ and Performance IQ significantly predicted the RMET-C score in absence of the Oral Language Composite score, the variance predicted by Verbal and Performance IQ was lost when Oral Language Composite was introduced. The

Variance Inflection Factor (VIF) was below 4 for all variables, thus the three predictors were not highly correlated.

Analysis across the five age groups. A correlation between the Oral Language Composite standard score and performance on the RMET-C was examined across the five age intervals using the Pearson Correlation Coefficient. Based on the findings, there was a strong positive relationship between the Oral Language Composite and performance on the RMET-C task in the 8-10 age group, $r = .67$, $p < 0.001$ and the 10-12 age group, $r = .75$, $p = 0.001$, as well as in the 12-16 age group, $r = .51$, $p = 0.02$, and the 16-20 age group, $r = .63$, $p = .02$. Although there was no significant correlation in the 6-8 age group, this result could be due to small sample size (see Table 8).

The relationship between Verbal IQ and performance on the RMET-C was also examined across the five age intervals using the Pearson Correlation Coefficient. There was a strong positive relationship between Verbal IQ and performance on the RMET-C task in the 8-10 age group, $r = .55$, $p = .002$ and the 10-12 age group, $r = .73$, $p = 0.001$ (see Table 8). The correlations were not significant in the 6-8, 12-16, and 16-20 age categories, thus underscoring that oral and receptive language skills as assessed by the Oral Language composite are more predictive of performance on the RMET-C task than Verbal IQ for the majority of age intervals.

Next, a relationship between the Performance IQ and performance on the RMET-C was also examined across the five age intervals using the Pearson Correlation Coefficient. There was a strong positive relationship between Performance IQ and performance on the RMET-C task in the 8-10 age group, $r = .62$, $p < .001$, the 10-12 age group, $r = .56$, $p = .002$, and the 12-16 age group, $r = .67$, $p = .001$ (see Table 8). The correlations were not significant in the 6-8, and 16-20 age categories.

Table 8

Correlations between mean performance on the RMET-C and Verbal IQ, Oral Language and Performance IQ across the five Age Groups

Age Groups	Standardized Measures	*N*	Pearson Correlations
6-8	Combined_verbal_IQ		NS
	Oral Language	8	NS
	Performance IQ		NS
8-10	Combined_verbal_IQ		.549**
	Oral Language	30	0.673**
	Performance IQ		0.619**
10-12	Combined_verbal_IQ		.728**
	Oral Language	16	.7482**
	Performance IQ		0.565**
12-16	Combined_verbal_IQ		NS
	Oral Language	21	0.513*
	Performance IQ		0.668**
16-20	Combined_verbal_IQ		NS
	Oral Language	14	0.630*
	Performance IQ		NS

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed)
NS = not significant

Next, a one-way between subjects ANOVA was conducted to determine the effect of age group on mean performance (in percentiles) on the Oral Language Composite across the five age groups (Table 9). All alpha levels were set at 0.05. Based on the findings, there was a significant effect of age group on the Oral Language Composite (*F*(4, 85) = 3.03, *p* = 0.002). Post hoc comparisons using the Sidak test indicated that the mean score for the 12-16 age interval (*M* = 21.16, *SD* = 24.10) was significantly lower than the 16-20 age interval (*M* = 55.67, *SD* = 39.31). There was no statistical difference between the remaining age intervals. Looking across table 5 and 9, it is evident that language skills in the 12-16 age category (mean of 21.16) are lower than the 10-12 (mean of 32.64) age category in presence of nearly linear increase of performance on the RMET-C across these age categories.

Table 9
Mean performance (in percentiles) on the Oral Language Composite across the five Age Groups

Age Goups	Mean	*SD*	*N*	*p*
6to 8	24.55	18.79	8	NS
8 to 10	37.31	31.46	30	NS
10 to 12	32.64	30.42	16	NS
12 to 16	21.16*	24.10	21	sig
16 to 20	55.67*	39.31	14	sig

sig = significant
NS = not significant

Objective 3. Valence and Individual Item Analysis

Vogindroukas et al. (2014) classified the correct responses to the 28 questions on the RMET-C based on valence, that is, positive feelings, positive mental states, negative feelings, and negative mental states. To date, this has been the only study with this type of classification; therefore, this classification system was adopted for the present study, with one exception. For two questions on the RMET-C, Vogindroukas et al. (2014) did not provide classification, specifically, “making somebody do something” and “upset”. For the purpose of the present study, these two items were classified as negative mental state and negative feeling, respectively. The following classification of responses were created: 1. positive feelings: “hoping” and “happy”; 2. positive mental states: “kind”, “friendly”, “interested”, “remembering”, “thinking about something”, and “sure about something”; 3. negative feelings: “sadness”, “worried”, “nervous”, “not pleased” and “upset”; and 4. negative mental states: “serious”, “not believing”, “making up her mind”, “a bit worried”, “thinking about something sad”, and “making somebody do something”. Individual item responses are provided in Table 10.

Table 10

Breakdown of responses for 28 RMET-C questions for all participants (correct responses are in bold)

Question 1: (correct=3)

Item	Response	Frequency	Percent
1	hate	18	19.8
2	surprised	2	2.2
3	**kind**	**23**	**25.3**
4	cross	48	52.7
	Total	91	100

Question 2: (correct=4)

Item	Response	Frequency	Percent
1	unkind	20	22
2	cross	12	13.2
3	surprised	5	5.5
4	**sad**	**54**	**59.3**
	Total	91	100

Question 3: (correct=1)

Item	Response	Frequency	Percent
1	**friendly**	**20**	**22**
2	sad	8	8.8
3	surprised	8	8.8
4	worried	54	59.3
	missing	1	1.1
	Total	91	100

Question 4: (correct=2)

Item	Response	Frequency	Percent
1	relaxed	4	4.4
2	**upset**	**83**	**91.2**
3	surprised	2	2.2
4	excited	1	1.1
	missing	1	1.1
	Total	91	100

Question 5: (correct=2)

Item	Response	Frequency	Percent
1	feeling sorı	20	22
2	**making somebody do something**	**56**	**61.5**
3	joking	4	4.4
4	relaxed	11	12.1
	Total	91	100

Question 6: (correct=3)

Item	Response	Frequency	Percent
1	hate	3	3.3
2	unkind	4	4.4
3	**worried**	**50**	**54.9**
4	bored	34	37.4
	Total	91	100

Question 7: (correct=3)

Item	Response	Frequency	Percent
1	feeling sorry	5	5.5
2	bored	11	12.1
3	**interested**	**61**	**67**
4	joking	14	15.4
	Total	91	100

Question 8: (correct=1)

Item	Response	Frequency	Percent
1	**remem bering**	**56**	**61.5**
2	friendly	4	4.4
3	angry	31	34.1
	Total	91	100

Question 9: (correct=4)

Item	Response	Frequency	Percent
1	annoyed	11	12.1
2	hate	3	3.3
3	surprised	4	4.4
4	**thinking about something**	**73**	**80.2**
	Total	91	100

Quesdtion 10: (correct=3)

Item	Response	Frequency	Percent
1	kind	13	14.3
2	shy	30	33
3	**not believing**	**42**	**46.2**
4	sad	6	6.6
	Total	91	100

Question 11: (correct=2)

Item	Response	Frequency	Percent
1	bossy	22	24.2
2	**hoping**	**51**	**56**
3	angry	2	2.2
4	disgusted	16	17.6
	Total	91	100

Quesiton 12: (correct=4)

Item	Response	Frequency	Percent
1	confused	8	8.8
2	joking	6	6.6
3	sad	1	1.1
4	**serious**	**76**	**83.5**
	Total	91	100

Question 13: (correct=1)

Item	Response	Frequency	Percent
1	**thinking about something**	**50**	**54.9**
2	upset	11	12.1
3	excited	6	6.6
4	happy	23	25.3
	missing	1	1.1
	Total	91	100

Question 14: (correct=2)

Item	Response	Frequency	Percent
1	happy	9	9.9
2	**thinking about something**	**41**	**45.1**
3	excited	10	11
4	kind	31	34.1
	Total	91	100

Question 15: (correct=1)

Item	Response	Frequency	Percent
1	**not believing**	**44**	**48.4**
2	friendly	2	2.2
3	wanting to play	24	26.4
4	relaxed	21	23.1
	Total	91	100

Question 16: (correct=1)

Item	Response	Frequency	Percent
1	**made up her mind**	**54**	**59.3**
2	joking	4	4.4
3	surprised	12	13.2
4	bored	21	23.1
	Total	91	100

Question 17: (correct=4)

Item	Response	Frequency	Percent
1	angry	17	18.7
2	friendly	6	6.6
3	unkind	5	5.5
4	**a bit worried**	**63**	**69.2**
	Total	91	100

Question 18: (correct=1)

Item	Response	Frequency	Percent
1	**thinking about something sad**	**60**	**65.9**
2	angry	3	3.3
3	bossy	11	12.1
4	friendly	17	18.7
	Total	91	100

Question 19: (correct=4)

Item	Response	Frequency	Percent
1	angry	14	15.4
2	day dreaming	36	39.6
3	sad	3	3.3
4	**interested**	**38**	**41.8**
	Total	91	100

Question 20: (correct=3)

Item	Response	Frequency	Percent
1	kind	5	5.5
2	surprise	2	2.2
3	**not pleased**	**80**	**87.9**
4	excited	4	4.4
	Total	91	100

Question 21: (correct=1)

Item	Response	Frequency	Percent
1	**interested**	**58**	**63.7**
2	joking	15	16.5
3	relaxed	8	8.8
4	happy	9	9.9
	missing	1	1.1
	Total	91	100

Question 22: (correct=4)

Item	Response	Frequency	Percent
1	playful	14	15.4
2	kind	8	8.8
3	surprised	4	4.4
4	**thinking about something**	**65**	**71.4**
	Total	91	100

Question 23: (correct=2)

Item	Response	Frequency	Percent
1	surprised	5	5.5
2	**sure about something**	**70**	**76.9**
3	joking	9	9.9
4	happy	7	7.7
	Total	91	100

Question 24: (correct=1)

Item	Response	Frequency	Percent
1	**serious**	**41**	**45.1**
2	ashamed	21	23.1
3	confused	21	23.1
4	surprised	8	8.8
	Total	91	100

Question 25: (correct=4)

Item	Response	Frequency	Percent
1	shy	14	15.4
2	guilty	42	46.2
3	day dreaming	11	12.1
4	**worried**	**24**	**26.4**
	Total	91	100

Question 26: (correct=3)

Item	Response	Frequency	Percent
1	joking	13	14.3
2	relaxed	10	11
3	**nervous**	**45**	**49.5**
4	sorry	23	25.3
	Total	91	100

Question 27: (correct=3)

Item	Response	Frequency	Percent
1	ashamed	24	26.4
2	excited	1	1.1
3	**not believing**	**51**	**56**
4	pleased	15	16.5
	Total	91	100

Question28: (correct=3)

Item	Response	Frequency	Percent
1	disgust	21	23.1
2	hate	7	7.7
3	**happy**	**31**	**34.1**
4	bored	32	35.2
	Total	91	100

In total, out of the 28 questions, there were 2 questions with correct answers pertaining to positive feelings and 11 pertaining to positive mental states for a total of 13 correct answers in the “positive category”. There were 6 correct answers pertaining to negative feelings and 9 pertaining to negative mental states for a total of 15 correct in the “negative category”. For all participants, the mean percentage of correct responses for the positive feelings and positive mental states was 45.05% and 55.44%, respectively for a total mean percentage of 53.85% for the entire “positive category”. This category encompasses both feelings and mental states and is equivalent to 15.08 mean number of correct responses per 28 questions. The mean percentage of correct responses for the negative feelings and negative mental states was 61.54% and 59.46%, respectively for a total mean percentage of 60.29% for the entire “negative category”. This category encompasses both feelings and mental states and translates to 16.88 mean number of correct responses out of 28. A z-test was conducted to compare the mean proportion of correct responses in the "positive category" (53.85%) to the "negative category" (60.29%) and the result was not statistically significant ($z = -.49$, $p = 0.62$, two-tailed). In other words, the participants were as likely to provide a response that corresponded with a positive as they were with a negative mental state or emotion.

