THE IMPACT OF EXECUTIVE FUNCTIONING ON THE MATHEMATICS ACHIEVEMENT OF AUTISTIC ADOLESCENTS

by

Nicole R.A. Eddy

A thesis submitted to the
School of Graduate Studies
in partial fulfilment of the
requirements for the degree of
(Master of Science)

Department of Psychology
Memorial University of Newfoundland

August 2022

Abstract

Executive Function (EF) involves the neurological processes behind how we organize environmental information, plan, and direct behaviour. EF components have been found to be related to mathematical performance in the general population. Autistic individuals often experience challenges with their executive functioning– what is referred to as “executive dysfunction”. Additionally, autistic individuals display a widespread pattern of mathematics achievement, with below and above IQ-expected performance, and are significantly more likely to be diagnosed with a math-based learning disability than the general population. The present study explored the relationship of EF and math performance in an autistic youth (around 12-18 years old) sample as compared to a non-autistic sample. Twenty autistic and 34 non-autistic individuals participated in individual remote, online, and synchronous testing sessions in which they completed a battery of measures– including an intellectual ability assessment, fractional (procedural and conceptual) and operational math measures, and EF component tasks. Results from a series of hierarchical linear regressions indicated that the only significant group difference in math skills was found for the fractional procedural math measure; however, none of the EF component task scores were able to account for this difference. This suggests that EF may not exert the same degree of influence on autistic math performance as it does for the general population; however, the small sample size acquired for this study largely limits the statistical power and scope of the results. Future research should explore other possible predictors of math performance in autistic samples.

Keywords: autism, executive function, mathematics, working memory, shifting, inhibition

Acknowledgements

First and foremost, I wish to thank my supervisor, Dr. Darcy Hallett, for without his continued guidance, support, and wealth of knowledge, this project would not be possible. His generosity and efforts have not gone unnoticed. Similarly, a huge thank you to my committee members, Drs. Carole Peterson and Sharon Penney, for lending their expertise to this project. Additional thanks to the members of the Research Center for the Development of Mathematical Cognition and lab staff for their help with recruitment and data collection. Lastly, thank you to my friends and family for their unwavering love and support. One of my greatest accomplishments in life is knowing that I have made my parents proud.

Table of Contents

Page

List of Tables

Introduction

Executive function (EF) is the system of neurological processes which support our ability to think and problem solve in a flexible and goal-directed manner– a system of planning and execution which receives incoming stimuli information, filters out that which is not relevant, and further directs our behavioural responses (Cragg & Gilmore, 2014; Ozonoff et al., 1991). Three primary components of EF have been identified: working memory (i.e., updating), shifting, and inhibition (Miyake et al., 2000). Each of these EF components have been implicated in math performance outcomes (see Cragg & Gilmore, 2014). While strong evidence exists in support of working memory having influence on math achievement (Andersson, 2007, 2008; see Cragg & Gilmore, 2014), and similarly for shifting (see Yeniad et al., 2013), results on the impact of inhibition are mixed (Andersson, 2008; Bull & Scerif, 2001; see Cragg & Gilmore, 2014).

It has been noted that EF challenges often present in those diagnosed with Autism Spectrum Disorder (ASD), what is commonly referred to as *executive dysfunction*. Again, there is stronger evidence for the existence of executive dysfunction in ASD for certain EF components (i.e., working memory and shifting) than for others (i.e., inhibition, see Hill, 2004). Research on math performance of autistic[1] individuals has demonstrated a large range of abilities, with reports of performance below average (Chiang & Lin, 2007; Oswald et al., 2016; Wei et al., 2015), comparable to average (Chiang & Lin, 2007; Mayes & Calhoun, 2008), and above average scores (Estes et al., 2011; Jones et al., 2009). Because of the impact which EF abilities have on math performance (see Cragg & Gilmore, 2014), it is of interest to know if the challenges in EF associated with ASD are linked to subsequent math achievement. The purpose

---

[1] Identity-first language (i.e., “autistic”), will be used throughout this paper, in accordance with recommendations from the academic and autistic community (e.g., Bottema-Beutel et al., 2021; Gernsbacher, 2017) to reduce applications of ableist language and the subsequent stigmatization of autism.

of this dissertation is to explore the extent to which EF difficulties might explain math ability in an ASD sample.

What is Executive Function?

One of the earliest conceptualizations of EF was the working memory model proposed by Baddeley and Hitch (1974). This model posits that working memory consists of a *central executive*, which monitors information processing, and two underlying “slave” functions: the *phonological loop* (storage and rehearsal of verbal information) and the *visuospatial sketch pad* (storage of visuo-spatial information and images), both of which are regulated by the central executive. EF has further been expanded to include three primary components: working memory (also referred to as updating), shifting, and inhibition. Other models of EF do exist (e.g., attentional control model, Norman & Shallice, 1980, Cognitive Complexity and Control Theory-revised, Zelazo et al., 2003, as cited in Sparrow & Hunter, 2012); however, the Baddeley and Hitch (1974) model, more specifically, the three aforementioned EF components, seemed most relevant to discussions of math achievement due to the extensive research on, and the strength of, these math-component relationships (further discussed below). Additionally, these three components have been repeatedly reported on in the literature evaluating autistic EF (see Hill, 2004) and in relation to behaviours associated with autism (Yerys et al., 2009).

Often associated with planning (Happé et al., 2006), the *working memory* process begins with the monitoring and subsequent coding of relevant information from the environment, followed by the manipulation and replacement of outdated information held in working memory with the newly acquired information (Morris & Jones, 1990). Working memory is composed of two dissociated parts: verbal working memory and visuospatial working memory (Jarvis & Gathercole, 2003).

Shifting (also known as flexibility) refers to the ability to shift attention between specific tasks or mental sets; “disengagement of an irrelevant task set and the subsequent engagement of a relevant task set” (Miyake et al., 2000, p. 55). This can include the switching of our mental focus from one type of operation to another.

Lastly, *inhibition* refers to our ability to inhibit prepotent responses (Miyake et al., 2000). By inhibiting our automatic behavioural responses, we reduce the potential for making errors, as is demonstrated by performance on the Stroop task (Stroop, 1992), in which participants must name the color of a colour word displayed without reading the word itself (e.g., responding “blue” for the word “red” written in blue ink). While the EF components of working memory, shifting, and inhibition are closely related, they are distinct processes which can be measured independently (Lehto, 1996; Miyake et al., 2000).

EF processes have a neurological basis in the frontal lobe. Research has revealed that patients who have sustained frontal lobe damage are challenged in their ability to regulate their behaviours and show impaired performance on measures of EF; however, these results are not universal. A meta-analytic review of studies comparing the performance on EF measures (i.e., Wisconsin Card Sorting Task [WCST], phonemic verbal fluency task, and Stroop task) of patients with frontal lobe damage to healthy controls found that while performance did vary with frontal lobe damage, not all patients demonstrated this impairment (Alvarez & Emory, 2006). Though not all-encompassing, this pattern of frontal lobe damage-EF impairment sets a foundation for understanding the neurological functions which underly EF, and how activation of these brain areas may change throughout development.

Executive Functioning across Development

EF performance seems to follow a predominantly linear developmental projection for typically developing individuals from early childhood onward into mid-adolescence, although the different components of EF do appear to develop asynchronously. A review of the development of EF from age 5 years through adolescence demonstrated that reliance on inhibition tends to diminish as children age, whereby inhibition is superseded as the primary predictor of task-related performance by other components of EF (i.e., working memory and shifting; see Best et al., 2009). Specific to inhibition, errors on A-not-B-task variations are reduced significantly in pre-school years. In this task, participants repeatedly retrieve (following a delay) an object which they witness several times as being hidden in location A. Once the object is relocated to location B, any continued searching in location A once hiding in location B is established is considered an error (Piaget, 1954, as cited in Miller & Marcovitch, 2011). Improvement on this and similar inhibitory tasks continues throughout ages 5 to 8 years, and ultimately plateaus in mid-adolescence into adulthood (though improvements have been reported until ages 15 and 21, see Best et al., 2009). These improvements in inhibitory control are concurrently matched with documented fine-tuning of prefrontal cortex activity and related networks (Best et al., 2009).

Improvements in working memory have been noted from ages 4 to 15 years, with variability in performance development depending on the difficulty level of the task presented (developmental differences emerge with increased task difficulty). Changes in brain activity in relation to working memory performance occur both in the location of activity as well as the amount of activity. For instance, in studying visuospatial working memory, children showed activation in the ventromedial brain areas (i.e., thalamus and basal ganglia), which in adolescence then shifts towards frontal regions (increased right dorsolateral prefrontal cortex activation with decreased ventromedial activation), and by adulthood had established as local activation in the left dorsolateral prefrontal cortex with considerable increased activation of the anterior cingulate (see Best et al., 2009).

Children as early as ages 3 and 4 years have demonstrated rudimentary shifting capability. Unlike those of inhibition and working memory, the developmental trajectory of shifting ability does not appear to follow a linear pattern. In terms of shifting response times, there appear to be increases from age 6 years into adulthood (presumed to be resultant of speed-accuracy trade off); however, attainment of normative adult level response times by the age of 15 have been reported (see Best et al., 2009). In comparing children and adolescents (10- to -17-year-olds) to adults (20- to 43-year-olds), Rubia et al. (2006) noted that adults displayed increased activation in the right mesial and inferior prefrontal cortex, parietal lobe, and putamen during performance of a set-shifting task. In evaluating typically developing individuals' EF skills, it is necessary to consider these development differences in relation to performance outcomes.

Executive Function-Math Performance Relationship

Previous research has supported the existence of a relationship between mathematic achievement and EF. Both EF as a composite ability as well as its component parts have been shown to influence performance in mathematics (see Cragg & Gilmore, 2014). Experimental studies have shown that EF overload (while performing dual-tasks, specifically those taxing the central executive) significantly undermines adult arithmetic processing ability (though these studies did not assess math performance per se, see Cragg & Gilmore, 2014). Additionally, this EF-math performance relationship appears to be bidirectional, as children with math difficulties have displayed significant impairments in EF as compared to those with typical math performance (Bull & Scerif, 2001).

EF has also been shown to predict math performance over and beyond other factors, including age (Andersson, 2007), IQ (Kroesbergen et al., 2009), and reading ability (Andersson, 2007). As well, studies have found that EF ability contributes to math skill attainment across numerous skill-building categories, from the earliest stages of numeracy (i.e., counting, see Kroesbergen et al., 2009) to more advanced operations (i.e., arithmetic problem solving, see Andersson, 2007). Predictability of math achievement from EF performance has been shown to span from just one year to many years past (Cragg & Gilmore, 2014); however, this may be task-dependent (LeFevre et al., 2013). Overall, EF appears to be a significant predictor of success on mathematical constructs.

EF Component and Math: Working Memory/Updating

Looking at the EF components individually, working memory appears to have the strongest relationship with math performance. Research evidence supports the notion that working memory (specifically the updating of information in working memory) predicts performance across a variety of math assessments, such as written arithmetic calculations (Andersson, 2008) and word problems (Agostino et al., 2010; Andersson, 2007; Passolunghi & Pazzaglia, 2005; but see Oswald et al., 2016), as well as across various ages (Cragg et al., 2017; see Cragg & Gilmore, 2014). Andersson (2007) found that working memory was able to account for the variance in word problem solving ability beyond that which is accounted for by fluid IQ, reading ability, and age. As well, the independent processes of working memory have been shown to be predictive of improvements in mathematical achievement over time. For example, in studying the growth of mathematics achievement from first to fifth grade, Geary (2011) found that the central executive and visuospatial working memory systems, but not the phonological loop, were significant predictors of later math performance.

Lower performance in working memory has been implicated in studies of children with math learning difficulties. A meta-analysis of research studies comparing children with math learning difficulties to those of average achievement revealed that children with math learning difficulties display a significant deficit in the central executive system of working memory (see David, 2012). Likewise, another meta-analysis of the literature comparing children with math disabilities to average achievers found that children with math disabilities had poorer performance on measures of both verbal working memory and visual-spatial working memory (see Swanson & Jerman, 2006).

The specific components of working memory seem to contribute to math performance in distinctive ways. For instance, the phonological loop has been implicated as being an important process involved in number articulation (Krajewski & Schneider, 2009). Additionally, research suggests that for children's non-symbolic and symbolic mathematical ability, performance may vary with disruption to the verbal and visuospatial working memory separately, depending on the math task at hand (see Cragg & Gilmore, 2014). Thus, the working memory system and its component processes appear to have a significant impact on mathematical achievement.

EF Component and Math: Shifting

Studies have reported on the existence of a relationship between the shifting ability of EF and mathematical performance. A meta-analysis exploring this relationship in children found that performance on shifting tasks is significantly related to mathematical achievement (Yeniad et al., 2013). Similarly, children with lower math ability also display poorer shifting ability (Bull et al., 1999). While it would seem that shifting is closely related to mathematical achievement, not all research results have supported this notion. In one study evaluating math achievement levels across ages ranging from childhood (8- to 9-years-old) to young adulthood (18- to 25-years-old), shifting failed to predict performance on all mathematical measures, including factual knowledge, procedural skills, conceptual understanding, and arithmetic calculations (Cragg et al., 2017). Additionally, shifting ability failed to predict grade three to six students' performance on one-step and multi-step multiplication word problems (Agostino et al., 2010). Therefore, further research evaluating the relationship of shifting ability to math performance is required.

EF Component and Math: Inhibition

In terms of the influence that inhibitory abilities have on mathematical achievement, the evidence is also inconclusive. While a number of studies have found that inhibition skills are related to mathematic ability in children (Bull & Scerif, 2001; St. Clair-Thompson & Gathercole, 2006), competing evidence suggests otherwise (see Cragg & Gilmore, 2014). Andersson (2008) found that performance on a Stroop task (Stroop, 1992) did not significantly predict written arithmetic operation performance in third and fourth grade students. As well, Agostino et al. (2010) found that while inhibition ability was related to performance on multiplication word problems, this relationship was mediated by updating. It may be the case that only particular types of inhibition are relevant to mathematical performance. For instance, in Cragg et al.'s (2017) research on math achievement across different age groups, results indicated that numerical inhibition, but not non-numerical inhibition, significantly predicted unique variance in scores of factual knowledge and procedural skills. Based on these mixed findings, it appears that further investigation of the inhibition-math performance relationship is needed.

