The Use of a Newly Developed Computer Game to Measure Executive Functioning in Young Neurotypical Children and Children with a Diagnosis of an Autism Spectrum Condition

Amber Rose Scott

A thesis submitted in partial fulfilment of the requirements of the University of East London for the degree of Professional Doctorate in Clinical Psychology

May 2022

ABSTRACT

Over the past decade, executive functions have been associated with positive outcomes in children, highlighting the importance of having suitable measures available for this age group. Now, research is increasingly looking do develop computerised measures using game-based formats to address the limitations of established measures currently used. This study aimed to trial Davis' (2020) newly developed game-based measure of executive function, Dragon Adventure, in populations that have previously been neglected in the literature: neurotypical children aged 6 to 8 years and children with a diagnosis of an Autism Spectrum Condition aged 6 to 11 years.

Using a cross-sectional correlational design, this study compared the performance of participants on Dragon Adventure to established measures of executive function and teacher-ratings, controlling for computer literacy and processing speed. Within-subjects means comparisons were used to assess whether Dragon Adventure was rated as more acceptable than the established measures. Qualitative data on how Dragon Adventure could be improved was also collected and analysed through a content analysis.

Through Spearman's rank correlations, Dragon Adventure was found to be a valid measure of inhibition in both the neurotypical and ASC sample and of working memory in the neurotypical sample. Dragon Adventure also demonstrated good ecological validity. Dragon Adventure was rated by participants as acceptable, however, not substantially more enjoyable than established measures.

With some amendments addressing the limitations identified in this study, Dragon Adventure has the potential to be a suitable measure of executive function for school-aged children. Future research should look to develop Dragon Adventure and continue to trial it in neurotypical, neurodiverse and clinical samples.

ACKNOWLEDGEMENTS

Firstly, I would like to thank Jenny Davis for the time and creativity she put into making Dragon Adventure and allowing me to continue with her invaluable research. I feel privileged to have been able to contribute to the ongoing development of Dragon Adventure.

I would like to say a huge thank you to all the children who participated in this research. Thank you for taking the time out of your own study and engaging so enthusiastically. Your suggestions were inspiring and will be so helpful in the future development of Dragon Adventure. I had great fun working with you and I hope this research goes some way in giving back to you all.

Thank you to the two schools for hosting me, and the staff members who made this all possible during uncertain times. Thank you for allowing me to recruit from your schools, supporting with the organisation and for being so accommodating during the data collection, it is greatly appreciated.

Thank you to my supervisor, Dr Matthew Jones-Chesters, for your continued support and input throughout the whole process. I really valued your feedback, insights and responsiveness. I always came out of our meetings feeling calm and contained.

Thank you to my wonderful family and friends for being there for me over the past three years. I am so grateful to have such supportive and loving people around me. Thank you for providing me with those moments of relief and fun, which were very much needed. I would not be in the position I am today without the continued care and love from my family, thank you for all the practical and emotional support you have provided throughout the process. Thank you to my partner, for all your words and actions of support and encouragement, which have kept me going, and for keeping me balanced.

I would not have been able to do this without any of you!

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

LIST OF ABBREVIATIONS

ADHD	Attention-Deficit Hyperactivity Disorder
ASC	Autism Spectrum Condition
ATEC	Active Test of Embodied Cognition
BADS-C	Behavioural Assessment of the Dysexecutive Syndrome in Children
BRIEF	Behaviour Rating Inventory of Executive Function
CANTAB	Cambridge Neuropsychological Test Automated Battery
CHEXI	Childhood Executive Function Inventory
CTUQ	Children's Technology Use Questionnaire
DCCST	Dimensional Change Card Sort Test
D-KEFS	Delis-Kaplan Executive Function System
EDA	Exploratory Data Analysis
EYT	Early Years Toolbox
LMIC(s)	Low- and Middle-Income Country(s)
NEPSY-II	Developmental NEuroPSYchological Assessment-II
NTCB	NIH Toolbox for Assessment of Neurological and Behavioural Function Cognitive Battery
PEBL	The Psychology Experiment Building Language
PFC	Prefrontal Cortex
RACER	Rapid Assessment of Cognitive and Emotional Regulation
TMT	Trail Making Test
ToL	Tower of London
ToM	Theory of Mind
WCST	Wisconsin Card Sorting Test
WM	Working Memory
WMTB-C	Working Memory Test Battery for Children
WNV-SS	Wechsler Nonverbal Scale of Ability Spatial Span

1. INTRODUCTION

This research concerns performance-based measures, for children, of what has come to be termed 'executive function', and their validity with those diagnosed with an 'Autism Spectrum Condition' (ASC). There is growing research and development in the area of executive function measurement, however, the literature has identified certain age groups who have been neglected. The current study aims to address this gap through furthering recent research using a game-based measure of executive function in both a neurotypical and ASC population.