A z-test was conducted to compare the mean proportion of correct responses for "negative feelings" (61.54%) and "positive feelings" (45.05%) and the result was statistically significant ($z = 2.72$, $p = 0.007$, two-tailed). In other words, the participants were more likely to provide a response that corresponded with a "negative feeling" than "positive feeling." A z-test was not conducted to compare the frequency of responses for positive and negative mental states since the means were very close, 55.44% and 59.46%, respectively, although the negative mental states category garnered more responses. Lastly, the mean frequency of response was similar between the feelings category (57.41%) and the mental states category (53.24%) although the first was somewhat higher.

In terms of the frequencies for individual questions, five questions with the lowest frequency of correct response included four from the positive category. In ascending order of correct responses, only 20/91 participants chose "friendly" as the correct answer for one of the questions, 23/91 chose "kind" as correct for another, 24/91 chose "worried" as correct, 31/91 chose "happy", and 38/91 chose "interested" for the fifth least correctly answered question. The four questions with the highest frequency of correct response were from the negative category, that is, 83/91 participants answered "upset" as correct for one of the questions, 80/91 answered "not pleased", 76/91 "serious", 73/91 "thinking about something sad", and fifth correctly answered question from the positive category was 70/91, "sure about something". Based on the answers for individual questions, individuals with ASD had highest frequencies of response for questions in the negative category and lowest in the positive category.

Objective 4. RMET-C Task and the ADOS

Correlation analyses were performed between individual items on the ADOS (Module 3 and Module 4) and performance on the RMET-C, as well as between composite scores on the ADOS (i.e., Social Affect and Restrictive and Repetitive Behaviour domains on Module 3 and Communication, Social Interaction and Stereotyped Behaviours and Restricted Interests domains on Module 4) and performance on the RMET-C. It is of note that lower scores on the ADOS items designate better performance. Most ADOS items are on a scale starting from "0" (little concern) and ending on "3" (significant concern). Generally, to obtain a score of "0", full criteria for the item must be met, a "1" is obtained when most of the criteria are met, a "2" when some of the criteria are met and a "3" when the criteria are not met or minimally satisfied. Other scores may be given such as "7" or "8" for such observations as unusual preoccupations or specific limitations. For participants in the current study, the following scores were assigned to all items: "0", "1", "2", and "3" with one exception. Specifically, in both Modules 3 and 4, there was one item for which a score other than 0, 1, 2, and 3 was assigned, namely a score of "8" on an item that pertains to combining verbal and nonverbal communication, specifically, "Language Production and Linked Nonverbal Communication". All entries with a score of "8" were transformed to a score of "3" for this item. I chose to assign "3" for this item because the description for 8 pertains to lesser performance than a "2". Specifically, a score of "2" involves "little or no verbal communication linked with vocalization" whereas a score of "8" pertains to absence of vocalizations or limited nonverbal communication". A score of 3 represents a point on the continuum of the severity of the pattern of response.

Prior to the analysis, for the purpose of clarity of the description of results, the items on the ADOS were reverse coded such that, 0 was coded as 4, 1 was coded as 3, 2 remained as 2, and 3 was coded as 1. Although multiple correlations create a possibility of a Type I error, the present study is exploratory in nature and provides foundation for a more focussed and in depth analyses in the future.

Verbally fluent children and adolescents – Module 3: RMET-C and composite scores and individual items on the ADOS. In terms of performance on the composite scores (domains), there was no significant correlation between the Social Affect score and performance on the RMET-C, $r(57) = 0.13$, $p > 0.05$, nor between the Restricted and Repetitive Behavior domain score and performance on the RMET-C task, $r(57) = .04$, $p > 0.05$ (see Table 11).

Table 11

Correlations between item and domain scores on the ADOS II (Module 3) and mean performance on the RMET-C

Description of items	*r*	*p*
A1. Overall level of non-echoed spoken language	**0.37****	**0.00**
A2. Speech abnormalities associated with autism (intonation/volume/rhythm/rate)	-0.09	0.50
A3. Immediate echolalia	0.20	0.13
A4. Stereotyped/idiosyncratic use of words or phrases	0.16	0.13
A5. Offers information	**0.28***	**0.03**
A6. Asks for information	-0.14	0.30
A7. Reporting of events	**0.48****	**0.00**
A8. Conversation	0.12	0.37
A9. Descriptive, conventional, instrumental, or informational gestures	0.06	0.63
B1. Unusual eye contact	-0.01	0.93
B2. Facial expressions directed to examiner	-0.17	0.20
B3. Language production and linked nonverbal communication	-0.17	0.21
B4. Shared enjoyment in interaction	0.03	0.81
B5. Comments on others' emotions/empathy	0.07	0.62
B6. Insight into typical social situations and relationships	0.22	0.09
B7. Quality of social overtures	-0.10	0.45
B8. Amount of social overtures/maintenance of attention	-0.11	0.42
B9. Quality of social response	-0.07	0.60
B10. Amount of reciprocal social communication	0.08	0.55
B11. Overall quality of rapport	0.06	0.64
C1. Imagination/creativity	0.13	0.34
D1. Unusual sensory interest in play material/person	-0.07	0.60
D2. Hand and finger and other complex mannerisms	-0.11	0.41
D3. Self-injurious behavior	**-0.26***	**0.05**
D4. Excessive interest in or references to unusual or highly specific topics or objects	-0.03	0.81
D5. Compulsions or rituals	-0.21	0.10
E1. Overactivity/agitation	0.24	0.07
E2. Tantrums, aggression, negative or disruptive behavior	0.05	0.74
E3. Anxiety	-0.12	0.37
Social Affect total	0.09	0.49
Restricted and Repetitive Behavior total	-0.03	0.80

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Since there was no significant correlation between Social Affect score and performance on the RMET-C, scores were also tabulated according to Module 3 ADOS I categories that separate the Social Affect domain into the Communication and Reciprocal Social Interaction domains. Subsequent analyses revealed a significant positive correlation between the "Communication" domain and performance on the RMET-C, $r(59) = .38$, $p < 0.001$. Specifically, higher performance on this Communication domain was related to higher performance on the RMET-C. There was, however, no significant correlation between Reciprocal Social Interaction and performance on the RMET-C, $r(59) = .26$, p = n.s.

A simple linear regression using the Communication domain as a predictor of performance on the RMET-C task was conducted. A significant regression equation was found ($F(1,57) = 9.65$, $p = 0.003$), with performance on the Communication domain explaining 14.5% of the variance in RMET-C performance ($\beta = 0.38$, $p = .003$).

In terms of individual items, on Module 3 (see table 11), in the area of Language and Communication, the Pearson Correlation analyses revealed a significant correlation between performance on the "Reporting of Events" item and RMET-C task, $r = .48$, $p < 0.001$ (see Table 11 for Module 3 correlations). As per the ADOS scoring system, for a participants to receive a "0" or a perfect score on this item, the examinee is required to "report a specific non-routine event (e.g., a holiday, a vacation, a shopping trip) that is not part of any preoccupations or intense interests and seems likely to be real." The participant is required to give a "reasonable account without specific probes, but may need to be asked a general question to get started." As a result, the more successful the participants were at reporting the events per this criteria, the higher was their achievement on the RMET-C.

There was also a significant correlation, $r = .37$ at the $p = 0.004$ level, on the item involving "Overall Level of Non Echoed Spoken Language" and performance on the RMET-C.

For a participant to receive a score of “0”, the examinee is required to use “sentences in a largely correct fashion (must use some complex speech)”. The better was the overall use of spoken language, the higher was the performance on the RMET-C. There was a modest significant correlation between performance on the “Offers Information” item and the RMET-C, $r = .28$, $p = .03$. On this task, in order to obtain a score of “0”, participant is required to “spontaneously offer information about his or her own thoughts, feelings, or experiences on several occasions.” The more the participants offered information, the higher was their performance on the RMET-C task.

In terms of individual items, the last finding of significance involved a modest negative correlation between “Self-Injurious Behaviour” and performance on the RMET-C, $r = -.26$, $p = 0.05$. Specifically, “any kind of aggressive act to self, even if not clearly harmful” was related to increased performance on the RMET-C. There were no significant findings on any individual items in the area of Reciprocal Social Interaction.

Verbally fluent older adolescents and adults – Module 4: RMET-C and composite scores and individual items on the ADOS.

Module 4, is the only module in the ADOS II edition that includes separate domain scores for Social Interaction and Communication as opposed to the remaining modules that amalgamate these two areas and provide one Social Affect score. Although there was no significant correlation between performance on the Social Interaction item and the RMET-C task, $r = .14$, $p > 0.05$, there was a significant correlation between performance on the Communication domain and the RMET-C task, $r = .42$, $p = .04$; that is, better communication skills were related to higher performance on the RMET-C task.

There was also a significant correlation between performance on the "Stereotyped Behaviors and Restricted Interests" domain and the RMET-C, r = .46, p = 0.02; that is, the less the participants engaged in the stereotypical behaviours, compulsions, or preoccupations with highly specific topics, the higher was their performance on the RMET-C task.

A simple linear regression using the Communication domain as a predictor of performance on the RMET-C task was conducted. A significant regression equation was found, $F(1,22) = 4.74$, p = 0.04, and the Communication domain was found to explain 17.7% of variance, with R^2 0.177. The Communication domain significantly predicted performance on the RMET-C task (β = 0.42, p = .04). A simple linear regression using the Stereotyped Behaviors and Restricted Interests domain as a predictor of performance on the RMET-C task was also conducted. A significant regression equation was found ($F(1,22) = 6.04$, p = .02) and the "Stereotyped Behaviors and Restricted Interests" domain was found to explain 18.0% of variance. Performance on the "Stereotyped Behaviors and Restricted Interests" domain significantly predicted performance on the RMET-C task (β = 0.46, p = 0.02).

In terms of individual items, For Module 4 (see Table 12), in the area of Language and Communication, the Pearson Correlation analyses were performed and significant correlations were found between the following three individual items on the ADOS and performance on the RMET-C task: 1. "Overall Level of Non-Echoed Spoken Language" (for a score of "0" the examinee is required to produce sentences in a correct manner and use complex speech), r = .43, p = .04, 2. "Immediate Echolalia" (for a score of "0" the participant does not repeat examiner's speech)", r = .47, p = .02, and 3. "Empathic or Emotional Gestures" (for a score of "0", the participant uses empathic gestures with speech or emotional gestures), r = .45, p = .03. In summary, the increased overall production of language, the less repetition of another person's speech, and the increased use of empathic or emotional gestures in conjunction with speech was related to higher performance on the RMET-C task.