What is Autism Spectrum Disorder?

ASD, or more simply autism, is categorized in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) as a pervasive neurodevelopmental disorder comprised of the following primary characteristics: social-communicative impairments and restricted and repetitive behaviours and interests (American Psychiatric Association, 2013). Autism is an umbrella term which encompasses a variety of spectrum disorders, including those previously diagnosed with Asperger syndrome (since the removal of this diagnosis from the DSM-5, American Psychiatric Association, 2013). Discourse on autism has more recently shifted towards a perspective of autism as a neurodiversity rather than a disorder; signifying differences rather than inequalities between autistic and non-autistic persons (e.g., Baron-Cohen, 2017; Nicolaidis, 2012; Robertson, 2010). This corresponds with the qualitative results of Kenny et al.'s (2016) study, in which the use of identity-first language was preferred by most autistic adults, some of whom considered their autism "as a different way of seeing the world" (p. 447) rather than a condition.

By nature of the spectrum– though not descriptive in a linear sense with defined boundaries, but rather illustrating autism heterogeneity– the degree and presentation of autism symptoms can vary by individual diagnoses (Wing & Gould, 1979). Cases may include those with below average IQ (<70) which may be accompanied by a number of learning disabilities (i.e., 25-40% of cases, see Baird et al., 2000; Chakrabarti & Fombonne, 2001) or co-occurring diagnoses (e.g., anxiety, ADHD, OCD, epilepsy; Government of Canada, 2018). Those with intact communication skills and average to above average IQ (minimum IQ >70) have been previously referred to as having "high-functioning autism" (Baron-Cohen, 2000a; Honda et al., 2005), though reference to specific support needs of individuals as opposed to function-based labels may be more appropriate. Autistic children who demonstrate average intellectual ability are generally integrated into mainstream schooling along with neurotypical peers (Estes et al., 2011; Siegel et al., 1996).

Presently, it is estimated that approximately one in every 160 children worldwide qualifies as having an ASD (Elsabbagh et al., 2012), an increasing figure over the past several decades (World Health Organization, 2019). In Canada specifically, the autism prevalence rate is one in 66 children, with males (1 in 42) outnumbering females (1 in 89; Autism Speaks, 2020). Individuals (in particular, males) within the general population have also demonstrated autistic-like traits without meeting the full criteria for an autism diagnosis (Baron-Cohen et al., 2001). For example, White et al. (2011) found that approximately .7 to 1.9 percent of a college undergraduate population (n = 667) qualified as having “high-functioning autism”, an estimate of about 1 in 130 to 1 in 53 students. Due to its pervasive nature (American Psychiatric Association, 2013), autism is a lifelong diagnosis that persists beyond childhood. The employability rate of autistic persons (aged 20-64) in Canada is 33 percent, with 35 percent having received less than a high school education (Government of Canada, 2020).

Executive Functioning in Autism

It has been theorized that executive dysfunction accompanies diagnoses of autism, specifically in relation to deficits of the frontal lobe in this population (Hill, 2004). Autistic individuals appear to have difficulty executing tasks that tax their executive functioning systems, such as those involved in working memory updating and shifting (Ozonoff & Jensen, 1999). Challenges in EF also appear to be linked to autism symptomology. For example, Mosconi et al. (2009) found that increased inhibitory error rates were associated with increased severity of higher-order repetitive behaviours. Likewise, errors in shifting ability have also been reported as being related to increased repetitive behaviour symptoms (Yerys et al., 2009).

The EF abilities of autistic individuals have been likened to those of the aforementioned frontal lobe patients (Hill, 2004). There has been notable variability in EF with autism, specifically across different levels of IQ (see Hill, 2004) and ages (Chen et al., 2016; Happé et al., 2006; van den Bergh et al., 2014), indicating that executive dysfunction is not ubiquitous across all autism diagnoses. As EF appears to play a role in mathematics achievement, it is of interest to know if the EF challenges associated with autism impact their performance on math constructs.

Development of EF in Autism

Severity of executive dysfunction across the components of EF in autism varies with age. In comparing youth (aged 8-12 years) and adolescent (aged 13-18 years) groups to age-matched neurotypical groups, Chen et al. (2016) reported greater performance on EF measures (including verbal and spatial working memory, planning, and shifting) for the adolescent autistic group compared to the youth autistic group. Age-stratified analyses revealed that significantly poorer performance on the planning and shifting subtests occurred only for the youth autistic group when compared to the neurotypical group, while working memory difficulties were present for both autistic age groups.

Another study looking at EF challenges in autistic 6- to 18-year-olds found significant age-related differences, with younger participants demonstrating greater difficulties in inhibition (6- to 8-year-olds) and shifting (9- to 11-year-olds and 12- to 14-year-olds). However, it appeared that planning difficulty increased with age, with 12- to 14-year-olds showing greater impairments compared to 9- to 11-year-olds (van den Bergh et al., 2014). Happé et al. (2006)

also reported differences between age groups, with the older autistic group (11- to 16-year-olds) outperforming the younger (8- to 10-years-old) autistic group on measures of inhibition, shifting, planning, and spatial working memory. Similarly, in comparing autistic individuals from ages 8 to 33 years, Luna et al. (2007) found that adult level inhibition skills appeared around age 15, while neurotypical controls seemed to acquire this skill level by age 14. Developmental changes in working memory performance, however, were not significant for the autistic group. Overall, these differences in executive functioning across various ages lends to more concrete understanding of the developmental changes that accompany EF in autism.

Autism-EF Relationship

Research has suggested that the primary impairments in EF for autistic individuals is rooted in their capacity for flexible thinking (i.e., shifting), working memory, and planning. The ability to inhibit information, however, appears relatively intact (Ozonoff & Jensen, 1999). In order to gain a complete understanding of EF in autism, it is important to break down EF into its individual processes.

EF Component: Working memory/Updating

Research focused on working memory ability have found relatively consistent impairments for autistic individuals. Using the Tower of Hanoi (Borys et al., 1982) or related Tower of London task, both of which measure planning and working memory, numerous studies have reported impaired performance from autistic individuals when compared to neurotypical individuals, as well as other neurotypes (e.g., ADHD, dyslexia; see Hill, 2004). For example, autistic individuals with a learning disability have been shown to display greater impairments on tasks of working memory and planning when compared to non-autistic individuals with a learning disability (Barnard et al., 2008). However, variable results of autistic executive dysfunction have suggested that impairments may arise from other factors, such as IQ level (see Hill, 2004), though evidence exists which suggests otherwise. For instance, autistic children labelled “high-functioning” have demonstrated below average scores on the WISC-IV working memory subtest (Mayes & Calhoun, 2008). Additionally, while some aspects of working memory may present as dysfunctional (e.g., visuospatial working memory, see Williams et al., 2005) others may not (e.g., verbal working memory, see Williams et al., 2005; Cui et al., 2010). Therefore, there appears to be a relatively consistent pattern of working memory impairment with autism, though which components of working memory are most affected remains indeterminate.

EF Component: Shifting

Shifting ability is also seemingly challenged in autism. Autistic individuals have demonstrated greater impairment in performance on measures of shifting, such as the Wisconsin Card Sorting Task (WCST, Grant & Berg, 1948), as compared to samples of neurotypicals, ADHD, and individuals with dyslexia. The primary outcome of these measures is that autistic individuals struggle to shift to a new rule of sorting, thus continuing to sort the cards by way of the initial rule (see Hill, 2004). Some studies, however, have failed to find impaired shifting ability in autism. In comparing autistic adults with normal IQ (>70) to neurotypical controls, Minshew et al. (1992) reported finding no group differences in error rates on the WCST. A review of the literature on flexibility in autism reported that inconsistencies exist across studies of shifting dysfunction, namely when considering alternative measures of shifting outside of the WCST (e.g., trail-making test of modified card sorting task, see Geurts et al., 2009). Thus, further investigation is required to fully comprehend the impairment of shifting ability in autism.

EF Component: Inhibition

In contrast to the other EF components, autistic individuals appear to perform within normative levels on measures of inhibition (see Hill, 2004), such as the Stroop task (Eskes et al., 1990) and stop-signal task (Ozonoff & Strayer, 1997). Additionally, Brian et al. (2003) reported no difference when comparing an autistic and control sample on performance of a negative priming task of inhibition. The only evidence for potential inhibitory control impairment comes from studies utilizing measures of inhibiting prepotent responses, results of which have been attributed to the arbitrary rules which constitute these measures (see Hill, 2004). For example, studies have reported inhibitory control challenges for autistic individuals when evaluated via antissacade measures (Goldberg et al., 2002; Luna et al., 2007; Mosconi et al., 2009). Despite a lack of impairment in the inhibition component of EF, it may still be worthwhile to investigate whether an autistic samples' performance on such measures is linked to their mathematical ability.

Autism and Academic Performance

In terms of general academic achievement, autistic individuals have been shown to demonstrate a vast range of academic success. For instance, Jones et al. (2009) found that among the autistic adolescents (N = 100; IQ range 50-119) sampled, four separate subgroups of achievement (above or below IQ-expected attainment levels) were identified: Reading Peak (n = 14), or reading skills above intellectual ability; Reading Dip, or reading skills below intellectual ability (n = 10); Arithmetic Peak, or arithmetic abilities above intellectual ability (n = 16); and Arithmetic Dip, or arithmetic abilities below intellectual ability (n = 6). This indicates that just under half the sample occupied at least one of these subgroups. As well, in another study by Wei et al. (2015), of the 130 autistic children (ages 6- to 9-years-old) tested, a considerable portion of the sample fell below (32 percent) or within (39 percent) the national average for reading and arithmetic scores. Likewise, achievement across measures of spelling, word reading, and basic number skills were variable in autistic 9-year-olds, with achievement extremities both lower and higher than anticipated based on participant IQ levels (Estes et al., 2011).

Growth in academic performance may vary depending on degree of autism symptoms. For instance, for autistic children, greater teacher-reported social skills, but not degree of parent-reported problem behaviours, at age 6 were related to greater success academically by age 9 (Estes et al., 2011). Another study looking at the reading and mathematical projection outcomes of autistic children identified four subgroups: higher-achieving (average reading and math scores and below average Rapid Letter Naming scores); hyperlexia (average letter/word identification, above average Rapid Letter Naming, and below average reading comprehension and math scores); hypercalculia (average calculation and below average reading and Rapid Letter Naming scores); and lower-achieving (below average scores on all measures). The higher-achieving and hyperlexia groups were reportedly greater in functional cognitive skills than the hypercalculia and lower-achieving groups; however, no differences existed between groups on social or conversational skills (Wei et al., 2015). Therefore, it remains unclear whether autism symptomology plays a role in academic achievement outcomes.

Extreme Male Brain Theory

Instances of savant-like abilities in math-based disciplines (e.g., González-Garrido et al., 2002) and systematic processing (e.g., Hughes et al., 2018) has led to the general assumption that autistic individuals possess superior systemizing skills (Baron-Cohen, 2002). *Extreme Male Brain Theory*, as proposed by Baron-Cohen (2002), posits that autistic individuals demonstrate a strengthened ability to systemize. More specifically, *systemizing* consists of “the drive to analyze the variables in a system, to derive the underlying rules that govern the behaviour of a system…

[allowing] you to predict the behaviour of a system and to control it" (Baron-Cohen, 2002, p. 248). In contrast, *empathizing* is the ability to recognize others' mental states and to respond accordingly using the appropriate affectation. From a gendered-binary standpoint, male brains are considered to be more systemizing than empathizing, possessing greater comprehension of rule-based systems, superior visual-spatial abilities, and finer attention to relevant details. The female brain, on the other hand, is more socially-oriented, with inclined sensitivity for others' emotional well-being (i.e., "mindreading"), superior communicative and cooperative skills, and assignment of greater value to social relationships.

This theory proposes that the brain of autistic individuals resembles an exaggerated version of the male brain. The ability to comprehend and empathize with others' mental states is seemingly compromised (i.e., 'mindblindness', see Baron-Cohen, 1995), as is demonstrated by performance on Theory of Mind tasks (Baron-Cohen, 1985; Happe, 1995). However, autistic individuals appear to be acutely sensitive to and partial towards systems which are based upon predictable, concrete rules, such as computers (Baron-Cohen, 2002). In specific cases, autistic individuals have demonstrated superior skills in mathematical computation (Baron-Cohen & Bolton, 1993). There is also a trend of autistic individuals occupying a greater proportion of academic positions within the science, technology, engineering, and mathematics disciplines (i.e., STEM; Baron-Cohen et al., 2007; Wei et al., 2013) than the general population (Chen & Weko, 2009). Additionally, rates of autism are higher amongst families of parents who work in mathematics-, engineering-, or physics-based disciplines (Baron-Cohen, 2002). While this theory fails to acknowledge the gender stereotypes and expectations which often constitute these more empathetic-female systematic-male mentalities (e.g., Baez et al., 2017), as well as predates perspectives of empathy as being bound to disjointed non-autistic and autistic social understanding (i.e., the “double empathy problem”, see Milton, 2012), it is necessary to take this theory into consideration when evaluating the more “systematic” math sense of autistic individuals.

Math Performance in Autism

There is a paucity of research looking at the math skills of autistic persons (Oswald et al., 2016) and even less exploring executive function challenges in relation to math achievement. That which does exist, however, has presented a similar widespread pattern of achievement levels as was noted for autistic academic success (Estes et al., 2011; Jones et al., 2009).

In reference to math achievement, a review of studies from 1986 to 2006 (Chiang & Lin, 2007) reported that most autistic individuals demonstrated average mathematical ability, though many do present significant (yet clinically modest) mathematical ineptitude. Additionally, some studies have reported instances of superior mathematical ability in their autism samples (Chiang & Lin, 2007). In evaluating autistic adolescents (ages 14-16 years, N = 100) with a large range of IQ scores (50-119), approximately 16 percent of the sample with average IQ presented superior mathematic ability. That said, the researchers also reported below IQ-expected arithmetic performance for six of the participants, with the remainder falling somewhere in between the two subgroups (Jones et al., 2009). In another study looking at achievement levels in thirty “high-functioning” autistic 9-year-olds, results identified 12 participants with significantly lower, and four with significantly higher, performance scores on a basic number skills subtest (Estes et al., 2011). Thus, there appears to be significant discrepancy in the IQ levels, and subsequent math-performance scores, of autistic individuals.