The term ‘neurotypical’ refers to children who are considered to be ‘typically developing’, in that they reach developmental milestones in an order and at a time that has been defined as the norm within Western culture. The term neurotypical is becoming more widely used as a result of the shifting views towards ‘neurodiversity’, and away from disorders and deficits. In line with this, ‘ASCs’ is an alternative term to ‘Autism Spectrum Disorder’, a classification defined by Western diagnostic criteria. The Diagnostic and Statistical Manual, Fifth Edition specifies that a person must have persistent deficits in social communication and interaction, and restrictive, repetitive patterns of behaviour to receive a diagnosis (American Psychiatric Association, 2013).

Both ‘neurotypical’ and ‘ASC’ are broad heterogeneous terms, which have developed from the Western medicalisation of human behaviour. The language used in the current literature base is adopted throughout this thesis; it is acknowledged that many of these concepts are socially contextualised ways of understanding, due to the absence of universal definitions. Using such terms can limit generalisability to other cultures and this should be taken into consideration when making interpretations of the current study.

This chapter will introduce executive function and the conceptual framework most widely used to understand it. The scope will then narrow to focus on executive function in childhood and the application of the conceptual framework to this population. Executive dysfunction will be introduced, with a further focus on the executive function profile of those with an ASC. The introduction of these topics will lead on to the literature review, which will summarise the research on measures of executive function for children and their use with children with an ASC. Literature on established measures of executive function will be summarised, followed by the challenges and limitations of using these measures and how these are exacerbated when used with children with an ASC. The focus will then turn to recent developments of measures, including the use of computerised tasks and game-based paradigms, and the benefits these developments may have on the populations being addressed. The chapter will conclude with the rationale for the current study, as derived from the literature, followed by the research aims, questions and hypotheses.

1.1. Executive Function

Executive function is an umbrella term used to describe higher-order cognitive processes, necessary for cognitive control in demanding tasks that involve novelty and goal-directed behaviour (Henry & Bettenay, 2010; Johann & Karbach, 2018; McCoy, 2019; Weintraub, Dikmen, et al., 2013). Executive function has been demonstrated to play a crucial role in everyday tasks, such as decision making, planning and evaluating, emotion regulation and social interaction (Diamond, 2013; Hofmann et al., 2012; Jurado & Rosselli, 2007).

Executive function has been associated with the prefrontal cortex (PFC), through research examining people with frontal lobe injuries. Frontal lobe injuries have been associated with difficulties such as attention, self-control, planning, reasoning, and problem-solving (Jurado & Rosselli, 2007; Miyake et al., 2000), all of which have been defined in the literature as executive functions. Executive functioning has also been found to be impacted by other areas of the cortex, such as posterior regions, due to the vast connections it has with other areas of the brain, along with being associated with lower-order functions (P. Anderson, 2002; Jurado & Rosselli, 2007).

Several models and frameworks have been developed as a way of understanding executive function. The majority of theories have argued and evidenced executive function as either a unitary construct (Baddeley, 1986, 1992; Dempster, 1992; Norman & Shallice, 1986; Shallice, 1988), or as a componential construct (Carlson & Moses, 2001; Diamond, 1991; Pennington, 1997; Welsh et al., 1991). The theory of executive function as a unitary construct has been supported by research finding intercorrelations between measures of executive function (Friedman & Miyake, 2004; Lehto et al., 2003; Miyake et al., 2000). Whereas, the componential view of executive function has been supported through research finding executive function processes to have different developmental trajectories (Archibald & Kerns, 1999; Carlson, 2005; Klenberg et al., 2001; Luciana & Nelson, 1998, 2002; Rosso et al., 2004; Welsh et al., 1991). One of the most influential models of executive function brought the two broad theories together through an integrative framework, the three-factor model (Miyake et al., 2000).

1.1.1. Miyake's Three-Factor Model

Since 2000, Miyake et al.'s three-factor model has been considered to be the most frequently evaluated model of executive function in the published literature and has shaped much of the understanding around executive function today (Karr et al., 2018). Based on a Confirmatory Factor Analysis in university students, Miyake et al. (2000) theorised that there are three key "separable but moderately correlated" components of executive function: inhibition, working memory and flexibility" (p. 87).