Table 12

Correlations between performance on the item and domain scores on the ADOS II-Module 4 and performance on the RMET-C

	r	*p*
A1. Overall level of non-echoed spoken language	**.43***	0.04
A2. Speech abnormalities associated with autism	0.31	0.14
A3. Immediate echolalia	**.47***	0.02
A4. Stereotyped/idiosyncratic use of words or phrases	0.30	0.16
A5. Offers information	-0.04	0.84
A6. Asks for information	**.46***	0.03
A7. Reporting of events	0.15	0.48
A8. Conversation	0.09	0.68
A9. Descriptive, conventional, instrumental, or informational gestures	0.34	0.11
A10. Emphatic or emotional gestures	**.45***	0.03
B1. Unusual eye contact	0.13	0.56
B2. Facial expressions directed to examiner	-0.15	0.48
B3. Language production and linked nonverbal communication	0.10	0.65
B4. Shared enjoyment in interaction	0.07	0.74
B5. Communication of own affect	0.21	0.32
B6. Comments on others' emotions/empathy	0.29	0.18
B7. Insight into typical social situations and relationships	0.23	0.29
B8. Responsibility	0.06	0.77
B9. Quality of social overtures	-0.21	0.33
B10. Amount of social overtures/maintenance of attention	-0.08	0.70
B11. Quality of social response	0.30	0.16
B12. Amount of reciprocal social communication	0.14	0.52
B13. Overall quality of rapport	0.31	0.13
C1. Imagination/creativity	0.36	0.08
D1. Unusual sensory interest in play material/person	0.15	0.50
D2. Hand and finger and other complex mannerisms	0.25	0.24
D4. Excessive interest in or references to unusual or highly specific topics	**.46***	0.02
D5. Compulsions or rituals	0.30	0.16
E1. Overactivity/agitation	-0.25	0.24
E2. Tantrums, aggression, negative or disruptive behavior	-0.18	0.40
E3. Anxiety	-0.09	0.67
Communication Total	**.42***	0.04
Social Interaction Total	0.14	0.51
Imagination/Creativity Total	0.36	0.08
Stereotyped Behaviors and Restricted Interests Total	**.46***	0.02

* Correlation is significant at the 0.05 level (2-tailed).

There was also a positive correlation between performance on the RMET-C task and "Asks for Information" item on the ADOS, $r = .45$, $p = 0.03$. To obtain a score of "0" on this item, the participant is required to ask "the examiner about his or her thoughts, feelings, or experiences on several occasions." In this situation, the less the participant asked about the examiner's thoughts, feelings, or interests, the better was their performance on the RMET-C task. Lastly, none of the participants exhibited any difficulties with self-injurious behaviour on this module.

Objective 5 Anxiety and Performance on the ADOS

On Module 3, there was no significant relationship between anxiety score on the ADOS and the RMET-C, $r(56) = 0.16$, $p > 0.05$. Similarly, on Module 4, there was no significant relationship between anxiety score on the ADOS and the RMET-C, $r(24) = 0.09$, $p > 0.05$.

Discussion

The purpose of the current study was to explore the development of the theory of mind (ToM) in children and adolescents with autism and to determine evidence for a relationship between ToM abilities and cognitive skills, language skills, as well as impairments associated with autism, specifically, social, communication and behaviour difficulties. This was accomplished by examining the accuracy and pattern of responses on a ToM measure across different age intervals and gender in individuals with ASD, both within the groups and compared to published data for typically developing youth. Secondly, associations were explored between performance on the ToM measure and standardized language and cognitive assessment measures. Additionally, qualitative analyses involving the response patterns related to valence of emotions on the ToM task were conducted. Furthermore, social, communicative and behavioural symptoms of autism, as well as, anxiety, in individuals with ASD were assessed relative to their performance on the ToM task. In the present discussion, each objective is aligned with the results and discussed in turn, including connections to empirical literature and implications for the specific theoretical paradigms.

Objective 1: Performance of ASD participants on the RMET-C.

Links to existing empirical literature. It was hypothesized that there is an increase in response accuracy on the RMET-C across ascending age intervals up to the age of 12 based on this trend in the typical populations (Baron-Cohen, 2001). Based on studies by Kaland et al. (2007) and Baron-Cohen et al. (2001), it was expected that participants with ASD in the present study would perform lower compared to the published data for typically developing youth. It was also hypothesized that there would be gender differences in the performance of participants with ASD consistent with Holt et al.'s (2014) neuroimaging study that indicated significantly lower performance on the RMET-C for males with ASD and a trend toward lower performance for females with ASD as compared to the typically developing group.

The effect of age and gender. Generally, there was a linear increase in mean performance across the five age groups, 6-8, 8-10, 10-12, 12-16, and 16-20. There was an increase in performance up to adulthood, although the 8-10 and 10-12 age groups had nearly identical means. Both Verbal and Nonverbal IQ were significant predictors of performance on the RMET-C. Prior to and after controlling for the effect of extraneous variables, there were no significant differences in the mean accuracy on the RMET-C task between the five age groups. The mean accuracy on the RMET-C task was not significantly different based on gender and remained insignificant after controlling for the effects of Verbal and Nonverbal IQ. These results are in contrast to the study by Holt et al. (2014) who reported that both males and females with autism performed significantly lower on the RMET-C than typically developing individuals, and their results reached statistical significance for males with a trend towards females performing lower. The lack of significance based on gender in the present study may be due to presence of fewer female participants in each of the five groups, roughly a 4:1 male to female ratio.

Comparison to published data for typically developing youth. Participants with ASD in the youngest age group (6-8 years) in the present study performed closely to the typically developing participants in the study of Baron-Cohen et al. (2001). This finding of comparable performances could be due to a small sample size of participants in the present study or due to a possibility that children in the 6-8 age group in general (i.e., typically developing) may not yet have developed the capacity to perform well on this task. Compared to the published data for typically developing youth for mean performance on the RMET-C, participants with ASD in the current study in the 8-10 and 10-12 age groups performed significantly lower than typically developing individuals in the same age groups in the Baron-Cohen et al. (2001) study. Similar to the findings in the present study, Baron-Cohen et al. (2001) reported that children in the Asperger's syndrome group (aged 8-14) who had at least average intellectual abilities performed significantly lower than the typically developing children in the 8-10 and 10-12 groups. However, the mean accuracy on the RMET-C for the 8-14 Asperger's participants in the Baron-Cohen (2001) study was about 3 points lower than performance for participants in the current study for 8-10 and 10-12 age groups, and about 4 points lower than participants in the 12-16 group.

In terms of the 12-16 age group, ASD participants in the current study performed significantly lower than typically developing participants in a similar age range (13-15 years) in the Muller and Gmunder (2014) study. Similarly, in a study by Holt et al. (2014), male adolescents (12-18) with ASD performed lower than typically developing individuals and a trend towards lower performance was observed in females with ASC as compared to controls. Notably, in the Holt et al. (2014) study, compared to participants with ASC, typically developing participants had increased activation on the MRI in the left inferior prefrontal gyrus, orbitofrontal cortex, temporopolar and middle temporal gyrus. Similarly, typically developing females had increased activation in the left orbitofrontal and temporopolar areas, as well as, the bilateral inferior and anterior prefrontal cortices. According to Holt et al. (2014), frontotemporal regions in the eyes task serve as potential neuroendophenotypic markers for ASC and these regions have also been associated with social cognition and language.

Lastly, participants with ASD in the highest age range (16-20 years) in the present study had similar performance to typically developing participants in the study of Moor et al. (2012) who were in the 19-23 age range, thus having achieved developmentally expected performance.

The present findings are consistent with results of Kaland et al. (2008) who compared the performance of children and adolescents with a diagnosis of Asperger Syndrome (AS) to typically developing controls. Compared to typically developing individuals, children and adolescents with AS demonstrated lower performance on the RMET when Full Scale IQ and age were controlled.

While the data presented is not longitudinal so the findings pertain to groups of different participants across time, there was generally a linear increase in mean response accuracy on the

RMET-C across ascending age intervals for participants with autism in the present study, while a slightly different trend was observed for typically developing individuals based on the published data for different age groups. Notably, after a steady increase in mean accuracy across the childhood, there was a decline in middle adolescence (13-15 years) with recovery in mean accuracy in the period of late adolescence to early adulthood (19 years to 23 years) in typically developing individuals.

In the Moor et al. (2014) neuroimaging study of typically developing participants, this "dip" in performance was reported in the 14-16 year old group. Moor et al's analysis of the brain areas during RMET-C revealed that posterior superior temporal sulcus was activated for participants across all age groups (10-23 years); however, one of the groups, namely 10-12 year olds, also had activated medial prefrontal cortex, the inferior frontal gyrus, and the temporal pole. Based on the data for typically developing children in the Baron-Cohen study, the 10-12 age group had the highest mean in performance on the RMET. In the present study, the ASD participants in the 10-12 age group had nearly similar performance to the 8-10 year old group, with a minimal increase between 8-10 and 10-12. In contrast, in typically developing individuals there was a two and a half point increase between mean performance of the 8-10 year olds and 10-12 year olds. This difference in the two age groups (8-10 and 10-12) may point to a lack of involvement of those brain regions in individuals with ASD at that point in development and thus a different developmental trajectory in relation to theory of mind acquisition. The findings indicate that typically developing individuals appear to obtain maximal performance by the time they are 12 years of age, whereas individuals with autism continue to acquire theory of mind skills into adolescence. This may coincide with delayed development of the frontotemporal areas of the brain in individuals with ASD.

ASD status and performance on the RMET-C. To date, no study has focused on different diagnostic criteria and patterns of performance on the RMET-C. The present study confirmed the hypothesis that individuals with Asperger's perform better than other groups due to individuals with Asperger's having stronger verbal skills compared to other groups on the spectrum. Although participants with Asperger's had a statistically significant higher mean on this task compared to individuals on the spectrum of autism from the autism, PDD-NOS and ASD PDD groups, this discrepancy only approached significance after verbal and nonverbal cognitive abilities were controlled. The unexpected finding is that nonverbal intelligence was a significant predictor of the RMET-C performance of participants with Asperger's rather than verbal intelligence since strong verbal skills have been documented to be predictive of performance on the RMET-C (e.g., Brent et al., 2004; Bennett et al., 2013). This finding underscores the role of nonverbal abilities and possible involvement of other brain areas in a subgroup of individuals with ASD who have highly developed verbal skills.

In summary, the present study revealed marked differences in the development of ToM in individuals with ASD (between 8 and 16 years of age) compared to typically developing children when verbal and nonverbal IQ were controlled. Compared to typically developing children, participants with ASD in the three age groups performed significantly more poorly. For individuals with ASD, there was a linear increase towards ToM skill acquisition until adulthood, whereas typically developing participants had a faster acquisition of ToM skills (by 12 years of age). This discrepancy suggests that some brain areas develop later in children with ASD than in typically developing children, especially the orbitofrontal areas that have been linked not only to ToM development but also executive functioning, language, and working memory. The results also point to possible involvement of nonverbal cognitive skills in the development of ToM in children with higher verbal skills. Although no gender differences emerged in the present study, future research with larger cohorts including both genders would provide important insights into the presence and nature of gender differences.

Links to existing theoretical literature: Eye Avoidance Theory, Affective Theory, Metarepresentation, and Alexithymia Hypothesis. The reduced performance on the RMET-C task and delayed acquisition of this skill could be explained by the “eye avoidance theories” of face processing in individuals with ASD. Tanaka and Sung (2016) explained the face recognition difficulties in individuals with autism based on the “eye avoidance hypothesis”, that is, people with autism tend not to look in the eye region of the face because they find it threatening. Since children with ASD spend less time looking at the eyes, it may take them longer to acquire the skills associated with ToM tasks as discovered in the current study because they have fewer learning opportunities to look at the eyes.