Results of another study have suggested that autistic individuals are far more likely to suffer from a mathematics learning disability than to possess exceptional mathematics ability

(Oswald et al., 2016). As well, the co-occurrence of mathematics learning disability with autism (22 percent, see Mayes & Calhoun, 2003) has been reported at rates of approximately three times higher than that of the general population (Oswald et al., 2016). Autistic children and adolescents have been shown to not only score lower in mathematical calculations compared to age-matched individuals with learning disabilities, but also demonstrate a significantly slower growth rate over a 3-year time period (Wei et al., 2013). However, a separate study reported normative math scores for autistic children (Mayes & Calhoun, 2008). Therefore, the heterogeneity of math achievement levels associated with autism suggests that other factors may serve to influence the math performance outcomes in this population.

EF-Math Performance Relationship in Autism

Very few studies have explored how executive functioning relates to mathematical achievement in the autistic population. In studying mathematical achievement in autistic adolescents (ages 12 to 17 years), Oswald et al. (2016) reported that, unlike perceptual reasoning, verbal ability, and test anxiety, working memory failed to predict unique variance in math problem solving. Another study which tested EF composite measures (i.e., A-not-B Invisible Displacement task, and Spatial Reversal task) in a sample of autistic children found that Spatial Reversal task scores, but not A-not-B Invisible Displacement scores, at age 6 predicted unique variance in age 9 performance on a basic number skills subtest. As well, age 9 EF did not significantly account for unique variance in age 9 basic number skills (St. John et al., 2018). Lastly, May et al. (2013) reported that for 7- to 12-year-olds (labelled as “high-functioning”), performance on an attention-shifting task was significantly correlated with mathematical achievement. To our knowledge, no studies have looked at how all three components of EF (working memory, shifting, and inhibition) individually relate to mathematical achievement in an autistic sample.

Why is Math Achievement Important?

One's mathematics ability can have profound influences on their life outcomes. For instance, skills in mathematics (unsurprisingly) predict university performance, such as student dropout rates (McCoy & Byrne, 2017), achievement in science, technology, engineering, and mathematics (STEM) courses, as well as the likelihood of obtaining a first-class honours degree (Delaney & Devereux, 2020). This in turn, can increase one's opportunities for high-wage employment (Feng & Graetz, 2017). Even knowledge of basic algebra can impact one's employability prospects beyond the influence of reading ability and IQ (Rivera-Batiz, 1992). Considering the low education and employment rates of autistic individuals in Canada (Government of Canada, 2020), consideration needs to be made regarding how mathematics skills contribute to these statistics.

Numeracy comprehension can also influence us outside of an academic setting. In our everyday, we rely on our numerical skills to make decisions about our health (Ancker & Kaufman, 2007) and finances (Banks & Oldfield, 2007). Therefore, it is vital that we understand the impact which EF has on mathematics abilities, including for autistic individuals.

The Present Study

The purpose of the present study was to investigate the influence of executive function components on mathematical achievement in autistic individuals. School-aged autistic adolescents (around 12- to 18-years-old) were tested on discrete measures of the three primary EF components–working memory, shifting, and inhibition–as well as mathematical achievement measures. A control measure (e.g., intellectual functioning) was also administered. Because mathematical achievement can vary with one's IQ level (Andersson, 2008; Mayes et al., 2009), only individuals of average intellectual ability were included in the study to limit this influence. A comparison group of non-autistic adolescents were also similarly tested.

Although the broad mathematical capabilities of autistic individuals have been well documented (e.g., Estes et al., 2011; Jones et al., 2009; Oswald et al., 2016), it remains unknown whether the EF challenges attributed to this population impact their mathematical performance outcomes; specifically, whether the tendency for math ineptitude associated with autism can be explained primarily by executive functioning. The limited preliminary research on the relation of EF to math achievement in autism which does exist (i.e., May et al., 2013; Oswald et al., 2016; St. John et al., 2018) has set the groundwork for further examination of this relationship. Based on findings from the general population which support the notion of an individual's EF capacity exerting influence on their math achievement scores (see Cragg & Gilmore, 2014), it seems reasonable to submit that a like influence would be anticipated for an autistic sample.

Method

Participants

Participants were separated into two groups: an autistic and a non-autistic group. As EF capabilities appear to vary considerably with age, adolescents around ages 12 to 18 years (i.e., grades 6 through 12, or recently graduated) were considered for this study. Participants were recruited across Canada. According to the guided reading levels outlined by Fountas and Pinnell (1996), children in grade 7 are recommended to be at a *Z* reading level (Fountas & Pinnell Literacy, 2012), whereby they should be capable of reading more complex passages. By recruiting participants at a sufficient reading level, it was anticipated that the risk of parental assistance in reading (and possibly on the measures tested) would be minimized.

As most participants were under the age of majority (i.e., 18 to 19 years, Government of Canada, 2013) parental consent was required. All parents were provided an informed consent form, which outlined the purpose and procedures of the study for parent/caregiver comprehension. Parents/caregivers who wished to allow their children to participate in the study were required to provide their written/digital consent and return the completed forms to the primary researcher. Parents who did not wish for their child to participate in the study simply did not fill out the consent form. As part of the consent form, parents also had the opportunity to fill out an entry submission to be included in a draw for one of ten gift cards valued at $50 each. This study was approved by Memorial University's Interdisciplinary Committee on Ethics in Human Research.

Non-Autistic Group

The non-autistic group consisted of adolescents between the ages of 12.09 and 18.04 years (M_{age} = 14.51, SD = 1.68). This group included students from schools under the English School District (in the Avalon Region of Newfoundland) as well as individuals recruited through social media and snowball recruitment. Individuals whose parents did not disclose an autism diagnosis were assigned to this group.

Autistic Group

This group was comprised of individuals (M_{age} = 14.63, SD = 2.12, range = 11.81-18.97 years) who were identified (by their parents/caregiver) as having an autism diagnosis. These participants were recruited through provincial autism societies and autism service centres across Canada, as well as social media pages (e.g., parent autism Facebook groups), English School District schools (Avalon region), Canadian universities, community research groups (e.g., NL Centre for Applied Health Research), and snowball recruitment.

Eligibility Criteria

All individuals had to fall within a 12-month range of the 12- to -18-year age limit to participate in the study. This permitted those individuals who were no longer enrolled in the formal K-12 school system to be eligible to participate.

To limit IQ-achievement discrepancies, only those individuals of average intellectual ability (as assessed by the Raven's 2, Raven et al., 2018; see below for more details) were considered for this study. Additionally, all participants had to be currently attending, or had previously attended, regular classroom-based schooling at least the majority of the time (i.e., they could participate even if they received additional educational supports at school). These criteria were set to ensure that both groups were relatively similar in terms of their academic achievement levels.

Lastly, individuals (within both samples) who experienced severe math-based learning difficulties and/or were diagnosed with a math-based learning disability (i.e., dyscalculia) were not eligible to partake in the study. Fulfillment of all eligibility criteria were determined by parent responses to questions on the demographic questionnaire.

Recruitment

Elementary, Junior High, and High Schools. Students of grades 6 through 12 attending schools in the Avalon Region of Newfoundland, Canada were invited to participate in the study. All schools were under the Newfoundland and Labrador English School District; therefore, school board approval was sought prior to institutional recruitment. Principals of these schools were contacted by the primary researcher for permission to recruit students. Once principal approval was acquired, recruitment posters were distributed to students to be brought home to their parents. Recruitment posters asked if any parents/caregivers of individuals (within the 12-

to 18-year age range) were interested in having their child participate in a study of executive functioning and math ability, and included brief descriptions of the study's purpose, participant eligibility (without specification of autism), testing procedures, participant incentive, and researcher contact information. Parents who were interested in having their child participate in the study were to contact the primary researcher directly for further information (at which point they received the parental consent form and demographic questionnaire). Due to COVID-19 restrictions, all testing was scheduled to take place remotely and outside of school hours. The majority of participants recruited from schools constituted the non-autistic sample, though some autistic individuals were identified (and thus allocated to the autistic group).

Provincial Autism Societies and Service Centres. The second method of recruitment occurred via autism societies and autism services centres located across Canada. The societies/centres were contacted by the primary researcher informing staff of the study purpose and intent to recruit. Staff were asked to distribute recruitment posters either directly to families within their community via email or through the centre's social media and/or website. Recruitment posters were identical to those distributed to schools, with the exception of specified autism identification. Parents who were interested in the study needed to contact the primary researcher via email, who then forwarded along the consent form and demographic questionnaire to the inquiring parent. Once parental consent was received, an online testing session was scheduled. Only those clients of the Autism Society NL were offered the option of in-person testing; this occurred only for one participant testing session. Participants recruited from autism societies and centres constituted the autism sample, so long as they fulfilled the additional eligibility criteria.

Canadian Universities. Recruitment through Canadian universities followed an identical procedure to that of the autism societies and services centres.

Social media pages. Recruitment posters, both identifying autistic participant recruitment as well as general recruitment, were shared via personal social media accounts (e.g.., Facebook) of the primary researcher and persons within their social network. Additionally, posters were shared on autism specific pages, namely parent and family Facebook groups. Those parents interested in having their child participate in the study either emailed or messaged the primary researcher for further information.

Snowball recruitment. Approximately one week post-testing, each participant and/or parents of the participant would receive a follow-up Thank You email to show appreciation for their participation. In this email, participants were also asked if they would be willing to share a copy of the recruitment poster (general poster as was distributed to schools) to any families of autistic/non-autistic children who might be interested in participating in the study. As well, "word-of-mouth" sharing of study details was employed by lab members and other individuals to encourage recruitment of persons from within their network.

Materials

As COVID-19 had limited in-person testing, all measures either already existed in an online format or were adapted by the researcher for online distribution. The measures were divided into three separate sections, each of which had its own link to be shared by the researcher to the participant. The first section contained the Raven's 2 (Raven et al., 2018), which was made available via the publisher's online testing platform, *Q-global* (Pearson, 2021a). A unique link was created for each participant tested. The second link directed participants to a survey via the *Qualtrics* Platform (Qualtrics, 2022). The survey was counter-balanced and contained the two math measures. The final section included all four EF measures which were integrated, in a counter-balanced order, into a survey created using the software *PsyToolkit* (Stoet, 2010; Stoet, 2017). With the exception of the Backwards Letter Span task (which was created specifically for the present study) all EF measures belonged to the *PsyToolkit* website (Stoet, 2010; Stoet, 2017). The validity of PsyToolkit software in measuring response times has been demonstrated (Kim et al., 2019).

For remote testing, participants required access to their own personal computer, generally located in the participant's home. In the case of one participant who was tested in-person at an alternative location (i.e., Autism Society NL), the primary researcher provided their own personal computer for participant use. Specific tasks required the use of a keyboard, mouse/track pad, and working speakers/headphones. As well, the Raven's 2 is a timed measure, therefore, the researcher used a stopwatch during administration of this measure.

Lastly, Zoom virtual communications software (Zoom Video Communications, Inc., 2022) was utilized for online testing in order to connect participant and researcher during the testing session. The use of synchronous testing allowed the researcher to maintain some degree of control over the testing environment (e.g., if parents were present during testing, the researcher was able to remind parents not to provide aid to their children) and permitted participant questions (e.g., instruction clarification) to be addressed in real time. Additionally, links could be shared with participants via the Zoom chat function. Some participants opted to have their webcams turned on, though this was not a requirement. The majority of participants had working microphones; however, for those who did not, the chat function was employed to relay messages/questions to the researcher.

Measures

Prior to testing, parents/caregivers were required to complete an informed consent form. Parents and caregivers of participants were asked (and agreed) in the consent form not to provide assistance to their child during testing, making an informed judgement of their child's ability to read through the instructions without assistance.

During completion of the consent form, all parents were also asked to fill out a demographic questionnaire on behalf of their child.

Demographic Questionnaire

The demographic questionnaire consisted of both yes/no and multiple choice questions pertaining to the adolescent participant which addressed the following topics: name; age (birth month and year); sex (i.e., male, female, non-binary/third gender, or prefer not to say); ethnicity/racial identity; school attendance (i.e., "Does your child attend school?") and classroom engagement (i.e., "Does your child attend regular classroom-based schooling?"); grade level (selecting from grades 6 through 12, or graduated/university); requirement of special educational support (i.e., "Does your child receive support in the classroom (i.e., instructional resource teacher/special education)?") and type of support (i.e., literacy, numeracy, social/emotional, and/or other); diagnoses of a learning difficulty(s)/disability(s) and type (i.e., reading, math, writing, and/or other); diagnosis of an Autism Spectrum Disorder and type (i.e., ASD, "High-functioning" or Asperger Syndrome, Pervasive Developmental Disorder-Not Otherwise Specified, or other); and additional diagnoses (e.g., depression, ADHD, schizophrenia, Tourette syndrome, etc.). Parents were also asked to indicate whether or not they wished for their child to be entered into the gift card draw as an appreciation for their participation.

Intellectual Ability

Raven’s 2 Progressive Matrices Clinical Edition (Raven’s 2; Raven et al., 2018). The Raven’s 2 is a measure of general cognitive functioning; it is used to test *eductive* (“meaning-making”) ability. The digital short form of this measure (available through the *Q-global* platform, Pearson, 2021a) was used in the current study as a metric of nonverbal intellectual ability. This version consisted of 24 visual problems of increasing difficulty to be completed within 20 minutes.

The participant first worked through the demonstration and sample items, while also attending to researcher verbal instruction, before beginning the measure. If the participant erred on one of the sample items, expanded directions were to be provided to ensure participant comprehension (this did not occur for any participant). The researcher was required to manually time the testing using a stopwatch. If the participant finished all problems before the allotted time period was up, they simply closed the browser tab and waited for the researcher to administer the next link. In the case that a participant did not complete all puzzles within the given time period, they were directed by the researcher to stop solving the problems and to close the tab on their browser.

In each of the 24 items, participants had to determine which image or shape, among a selection, belongs to the pattern of images/shapes presented on the screen. Prior to starting the assessment, the researcher informed the participant that they were under a time limit and therefore should take their “best guess” if unsure of which image to select on any specific item. Items could not be skipped, nor could a response be changed once submitted. A discontinue rule exists whereby, starting at item 11, the test would end following six incorrect responses.