Inhibition refers to the ability to deliberately supress a well-learned or highly motivated prepotent response (Diamond, 2006; Józsa et al., 2017; Mischel & Ebbesen, 1970; Miyake et al., 2000; Miyake & Friedman, 2012). Working memory, sometimes referred to as 'updating', refers to the ability to hold in mind and manipulate information which is no longer present in the environment for a short period of time, while completing a task and/or problem solving (Józsa et al., 2017; Miyake & Friedman, 2012). Flexibility, also referred to as 'set-shifting' or 'switching', is the ability to rapidly change or adapt mental set towards a goal or action, responding to changing contextual demands in a responsive manner (Henry & Bettenay, 2010; Józsa et al., 2017; Miyake & Friedman, 2012). Flexibility has been deemed essential for adapting to a changing environment (Diamond, 2013).

Miyake et al.'s (2000) model of executive function has received much support (Duan et al., 2010; Fisk & Sharp, 2004; Vaughan & Giovanello, 2008). Research demonstrating modest correlations between different executive function tasks measuring inhibition, working memory and flexibility provides evidence of the theory that executive functions are made up of distinct but related cognitive functions (Gruber & Goschke, 2004). On the other hand, research challenging this model may be missing, due to conceptual replication failures rarely being published (Karr et al., 2018; Makel et al., 2012).

Miyake et al. (2000) did note that inhibition, working memory and flexibility were not the only components of executive function, rather the most common components in the literature. Other discrete sub-skills of executive function have been identified and empirically supported, including planning, concept formation, self-monitoring and fluency (Pennington & Ozonoff, 1996). However, reviews addressing our understanding of the full range of executive functions are limited as a result of methodological differences across studies (Jurado & Rosselli, 2007). Contributing to the broad nature of executive function, definitions have not just related to the concepts and process it encompasses, but also to the type of situations requiring executive function abilities.

1.1.2. "Hot" and "Cool" Executive Functions

Zelazo & Müller (2002) differentiated between "hot" and "cool" executive functions, based on the observation that executive function abilities vary depending on the motivational significance of a situation. Hot versus cool executive functions refer to the application of executive function skills in "motivationally and emotionally significant high-stakes situations", or in the absence of these demands, respectively (Zelazo & Carlson, 2012, p. 355). Hot and cool executive functions have been considered to be associated with different parts of the PFC, hot executive functions have been associated with ventral and medial regions, and cool executive functions with dorsolateral regions (Zelazo & Müller, 2002).

1.2. Executive Function in Childhood

Previously, it was believed that executive functions did not emerge until early adulthood, based on views that the frontal lobes do not develop and become integrated with other lobes until around this time (Culbertson & Zillmer, 1998; Dennis, 1991). Therefore, despite executive functions being extensively studied in adults, there has been less research on the development of executive function in children (Lan et al., 2011).

Recent research has demonstrated executive function development happening throughout childhood, in line with the development of the PFC (P. Anderson, 2002; Best & Miller, 2010; Roberts Jr & Pennington, 1996; Zelazo et al., 2008). It has been suggested that executive functions have three distinct growth periods: birth to two years; six to nine years; and adolescence to early twenties (V. Anderson, 1998, 2001; Hudspeth & Pribram, 1990; Romine & Reynolds, 2005). V. Anderson (1998, 2001) summarised that executive function skills have been evidenced in children aged 6 years, when “developmentally appropriate assessment tools are employed” (p. 119).

Research findings have demonstrated the importance of executive function processes in development (P. Anderson, 2002; Best & Miller, 2010). As a result, much focus has been placed on childhood executive function in both clinical and research settings due to the association with academic achievement (Allan et al., 2015; Jacob & Parkinson, 2015; St Clair-Thompson & Gathercole, 2006), school readiness (Blair, 2002), wellbeing (Pauli-Pott & Becker, 2011; Schoemaker et al., 2013) and physical health (Reinert et al., 2013).

1.2.1. Miyake’s Three-Factor Model in Childhood

Research has generally supported Miyake et al.'s (2000) tripartite model of executive function in children from the age of 6 years, with the three subcomponents being less distinct in early childhood (P. Anderson, 2002; Garon et al., 2008; Lehto et al., 2003; Wiebe et al., 2008, 2011). Although, van der Sluis et al. (2007) found that switching and working memory, however not inhibition, were distinguishable in children aged 9 to 12 years after controlling for non-executive variance.

In everyday life, inhibition helps children achieve goals through preventing them from attending to irrelevant information and exhibiting an undesired dominant/automatic response (Józsa et al., 2017). As children get older, they more accurately and quickly inhibit the prepotent response (Lagattuta et al., 2011). Flexibility helps children to problem solve through following changing or conflicting rules (Józsa et al., 2017). Working memory helps children to retain information, as their short-term retention span has been found to double in capacity between the ages of 5 and 10 years (Riggs et al., 2003).

Initially, distinctions were also made been made between hot and cool executive functions in children (Zelazo & Müller, 2002). However, it has since been argued that "current research does not make a compelling case for the separability of the hot and cool forms of executive function in childhood", especially through performance-based measures (Welsh & Peterson, 2014, p. 153).