For example, if a person sees someone retracting the corner of the mouths and perceives that this person is happy then they relying on both internal and observable experience of happiness from the past of what a happy person looks like. Similarly, if a person looks at someone’s eyes and determines that he/she “knows” or “believes” then this person is making inferences based on social learning from the past that a specific facial expression implies these mental states. The current findings indicate that many individuals with ASD are able to “catch-up” on their ToM skills by middle adolescence; however, they are delayed on this skill during most of their development during childhood likely due to limited learning opportunities as a result of reduced eye contact.

The absence of eye contact and reduced attention to the eyes of another person is an early warning sign of autism. By the first year of life, children who are later diagnosed with ASD exhibit a lack of attention to faces (Osterling, Dawson, & Munson, 2002) and diminished eye contact (Zwaigenbaum et al. 2005). According to these authors, the eye tracking studies to date show that whether viewing a static image or a dynamic video of faces, individuals with autism show a preference for the mouth features and avoidance of the eye features.

Jones, Carr, and Klin (2008) compared fifteen 2-year-old children with ASD to 36 typically developing children and to 15 developmentally delayed children without the diagnosis of autism. When presented with videos of childhood games, the 2-year-old children with autism exhibited a significant increase in time looking at the mouth region and decrease in time looking at the eyes in comparison to both control groups. Furthermore, their fixation time on the eyes was correlated with their level of social competence, such that less fixation on eyes predicted greater levels of social disability.

Based on Tanaka and Sung (2016), direct eye contact elicits an increased physiological response as indicated by heightened skin conductance and amygdala activity in individuals with ASD so individuals with autism adapt a strategy of avoiding eyes. Instead, they focus on other features of the face as well as parts of the body in order to feel less discomfort when they look in the eyes. This in turn interferes with processing of facial expressions and intentions, and affects their social cognition.

This ties to Hobson's (1986) theory of a lack of personal relatedness. Specifically, according to Hobson, children with autism lack the innate ability to both perceive and respond to the affective expressions of other people. This, in turn, renders the children incapable of receiving the social experiences necessary to develop the cognitive structures underlying social understanding. Impairments in language and cognition are viewed as secondary to these social deficits.

According to Hobson's (1986) affective theory, children with ASD have an inborn deficit in perceiving emotions and responding to emotions of other people and thus lack the social experiences that are necessary for the development of social cognition and communication. This deficit in expressing and perceiving emotions is viewed as primary and difficulties in language and cognition are viewed as secondary.

The impaired processing of facially expressed emotions in persons with ASD has been well documented and observed to interfere with an ability to interpret emotional process of others (Deutsch and Raffaele, 2019). This deficiency, in turn, may interfere with the development of the understanding of mental states of others and be linked to difficulties in deciphering both emotions and unobservable mental states of others among children and adolescents with ASD.

In effect, Baron-Cohen (1988) extends Hobson's theory by arguing that the affective deficits do not account for difficulties in individuals with ASD in understanding beliefs. Rather, according to Baron-Cohen's ToM hypothesis of autism, mental states such as knowing and believing are not directly observable and need to be inferred and as a result require higher order understanding, which individuals with autism have difficulty with.

Based on these theories, one can infer that the ability to perceive emotions is not the same as the ability to perceive mental states such as knowing or believing. In other words, to have a well-developed theory of mind, it is not enough to perceive what the other person is feelings, rather one has to perceive what the other person is e.g., thinking, knowing, believing, and understanding.

To that end, Oakley, Brewer, Bird and Catmur (2016, p. 819) indicated that "The RMET is unusual among ToM tasks in that it includes emotional states and relies on the detection of subtle facial cues, features typically used to test emotion recognition." As a result, the authors examined whether the RMET task (adult version) is a measure of ToM or whether it assesses emotion recognition. The rationale for this study came from alexithymia hypothesis of emotion deficiencies in ASD (Cook, Brewer, Shah & Bird , 2013). The underlying assumption for this hypothesis is that observable deficits in emotion recognition in ASD are not due to the ASD disorder itself but to the co-morbid presence of alexithymia (inability to perceive emotions in oneself).

The authors assessed whether alexithymia or ASD was the better predictor of RMET performance. They found that alexithymia significantly predicted RMET performance but not the ASD diagnosis. Additionally, alexithymia did not predict performance on a ToM task called MASC (Movie for the Assessment of Social Cognition) that relies on emotion recognition to a lesser degree and more on dynamic nature of social interaction. Consequently, Oakley et al. (2016) concluded that RMET measures recognition of emotion and not ToM skills, thus lending support to the alexithymia hypothesis of emotion-related deficits in ASD. Although the study involved adults and a relatively small sample size, it raised an important question of definitions related to measurement of ToM and the need to decipher these constructs.

Along these lines, Zainal and Newman (2018) argued that RMET is a “ToM decoding" task rather than reasoning task because it relies on one mode of social cognition, that is, recognition of emotion, based on eyes. In contrast, according to the authors “ToM reasoning” tasks use several “channels” of information (e.g., visual, auditory, body movements) to understand emotions and intentions of other people. Zainal and Newman (2018) described RMET as a basic ToM decoding task on which ToM reasoning skills are built. These reasoning skills, in turn, encompass both an individual’s belief system and experiences and processing of multimodal and dynamic information from environment. According to Zainal and Newman (2018), both neuroimaging and lesion studies have shown that the decoding and reasoning components of ToM are part of different neural networks.

In contrast to Baron-Cohen’s statement that mental states cannot be observed and need to be inferred in the RMET, the present study was based on an assumption that the mental states can be both, observed and inferred based on social experiences on the RMET task. Moreover, neuroimaging studies point to involvement in the orbitofrontal areas of the brain during completion of the RMET, thus lending support to a ToM hypothesis. At the same time perception and expression of emotion likely play a crucial role in early development and may be the precursors to the development of the ToM.

On the RMET-C, mental states and feelings are represented collectively under the ToM construct. In the present study and in one other study, these two constructs were examined separately in addition to being examined together. The descriptions involving RMET vary in existing literature so clarification of these constructs will make the comparisons more equitable. Moreover, further clarification of what constitutes ToM will further clarify the constructs that are intended to be measured. The developmental analysis of acquisition of ToM skills as presented in the current study is a step in the direction of delineating the constructs related to mental and emotional state and ToM.

Objective 2: Cognitive and Language Measures and Performance on the RMET-C

Links to existing empirical literature. It was hypothesized that there would be a strong correlation between expressive and receptive language skills and performance on the RMET-C. This hypothesis was built upon research by Brent et al. (2004) and Bennett et al. (2013) who found strong associations between language skills and performance on the RMET. In line with findings of Brent et al. (2004) and Rueda et al. (2015), it was expected that the RMET-C performance of individuals' with ASD would be independent of intellectual abilities, that is, at least average performance on the cognitive measures would be unrelated to higher performance on the RMET-C.

Consistent with Brent et al.'s (2004) research that also included both receptive and expressive language measure, the present findings revealed a significant correlation between standardized scores on the Oral Language Composite of the OWLS and performance on the RMET-C, that is, higher language skills were related to higher performance on the RMET-C. Brent and colleagues found significant links between language age and performance on the RMET in a group of children with ASD; however, they did not find a significant correlation between these two variables for typically developing children matched for chronological and language age, and a Full Scale IQ. Of note, the mean language age of participants in Brent et al.'s study was one year lower than their mean chronological age (no numbers). For both the ASD and typically developing groups, there were no significant correlations between IQ and performance on the RMET-C task for both groups.

The present findings are also consistent with the study by Peterson et al. (2015) who reported that children with ASD performed lower on the RMET-C task than typically developing participants who were matched on receptive language skills. The present study extends the research by Peterson et al. (2015) by controlling for IQ and obtaining the same finding. Similarly, Philpott et al. (2013) reported that although initially, children with ASD performed on par with typically developing children on the RMET-C task, once IQ was accounted for, their findings were similar to those of Baron-Cohen (2001) with ASD children performing lower than the controls.

The present study is unique in providing findings across separate age intervals in youth with ASD (6 years to adulthood). The relationship between RMET-C and language skills was strongest in the 10-12 age group (consistent with orbitofrontal brain regions in typically developing individuals as described by Moor et al. 2014), followed by the 8-10, 16-20, and 12-16 age groups. Although no association was found in the 6-8 age group, this could be attributable to the small sample size or delayed development of this connection in this age group. A relationship between the Verbal IQ and performance on the RMET-C was also examined across the five age groups and similarly to the Language Composite findings, the strongest relationship was in the 10-12 age group followed by 8-10 age group. The correlations were not significant in the 6-8, 12-16, and 16 and up age categories, underscoring that across the majority of age intervals oral and receptive language skills as assessed by the Oral Language composite are more predictive of performance on the RMET-C task than Verbal IQ.

Lastly, the relationship between Performance IQ and performance on the RMET-C was also examined across the five age intervals and the strongest positive relationship was in the 12-16 age group, followed by 8-10 and 10-12 age groups. The correlations were not significant in the 6-8 and 16-20 age categories. Even though the data in the present study indicate a decrement in language skills in the 12-16 age interval, children with ASD continue to acquire ToM skills through the late adolescent years. The mean oral language score for the 12-16 age interval was 21.16, and it was significantly lower than the 16-20 age interval (mean = 55.67) thus underscoring the drop in language skills in the 12-16 age category from a mean of 32.64 to 21.16 in the 10-12 age range in presence of nearly linear increase of performance on the RMET-C. These differences in rate of acquisition at different developmental periods may point to involvement of different brain areas at those developmental points in individuals with ASD as compared to controls.

In summary, the present study extends previous research by including a standardized measure of language skills in addition to verbal and nonverbal IQ. In addition, the present study provided novel findings by examining these variables among youth with ASD across several age intervals, thus taking developmental trajectories into an account. The present research revealed that when language skills (expressive and receptive) are taken into an account, verbal and nonverbal IQ ceases to be predictive of performance on the RMET-C. This finding underscores that verbal intellectual abilities are not sufficient in making predictions about performance on the ToM tasks such as RMET-C, rather intellectual abilities (verbal and nonverbal) along with language skills provide a more accurate forecast of these types of tasks.

Language skills were significantly correlated with RMET-C performance across four age groups (all except 6-8 year olds), whereas verbal IQ was significantly correlated in only two of the age groups, 8-10 and 10-12. Non-verbal IQ was significantly correlated with RMET-C performance in three age groups, 8-10, 10-12, and 12- 16. Although there was a drop in language skills among individuals with ASD in the 12-16 age group, their RMET-C performance was higher than in the earlier age groups, suggesting the potential involvement of nonverbal abilities. Nonverbal skills seem to play an independent role in the development of the theory of mind as measured by the RMET-C since these skills also predicted a significant amount of variance in performance of individuals with ASD. In future research, more detailed deciphering of the components of language and nonverbal skills will elucidate the connection between ToM, language and nonverbal skills.

Links to existing theoretical literature: Language theories. In the area of language, Levy (2007) noted that impairments in social communication and deficits in pragmatic language in individuals with ASD can be explained by ToM. Specifically, Walensky, Tager-Flusberg, and Ullman (2006) hypothesized that individuals with ASD have deficits in making causal links between their own thoughts and behaviours and those of other people, as evidenced in difficulties understanding emotion and false belief. In the present study, this is supported by the findings that the impaired development of ToM as measured by RMET-C in individuals with ASD is predicted by reduced verbal communication or the language measure, that is, the pragmatics of the language. For individuals with ASD who do develop verbal skills, speech often lacks communicative intent (Johnson & Myers, 2007). In effect, it is the use of these verbal skills in communicative context that is relevant to ToM, given that individuals with ASD continued to have impaired performance on the RMET-C, even in presence of intact verbal intellectual abilities. According to Eigsti et al. (2007), pragmatics is the use of language for the purpose of communication during social interactions and it involves the use of both verbal and nonverbal cues (head shaking, eye contact). Bennett et al. (2013) noted that children who develop language abilities earlier will likely have a better ability to understand social situations because of wider exposure to different interactions and therefore will develop ToM (including recognition of facial expressions) earlier compared to those who have difficulty with language skills.