Each participant’s scores (raw scores, ability scores, and standard scores) are generated automatically as part of the *Q-global* scoring. Additionally, the reports include a Descriptive

Classification for each participant based on their standard scores. The classifications and associated standard scores, as outlined in the Raven's 2 manual (Pearson, 2018), are as follows: Extremely High (> = 130); Very High (120-129), High Average (110-119); Average (90-109); Low Average (80-89); Very Low (70-79); and Extremely Low (< = 69). Descriptive Classifications were used to eliminate any participants whose performance on the Raven's fell into the "Extremely Low" classification. For the purposes of this project, only the ability scores (which are not age-dependent but based on items sets B-E, which all participants encountered) were considered for the analysis. Standard scores were not used for the analysis as comparisons were to be drawn only between the two sample groups' scores, therefore comparison to the general population scores was not deemed relevant for the purposes of this study.

Individuals from ages 4 to 90 years 11 months can be tested with the Raven's 2 (Pearson, 2021b). The reliability rate for the digital short form assessment is .80 (Pearson, 2018).

Math Performance

Math fluency measure. Inspired by the Math Fluency subtest of the Wechsler Individual Achievement Test III (WIAT-III; Wechsler, 2009) and the Math Facts Fluency subtest of the Woodcock Johnson IV Tests of Achievement (Schrank et al., 2014), the math fluency measure used in the present study consisted of 160 simple math calculations of which participants needed to solve as many as possible within a 2-minute time limit. Participants were hinted to use the TAB button on their keyboard to move quickly through the calculation questions. Items included single-digit/single-digit and single-digit/double-digit addition, subtraction, and multiplication equations stacked vertically. The timer began when participants made their first click on the page displaying instructions and questions. The following page would automatically be displayed once the timer had run out. Use of a calculator was not permitted. A sum score of correct responses was generated for each participant.

The Math Fluency composite of the WIAT-III (Wechsler, 2009) is appropriate for testing grades 1 through 12 (Maccow, 2011). Exceptional reliability has been reported for all of the WIAT-III composite scores, ranging from .90 to .98 (McCrimmon & Climie, 2011).

Conceptual and Procedural Knowledge of Fractions Measure (Hallett et al., 2012). Borrowing items from the Chelsea Diagnostic Test on fractions (Brown et al., 1984), the Conceptual and Procedural Knowledge of Fractions Measure test (Hallett et al., 2012) was designed to assess both conceptual and procedural knowledge of fractions. The test contains 23 fraction-based questions which measure a broad range of math skills (e.g., computational, fractional, etc.). Eleven items test *conceptual* knowledge, being the understanding of relationships between informational pieces, while the remaining 12 items test one's *procedural* knowledge, or the understanding of steps towards achieving a correct response (Hallett et al., 2010) – in this case, to calculate a correct value. Questions included visual representation of fractions (e.g., item 1), arithmetic operations (e.g., item 2), magnitude comparisons (e.g., item 3) and ordering of fraction values (e.g., item 14), solving for a missing numerator/denominator (e.g., item 11), and word problems (e.g., item 22). This measure is not timed; however, it is estimated to take approximately 15-20 minutes to complete. Participants were advised to ask questions when uncertain and to skip any items with which they were not familiar. Use of a calculator was not permitted, though participants were informed that they could work through the questions by hand. Separate sum scores for procedural- and conceptual-based items were generated for each participant.

Students in grades 6 (Hallett et al., 2012), 7 (Eddy, 2020), and 8 (Hallett et al., 2012) have been tested with this measure.

Working Memory

Backward Corsi Span of the PsyToolkit Software (Stoet, 2010; Stoet, 2017). The Corsi test (Corsi, 1972) backwards span measures visual-spatial working memory capacity (Kessels et al., 2008). Participants view an array of boxes, some of which light up in sequence. On hearing "Go", participants are required to repeat back the sequence in *reverse* order by clicking on the boxes. Participants are presented with a smiling or frowning face image for each correct or incorrect response, respectively. If the sequence is correct, the following trial sequence increases by one block; sequences begin at two blocks and can increase to a maximum of nine blocks. After two incorrect reverse sequence responses, the task ends, and the last correct sequence span is recorded as the participant's Corsi backward span score. An average span for this task is approximately six blocks (PsyToolkit, 2021a). The task takes approximately one minute to complete.

Internal consistency for the traditional Corsi task (using blocks) has been reported as .78 in an older sample with neurocognitive disorders (de Paula et al., 2016). The Corsi test has also been applied in both child, adolescent (Isaacs & Vargha-Khadem, 1989) and young adult (Pagulayan et al., 2007) samples.

Backwards Letter Recall Task. This measure was designed specifically for this study. It combines elements of the backwards digit span task (Working Memory Test Battery for Children, WMTB-C; Pickering & Gathercole, 2001; Weschler Intelligence Scale for Children-Fifth Edition, WISC-V; Wechsler, 2014) and the Letter Memory task (Morris & Jones, 1990), both of which have been used to evaluate working memory capacity. The Letter Memory task in particular is considered useful for measuring updating (Agostino et al., 2010; Miyake et al., 2000; St. Clair-Thompson & Gathercole, 2006).

In the backwards digit span task (WMTB-C; Pickering & Gathercole, 2001; WISC-V; Wechsler, 2014), participants listen to a series of digits and are asked to recall them in reverse order. In adaptations of the Letter Memory task (Morris & Jones, 1990), a series of letters are presented serially (2000 ms per letter), of which participants are required to rehearse aloud, and then recall, the last three (Agostino et al., 2010) or four (Miyake et al., 2000; St. Clair-Thompson & Gathercole, 2006) letters of the list. The length of the letter sequences varies across a number of trials. Backwards digit recall tasks have been tested with children ages 6- to 16-years old (WISC-V; Wechsler, 2014), while the Letter Memory Task has been tested with children as young as 8-years of age (Agostino et al., 2010) and young adults (Friedman et al., 2006; Morris & Jones, 1990). Reliability estimates for backwards digit recall tasks have reached .71 for ages 9- to 11-years (WMTB-C; Pickering & Gathercole, 2001) and .80 (for the WISC-V; Wechsler, 2014; see Canivez & Watkins, 2016).

For the present study, a novel version of the backwards recall task was developed. As in the Letter Memory task (Morris & Jones, 1990), participants were presented visually with a series of letters at a rate of 2000 ms per letter. Following presentation, participants were required to type out the correct backwards order of all letters presented in the sequence, inputting one letter at a time and hitting the ENTER button between letters. Participants receive feedback (“Correct!” or “Wrong!”) depending on whether or not they input the correct sequence. Sequences increased in length by one letter with each correct response– starting at a span of two letters and increasing to a maximum span of nine. After two incorrect responses, the task ends and the last correct sequence span is recorded as the participant’s backwards letter span score.

This task takes approximately 5 minutes to complete, and participants were scored based on the accuracy of their recall (i.e., span score).

Shifting

Wisconsin Card Sorting Inspired Task (WCSIT) of the PsyToolkit Software (Stoet, 2010; Stoet, 2017). The WCSIT is a computerized measure of one's mental flexibility. Adapted from the original WCST (Grant & Berg, 1948), the WCSIT requires shifting from one rule of sorting to the next. Participants are presented with a card containing a number of coloured shapes which they need to match to one of four other cards based on a sorting rule (i.e., colour, shape, or number of shapes). The rule for sorting changes at random without notice, and participants are expected to adopt and apply the new sorting rule. Participants receive feedback ("GOOD!" or "WRONG!") depending on whether they applied the correct rule. The stimuli time out after ten seconds. The task contains 60 trials and requires approximately 2-3 minutes to complete. Participants were scored based on perseveration errors (i.e., continuing to apply the old rule) and non-perseveration errors (i.e., incorrect sorting selection) (PsyToolkit, 2021c).

Variations of the WCST–both as traditional in-person (Kopp et al., 2021) and self-administered digital (Steinke et al., 2021) forms–have achieved split-half reliability estimates above .90 for perseveration errors. The WCST has been tested with children (Ozonoff & Jensen, 1999; see Yeniad et al., 2013), adolescents (Lehto, 1996; Ozonoff & Jensen, 1999), and young adults (Miyake et al., 2000; Steinke et al., 2021).

Inhibition

Go/No-Go task of the PsyToolkit Software (Stoet, 2010; Stoet, 2017). The Go/No-Go task was administered as a measure of response inhibition. It requires participants to withhold responses when presented with a specific stimulus. When green "GO" stimuli appear, participants must press the SPACE BAR on their keyboard as quickly as possible; however, if red “NOGO” stimuli appear, participants are expected to do nothing (i.e., inhibit the impulse to respond). Stimuli are displayed for a maximum of 1000ms (i.e., timeout). There are 50 trials in total (40 go trials and 10 no-go trials) and takes approximately one minute to complete. Participants were scored based on response time and the number of errors (i.e., responses on the no-go trials) committed (PsyToolkit, 2021b). The Go/No-Go task has been tested previously with ages 8- to 16-year-olds (Happé et al, 2006).

Procedure

Once parental consent was obtained and the demographic questionnaire returned, parent involvement in the study was largely complete (aside from the possible need of receiving the Zoom link setting up the meeting for their child). Due to COVID-19 restrictions, all testing sessions (except one) were completed remotely (via Zoom meeting) and synchronously at a time which worked best for the participant’s schedule. Testing procedures were identical for the single case of in-person data collection with the exception that the researcher and participant were in the same room during testing. On the morning of the testing date, the researcher sent along a brief reminder of the testing session as well as the zoom link for the session. Participants (or parents of participants) were also be reminded to have a pencil and paper available for convenience of completing one of the math measures.

Participants were to complete the measures individually during their designated testing period, without the aid of a parent/caregiver. Parents and caregivers of participants were asked (and agreed) in the consent form not to provide assistance to their child during testing. Though most parents left their children alone to complete the study, those parents who chose to remain present during testing were reminded of this rule. A researcher (either the primary researcher, or a researcher of the RCDMC lab) was available during the entirety of all testing sessions to facilitate testing and address any participant questions or concerns.

The majority of sessions were 1-on-1 with a single participant and researcher; however, four pairs of siblings were tested in paired sessions (two participants and one researcher). Under these circumstances, a single Zoom meeting was set up, and each sibling completed the study from separate computers and separate rooms of the home. Interference was reduced as much as possible (no more than would occur within a single participant session); for example, both participants start the Raven's 2 assessment at the same time, and when one sibling finished the measure before the other, they would simply proceed to the next link and be provided instruction via chat.

At the beginning of each session, the researcher greeted the participant(s) and gave a brief description of what their participation in the study would involve (i.e., that they would be completing a few math-based measures along with some cognitive tasks/games). Participants were reminded that the testing was not meant to be stressful, that their performance was not a reflection of how smart they are and to simply try their best, and that they could ask questions or take a break at any point throughout the study. Once rapport was established, the researcher received verbal or written (i.e., typed in chat) assent to administer the first section of testing.

Link #1: Q-global (Raven's 2)

Using the Zoom chat function, the researcher pasted and sent the first link (containing the Raven's 2 assessment) to the participant, along with their unique session code (generated within *Q-global*). Participants were required to input their name and date of birth and to verify this information before beginning the assessment.

When a listening symbol (coloured background with a centred hexagon containing an ear-like silhouette) appeared on screen, the participant was not to proceed and to instead inform the researcher to read instructions. The first set of instructions indicated to participants that upon proceeding to the next page (by clicking the blue arrow) they would view the computer solving samples of visual problems, and to pay attention to how the computer solved the puzzles. They were then informed to click the arrow to continue. A video of three problems being solved by the computer were played in sequence; once completed, another listening symbol would appear, and the researcher read further instructions.

The second set of instructions informed the participant that they would now try some practice puzzles like those solved by the computer. Participants were instructed to select the piece that finished the picture, and to click the arrow to continue. As no participants failed the practice items, upon completion, another listening symbol appeared to initiate the final set of instructions.

The researcher assured the participants (as the instructions stated) that they now knew what they were doing and that more puzzles were to follow. Participants were given the same prompt of selecting the piece that finished the puzzle, and to click the arrow to continue to the next puzzle. Additionally, participants were informed to take their best guess when they were uncertain of the correct answer. Finally, before beginning, the researcher stated that they would have 20 minutes to complete all the puzzles, that the puzzles would get a bit trickier throughout the measure (this transparency was established to set participant expectation for the assessment in order to reduce distress), and that it was OK if they were unable to complete them all in the allotted time frame; just try their best. The researcher then did a countdown to synchronize the timer and the participant beginning the assessment. If at any point the participant appeared or sounded distressed (e.g., sighing) during the assessment, the research reminded them to just take their best guess. The timer was not displayed for participant viewing; though if asked, the researcher would inform the participant of how much time remained.

Once the Raven's assessment was complete (either because the participant finished within the timeframe or the 20-minute limit was up), participants were advised to close the browser tab for the assessment. Between each link, the researcher did a quick debrief with participants to gage their stress level; the researcher would ask something related to "how did you find that?" and, if participants seemed especially distressed, would provide verbal assurance (e.g., "It's impressive that you completed all those puzzles" or "The puzzles do get pretty tricky towards the end, but good job for sticking through it"). The researcher would also ask the participant if they required a quick break before proceeding to the next section.

Link #2: Qualtrics Survey (Math Fluency and Fraction Measures)

When the participant confirmed they were ready to continue, the researcher sent the second link directing the participant to the Qualtrics survey (containing the math fluency and fraction measures). The default font and font size were used throughout the survey, with the exception of text which were slightly enlarged for emphasis (e.g., math fluency measure heading). The researcher informed the participant that they would now be completing the math portion of the testing, which would involve two separate measures: one was a timed measure in which the participant would have 2 minutes to complete as many addition, subtraction, and multiplication questions within that timeframe (and would be automatically progressed to the next page once the 2-minutes were up), and the other being a series of fractional questions which would not be timed. For the latter assessment, the participant was assured that they could work through those items by hand if they wished and could skip any questions with which they were unfamiliar or uncertain. Additionally, the participant was reminded that they could ask any questions throughout the survey, as well as skip any questions which they could not solve.