Holt et al. (2014) reported that the frontotemporal brain regions activated during the RMET-C task have also been associated with social cognition and language processing. Milligan, Astington, and Dack (2007) reviewed over 100 studies assessing the relationship between ToM and language and found a strong relation between false-belief understanding and language ability, which was consistent across different language measures.

Objective 3: Valence and Individual Item Analysis

Links to existing empirical literature. Consistent with the study by Vogindroukas et al. (2014) who revealed that children with High Functioning Autism (HFA) identified more negative than positive feelings, and slightly more negative mental sates than feelings, it was hypothesized that participants in the current study would identify more negative emotions and mental states than positive emotions and mental states. The current findings were in line with Vogindroukas et al. (2014) who reported lower means of correct percentage scores across the four domains of emotional valence. Specifically, the mean percentage of correct responses in the group of participants with high functioning autism in the Vogindroukas study was 16.96%

and 22.14% for positive feelings and positive mental states, respectively, compared to 25.36% and 25.71% for negative feelings and negative mental states.

Although these mean percentages were lower than the ones in the current study: 45.05% and 55.44% for positive feelings and mental states, respectively, and 61.54% and 59.46%, for negative feelings and negative mental states, respectively, the overall trend of lower performance on items assessing positive valence was observed across all domains in both the Vogindroukas and the current study.

In terms of the overall mean correct scores, children with High Functioning Autism (HFA) had a mean score of 13.6 out of 28 (mean age = 10.06 years) questions correct in the Vogindroukas study, whereas children in the current study in the age groups of 8-10 and 10-12 had a mean score of 15.8 and 15.7 out of 28, respectively.

In terms of responses to individual questions, the results of the present study also appear aligned with those reported by Vogindroukas et al et al. Specifically, Vogindroukas et al. reported that children with HFA had the most correct responses on the 'not pleased' image (85.7% answered correct), and they had the lowest number of correct responses for the woman who appears 'kind' (18.5%). In the current study, the "kind' question was second lowest in terms of frequency of correct responses (25.27%) and the lowest was "friendly" (21.99%). Similarly to Vogindroukas et al., in the current study, the highest frequency of correct responses was for "upset" response (83/91or 91.21%) followed by "not pleased" (87.91%).

One of the shortcomings of the study by Vogindroukas and colleagues is that language skills were not assessed using a standardized measure of language; therefore, the reduced performances compared to the results of the present study may be due to reduced language skills in presence of an otherwise generally intact cognitive abilities. The present study is generally consistent with the studies on perception of emotion where individuals with ASD perform lower on tasks assessing recognition of emotions. For example, a number of studies have reported deficits in identification of negative emotions, as well as, happiness and surprise in children and adolescents with ASD (Griffiths et al. 2019; Yeung, Han, Sze, & Chan, 2014; Yeung, Lee, & Chan, 2019).

Rueda et al. (2015) found that children with ASD identified fewer positive emotions and mental states than typically developing individuals on the RMET-C task. The current study did not encompass a typically developing comparison group and used a different classification system for the RMET-C than Rueda and colleagues (2015) who divided the items to positive, negative and neutral categories. Nevertheless, the findings of the present study are partly consistent with Rueda et al. who noted that compared to typically developing youth, children with AS identified significantly fewer positive emotions on the RMET-C task.

In summary, overall, participants with ASD appear to be less efficient in identifying positive feelings and slightly less efficient at identifying positive mental states than negative feelings and negative mental states, although there were differences in frequencies of responses to specific items. For example, the response involving “happiness” garnered only 20 correct responses out of 91 participants whereas identifying someone as “upset” garnered 83 corrects responses out of 91 participants. These findings contribute to research on emotion and ToM, as well as to theories such as Darwin’s (1879) theory of innateness of emotion, Hobson’s theory of deficits in intersubjectivity and Baron-Cohen’s theory of mind (2001). As Krzeminska (2001) noted, the lack of facial expression is a bi-directional process. For example, lack of expression of emotion early in life (such as “happiness”) may elicit less expressed emotion of this type from other people and lack of exposure to this emotion may affect children’s own experiences of emotion of “happiness”. Emotion serves as a mediator for learning about things in the environment so it plays a vital role in both social and cognitive development.

Links to existing theoretical literature: Evolutionary Theory. Recognition of emotions in the face has been attributed to inborn and reflexive mechanisms, specifically. Darwin (1872) posited that there is a universal and innate basis to our emotional expressivity and perception. He was the first to document a connection between the infants reaction to others ’ facial emotion and the instinctive recognition of that emotion. Darwin described his 6-month-old son who saw his nurse pretend to cry: "his face instantly adopted a melancholy expression, with the comers of the mouth strongly depressed." Since then studies on the expression and perception of emotion in typically developing infants confirmed that young infants can spontaneously produce, respond to, and recognize several distinct facial expressions, thus not requiring higher order brain development.

Deficits in recognition of emotion in individuals with autism have been well documented (e.g., see Uljarevic & Hamilton 2013 for review); however, there remains a question as to whether these deficits pertain to only certain type of emotions or whether this impairment is across all categories of emotion. To that end, variable findings have been reported. Philip et al., (2010) reported deficits in recognizing a broad range of emotions in adults with ASD, however, more recently, Yeung, Lee and Chan (2019) reported deficits in recognizing only negative emotions in adults and adolescents with ASD. Similarly, Aschwin, Chapman, Colle, and Baron-Cohen (2006) noted that participants with Autism Spectrum Conditions had more difficulty recognizing some of the negative emotions compared to controls whereas Lacroix (2009) reported no impairments in recognition of negative emotion among participants with Asperger's syndrome. In another study, Aschwin, Wheelright, and Baron-Cohen (2006) reported that participants with High Functioning Autism and Asperger Syndrome performed similarly to controls on a task involving schematic depiction of emotions but not when faces were inverted and when they used varying crowd sizes.

Along the lines of evolutionary-developmental perspective, Izard, Fine, Mostow, Trentacosta, and Campbell (2002) reported that recognition of negative emotions is essential for survival (e.g., in a situation when one needs to protect oneself from danger) and the recognition of another person's anger facilitates that process. These findings and theoretical interpretation would be consistent with an emotion-specific recognition deficit in individuals with ASD.

Although research in the area of expression of emotion in individuals with autism is less abundant than research on the perception of emotion, there is evidence that children with autism express less affect overall, specifically, less enjoyment; however, they express more neutral, negative, and atypical facial expressions (Krzeminska, 2001). In summary, there is some evidence that children with ASD may be better at recognizing and expressing negative emotions, but more deficient in identifying and recognizing positive ones. Moreover, there is evidences of reduced expressivity in facial muscles. Specifically, Czapinski and Bryson (2003), examined facial muscle movements during the expression of emotion in young children with autism. Compared to both language-delayed and typically developing children, children with autism exhibited reduced and weak muscle movements in the eye and mouth regions, but not the brow region of the face. These findings were interpreted as being consistent with the possibility that faulty innervation of the face in children with autism results from impairment in the neuromuscular pathway of the cranial nerves from the brainstem as proposed by Rodier, Ingram, Tisdale, Nelson, & Romano (1996).

The impairment in the ability to produce facial expressions may result in deficient interactions with other people (Hobson, 1986) and in turn may contribute to difficulties reading feelings and mental states of others (Baron-Cohen, 2001). Krzeminska (2001) postulated that difficulties with expressing emotions may lead to difficulties in recognizing emotions of others and sharing emotional experiences with others, or impaired intersubjectivity (Hobson, 1986). Specifically, not having the experience of displaying appropriate emotions may first lead to difficulties in identifying emotions and then lead to difficulties in higher order thinking, such as, identification of mental states in other people (Baron-Cohen, 2001). Krzeminska (2001) hypothesized that movements in the face that correspond to specific emotions may also "stimulate" the physiological part of the brain responsible for the subjective experience of that emotion and, in turn, facilitate learning about emotions in relation to the self and other people. As demonstrated in the present study, participants were better at detecting negative emotions, and as Krzeminska (2001) reported, children with autism also express negative emotions more frequently. As a result, children with ASD may be more attuned to the negative emotions from birth and thus perceive and express them more efficiently.

Objective 4: RMET-C Task and the ADOS

Links to existing empirical literature. It was hypothesized that increased socio-communicative symptoms scores on the ADOS would be related to reduced performance on the RMET-C. Although there is limited literature on ToM and ADOS, it was also hypothesized that performance on the RMET-C would be correlated with Restrictive and Repetitive Behaviors (RRB) on the ADOS, based on Romero et al. (2018) findings that ToM along with RRB and body movements account for the same amount of variance.

Verbally fluent children and adolescents (Module 3).

Composite Scores. Based on ADOS-II classification, there was no association between the increased symptoms of autism on the Social Affect scale and performance on the RMET-C. The results were contrary to the prediction that increased socio-communicative symptoms on the ADOS would be related to reduced performance on the RMET-C based on the findings by Kamp-Becker et al. (2009) who found that socio-communicative deficits were associated with lower ToM. The findings were also contrary to findings by Li et al. (2012) who noted that the severity of autism may be associated with lower performance on the ToM task. Similarly, based on the ADOS-II classification, the findings were contrary to the hypothesis that performance on the RMET-C would be correlated with Restrictive and Repetitive Behaviors (RRB) based on Romero et al. (2018) who reported that ToM along with RRB and body movements account for the same amount of variance. In the present study, using the ADOS II classification, there was no association between symptoms on the RRB scale and performance on the RMET-C.

Different results were found based on the ADOS I classification, which includes two categories, Social Communication and Reciprocal Social Interaction domains. Specifically, increased performance (better skills) on the Communication domain was related to increased performance on the RMET-C. At the same time, there was no significant correlation between Reciprocal Social Interaction and performance on the RMET-C, nor between performance on the RRB scale and RMET-C. These findings based on the ADOS-I classification are partly consistent with those of Kamp-Becker et al. (2009) and Li et al. (2012) in that communicative deficits associated with autism were associated with lower performance on the RMET-C. Performance on the communicative domain explained 14% of variance on the RMET-C task and although this number is lower than the amount of variance language skills represent on standardized assessment measures, it underscores the importance of communication as a separate entity linked to performance on the RMET-C and thus the theory of mind.

Individual items. On the ADOS I, the Social Communication scale encompasses four items: “Stereotyped/Idiosyncratic Use of Words and Phrases”, “Reporting of Events”, “Conversation”, and “Descriptive Conventional, Instrumental, or Informational Gesture.” Together, under the Communication domain, these items were correlated significantly with performance on the RMET-C but not individually.

Of note, two individual items (“Reporting of Events” and “Overall Level of Non-Echoed Spoken Language”) were strongly correlated with performance on the RMET-C task. The better was the overall use of spoken language, the higher was the performance on the RMET-C. The more successful participants were at reporting events, the higher was their performance on the RMET-C. Additionally, there was a modest correlation on the “Offers Information task”, that is, the more the participants offered information, the higher was their performance on the RMET-C task. There were no significant correlations for individual items in the domain of Reciprocal Social Interaction. The “Reporting of Events” item was excluded from the Social Affect domain of ADOS II but was part of the ADOS I domain of Communication domain.