At the start of the survey, a welcome screen appeared, followed by a prompt reminding participants that the survey is to be completed independently without aide of a parent and to read through the instructions carefully. It also noted that participants could quit the survey at any time without repercussion by simply closing the survey tab. Participants were then directed to press the arrow on the bottom of the survey to begin.

Before starting the measures, participants were required to fill in some brief demographic information that would be used to generate a unique seven-character long ID code (with assurance that none of the information provided would be shared). The ID code would be needed for later data analyses to link a participant's survey submission to their data from the PsyToolkit measures.

Progress through the survey could be tracked in real-time by the researcher. If at any point a participant appeared to be stuck on a single question for a considerable amount of time, or if a participant expressed any sign of distress throughout completion of the survey, the researcher would prompt the participant with a check-in (e.g., "everything still going okay?") to ensure the participant was alright and to remind the participant that the researcher was present to answer any questions.

The math fluency and fraction measures appeared in a randomized order. At the start of the math fluency measure, an instruction page appeared with the heading "*WELCOME TO THE MATH FACTS TASK.*" The instructions were as follows: "For this section, you will be asked to complete as many math facts as you can. You have 2 minutes. *The timer will start as soon as you click next*. Tip: Use the TAB key to move between questions. Please answer as many questions as possible to the best of your ability. **[DO NOT USE A CALCULATOR]**". Participants then needed to click the arrow on the bottom of the page to move onto the assessment. Though not visible to the participant, the timer would initiate countdown with the participant's first click. Items 1 through 59 were a blend of addition and subtraction questions. The first multiplication question was introduced at item 60, with the remainder of the questions being a mix of the three operation types. Following the 2-minute time limit, the participant was automatically progressed to the next page.

An instruction page also preceded the fractions assessment, which read the following: "In this section, you will be given math problems to solve. Do your best to solve each of the questions. If you wish to input a **fraction value**, use the **forward slash "/"** to indicate a fraction (example: **1/2**). Press the arrow to start.". Among the 23 items in the measure were a mix of text input and selection-based response types. Answers in mixed, whole, and reduced fractions, as well as decimal value form were permitted.

At the end of the survey, text appeared informing the participant that they would now be starting the next tasks, that they would need their ID code (shown on-screen), and to notify the researcher that they were ready to begin the next part. Clicking the arrow on the bottom of the screen would bring the participant to a text prompt thanking the participant for completing the survey and notifying them that their responses had been recorded. Again, the researcher would do a check-in to see how the participant was feeling and if they needed to take a quick break before proceeding.

Link #3: PsyToolkit Survey (Executive Function Measures)

The final link directed participants to the PsyToolkit survey which was comprised of the four EF tasks in a randomized order. The screen size was 800 x 600 pixels and centred for all tasks. The participant could toggle to full screen if so desired. The researcher explained to the participant that this final section contained four cognitive games, each of which had their own set of instructions. The researcher advised participants to read through the instructions carefully and to ask for clarification if needed before beginning the tasks.

The survey began with a welcome screen titled “Part of research study “Executive Function in adolescents”,” which detailed the following instructions as bulleted points: “Instructions will appear before each task”; “Read the instructions carefully”; “You will need a **keyboard** to complete two of the tasks”; “You will need **headphones/speakers** to complete another one of the tasks”. A button at the bottom of the screen was to be clicked when the participant was ready to continue. The following page asked the participant to input their ID code and to once again press the continue button, which thus began the task portion of the survey.

Each task was preceded by a screen displaying a centred red square which included text instructing the participant to “click to start”. The instructions for the Backwards Letter Span task (referenced as “Letter Memory Test” to participants) informed participants that they would be presented with a series of blue letters shown one at a time and to try their best to remember the order of the letters. When letters were done displaying, a yellow block would appear on screen, in which the participant would need to type in the letters in reverse sequence one at a time by typing in the most recent letter displayed, hitting the ENTER button, typing the next letter in sequence, and so on. The instructions stated that the participant would receive feedback with each trial, that the sequence would increase by one letter with each trial correct, and that they would not be timed (they were advised to take their time responding in order to be as accurate as possible). All instructions were accompanied by screenshots of the task and images of key input (i.e., letter, ENTER key, letter) for illustrative purposes. After two failed trials, or upon achieving the ceiling span score (i.e., 9 letters), a display indicating that the task had ended and directing participants to hit SPACE BAR to continue would appear on-screen.

For the Backward Corsi Span task, the instructions began with the heading "Memory Test Instructions" and contained bulleted points which read as follows: "You will need speakers/headphones for this task. Make sure the **sound** on your computer is **ON**."; "A number of boxes will appear on the screen. Some of the boxes will light up in sequence, one at a time."; "Try to remember the order in with the boxes light up."; "When the boxes are done lighting up, you will **hear** a voice say "**GO**". Repeat back the sequence in which the boxes lit up in **REVERSE** order by **CLICKING** on each of the boxes."; "Clicking on a box will cause a CHECK MARK to appear. You cannot undo a box once it has been clicked."; "When you think you have chosen the correct REVERSE order, click DONE (bottom right corner of screen)."; "A happy face will appear if the order was CORRECT, otherwise, a sad face will appear."; "Each trial, the sequences will increasingly get longer". A screenshot of the task was also included in the instructions for visual clarity. The participant then pressed the SPACE BAR when they were ready to begin the task. A 3-second visual countdown signified the commencing of the task. Once two subsequent trials were failed, or the participant achieved the ceiling span score (9 blocks), a screen appeared displaying the participant's Corsi backward span (i.e., "Your Corsi backward span is __ items"). Participants were also directed to press the SPACE BAR to continue to the next task.

The WCSIT instructions opened with a screen stating, "Wisconsin Card Sorting Test", the PsyToolkit logo, the website URL (www.psytookit.org) and instruction to press the SPACE BAR to start. The instructions informed participants that they would need to match a card to one of the four cards presented across the top of the screen by selecting one of the four, but that the rule for matching would not be given. Thus, participants would need to use trial and error to discover the rule, and that a new card would appear in the bottom left corner of the screen after each match. An example (with screenshot) was provided to explain the three ways of sorting (i.e., according to shape, number of shapes, or colour). Participants were also notified that they would receive feedback on each match: if they were correct, they would continue to use the same rule, otherwise, they must try sorting by a different rule. Lastly, the instructions explained that the rule for matching would change at some point throughout the trials without notice so participants should be careful to monitor their feedback. As well, participants were informed that there were 60 trials in total. After all 60 trials were complete, the participant was presented with a screen displaying their error count, number of perseverations errors, and number of non-perseveration errors committed, along with the percentage value calculated from the trial total. The screen also indicated to press the SPACE BAR to continue.

Lastly, for the Go-/No-Go task, the instructions (with accompanying screenshots of stimuli) read as follows: "When you see the <u>**Green GO**</u> signal, press the **SPACE BAR** <u>AS QUICKLY AS POSSIBLE</u>. **DO NOT** press the SPACE BAR if a <u>**Red NO-GO**</u> signal appears. Press the **SPACE BAR** to quit instructions and begin the test.". Once the task was complete, participants were moved on to the subsequent task.

Once all four tasks were completed, a final survey page appeared informing the participant that they could now close the tab in the browser. The researcher again checked-in with the participant and notified them that they had now completed the testing session. The participant was given the opportunity to ask any questions at this time. Additionally, the researcher read a debriefing script disclosing the intent of the study, describing the data de-identification process, and reminding the participant of their right to withdraw their data up to one week post-testing. Following any questions, the participant was thanked for their time, and the Zoom meeting was ended. Testing overall required approximately 50 minutes to complete.

Data Analysis

All data were anonymized and amalgamated, meaning no individual participant data was included in the analysis output. Reported statistics are based on group assignment (i.e., autistic vs. non-autistic) rather than the individual.

Data were compiled via RStudio, version 2021.9.0.351 (RStudio Team, 2021). Statistical analyses were conducted using Jamovi statistical analysis software, version 2.3.0.0 (The Jamovi Project, 2022). Before investigating whether math differences are explained by EF differences, it is useful to determine whether or not EF differences exist between the autistic and non-autistic group. To do so, Pearson's *r* partial correlations were conducted to determine if group differences were independently related to each of the EF task performance scores, while controlling for age, intellectual ability, and gender. A series of multiple regressions were then conducted to illustrate group differences on the math measures while controlling for those additional factors listed. Finally, for those differences which were significant, mediations were evaluated using the *medmod* module in Jamovi to test if any of the individual EF component task scores helped to explain these group differences in math performance, Alpha was set at .05 for all analyses.

Results

Demographics

Full demographic information can be found in Table 1**.** Twenty autistic and 34 non-autistic individuals participated in this study, for any overall sample size of 54 participants. One non-autistic participant's data were removed from the study for scoring "Extremely Low" on the

Raven’s 2 assessment, leaving a sample size of 53 (20 autistics and 33 non-autistics). The autistic group was largely male-identifying (n = 14), whereas the non-autistic group contained mostly females (n = 19). Two participants identified as non-binary/third gender.

In terms of supports, 10 autistic and 8 non-autistic participants were identified as requiring some form (e.g., writing, literacy, social/emotional, numeracy, and/or other) of additional support(s) at school. Five autistic and four non-autistic participants were indicated as having some type of learning difficulty (e.g., writing, math, reading, working memory/processing, and/or other). Lastly, it was identified that 11 autistic and 9 non-autistic participants had at least one of the following additional diagnoses: ADHD/ADD, anxiety, depression, Tourette syndrome, sensory processing disorder, general learning disorder, and/or developmental coordination disorder.

Table 1

Demographic Information of Participants, Separated by Group ($n_{autistic}$ = 20, $n_{non\text{-}autistic}$ = 33)

Characteristic	Autistic group	Non-Autistic group
Age (years):		
Mean (SD)	14.63 (2.12)	14.51 (1.68)
Range	11.81-18.97	12.09-18.04
Gender:		
Male	14	14
Female	4	19
Non-Binary/Third gender	2	0
Identity (Ethnic, Racial, etc.):		
White	15	28
Black	0	2
South-East Asian	0	1
Bi-Racial	0	1
Indigenous	2	0
(First Nations)	2	0
(Métis)	0	2
South Asian Indian	1	0
Canadian/Lebanese	1	0
Hindu	0	1

Demographic Information of Participants, Separated by Group (continued)

Grade Level:		
6th	3	2
7th	2	8
8th	5	4
9th	2	7
10th	4	8
11th	2	2
12th	0	2
Graduated	2	0
School Support?		
Yes	10	8
No	10	25
Type of Support:		
Literacy Support (i.e., reading and writing)	5	3
Numeracy Support (i.e., mathematics)	2	0
Social/Emotional Support	7	2
Other	3	4
Learning Difficulties/Disabilities?		
Yes	5	4
No	14	29
Type of Learning Difficulties/Disabilities:		
Dyslexia or difficulty reading	1	3
Dyscalculia or difficulty with math	2	0
Dysgraphia or difficulty with writing	2	1
Other	1	1
Autism Diagnosis?		
Yes	20	0
No	0	33
Additional Diagnosis?		
Yes	11	9
No	4	19
Type of Additional Diagnosis:		
Attention Deficit Disorder/Attention Deficit Hyperactivity Disorder	9	7
Obsessive Compulsive Disorder	0	0
Schizophrenia	0	0
Depression	0	1
Anxiety	5	2
Tourette Syndrome	1	0
Down Syndrome	0	0
Other	2	1

Partial Correlations: Determining EF Differences Between Groups

Performance scores and correlation values are summarized in Tables 2 and 3, respectively. As illustrated by Table 3, there were no significant differences in EF performance scores between the autistic and non-autistic groups while controlling for the additional factors.

Table 2

Group Means and SD on Measures of Intellectual Ability, Math Performance, and Executive Functioning Component Tasks

Measure	Mean (SD)	
	Autistic	Non-Autistic
Intellectual Ability: Raven's 2	529.90 (26.10)	530.36 (31.31)
Math Fluency	37.25 (15.46)	40.45 (11.27)
Fractional Conceptual	14.45 (5.63)	15.91 (5.26)
Fractional Procedural	4.80 (3.97)	6.85 (3.95)
WM: Corsi Backward Span	3.58 (2.76)	3.00 (2.74)
WM: Backward Letter Span	1.16 (2.01)	1.09 (2.05)
Shifting: WCSIT Perseveration Errors	10.56 (4.02)	10.34 (4.45)
Shifting: WCSIT Correct	43.11 (7.12)	43.78 (7.44)
Inhibition: Go/No-Go Errors	2.22 (2.07)	2.22 (1.41)
Inhibition: Go/No-Go RT	329.67 (36.20)	329.66 (48.45)

Table 3

Partial Correlations Between Group and Executive Functioning Component Task (while Controlling for Age, Intellectual Ability, and Gender)

EF: Performance	r_s	*p*-value
WM: Corsi Backward Span	-0.11	.462
WM: Backward Letter Span	-0.01	.966
Shifting: WCSIT Perseveration Errors	-0.03	.865
Shifting: WCSIT Correct	0.05	.729
Inhibition: Go/No-Go Errors	-0.01	.960
Inhibition: Go/No-Go RT	-0.02	.917

Despite the lack of variability in group scores, the regression analyses were still conducted to explore group differences in math performance.

Regression Analyses

Model 1: Group Differences in Math Fluency Performance

The results of the regression analyses are summarized in Table 4. To begin building the model, the grouping variable (autistics vs. non-autistics) was entered as a fixed factor, along with covariates age, intellectual ability, and gender to predict performance on math fluency questions. The autistic group served as the reference level.

Though the overall model was significant, $R^2 = .307$, $F(4, 48) = 5.31$, $p = .001$, the grouping variable itself was not a significant predictor of math fluency scores; therefore, no further analyses were conducted on this model.

Model 2: Group Differences in Fractional Conceptual Math Performance

This model followed the same build as Model 1, except with fractional conceptual performance as the dependent variable. As in the previous model, the overall model was significant, $R^2 = .382$, $F(4, 48) = 7.40$, $p < .001$, however, the grouping variable was not a significant predictor of fractional conceptual scores (see Table 4). No further analyses were conducted on this model.

Model 3: Group Differences in Fractional Procedural Math Performance

This final model build mirrored that of the previous models, but instead predicted fractional procedural performance. The overall model was significant, $R^2 = .458$, $F(4, 48) = 10.15$, $p < .00001$. The grouping variable was significantly predictive of fractional procedural scores over and above the variance accounted for by age, intellectual ability, and gender, indicating greater performance for the non-autistic group. A medium to large effect size was also found (see Table 4).