An unexpected finding was that self-injurious behaviour was related to increased performance on the RMET-C on Module 3. It is possible that injurious behaviour is performed in order to reduce anxiety in individuals with ASD. Moreover, Moss et al. (2000) reported that adults with autism who displayed Severe Injurious Behaviour (SIB) were most commonly diagnosed with anxiety as a comorbid disorder. Since higher levels of anxiety were found to be related to lower performance on the RMET-C in individuals with autism (Kamp-Becker et al., 2009 & Livingston et al., 2018), individuals in the present study who performed behaviours associated with the reduction of anxiety may have been also indirectly contributing to better development of the ToM.

Verbally fluent older adolescents and adults – Module 4.

Composite scores. Social Interaction and Communication domains are included as separate scales on Module 4. In the present study, communicative deficits associated with autism (Communication domain) were associated with lower performance on the RMET-C. This finding is consistent with those Kamp-Becker et al. (2009) and Li et al. (2012). The performance on the Communication domain accounted for about 18% of variance on the RMET-C task. Similarly to Module 3, there was no significant correlation between Reciprocal Social Interaction and performance on the RMET-C. There was, however, a significant correlation between Stereotypical Behaviours and Restricted Interest (SBRT) scale and performance on the RMET-C, on Module 4, that is, the less the participants engaged in the stereotypical behaviours, compulsions, or preoccupations with highly specific topics, the higher was their performance on the RMET-C task. Similarly to the Communication domain, the score on the SBRT scale accounted for about 18% of variance on the RMET-C.

Individual items. As indicated above, In terms of individual items on Module 4, similarly to Module 3, the increased overall production of language was related to higher performance on the RMET-C task. The less the participant asked about the examiner's thoughts, feelings, or interests, the better was their performance on the RMET-C. Additionally, the less participants engaged in repetition of another person's speech, the better was their performance on the RMET-C task. In general, the repetitive and stereotypical behaviours, either verbal or nonverbal seem to interfere with reading mental states of others in adolescents and adults that Module 4 is usually administered to. Additionally, an increased use of empathic or emotional gestures in conjunction with speech was related to higher performance on the RMET-C task.

In summary, in Module 3, for younger children and adolescents, communication plays a role in being able to recognize the mental states and feelings of others. The more successful participants were at the overall us of spoke language, reporting of events and the more the participants offered information, the higher was their performance on the RMET-C. An unexpected finding was that self-injurious behaviour was related to increased performance on the RMET-C.

In Module 4, used for adolescents and adults, in addition to increased communication skills being related to better performance on the RMET-C, the increased production of language, the less an individual engaged in stereotypical behaviours or perseverated by repeating another person's speech, the better was their performance on the RMET-C. Lastly, asking less for information and performing more of empathic or emotional gestures in conjunction with speech was related to higher performance on the RMET-C task in this module. It is thus the communicative aspect of social interaction either through language or gestures that may facilitate learning about emotional and mental states of others whereas behaviours that are not intended for communication such as echolalia may interfere with the process of communication and in turn the development of ToM.

Objective 5: Anxiety and Performance on the ADOS

It was hypothesized that higher anxiety is predictive of lower scores on the RMET-C, consistent with the findings of Kamp-Becker et al. (2009) and Livingston et al., (2018). In the present study, there was no significant relationship between the anxiety score on the ADOS, both Module 3 and Module 4, and performance on the RMETC. This is in contrast to findings by Kamp-Becker et al. (2009) and Livingston et al., (2018) who reported that higher anxiety is predictive of lower scores on the RMET-C. It is of note that in the present study, the anxiety scores were based on observation rather than on self-report as in Kamp-Becker et al. (2009) and Livingston et al., (2018) research, which could have contributed to the discrepancy in findings. The role of anxiety needs to be investigated further at it pertains not only to performance on ToM tasks but also as it relates to repetitive and self- injurious behaviours as a possible means to reduce anxiety.

Objective 4 and 5: Connection to Existing Theoretical Literature – The Missing Link of ToM and RRB's and the Role of Anxiety.

Based on the findings of the present study, rigid and stereotypical behaviours seem to interfere with the development of ToM because the less participants engaged in these behaviours, the higher was their performance on the ToM task. Taken together, if speech is largely egocentric and repetitive and not intended for communication (including echolalia), then it appears to interfere with performance on the RMET-C task. One of the limitations of research on the ToM of children with autism has been that it does not explain the second core feature of this disorder, namely, the stereotypical behaviours. Specifically, Frith and Happe (1994) reported that repetitive behaviour, including restricted repertoire of interests and obsessive desire for sameness, cannot be explained in terms of impaired "mentalizing." Instead, these behaviours have been explained by faulty neural connectivity (Levy, 2007) and as a coping mechanism (Turner, 1997). Based on the results of the current study, stereotypical behaviours appear to interfere with RMET-C task and thus with the recognition of mental states and emotions. Turner (1997) noted that according to ToM, repetitive behaviour will increase when a person is in a new or unpredictable social environment and it will occur less frequently in a familiar situation (Turner, 1997). According to Levy, however, the majority of studies indicate that repetitive behaviour is lowest during social interaction.

If stereotypical behaviours are a coping mechanism, then they may be performed to reduce anxiety. Parents frequently report that their children increasingly exhibit these behaviours in situations that are likely to exhibit anxiety such as, transitions and new settings. Although there was no relationship between observed anxiety on the ADOS and performance on the RMET-C in the present study, this relationship needs to be investigated further with self-report studies of anxiety. On the other hand, self-injurious behavior that was reported to occur in the present study may have been performed to reduce anxiety and thus the performance on the RMET-C increased, as self-injurious behaviours increased as well.

Research by Zainal and Newman (2018) sheds light on the finding of anxiety enhancing performance on ToM tasks. These authors found that college participants with a Generalized Anxiety Disorder (and no other diagnosis) who were "in a state of worry" performed better than controls on different types of ToM reasoning tasks. There were, however, no differences between these two groups when they were in a state of relaxation. Moreover, participants with GAD who worried, but did not relax, were also significantly better than the controls at deciphering negative signals which is consistent with findings in the present study of youth with ASD being able to decipher better negative mental states rather than positive. The link between anxiety and ToM performance needs further investigation with self-report measures of anxiety across different developmental stages.

Limitations and Outstanding Questions

A limitation of the present study was the absence of a normative comparison group. Although data for typically developing youth were drawn from existing literature, in the future a comparison group of typically developing participants would strengthen the findings, provide greater generalizability, and allow for a more sophisticated statistical analysis. Another limitation of the present study is that different participants were included in the five age groups. As such, longitudinal rather than cross sectional analysis would account for individual differences in maturation and variability in development within each participant. The benefit of the current study is that it included a large number of ASD participants, which enabled quantifiable classification and predictive analyses. In the present study, the number of participants varied across the developmental groups; a larger sample size in future research would make the comparison across the different groups more comparable and thus account for heterogeneity within each developmental stage. Additionally, an equivalent number of participants from both genders would enable comparisons between males and females; however, this is difficult given the higher prevalence of males with ASD. In the future, studies should include populations from all developmental stages up to adulthood, with longitudinal designs to follow participants through diverse developmental stages.

When statistical analyses involved performing multiple tests, there was a risk of finding significant results by chance so these findings should be interpreted with caution. At the same time, the study was exploratory and some of these results may be used to generate new hypothesis and conduct a more in depth analyses. Lastly, the participants in the current study came from families who were taking part in the genetics study and as such may not have been representative of youth with autism in the general population. In the future, including demographics such as socioeconomic status would be helpful in order to determine the representativeness of the participants in the sample.

Many questions remain from both practical and theoretical standpoints. Specifically, the operationalization of the concept of the ToM needs to be studied further. For example, there is still lack of consensus as to whether the ability to infer what another person is thinking is the same as deciphering what another person is feeling. Is thinking that someone is happy is the same as thinking that an individual feels happy? The expression and perception of emotion as it is connected to the concept of ToM from both behavior and neuroanatomical needs to be studied further. Are language and emotion distinct constructs with separate neural substrates or are these processes on the continuum of functions and behaviours? Unravelling these complex ideas will contribute to the understanding of not only children, adolescents and adults with ASD, but also of typically developing individuals.

Main Contributions, Implications for Intervention and Policy and Future Directions

The present study contributes to the understanding of the neurodevelopmental, theoretical, empirical and clinical underpinnings of autism. One of the main findings of the present study across different measures is the variability in performance of youth with autism across the developmental stages. To that end, this study provides a significant contribution by describing performance on the RMET-C task of individuals with autism ranging from 6 years of age to adulthood. Based on the results from different age cohorts, youth with ASD continue to acquire skills on RMET-C from childhood through to early adulthood, in contrast to typically developing individuals who achieve the same skills by the age of 12.

This finding underscores the need for therapeutic interventions for individuals with ASD across childhood, adolescence and up to early adulthood thus promoting metacognitive abilities throughout these developmental stages. For example, learning tools like social stories could be used at different developmental stages to teach appropriate responses, perspective taking, and empathy during social situations. These interventions will in turn address impairments in both processing of expressed emotion and deficits in making inferences about another person's thoughts and feelings.

The findings provide implications for both practice and policy for funding of programmes for youth with ASD. Given that children and adolescents with ASD are developing capacity for the theory of mind at a different rate than their typically developing peers and this coincides also with the development of language and working memory, policies should include funding of programmes that address these skills until at least early adulthood in order to maximize the development of metacognition which in turn plays an integral part in the development of social and relational skills.

The second major implication is presentation of performance of individuals with ASD on tasks involving cognitive and language abilities (e.g., verbal and nonverbal IQ and language skills) across different developmental stages, as well as, the association of these variables with the development of ToM. To that end, therapeutic interventions can incorporate these findings by placing more emphasis on the development of verbal and nonverbal skills at specific stages of development to enhance the development of ToM and social skills. For example, nonverbal cognitive abilities are highly related to performance on RMET-C between 8 and 16 years of age; therefore, an emphasis should be placed on incorporating nonverbal and visual spatial skills during that time in development. Between 12 and 16 years of age language skills seem to play a lesser role in the development of RMET-C, however, adolescents with ASD continue on their path of developing the ToM skills on RMET-C and eventually acquire them sometime between 16 and 20 years of age. This finding points to a possible involvement of brain areas responsible for nonverbal cognitive potential during the acquisition of ToM skills during that developmental interval.

The third main contribution of the present study is the differentiation between verbal IQ and language skills (expressive and receptive) and their relationship to performance on the RMET-C across different developmental stages. Specifically, the communicative aspect of verbal skills as assessed by the language measure (perception and expression of language) should be promoted from the early childhood years through to adulthood. Therapeutic interventions should continue to place emphasis on the communicative aspect of language throughout the lifespan.

The importance of the intervention involving support for expressive and receptive language skills in youth with ASD cannot be overemphasized. These skills appear to take precedence over verbal intellectual ability in relation to ToM. To that end, more research is necessary to decipher the differences between verbal intellectual abilities and specific language skills that are used for communication. Although the overlap between the verbal cognitive abilities and receptive and expressive skills is high for the participants in the present study (r = 0.78), in presence of the receptive and expressive language skills, verbal intellectual ability ceased to be predictive of performance on the RMET-C task. It may be that the communicative aspect of language use as represented on the oral and receptive language tasks rather than retrieval of verbal knowledge as represented on the IQ tasks, plays a role in the development of theory of mind in both typically developing and in individuals with ASD.

One of the main goals of therapeutic interventions for children with autism is to increase verbal communication. To date, different techniques aimed at increasing vocalizations have been tested. For example, joint attention training increased positive affect, imitation, social initiation, play, and language have been a focus on intervention (Whalen, Schreibman, & Ingersoll, 2006). In the area of Applied Behavioural Analysis, various strategies, including discrete trial training, script fading, time delay/prompt fading, and others have been used to increase verbal output in individuals with autism (Duffy & Healy, 2011). The current study adds significantly to the area of intervention by identifying specific developmental stages during which optimal verbal and nonverbal output should be reinforced for children and adolescents with ASD to achieve therapeutic goals of ToM skill acquisition.

The ability to identify more accurately at least some negative rather than positive emotions by individuals with ASD has significant consequences for interactions with other people. If individuals with ASD are more attuned to specific facial expressions displaying negative affect, they may be more likely to respond to those negative displays of affect by other people. As evidenced by Krzeminska (2001), children with autism express more negative affect and less "joyful" facial expressions. Based on the current study, at least some negative emotional reactions displayed by a caregiver, teacher, or peer are more likely to be recognized by a young person with autism than a display of positive affect. Taken together, individuals with ASD may produce and recognize more frequently some expressions of negative rather than positive emotion. To that end, in a therapeutic context, efforts should be made to display positive emotional reactions in a salient manner by caregivers or teachers who work with youth with ASD so that they may learn to recognize these reactions better and in turn display more positive emotion in order to achieve joyful and synchronized reciprocity during social exchanges.

The current study also contributed to the understanding of the difficulty experienced by participants with ASD in deciphering mental states rather than emotions on a ToM task. In this regard, the precise delineation of the definition of what constitutes ToM and whether it encompasses both perception of feelings and mental states needs to be explored further. In the RMET-C, the ToM task encompasses both recognition of emotion as well recognition of mental states. For example, while being "happy" or "upset" is an emotion, "making someone do something" is classified as a mental state. The finding of slightly lower performance of individuals with ASD on the mental states category appears consistent with ToM, since recognition of mental states may involve higher cognitive neural structures than decoding of discrete emotions that are generally described as being more automatic.

One of the challenges associated with comparing outcomes on the RMET-C task is the variable classification system of negative, positive and neutral mental states and feelings used within different studies; therefore, working towards a unified classification of these constructs would provide increased consistency for comparisons across future studies. At this point, the most consistent comparison can only be conducted by qualitative analysis of responses to individual questions on the RMET-C task.

The discovery of the pattern of relationships between specific symptoms of autism on the ADOS and performance on the RMET-C task is a novel contribution of this study. To that end, on Module 3 (for verbally fluent children and adolescents), items pertaining to communication of language such as: reporting of events, overall level of non-echoed spoke language, and offering of language (all on ADOS-II), as well as, the Communication domain on ADOS-I were predictive of performance on the RMET-C, that is, the increased performance on these items was related to higher performance on the RMET-C.

On Module 4 (for verbally fluent older adolescents and adults), ADOS-II, in addition to the Communication domain, the Restrictive and Repetitive Behaviors (RRB) domain was also correlated with performance on the RMET-C, that is, the increased performance on the Communication domain and the fewer repetitive and restricted interests the individual engaged in, the better was their performance on the RMET-C. Moreover, the better was the participants' use of gestures on Module 4, the higher was their performance on the RMET-C.

To that end, interventions should focus on communicative aspects of language as well as nonverbal communication, such as gestures, facial expressions, body movements, and tone of voice. Repetitive behaviours, restricted interests, and generally getting stuck (e.g., echolalia) appear to interfere with the development of ToM and to that end, the therapy could focus on redirection of some of these behaviours or substituting them with more functional tasks such as: drawing, constructing, listening to music or reading. The capacity to signal different emotional states may be reduced by the presence of these stereotypical behaviors. At the same time, some of these behaviours may function as a way to soothe or diminish anxiety so the function of these behaviours, not only in relation to ToM, but also to the study of autism in general warrants further investigation.

References

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

Anagnostou, E., Zwaigenbaum, L., Szatmari, P., Fombonne, E., Fernandez, B. A., Woodbury-Smith, M., . . . Scherer, S. W. (2014). Autism spectrum disorder: Advances in evidence-based practice. *Canadian Medical Association Journal, 186*(7), 509-519.

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

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

Ashwin, C., Chapman, E., Colle, L., & Baron-Cohen, S. (2006). Impaired recognition of negative basic emotions in autism: A test of the amygdala theory. *Social Neuroscience, 1*(3-4), 349-363.

Assumpcao, F. B. Jr., Sprovieri, M. H., Kuczynski, E., & Farinha, V. (1999). Facial recognizing and autism. *Arquivos-de-Neuro-Psiquiatria. 57* (4), 944-949.

Astington, J. W., & Baird, J. A. (2005). Introduction: Why language matters. In J. W. Astington, & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 3-25). New York, NY: Oxford University Press.

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother–infant and peer–infant interaction. *Child Development, 55*(4), 1278-1289.

Baron-Cohen, S. (1988). Social and pragmatic deficits in autism: Cognitive or affective? *Journal of Autism and Developmental Disorders, 18*(3), 379-402.

Baron-Cohen, S. (1990). Autism: a specific cognitive disorder of "mindblindness". *International Review of Psychiatry, 2*, 81-90.

Baron-Cohen, S., Wheelwright, S., Spong, A., Scahill, V., & Lawson, J. (2001). Are intuitive physics and intuitive psychology independent? A test with children with Asperger syndrome. *Journal of Developmental & Learning Disorders, 5*, 1-58.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Child Psychology & Psychiatry & Allied Disciplines, 38*(7), 813-822.

Bennett, T. A., Szatmari, P., Bryson, S., Duku, E., Vaccarella, L., & Tuff, L. (2013). Theory of mind, language and adaptive functioning in ASD: A neuroconstructivist perspective. *Journal of the Canadian Academy of Child and Adolescent Psychiatry, 22*(1), 13-19.

Benton, A., Sivan, A., Hamsher, K., Varney, N., & Spreen, O. (1983). *Contribution to neuropsychological assessment*. New York: Oxford University Press.

Bieberich, A. A., & Morgan, S. B. (1998). Brief Report: Affective expressions in children with autism. *Journal of Autism and Developmental Delays, 28* (4), 333-338.

Brent, E., Rios, P., Happé, F., & Charman, T. (2004). Performance of children with autism spectrum disorder on advanced theory of mind tasks. *Autism, 8*(3), 283-299.

Bryson, S. E., & Smith, I. M. (1998). Epidemiology of autism: Prevalence, associated characteristics, and implications for research and service delivery. *Mental Retardation and Developmental Disabilities Research Reviews, 4*(2), 97-103.

Capps, L., Kehres, J., & Sigman, M. (1998). Conversational abilities among children with autism and children with developmental delays. *Autism, 2*(4), 325-344.

Carrow-Woolfolk, E. (2011). *Oral and Written Language Scales*, 2nd Edn. Minneapolis, MN: Pearson Psychcorp.

Chakrabarti, S., & Fombonne, E. (2005). Pervasive developmental disorders in preschool children: Confirmation of high prevalence. *The American Journal of Psychiatry, 162*(6), 1133-1141.

Chawarska, K., & Volkmar, F. R. (2005). Autism in infancy and early childhood. In F. R. Volkmar, R. Paul, A. Klin & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders: Diagnosis, development, neurobiology, and behavior* (pp. 223-246). Hoboken, NJ: John Wiley & Sons Inc.

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, not autism, predicts poor recognition of emotional facial expressions. *Psychological Science, 24*(5), 723-732.

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*, 12.

Czapinski, P. K., & Bryson, S.E. (2003). Reduced Facial Muscle Movements in Autism: Evidence for the Dysfunction in the Neuromuscular Pathway? *Brain and Cognition, 51*(2), 177-179.

Darwin, C. (1872). The expression of the emotions in man and animals. London: Murray.

De Giacomo, A., & Fombonne, E. (1998). Parental recognition of developmental abnormalities in autism. *European Child & Adolescent Psychiatry, 7*(3), 131-136.

Derksen, D. G., Hunsche, M. C., Giroux, M. E., Connolly, D. A., & Bernstein, D. M. (2018). A systematic review of theory of mind's precursors and functions. *Zeitschrift Für Psychologie, 226*(2), 87-97.

Deutsch, S. I., & Raffaele, C. T. (2019). Understanding facial expressivity in autism spectrum disorder: An inside out review of the biological basis and clinical implications. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 88*, 401-417.

Duffy, C., & Healy, O. (2011). Spontaneous communication in autism spectrum disorder: A review of topographies and interventions. *Research in Autism Spectrum Disorders, 5*(3), 977-983.

Eigsti, I., Bennetto, L., & Dadlani, M. B. (2007). Beyond pragmatics: Morphosyntactic development in autism. *Journal of Autism and Developmental Disorders, 37*(6), 1007-1023.

Fisher, N., Happé, F., & Dunn, J. (2005). The relationship between vocabulary, grammar, and false belief task performance in children with autistic spectrum disorders and children with moderate learning difficulties. *Journal of Child Psychology and Psychiatry, 46*(4), 409-419.

Fodstad, J. C., Matson, J. L., Hess, J., & Neal, D. (2009). Social and communication behaviours in infants and toddlers with autism and pervasive developmental disorder-not otherwise specified. *Developmental Neurorehabilitation, 12*(3), 152-157.

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind." *Cognition, 50(1-3)*, 115-132.

Griffiths, S., Jarrold, C., Penton-Voak, I., Woods, A. T., Skinner, A. L., & Munafò, M. R. (2019). Impaired recognition of basic emotions from facial expressions in young people with autism spectrum disorder: Assessing the importance of expression intensity *Journal of Autism* *and Developmental Disorders, 49*(7), 2768-2778.

Hayward, E. O. (2012). *Measurement of advanced theory of mind in school-age children: Investigating the validity of a unified construct* (Doctoral dissertation). Retrieved from http://ezproxy.library.yorku.ca/login?url=https://search-proquest-com.ezproxy.library.yorku.ca/docview/1269430795?accountid=15182.

Hill, E. L. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences, 8*(1), 26-32.

Hobson, R. P. (1993). Understanding persons: The role of affect. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. C ohen (Eds.), *Understanding other minds: Perspectives from Autism* (pp. 204-224 or 204-227). Oxford: Oxford University Press.

Hobson, R. P. (1992). Social perception in high-level autism. In E. Schopler, & G. B. Mesibov (Eds.). *High-functioning individuals with autism; high-functioning individuals with autism* (pp. 157-184). New York, NY: Plenum Press.

Hobson, R. P. (1986). The autistic child's appraisal o f expressions o f emotion. *Journal of Child Psychology and Psychiatry, 27*, 321-342.

Holt, R. J., Chura, L. R., Lai, M. -., Suckling, J., von, d. H., Calder, A. J., . . . Spencer, M. D. (2014). 'Reading the mind in the eyes': An fMRI study of adolescents with autism and their siblings. *Psychological Medicine, 44*(15), 3215-3227.

Howlin, P. (2003). Outcome in high-functioning adults with autism with and without early language delays: Implications for the differentiation between autism and Asperger syndrome. *Journal of Autism and Developmental Disorders, 33*(1), 3-13.

Izard, C. E., Fine, S., Mostow, A., Trentacosta, C., & Campbell, J. (2002). Emotion processes in normal and abnormal development and preventive intervention. *Development and Psychopathology, 14*(4), 761-787.

Johnson, C. P., & Myers, S. M. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics, 120*(5), 1183-1215.

Jones, E. A., Carr, E. G., & Feeley, K. M. (2006). Multiple effects of joint attention intervention for children with autism. *Behavior Modification, 30*(6), 782-834.

Jones, W., Carr, K., & Klin, A. (2008). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiatry, 65*(8), 946-954.

Joseph, R. M., & Tager-Flusberg, H. (2004). The relationship of theory of mind and executive functions to symptom type and severity in children with autism. *Development and Psychopathology, 16*(1), 137-155.