Table 4

Grouping Variable (Non-Autistic – Autistic) as a Predictor of Math Performance Scores, while Controlling for Age, Intellectual Ability, and Gender

Math Measure	β	t-value (df)	p-value	Cohen's d
Math Fluency	3.19	0.99 (48)	.328	0.284
Fractional Conceptual	1.50	1.18 (48)	.243	0.340
Fractional Procedural	2.16	2.43 (48)	.019*	0.697

Note. * $p = .05$

Each of the EF component task scores were added into Model 3 separately to evaluate their individual contributions to the model (see Table 5 for a summary).

Working Memory: Corsi Backward Span scores. The addition of Corsi span scores into the model as a separate block resulted in a nonsignificant change in variance accounted for, $\Delta R^2 = .039$, $F(1, 44) = 3.37$, $p = .073$. The overall model, however, maintained significance, $R^2 = .491$, $F(5, 44) = 8.49$, $p < .0001$. Additionally, the grouping variable remained a significant predictor despite adding the Corsi span scores into the model, indicating that the Corsi span scores were not able to account for the math performance difference between the two groups.

Working Memory: Backwards Letter Span scores. The addition of the Letter span scores into the model as a separate block resulted in a nonsignificant change in variance accounted for, $\Delta R^2 = .008$, $F(1, 46) = .657$, $p = .422$. The overall model, however, maintained significance, $R^2 = .462$, $F(5, 46) = 7.88$, $p < .0001$. The grouping variable remained a significant predictor despite adding the Letter span scores into the model, indicating that the Letter span scores were not able to account for the math performance difference between the two groups.

Shifting: WCSIT perseveration errors. The addition of the WCSIT perseveration errors into the model as a separate block resulted in a nonsignificant change in variance accounted for, $\Delta R^2 = .00009$, $F(1, 44) = .007$, $p = .933$. The overall model, however, maintained significance, $R^2 = .437$, $F(5, 44) = 6.83$, $p < .0001$. Additionally, the grouping variable as a predictor became marginally not significant when adding the perseveration errors into the model. A medium effect size was also found. For this mediation (i.e., perseveration errors mediating the group-procedural relationship), the indirect effect was not significant, $Z = .169$, $p = .866$, indicating that the change in the math performance group differences produced by adding perseveration errors into the model was not significant.

Shifting: WCSIT correct responses. The addition of the WCSIT correct responses into the model as a separate block resulted in a nonsignificant change in variance accounted for,

$\Delta R^2 = .004$, $F(1, 44) = .317$, $p = .576$. The overall model, however, maintained significance, $R^2 = .441$, $F(5, 44) = 6.94$, $p < .0001$. The grouping variable as a predictor became marginally not significant when adding the WCSIT correct responses into the model. A medium effect size was also found. For this mediation (i.e., correct responses mediating the group-procedural relationship), the indirect effect was not significant, $Z = .310$, $p = .756$, indicating that the change in the math performance group differences produced by adding WCSIT correct responses to the model was not significant.

Inhibition: Go/No-Go errors. The addition of the Go/No-Go errors into the model as a separate block resulted in a nonsignificant change in variance accounted for, $\Delta R^2 = .021$, $F(1, 44) = 1.69$, $p = .201$. The overall model, however, maintained significance, $R^2 = .458$, $F(5, 44) = 7.42$, $p < .0001$. The grouping variable as a predictor became marginally not significant when adding the Go/No-Go errors into the model. A medium effect size was also found. For this mediation (i.e., Go/No-Go errors mediating the group-procedural relationship), the indirect effect was not significant, $Z = .007$, $p = .994$, indicating that the change in the math performance group differences produced by adding Go/No-Go errors to the model was not significant.

Inhibition: Go/No-Go response times. The addition of the Go/No-Go response times into the model as a separate block resulted in a nonsignificant change in variance accounted for, $\Delta R^2 = .002$, $F(1, 44) = .190$, $p = .665$. The overall model, however, maintained significance, $R^2 = .439$, $F(5, 44) = 6.89$, $p < .0001$. The grouping variable as a predictor became marginally not significant when adding the Go/No-Go response times into the model. A medium effect size was also found. For this mediation (i.e., Go/No-Go response times mediating the group-procedural relationship), the indirect effect was not significant, $Z = .0008$, $p = .999$, indicating that the change in the math performance group differences produced by adding Go/No-Go response times to the model was not significant.

Table 5

Value of Grouping Variable in Predicting Fractional Procedural Math Scores, while Controlling for Age, Intellectual Ability, Gender, and Executive Functioning Component Task Performances

EF: Performance	*β*	*t*-value (df)	*p*-value	Cohen's *d*
WM: Corsi Backward Span	2.21	2.41 (44)	.020*	0.718
WM: Backward Letter Span	2.01	2.22 (46)	.031*	0.647
Shifting: WCSIT Perseveration Errors	1.85	1.96 (44)	.056[a]	0.585
Shifting: WCSIT Correct	1.83	1.94 (44)	.059[b]	0.579
Inhibition: Go/No-Go Errors	1.85	1.99 (44)	.052[c]	0.594
Inhibition: Go/No-Go RT	1.85	1.96 (44)	.056[d]	0.585

Note. * $p = .05$

[a,b,c,d]Mediation of EF performance was not significant

Power Analyses

Given the small sample size, I conducted a sensitivity Power Analysis using *G*Power* (version 3.1, Faul et al., 2009) to determine the size of effect that there would need to be in the population in order for us to have an 80 percent chance of finding statistical significance. With the present sample size and a total of five predictors– three control variables (intellectual ability, gender, age), one fixed variable (group) and one EF predictor of interest– a medium effect size ($f^2 = .154$) would be needed to detect an effect 80% of the time.

Discussion

The purpose of the present study was to expand on preliminary research (i.e., Oswald et al., 2016; St. John et al., 2018; May et al., 2013) looking at if – and how – EF influences math performance in an autistic sample, as it has been shown to within the general population (e.g., see Cragg & Gilmore, 2014). Each of the three primary EF components– working memory, shifting, and inhibition– were assessed independently to gage each component's unique contribution to autistic math performance on fractional (procedural and conceptual) and fluency measures. The results indicated that none of the EF components were able to account for the apparent group differences in procedural math scores. Mixed findings on the predictive nature of EF on mathematics performance in autistic samples have been reported previously in the literature (i.e., May et al., 2013; Oswald et al., 2016; St. John et al., 2018). Additionally, there were no significant differences in EF task performance found between the autistic and non-autistic groups. This finding was largely unexpected based on previous research which has suggested EF challenges associated with autism (e.g., Barnard et al., 2008; Mayes & Calhoun, 2008).

Executive Function, Math Performance, and Autism

One possible explanation for these mixed findings is that EF might not exert the same influence on autistic math ability as it does in the general population; however, more robust research is required to determine if this is truly the case. As a fuller consensus of autism as a distinct neurotype – rather than a disorder – is reached within the scientific community, recognition of variability in individuals' cognitive reliance should follow suit. Instead of focussing on the so-called deficits associated with autism through the viewpoint of a neurotypical lens, the scientific community will come to acknowledge differences between these neurotypes as merely diverse means of thought. That is to say, how one neurotype derives meaning from their external surroundings may differ from that of another neurotype. For example, one could argue that the math fluency measure, with items which vary in operational form (i.e., addition, subtraction, and multiplication) would require an astute shifting ability in order to solve items both quickly and accurately. It is possible, however, that recognition of the operational symbols themselves is a more useful cognitive skill; keen perception for localized details has been noted previously in autistic samples (e.g., see Happé & Frith, 2006). Therefore, cognitive dependency may vary with neurodiversity.

Additionally, the comparable math abilities between the two samples of this study speaks to the heterogeneity of autistic mathematical achievement. Previous research has documented the widespread math achievement of autistic samples (e.g., Chiang & Lin, 2007; Estes et al., 2011; Jones et al., 2009). Though other results have highlighted the higher rate of mathematics learning disability in autistic individuals (Oswald et al., 2016), only two participants in the present autistic sample were indicated as having math-based learning challenges (as reported in demographics). On a larger scale, these results may also speak to the variability in autistic abilities in general. Autism is not a one-size-fits-all diagnosis; it is as complex as the individuals themselves. The challenges associated with being autistic are described by Baron-Cohen (2017) as just that – associations of the neurotype and not definers of the neurotype. He shares, in respect to criticisms of neurodiversity, "[o]thers may say that a child who has language delay or severe learning difficulties is not an example of neurodiversity but has a disorder, and I would support their demand for treatments to maximize the child's potential in both language and learning. But again, although commonly occurring, these [language delay or learning difficulties] are not autism itself" (p. 746). Just as not all autistic individuals demonstrate a superior math ability, it is insufficient to explain any difficulties in mathematics as being the result of their neurology. It can be concluded, thus, that the present sample is simply one of autistics who are (overall) of average mathematical capabilities.

Only one group difference in math scores appeared across the three forms of math questions used in this study. More specifically, there were no significant differences in scores on the math fluency measure and the fractional conceptual questions between the autistic and non-autistic groups. This aligns with previous findings which support comparable math performances between these two groups (e.g., Chiang & Lin, 2007). It is interesting that the significant difference in group scores were on the fractional procedural questions (with non-autistics outperforming autistic participants), yet no such difference was present on the math fluency questions. The procedures used to solve fraction calculations often depend on the fluent use of the math facts that are assessed in the fluency measure (especially the multiplication questions). This means that differences in math fact knowledge do not explain why the autistic group does not perform as well on these questions, so it may be some other aspect of executing procedures that may pose a relative difficulty. It may have to do with knowing how and when to use this math fact knowledge. Whereas the math fluency measure uses whole numbers, the fraction measure items were solely fractional format, which is arguably more abstract and challenging to compute (e.g., in addition, knowing that numerators combine with numerators, but denominators are not summed). Additionally, the rules for computation with fractions are more complex than those of whole number-based arithmetic and require manipulation of the individual fractional values. For example, to add two whole number values together, one simply needs a knowledge of sums and such a calculation can be completed in a minimum of one step. In order to add fractions together, however, one must recognize that a common denominator is needed, and to achieve this the individual must multiply at least one of the fraction values by some factor to generate like denominator values. Then, one would need to sum only the numerators together to merge two separate fractions into one fraction value. A subsequent step involves reducing the fraction to its smallest term which, though not required for the fraction measured used in this study, was undertaken by several participants in the present sample. Differences in group scores, therefore, may have been more evident when solving these more challenging computations.

Group Differences in Executive Functioning

More surprising than the lack of evidence for EF as influencing autistic math performance was the fact that there were no differences in EF performance scores between the two groups. Countless research studies have reported on the EF challenges of autistic individuals, specifically in terms of their working memory system (see Hill, 2004). Though autistic EF abilities can vary with IQ (see Hill, 2004) and age (Chen et al., 2016; Happé et al., 2006; van den Bergh et al., 2014), even those individuals of higher IQ have demonstrated WM challenges (Mayes & Calhoun, 2008).

Research has linked difficulties with EF to characteristics of autism, namely repetitive behaviours (Mosconi et al., 2009; Yerys et al., 2009). There is a possibility that individuals who display more “recognizable” or “stereotypical” autistic behaviours are viewed as having greater needs and thus are allocated different supports in school, such as being removed from the mainstream classroom setting and placed into a specialized learning environment. If this were to be the case, then our autistic sample–who were considered eligible for the study on the condition that they attend regular classroom learning most of the time–may not possess the EF difficulties as have been noted for autistic individuals who demonstrate more stereotypical diagnostic traits.

No demographic information on participants' specific autistic behaviours or traits were collected for this project; therefore, this explanation is merely speculative.

An alternative explanation (and one to which I am more partial), is that the lack of evidence for autistic EF challenges in the present study merely speaks to the variability in autism presentation. Though, again, the particulars of the present sample's diagnoses were not explored, previous research has reported on differences in how individuals demonstrate autistic traits (Wing & Gould, 1979). It is possible, therefore, that the sample of participants tested in the present study did not display the same degree of EF challenges as has been demonstrated by autistic samples of past studies (see Hill, 2004).

Limitations

The present study is not without shortcomings. To start, the two sample sizes are quite small, with only 33 non-autistic and 20 autistic participants. One of the main concerns with a sample size this small is a lack of power. Given that with a power of 80 percent our sample size would only be able to detect a medium effect size, we run the risk of failing to detect a small or even medium to small effect in the population (i.e., committing a Type II error). This limits the representativeness of the sample results. Additionally, though participants were recruited from provinces across Canada, much of the non-autistic sample were recruited from schools within Newfoundland.

Another limitation of this study is that participation occurred outside school hours on weekday evening or weekend. The voluntary nature of such participation might have resulted in a sample who were highly motivated and/or who felt confident in their mathematical ability. On some occasions, participants had expressed even a liking for math, thus potentially skewing the results. Had the COVID-19 pandemic not limited testing to online participation, the researchers would likely have been able to test in-person at participating schools (as done with previous research studies). Under those circumstances, it is expected that there would have been both a greater number of student participants and as a sample more varied in their math abilities.

Furthermore, online testing is limiting in itself in that the researchers were not able to control the external environment and circumstances in which participants were tested. This issue was somewhat mitigated with participants who used their video feature during the Zoom meeting; however, webcams were not required. Additionally, the researchers were not able to view participants' screens (screensharing often interfered with Zoom meeting connection), which created the added challenge of troubleshooting any technical issues.

Lastly, autistic diagnosis was solely based on parent reporting and not supplemented using medical documentation or an autistic screening assessment. This is limiting in two ways. The first is the accuracy of parent accounts and lack of self-disclosure of diagnosis. Secondly, it is possible that individuals who were allocated to the non-autistic sample may have autistic traits and were undiagnosed. Some autistic (and other neurodiverse and neurotypical) individuals are accustomed to camouflaging or *masking* their traits to pass as more neurotypical, which, though potentially damaging to one's sense of identity (e.g., Hull et al., 2017, Miller et al., 2021), can lead to traits going unrecognized. This may also help to explain the larger proportion of males to females in our autistic group, as females engage in masking their autistic traits more than their male counterparts (McQuaid et al., 2022).

Future Directions

If and when in-person testing is permitted, similar research to the present study should be conducted to build the scientific understanding of autistic mathematical performance. Additionally, future studies should be mindful of confirming autism diagnosis in their participants, either with medical documentation or autism screening assessment. Math measures which are both leveled to an age/grade-expected difficulty, and that assess a broader scope of mathematics constructs (e.g., algebra, word problems), should also be utilized. Lastly, given the results of the present study – that EF does not appear to account for autistic math performance – other potential influencing variables should be considered.

Conclusion

Autism is a complex diagnosis with its own associated challenges and benefits, the variability of which is partially the reason I (and presumably other researchers) have chosen to study it. Understanding of autism is ever-changing and developing, and thus how we research and discuss autism must change accordingly. The overarching goal of such research, however, should always be to benefit the community. Obtaining a greater understanding of what variables negatively impact academic performance in the autistic population allows for the opportunity to implement interventions which can help mitigate such issues. Though comparisons between autistic and the non-autistic populations may provide certain insights, considering the diversity of autism diagnoses, autistic-based research may not always be a cut-and-dried procedure with wholly conclusive results. While the present research did not necessarily produce any extraordinary results with regards to promoting autistic academic success, the findings do provide substantive grounds for further evaluating the factors which influence autistic math ability.

References

Agostino, A., Johnson, J., & Pascual-Leone, J. (2010). Executive functions underlying multiplicative reasoning: Problem type matters. *Journal of Experimental Child Psychology, 105,* 286-305. doi: 10.1016/j.jecp.2009.09.006

Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review, 16,* 17-42. doi: 10.1007/s11065-006-9002-x

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Ancker, J. S., & Kaufman, D. (2007). Rethinking health numeracy: A multidisciplinary literature review. *Journal of the American Medical Informatics Association, 14*(6), 713-721. doi: 10.1197/jamia.m2464

Andersson, U. (2007). The contribution of working memory to children’s mathematical word problem solving. *Applied Cognitive Psychology, 21*, 1201-1216. doi: 10.1002/acp.1317

Andersson, U. (2008). Working memory as a predictor of written arithmetical skills in children: The importance of central executive functions. *British Journal of Educational Psychology, 78,* 181-203. doi: 10.1348/000709907X209854

Autism Speaks. (2020). *What is autism?* https://www.autismspeaks.ca/about/about-autism/#:~:text=The%202018%20National%20Autism%20Spectrum,based%20autism%20services%20%E2%80%93%20each%20year.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *Psychology of Learning and Motivation, 8,* 47-89. doi: 10.1016/S0079-7421(08)60452-1

Baez, S., Flichtentrei, D., Prats, M., Mastandueno, R., García, A. M., Cetkovich, M., Ibáñez, A., & Pavlova, M. A. (2017). Men, women…who cares? A population-based study on sex differences and gender roles in empathy and moral cognition. *PLoS One, 12*(6), e0179336. doi: 10.1371/journal.pone.0179336 J

Baird, G., Charman, T., Baron-Cohen, S., Cox, A., Swettenham, J., Wheelwright, S., & Drew, A. (2000). A screening instrument for autism at 18 months of age: A 6 year follow-up study. *Journal of the American Academy for Child and Adolescent Psychiatry, 39,* 694-702. doi: 10.1097/00004583-200006000-00007

Banks, J., & Oldfield, Z. (2007). Understanding pensions: Cognitive function, numerical ability and retirement saving. *Fiscal Studies, 28*(2), 143-170. doi: 10.1111/j.1475-5890.2007.00052.x

Barnard, L., Muldoon, K., Hasan, R., O'Brien, G., & Stewart, M. (2008). Profiling executive dysfunction in adults with autism and comorbid learning disability. *Autism, 12*(2), 125-141. doi: 10.1177/1362361307088486

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind.* MIT Press.

Baron-Cohen, S. (2000a). Is Asperger syndrome/higher-functioning autism necessarily a disability? *Development and Psychopathology, 12,* 489-500. doi: 10.1017/S09554579400003126

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences, 6*(6), 248-254. doi: 10.1016/s1364-6613(02)-01904-6

Baron-Cohen, S. (2017). Editorial perspective: Neurodiversity- a revolutionary concept for autism and psychiatry. *Journal of Child Psychology and Psychiatry, 56*(6), 744-747. doi: 10.1111/jcpp.12703

Baron-Cohen, S., & Bolton, P. (1993). *Autism: The facts.* Oxford University Press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition, 21,* 37-46. https://doi.org/10.1016/0010-0277(85)90022-8

Baron-Cohen, S., Wheelwright, S., Burtenshaw, A., & Hobson, E. (2007). Mathematical talent is linked to autism. *Human Nature, 18,* 125-131. doi: 10.1007/s12110-007-9014-0

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31,* 5–17. doi:10.1023/A:1005653411471

Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review, 29,* 180-200. doi: 10.1016/j.dr.2009.05.002

Borys, S. V., Spitz, H. H., & Dorans, B. A. (1982). Tower of Hanoi performance of retarded young adults and nonretarded children as a function of solution length and goal state. *Journal of Experimental Child Psychology, 33,* 87-110. doi: 10.1016/0022-0965(82)90008-X

Bottema-Beutel, K., Kapp, S. K., Lester, J. N., Sasson, N. J., & Hand, B. N. (2021). Avoiding ableist language: Suggestions for autism researchers. *Autism in Adulthood, 3*(1), 18-29. doi: 10.1089/aut.2020.0014

Brian, J. A., Tipper, S. P., Weaver, B., & Bryson, S. E. (2003). Inhibitory mechanisms in autism spectrum disorders: Typical selective inhibition of location versus facilitated perceptual processing. *Journal of Child Psychology and Psychiatry, 44*(4), 552-560. doi: 10.1111/1469-7610.00144

Brown, M., Hart, K., & Küchemann, D. (1984). *Chelsea diagnostic mathematics tests.* NFER-Nelson.

Bull, R., Johnston, R. S., & Roy, J. A. (1999). Exploring the roles of the visual-spatial sketch pad and central executive in children's arithmetical skills: Views from cognition and developmental neuropsychology. *Developmental Neuropsychology, 15,* 421-442. doi: 10.1080/87565649909540759

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology, 19*(3), 273-293. doi: 10.1207/s15326942dn1903_3

Canivez, G. L., & Watkins, M. W. (2016). Review of the Wechsler intelligence scale for children- fifth edition: Critique, commentary, and independent analyses. In A. S. Kaufman, S. E. Raiford, & D. L. Coalson (Eds.), *Intelligence testing with the WISC-V* (pp. 683-702). Wiley.

Chakrabarti, S., & Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *Journal of the American Medical Association, 285*, 3093-3099. doi: 10.1001/jama.285.24.3093

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

Chen, X., & Weko, T. (2009). *Stats in brief: Students who study science, technology, engineering, and mathematics (STEM) in postsecondary education*. DC: National Centre for Education Statistics. Retrieved from https://nces.ed.gov/pubs2009/2009161.pdf

Chiang, H.-M., & Lin, Y.-H. (2007). Mathematical ability of students with Asperger syndrome and high-functioning autism: A review of literature. *Autism, 11*, 547-556. doi: 10.1177/1362361307083259

Corsi, P. M. (1972). Human memory and the medial temporal region of the brain. *Dissertation Abstracts International, 34*, 819B.

Cragg, L. & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education, 3,* 63-68. doi: 10.1016/j.tine.2013.12.001

Cragg, L., Keeble, S., Richardson, S., Roome, H. E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. *Cognition, 162,* 12-26. https://doi.org/10.1016/j.cognition.2017.01.014

Cui, J., Gao, D., Chen, Y., Zou, X., & Wang, Y. (2010). Working memory in early-school-age children with Asperger's syndrome. *Journal of Autism and Developmental Disorders, 40*(8), 958-967. doi: 10.1007/s10803-010-0943-9

David, C. V. (2012). Working memory deficits in math learning difficulties: A meta-analysis. *International Journal of Developmental Disabilities, 58*(2), 67-84. doi: 10.1179/2047387711Y.0000000007

Delaney, J. M., & Devereux, P. J. (2020). Math matters! The importance of mathematical and verbal skills for degree performance. *Economics Letters, 186,* 108850. https://doi.org/10.1016/j.econlet.2019.108850

de Paula, J. J., Malloy-Diniz, L. F., & Romano-Silva, M. A. (2016). Reliability of working memory assessment in neurocognitive disorders: A study of the digit span and Corsi block-tapping tasks. *Revista Brasileira de Psiquiatria, 38*(3), 262-263. doi: 10.1590/1516-4446-2015-1879

Eddy, N. (2020). *The relation between math experience and math anxiety in grade 7 students* [unpublished undergraduate honours thesis]. Memorial University of Newfoundland.

Elsabbagh, M., Divan, G., Koh, Y. J., Kim, Y. S., Kauchali, S., Marcín, C., Montiel-Nava, C., Patel, V., Paula, C. S., Wang, C., Yasamy, M. T., & Fombonne, E. (2012). Global prevalence of autism and other pervasive developmental disorders. *Autism Research, 5*(3), 160–179. doi: 10.1002/aur.239

Eskes, G. A., Bryson, S. E., & McCormick, T. A. (1990). Comprehension of concrete and abstract words in autistic children. *Journal of Autism and Developmental Disorders, 20,* 61-73. doi: 10.1007/BF02206857

Estes, A., Rivera, V., Bryan, M., Cali, P., & Dawson, G. (2011). Discrepancies between academic achievement and intellectual ability in higher-functioning school-aged children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 41,* 1044-1052. doi: 10.1007/s10803-010-1127-3

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Feng, A., & Graetz, G. (2017). A question of degree: The effects of degree class on labor market outcomes. *Economics of Education Review, 61,* 140-161. doi: 10.1016/j.econedurev.2017.07.003

Fountas, I. C., & Pinnell, G. S. (1996). *Guided reading: Good first teaching for all children*. Heinemann.

Fountas & Pinnell Literacy. (2012). *F&P text level gradient™ & guided reading levels, a continuum of progress for readers.* https://www.fountasandpinnell.com/textlevelgradient/

Friedman, N., Miyake, A., Corley, R. P., Young, S., E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17*(2), 172-179. doi: 10.1111/j.1467-9280.2006.01681.x

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47*(6), 1539-1552. doi: 10.1037/a0025510

Gernsbacher, M. A. (2017). Editorial Perspective: The use of person-first language in scholarly writing may accentuate stigma. *Journal of Child Psychology and Psychiatry, 58*(7), 859-861. doi: 10.1111/jcpp.12706

Geurts, H. M., Corbett, B., & Solomon M. (2009). The paradox of cognitive flexibility in autism. *Trends in Cognitive Sciences, 13*(2), 74-82. doi: 10.1016/j.tics.2008.11.006

Goldberg, M. C., Lasker, A. G., Zee, D. S., Garth, E., Tien, A., & Landa, R. J. (2002). Deficits in the initiation of eye movements in the absence of a visual target in adolescents with high functioning autism. *Neuropsychologia, 40*(12), 2039-2049. doi: 10.1016/S0028-3932(02)00059-3

González-Garrido, A. A., Ruiz-Sandoval, J. L., Gómez-Velázquez, F. R., Oropeza de Alba, J. L., & Villaseñor-Cabrera, T. (2002). Hypercalculia in savant syndrome: Central executive failure? *Archives of Medical Research, 33,* 586-589. doi: 10.1016/S0188-4409(02)00404-6

Government of Canada. (2013). *Provincial definitions of a minor.* Canada. https://www.canada.ca/en/immigration-refugees-citizenship/corporate/publications-manuals/operational-bulletins-manuals/refugee-protection/canada/processing-provincial-definitions-minor.html

Government of Canada. (2018). *Autism spectrum disorder (ASD) and co-occurring health conditions.* Canada. https://www.canada.ca/en/public-health/services/diseases/autism-spectrum-disorder-asd/health-effects-conditions-associated-with-autism-spectrum-disorder-asd.html

Government of Canada. (2020). *Infographic: Autism spectrum disorder- Highlights from the Canadian survey of disability.* Canada. https://www.canada.ca/en/public-health/services/publications/diseases-conditions/infographic-autism-spectrum-disorder-highlights-canadian-survey-disability.html

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology, 38*(4), 404-411. doi: 10.1037/h0059831

Hallett, D., Nunes, T., & Bryant, P. (2010). Individual differences in conceptual and procedural knowledge when learning fractions. *Journal of Educational Psychology, 102*, 395-406.

Hallett, D., Nunes, T., Bryant, P., Thorpe, C. M. (2012). Individual differences in conceptual and procedural fraction understanding: The role of abilities and school experience. *Journal of Experimental Child Psychology, 113,* 469-486. doi: 10.1016.j.jecp.2012.07.009

Happe, F. G. E. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development, 66,* 843-855. doi: 10.2307/1131954

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

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36,* 5-25. doi: 10.1007/s10803-005-0039-0

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

Honda, H., Shimizu, Y., Imai, M., & Nitto, Y. (2005). Cumulative incidence of childhood autism: A total population study of better accuracy and precision. *Developmental Medicine and Child Neurology, 47*, 10-18. doi: 10.1017/S0012162205000034

Hughes, J. E. A., Ward, J., Gruffydd., E., Baron-Cohen, S., Smith, P., Allison, C., & Simner, J. (2018). Savant syndrome has a distinct psychological profile in autism. *Molecular Autism, 9*, 1-18. doi: 10.1186/s13229-018-0237-1

Hull, L., Petrides, K. V., Allison, C., Smith, P., Baron-Cohen, S., Lai, M-C., & Mandy, W. (2017). "Putting on my best normal": Social camouflaging in adults with autism spectrum conditions. *Journal of Autism and Developmental Disorders, 47*(8), 2519-2534. doi: 10.1007/s10803-017-31660-5

Isaacs, E. B., & Vargha-Khadem, F. (1989). Differential course of development of spatial and verbal memory span: A normative study. *British Journal of Developmental Psychology, 7*, 377-380. doi: 10.1111/j.2044-835X.1989.tb00814.x

Jarvis, H. L., & Gathercole, S. E. (2003). Verbal and nonverbal working memory and achievements on national curriculum tests at 11 and 14 years of age. *Educational and Child Psychology, 20,* 123-140.