Joseph, R. M., & Tager-Flusberg, H. (1997). An investigation of attention and affect in children with autism and Down syndrome. Journal of Autism and Developmental Disorders, 27, 385-396.

Kaland, N., Callesen, K., Møller-Nielsen, A., Mortensen, E. L., & Smith, L. (2008). Performance of children and adolescents with Asperger syndrome or high-functioning autism on advanced theory of mind tasks. *Journal of Autism and Developmental Disorders, 38*(6), 1112-1123.

Kamp-Becker, I., Ghahreman, M., Smidt, J., & Remschmidt, H. (2009). Dimensional structure of the autism phenotype: Relations between early development and current presentation. *Journal of Autism and Developmental Disorders, 39*(4), 557-571.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2,* 217-250.

Kasari, C., Sigman, M., Mundy, P., & Yirmiya, N. (1990). Affective sharing in the context of joint attention interactions of normal, autistic, and mentally retarded children. *Journal of Autism and Developmental Disorders, 20*, 87-100.

Kelley, E., Paul, J. J., Fein, D., & Naigles, L. R. (2006). Residual language deficits in optimal outcome children with a history of autism. *Journal of Autism and Developmental Disorders, 36*(6), 807-828.

Kielinen, M., Rantala, H., Timonen, E., Linna, S-L, and Moilanen, I (2004). Associated medical disorders and disabilities in children with autistic disorder: A population based study. *Autism, (8)*, 249-260.

Klin, A., & Jones, W. (2006). Attributing social and physical meaning to ambiguous visual displays in individuals with higher-functioning autism spectrum disorders. *Brain and Cognition, 61*(1), 40-53.

Krzeminska, P. (2001). *Facial Emotion in Autism: Evidence for atypical expressions and reduced muscle movements*. (Unpublished master's thesis). York University, Toronto, Canada.

Lacroix, A., Guidetti, M., Roge, B., & Reilly, J. (2009). Recognition of emotional and nonemotional facial expressions: A comparison between Williams syndrome and autism. *Research in Developmental Disabilities, 30*, 976–985.

Moss, S., Emerson, E., Kiernan, C., Turner, S., Hatton, C., & Alborz, A. (2000). Psychiatric symptoms in adults with learning disability and challenging behaviour. *The British Journal of Psychiatry, 177*, 452-456.

Piggot, J., Kwon, H., Mobbs, D., Blasey, C., Lotspeich, L., Menon, V., et al. (2004). Emotional attribution in high-functioning individuals with autistic spectrum disorder: A functional imaging study. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*, 473–480.

Levy, F. (2007). Theories of autism. *Australian and New Zealand Journal of Psychiatry, 41*(11), 859-868.

Li, X., Liu, J., Yang, W., Cao, B., He, X., Li, Z., . . . Guo, Y. (2012). Executive function, theory of mind, and symptom in children with high functioning autism. *Chinese Mental Health Journal, 26*(8), 584-589.

Livingston, L. A., Colvert, E., Bolton, P., & Happé, F. (2019). Good social skills despite poor theory of mind: Exploring compensation in autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 60*(1), 102-110.

Lord, C., Luyster, R. J., Gotham, K., & Guthrie, W. (2012). *Autism diagnostic observation schedule, second edition (ADOS-2) manual (Part II): Toddler module*. Torrance, CA: Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism Diagnostic Observation Schedule—WPS*. Los Angeles, CA: Western Psychological Services.

Mazzone, L., Ruta, L., & Reale, L. (2012). Psychiatric comorbidities in Asperger syndrome and high functioning autism: diagnostic challenges. *Annals of General Psychiatry, 11*, 16.

McGonigle-Chalmers, M., Alderson-Day, B., Fleming, J., & Monsen, K. (2013). Profound expressive language impairment in low functioning children with autism: An investigation of syntactic awareness using a computerised learning task. *Journal of Autism and*

Developmental Disorders, 43(9), 2062-2081.

McCrimmon, A., & Rostad, K. (2014). Review of autism diagnostic observation schedule, second edition (ADOS-2) manual (part II): Toddler module and autism diagnostic observation schedule, second edition. *Journal of Psychoeducational Assessment, 32*(1), 88-92.

Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychological Bulletin, 135*(5), 749-773.

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78*(2), 622-646.

Moor, B. G., Op, d. M., Güroğlu, B., Rombouts, S. A. R. B., Van, d. M., & Crone, E. A. (2012). Neurodevelopmental changes of reading the mind in the eyes. *Social Cognitive and Affective Neuroscience, 7*(1), 44-52.

Müller, C. M., & Gmünder, L. (2014). An evaluation of the "Reading the mind in the eyes-test" with seventh to ninth graders. *Journal of Mental Health Research in Intellectual Disabilities, 7*(1), 34-44.

Mundy, P., Sigman, M. D., Ungerer, J., & Sherman, T. (1986). Defining the social deficits of autism: The contribution of non-verbal communication measures. *Child Psychology & Psychiatry & Allied Disciplines, 27*(5), 657-669.

Mundy, P., Sigman, M. D., Ungerer, J., & Sherman, T. (1986). Defining the social deficits of autism: The contribution of non-verbal communication measures. *Child Psychology & Psychiatry & Allied Disciplines, 27*(5), 657-669.

Oakley, B. F. M., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not theory of emotion: A cautionary note on the reading the mind in the eyes test. *Journal of Abnormal Psychology, 125*(6), 818-823.

Osterling, J. A., Dawson, G., & Munson, J. A. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Development and Psychopathology, 14*(2), 239-251.

Overgaauw, S., van Duijvenvoorde, Anna C. K., Moor, B. G., & Crone, E. A. (2015). A longitudinal analysis of neural regions involved in reading the mind in the eyes. *Social Cognitive and Affective Neuroscience, 10*(5), 619-627.

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, 42*, 1081-1104.

Peterson, C. C., Slaughter, V., & Brownell, C. (2015). Children with autism spectrum disorder are skilled at reading emotion body language. *Journal of Experimental Child Psychology, 139*, 35-50.

Pellicano, E. (2010). The development of core cognitive skills in autism: A 3-year prospective study. *Child Development, 81*(5), 1400-1416.

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental Psychology, 43*(4), 974-990.

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.

Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., . . . Hall, J. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological Medicine, 40*(11), 1919-1929.

Philpott, A. L., Rinehart, N. J., Gray, K. M., Howlin, P., & Cornish, K. (2013). Understanding of mental states in later childhood: An investigation of theory of mind in autism spectrum disorder and typical development with a novel task. *International Journal of Developmental Disabilities, 59*(2), 108-117.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515-526.

Rodier, P. M., Ingram, J. L., Tisdale, B., Nelson, S., & Romano, J. (1996). Embryological origin for autism: Developmental anomalies of the cranial nerve motor nuclei. *Journal of Comparative Neurology, 370,* 247-261.

Romero, V., Fitzpatrick, P., Roulier, S., Duncan, A., Richardson, M. J., & Schmidt, R. C. (2018). Evidence of embodied social competence during conversation in high functioning children with autism spectrum disorder. *PLoS ONE, 13*(3), 27.

Rueda, P., Fernández-Berrocal, P., & Baron-Cohen, S. (2015). Dissociation between cognitive and affective empathy in youth with Asperger syndrome. *European Journal of Developmental Psychology, 12*(1), 85-98.

Snow, M. E., Hertzig, M. E., & Shapiro, T. (1987). Expression of emotion in young autistic children. Journal of American Academy of Child and Adolescent Psychiatry, 26, 836-838.

Tager-Flusberg, H., & Joseph, R. M. (2005). How language facilitates the acquisition of false-belief understanding in children with autism. In J. W. Astington, & J. A. Baird (Eds.), *Why language matters for theory of mind; why language matters for theory of mind* (pp. 298-318). New York, NY: Oxford University Press.

Tager-Flusberg, H., & Anderson, M. (1991). The development of contingent discourse ability in autistic children. *Child Psychology & Psychiatry & Allied Disciplines, 32*(7), 1123-1134.

Tanaka, J. W., & Sung, A. (2016). The "eye avoidance" hypothesis of autism face processing. *Journal of Autism and Developmental Disorders, 46*(5), 1538-1552.

Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Eds.), *Before Speech* (pp. 321-347). Cambridge: Cambridge University Press.

Turner, M. (1997). Towards an executive dysfunction account of repetitive behaviour in autism. In J. Russell (Ed.), *Autism as an executive disorder; autism as an executive disorder* (pp. 57-100). New York, NY: Oxford University Press.

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders, 43*(7), 1517-1526.

Vogindroukas, I., Chelas, E., & Petridis, N. E. (2014). Reading the mind in the eyes test (Children's version): A comparison study between children with typical development, children with high-functioning autism and typically developed adults. *Folia Phoniatrica Et Logopaedica:International Journal of Phoniatrics, Speech Therapy and Communication Pathology, 66*(1-2), 18-24.

Volden, J., & Lord, C. (1991). Neologisms and idiosyncratic language in autistic speakers. *Journal of Autism and Developmental Disorders, 21*(2), 109-130.

Walenski, M., Tager-Flusberg, H., & Ullman, M. T. (2006). Language in autism. In S. O. Moldin, & J. L. R. Rubenstein (Eds.), *Understanding autism: From basic neuroscience to treatment; understanding autism: From basic neuroscience to treatment* (pp. 175-203). Boca Raton, FL: CRC Press.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655-684.

Whalen, C., Schreibman, L., & Ingersoll, B. (2006). The collateral effects of joint attention training on social initiations, positive affect, imitation, and spontaneous speech for young children with autism. *Journal of Autism and Developmental Disorders, 36*(5), 655-664.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103-128.

Wechsler, D. (2004). *Wechsler Intelligence Scale for Children–Fourth Edition: Canadian Technical Manual.* Toronto, Ontario, Canada: PsychCorp.

Wechsler, D. (2008b). *Wechsler Adult Intelligence Scale–Fourth Edition: Canadian Technical Manual.* Toronto, Ontario: Pearson Canada

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence.* New York, NY: The Psychological Corporation: Harcourt Brace & Company.

Wechsler, D. (2012). *Wechsler Preschool and Primary Scale of Intelligence: Canadian Technical Manual.* Toronto, Ontario: Pearson Canada

Yeung, M. K., Han, Y. M. Y., Sze, S. L., & Chan, A. S. (2014). Altered right frontal cortical connectivity during facial emotion recognition in children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 8*(11), 1567-1577.

Yeung, M. K., Lee, T. L., & Chan, A. S. (2019). Impaired recognition of negative facial expressions is partly related to facial perception deficits in adolescents with high-functioning autism spectrum disorder. *Journal of Autism and Developmental Disorders, 8*(11), 167-1577. doi:http://dx.doi.org.ezproxy.library.yorku.ca/10.1007/s10803-019-03915-3.

Yirmiya, N., Kasari, C., Sigman, M., & Mundy, P. (1989). Facial expressions of affect in autistic mentally retarded and normal children. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 30*, 725-735.

Zainal, N. H., & Newman, M. G. (2018). Worry amplifies theory-of-mind reasoning for negatively valenced social stimuli in generalized anxiety disorder. *Journal of Affective Disorders, 227*, 824-833.

Zwaigenbaum, L., Bryson, S. E., Szatmari, P., Brian, J., Smith, I. M., Roberts, W., . . . Roncadin, C. (2012). Sex differences in children with autism spectrum disorder identified within a high-risk infant cohort. *Journal of Autism and Developmental Disorders, 42*(12), 2585-2596.

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience, 23*(2-3), 143-152.