Jones, C. R. G., Happé, F., Golden, H., Marsdeon, A. J. S., Tregay, J., Simonoff, E., Pickles, A., Baird, G., & Charman, T. (2009). Reading and arithmetic in adolescents with autism spectrum disorders: Peaks and dips in attainment. *Neuropsychology, 23*(6), 718-728. doi: 10.1037/a0016360

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism, 20*(4), 442-462. doi: 10.1177/1363261315588200

Kessels, R. P. C., van den Berg, E., Ruis, C., & Brands, A. M. A. (2008). The backward span of the Corsi block-tapping task and its association with the WAIS-III digit span. *Assessment, 15*(4), 426-434. doi: 10.1177/1073191108315611

Kim, J., Gabriel, U., & Gygax, P. (2019). Testing the effectiveness of the internet-based instrument PsyToolkit: A comparison between web-based (PsyToolkit) and lab-based (E-Prime 3.0) measurements of response choice and response time in a complex psycholinguistic task. *PLoS ONE, 14*(9), p.e0221802. doi: 10.1371/journal.pone.0221802

Kopp, B., Lange, F., & Steinke, A. (2021). The reliability of the Wisconsin card sorting test in clinical practice. *Assessment, 28,* 248-263. doi: 10.1177/1073191119866257

Krajewski, K., & Schneider, W. (2009). Exploring the impact of phonological awareness, visual-spatial working memory, and preschool quantity-number competencies on mathematics achievement in elementary school: Findings from a 3-year longitudinal study. *Journal of Experimental Child Psychology, 103,* 516-531. doi: 10.1016/j.j-ecp.2009.03.009

Kroesbergen, E. H., van Luit, J. E. H., van Lieshout, E. C. D. M., van Loosbroek, E., & van de Rijt, B. A. M. (2009). Individual differences in early numeracy: The role of executive functions and subitizing. *Journal of Psychoeducational Assessment, 27*(3), 226-236. doi: 10.1177/0734282908330586

LeFevre, J.-A., Berrigan, L., Vendetti, C., Kamawar, D., Bisanz, J., Skwarchuk, S.-L., & Smith-Chant, B. L. (2013). The role of executive attention in the acquisition of mathematical skills for children in grades 2 through 4. *Journal of Experimental Child Psychology, 114*, 243-261. https://doi.org/10.1016/j.jecp.2012.10.005

Lehto, J. (1996). Are executive function tests dependent on working memory capacity? *The Quarterly Journal of Experimental Psychology, 49,* 29-50. doi: 10.1080/713755616

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

Maccow, G. (2011, September 8). *Administration, scoring, and interpretation of Wechsler individual achievement test-III* [PowerPoint slides]. Pearson. https://images.pearsonassessments.com/Images/PDF/Webinar/WIAT-III_Administration_and_Scoring_Handout_September_2011.pdf

Maccow, G. (2018, February 20). *Raven’s 2: Assessing general ability nonverbally (CAN)* [Webinar]. Pearson. https://www.brainshark.com/1/player/pearsonassessments?pi=zG0zY1NtDzC5WMz0&r3f1=350f71222e237b6939731072782e352034656b09663a373e2825776849663771&fb=0

May, T., Rinehart, N., Wilding, J., & Cornish, K. (2013). The role of attention in the academic attainment of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 43*(9), 2147-2158. doi: 10.1007/s10803-013-1766-2

Mayes, S. D., & Calhoun, S. L. (2003). Analysis of WISC-III, Stanford-Binet: IV, and academic achievement test scores in children with autism. *Journal of Autism and Developmental Disorders, 33,* 329-341. doi: 10.1023/A:1024462719081

Mayes, S. D., & Calhoun, S. L. (2008). WISC-IV and WIAT-II profiles in children with high-functioning autism. *Journal of Autism and Developmental Disorders, 38,* 428-439. doi: 10.1007/s10803-007-0410-4

Mayes, S. D., Calhoun, S. L., Bixler, E. O., & Zimmerman, D. N. (2009). IQ and neuropsychological predictors of academic achievement. *Learning and Individual Differences, 19*(2), 238-241. doi:10.1016/j.lindif.2008.09.001

McCoy, S., & Byrne, D. (2017). Student retention in higher education. In J. Cullinan & D. Flannery (Eds.), *Economic insights on higher education policy in Ireland: Evidence from a public system* (pp. 111-141). Springer. doi: 10.1007/978-3-319-485530-9_5

McCrimmon, A. W., & Climie, E. A. (2011). Test review: D. Wechsler individual achievement test- third edition. *Canadian Journal of School Psychology, 26*(2), 148-156. doi: 10.1177/0829573511406643

McQuaid, G. A., Raitano Lee, N., & Wallace, G. L. (2022). Camouflaging in autism spectrum disorder: Examining the roles of sex, gender identity, and diagnostic timing. *Autism, 26*(2), 552-559. doi: 10.1177/13623613211042131

Miller, S. E., & Marcovitch, S. (2011). Toddlers benefit from labeling on an executive function search task. *Journal of Experimental Child Psychology, 108*(3), 580-592. doi: 10.1016/j.jecp.2010.10.008

Miller, D., Rees., J., & Pearson, A. (2021). “Masking is life”: Experiences of masing in autistic and nonautistic adults. *Autism in Adulthood, 3*(4), 330-338. doi: 10.1089/aut.2020.0083

Milton, D. E. M. (2012). On the ontological status of autism: The ‘double empathy problem’. *Disability and Society, 27*(6), 883-887. doi: 10.1080/09687599.2012.710008

Minshew, N. J., Goldstein, G., Muenz, L. R., & Payton, J. (1992). Neuropsychological functioning in nonmentally retarded autistic individuals. *Journal of Clinical and Experimental Neuropsychology, 14*(5), 749-761. doi: 10.1080/01688639208402860

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49-100. doi: 10.1006/cogp.1999.0734

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81*, 111–121. doi: 10.1111/j.2044-8295.1990.tb02349.x

Mosconi, M. W., Kay, M., D’Cruz, A.-M., Seidenfeld, A., Guter, S., Stanford, L. D., & Sweeney, J. A. (2009). Impaired inhibitory control is associated with higher-order repetitive behaviors in autism spectrum disorders. *Psychological Medicine, 39,* 1559-1566. doi: 10.1017/S0033291708004984

Nicolaidis, C. (2012). What can physicians learn from the neurodiversity movement? *Virtual Mentor: American Medical Association Journal of Ethics, 14*(6), 503-510. doi: 10.1001/virtualmentor.2012.14.6.oped1-1206

Oswald, T. M., Beck, J. S., Iosif, A., McCauley, J. B., Gilhooly, L. J., Matter, J. C., & Solomon, M. (2016). Clinical and cognitive characteristics associated with mathematics problem solving in adolescents with autism spectrum disorder. *Autism Research, 9*, 480-490. doi: 10.1002/aur.1524

Ozonoff, S., & Jensen, J. (1999). Brief report: Specific executive function profiles in three neurodevelopmental disorders. *Journal of Autism and Developmental Disorders, 29*(2), 171-177. doi: 10.1023/A:10230052913110

Ozonoff, S., & Strayer, D. L. (1997). Inhibitory function in nonretarded children with autism. *Journal of Autism and Developmental Disorders, 27,* 59-77. doi: 10.1023/A:1025821222046

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal Child Psychology, 32*(7), 1081-1105. doi: 10.1111/j.1469-7610.1991.tb00351.x

Pagulayan, K. F., Busch, R. M., Medina K. L., Bartok, J. A., & Krikorian, R. (2007). Developmental normative data for the Corsi block-tapping task. *Journal of Clinical and Experimental Neuropsychology, 28*(6), 1043-1052. doi: 10.1080/13803390500350977

Passolunghi, M. C., & Pazzaglia, F. (2005). A comparison of updating processes in children good or poor in arithmetic word problem-solving. *Learning and Individual Differences, 15,* 257-269. doi: 10.1016/j.lindif.2005.03.001

Pearson. (2018). *Raven's 2 progressive matrices clinical edition*. NCS Pearson, Inc. https://qglobal.pearsonclinical.com/qg/viewRestrictedPdfS.seam

Pearson. (2021a). *Q-global* [Online testing platform]. Pearson. https://www.pearsonclinical.ca/en/q-global/main.html

Pearson. (2021b, March 19). *Raven's 2 - Raven's Progressive Matrices Second Edition*. Pearson Clinical. https://www.pearsonclinical.ca/store/caassessments/en/raven%27s/Raven%27s-Progressive-Matrices-%7C-Clinical-Edition/p/P100008213.html

Pickering, S. J., & Gathercole, S. E. (2001). *Working memory test battery for children.* Psychological Corporation.

PsyToolkit. (2021a, March 2). *Backward Corsi task.* PsyToolkit. https://www.psytoolkit.org/experiment-library/backward_corsi.html

PsyToolkit. (2021b, March 2). *Go/no-go task.* PsyToolkit. https://www.psytoolkit.org/experiment-library/go-no-go.html

PsyToolkit. (2021c, March 17). *Wisconsin card sorting inspired task (WCST).* PsyToolkit. https://www.psytoolkit.org/experiment-library/wcst.html

Qualtrics. (2022). *Qualtrics survey* (Version October 2021- July 2022) [Computer software]. Qualtrics. https://www.qualtrics.com/

Raven, J., Rust, J., Chan, F., & Zhou, X. (2018). *Raven's 2 progressive matrices, clinical edition (Raven's 2).* Pearson.

Rivera-Batiz, F. (1992). Quantitative literacy and the likelihood of employment among young adults in the United States. *Journal of Human Resources, 27,* 313-328. doi: 10.2307/145737

Robertson, S. M. (2010). Neurodiversity, quality of life, and autistic adults: Shifting research and professional focuses onto real-life challenges. *Disability Studies Quarterly, 30*(1). doi: 10.18061/dsq.v30i1.1069

RStudio Team (2021). *RStudio: Integrated Development Environment for R.* (Version 2021.9.0.351) [Computer Software]. RStudio, PBC. Retrieved from http://www.rstudio.com/

Rubia, K., Smith, A. B., Woolley, J., Nosarti, C., Heyman, I., Taylor, E., & Brammer, M. (2006). Progressive increase of frontostriatal brain activation from childhood to adulthood during event-related tasks of cognitive control. *Human Brain Mapping, 27,* 973-994. doi: 10.1002/hbm.20237

Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnson IV Tests of Achievement*. Riverside.

Siegel, D. J., Goldstein, G., & Minshew, N. J. (1996). Designing instruction for the high-functioning autistic individual. *Journal of Developmental and Physical Disabilities, 8*, 1-19. doi: 10.1007/BF02578436

Sparrow, E. P., & Hunter, S. J. (2012). *Executive function and dysfunction: Identification, assessment and treatment.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511977954

St. Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology, 59*(4), 745-759. doi: 10.1080/17470210500162854

Steinke, A., Kopp, B., & Lange, F. (2021). The Wisconsin card sorting test: Split-half reliability estimates for a self-administered computerized variant. *Brain Sciences, 11*(5), 529. doi: 10.3390/brainsci110550529

St. John, T., Dawson, G., & Estes, A. (2018). Brief report: Executive function as a predictor of academic achievement in school-aged children with ASD. *Journal of Autism and Developmental Disorders, 48,* 276-283. doi: 10.1007/s10803-017-3298-9

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42*(4), 1096-1104. doi: 10/3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology, 44*, 24-31. doi: 10.1177/00986283166777643

Stroop, J. R. (1992). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 121,* 15-23. doi: 10.1037//009603445.121.1.15

Swanson, H. L. & Jerman, O. (2006). Math disabilities: A selective meta-analysis of the literature. *Review of Educational Research, 76*(2), 249-274. doi: 10.3102/00346543076002249

The Jamovi Project (2022). *Jamovi* (Version 2.3) [Computer Software]. Jamovi. Retrieved from https://www.jamovi.org.

Van den Bergh, S. F. W. M., Scheeren, A. M., Begger, S., Koot, H. M., & Guerts, H. M. (2014). Age related differences of executive functioning problems in everyday life of children and adolescents in the autism spectrum. *Journal of Autism and Developmental Disorders, 44,* 1959-1971. doi: 10.1007/s10803-014-2071-4

Wechsler, D. (2009). *Wechsler individual achievement test (3rd ed.).* Psychological Corporation.

Wechsler, D. (2014). *Technical and interpretive manual for the Wechsler intelligence scale for children- fifth edition.* Pearson.

Wei, X., Christiano, E. R. A., Yu, J. W., Wagner, M., & Spiker, D. (2015). Reading and math achievement profiles and longitudinal growth trajectories of children with an autism spectrum disorder. *Autism, 19*(2), 200-210. doi: 10.1177/1362361313516549

Wei, X., Lenz, K. B., & Blackorby, J. (2013). Math growth trajectories of students with disabilities: Disability category, gender, racial, and socioeconomic status differences from ages 7 to 17. *Remedial and Special Education, 34*, 154-165. doi: 10.1177/0741932512448253

White, S. W., Ollendick, T. H., & Bray, B. C. (2011). College students on the autism spectrum: Prevalence and associated problems. *Autism, 15*(6), 683-701. doi: 10.1177/136236131039

Williams, D. L., Goldstein, G., Carpenter, P. A., & Minshew, N. J. (2005). Verbal and spatial working memory in autism. *Journal of Autism and Developmental Disorders, 35*(6), 747-756. doi: 10.1007/s10803-005-00210x

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders, 9*, 11-29. doi: 10.1007/bf01531288

World Health Organization. (2019). *Autism spectrum disorders.* https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders#:~:text=Epidemiology,figures%20that%20are%20substantially%20higher

Yeniad, N., Malda, M., Mesman, J., van Ijzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences, 23,* 1-9. doi: 10.1016/j.lindif.2012.10.004

Yerys, B. E., Wallace, G. L., Harrison, B., Celano, M. J., Giedd, J. N., & Kenworthy, L. E. (2009). Set-shifting in children with autism spectrum disorders: Reversal shifting deficits on the intradimensional/extradimensional shift test correlate with repetitive behaviors. *Autism, 13*(5), 523-538. doi: 10.1177/1362361309335716

Zoom Video Communications, Inc. (2022). *Zoom meetings* (Version 5.10.4) [Computer software]. Zoom. https://zoom.us/