1.3. Executive Dysfunction

Executive dysfunction refers to "deficits in the ability to inhibit well-learned patterns of behaviour, derive new ways of solving problems", and adapt behaviour to novel situations (Henry & Bettenay, 2010, p. 3). Executive dysfunction can present in a variety of ways: for example, an inability to avoid distractions and control impulsive responses or behaviour; this would be related to difficulties with inhibition (Diamond et al., 2002; Hofmann et al., 2012). Difficulty with flexibility can present itself as repeated or prolonged action and an inability to switch between tasks (Crone et al., 2004). Problems with working memory can present as difficulties following instructions and weighing-up information for decision making and learning (Diamond, 2013). These difficulties can extend into a child's everyday functioning, impacting both home life and academic achievement through problems with social adaptation, the organisation of daily activities, planning, initiating and completing tasks and emotional control (Hofmann et al., 2012; Stern & Prohaska, 1996).

Executive dysfunction is typically seen in those who have experienced brain injury to or neurodegeneration of the frontal lobes. However, due to the complex neural networks associated with executive function, it is not limited to these populations (Stuss & Alexander, 2000). It has been suggested that particular executive dysfunction profiles are associated with different developmental conditions (Henry & Bettenay, 2010) and other conditions, such as epilepsy (Høie et al., 2006). This research will focus on children diagnosed with an ASC.

1.4. Executive Function and Autism Spectrum Conditions

The executive function profile of children with an ASC is a contested topic (Hill, 2004). Historically, research has consistently reported children with an ASC to present with deficits in their ability to understand and reflect on others mental states and the links between mental state and action, referred to as Theory of Mind (ToM; Kimhi, 2014). Executive function and ToM have been linked in both neurotypical and ASC populations (Kimhi et al., 2014). In a longitudinal study, Pellicano (2010) found executive function to be predictive of ToM performance in children with an ASC, however, these results were limited by the presence of only two time points and an unrepresentative high-functioning sample. Much of the research investigating the relationship between executive function and ToM is limited by the heterogeneity between the executive function measures used and the absence of comprehensive task batteries.

Executive dysfunction was initially suggested as a theory of ASCs due to similarities identified between those diagnosed with an ASC and those with frontal lesions (Damasio & Maurer, 1978). Some have argued that it is executive dysfunction that gives rise to ASC presentations (Griffith et al., 1999; Russell, 1997), whereas others have argued that executive dysfunction exacerbates ASC presentations (Geurts et al., 2014; Russell, 1997). Characteristics of ASCs associated with executive dysfunction include high levels of impulsivity, the need for sameness, a strong liking for repetitive behaviours, difficulties in initiating new non-routine actions, difficulty in switching between tasks, a lack of self-regulation and self-monitoring, and inflexibility in thought and action (Robinson et al., 2009). ASCs are often diagnosed on the basis of these characteristics (Wing et al., 2011).

Associations have been evidenced between ASC presentations and executive dysfunction (Chan et al., 2009; Czermainski et al., 2014; Landa & Goldberg, 2005; Robinson et al., 2009). However, there is much heterogeneity among this population and there has been no consensus in the research around which executive function components are impaired in ASC presentations (Czermainski et al., 2015). Some researchers have argued that there is a general executive dysfunction in children with an ASC (De Vries et al., 2015; Demetriou et al., 2018; Kenworthy et al., 2008; Vanegas & Davidson, 2015), others have identified specific executive function deficits in this population. Ozonoff & Jensen (1999) viewed cold executive dysfunction as the primary deficit experienced by children with an ASC. Whereas others have argued that executive dysfunction only becomes apparent when situational demands require coordination of executive function abilities; which would be described as hot executive function (Gardiner et al., 2017). It has also been argued that ASCs are only associated with working memory and flexibility, due to deficits in inhibition being less pronounced (Christ et al., 2007; Hill, 2004; Liss et al., 2001; Mackinlay et al., 2006; Ozonoff & Strayer, 2001).

Supporting the argument that executive function abilities are present from the age of 6 years and are impaired in children with an ASC, research has noted increased difficulty differentiating children with an ASC from neurotypical children, in relation to executive function abilities, in preschool-aged children compared to school-aged children (Dawson et al., 2002; Gardiner et al., 2017; Yerys et al., 2007). On the other hand, some researchers have not found any differences between children with an ASC and neurotypical children on performance-based measures of executive function (Gómez-Pérez et al., 2016).

These inconsistencies have been attributed to the heterogeneity within the measures that are used to assess executive function. Therefore, the apparent executive function profile of an ASC population can vary depending on assessment methodology.

1.5. Literature Review

1.5.1. Rationale

Historically, there has been a lack of executive function measures, with most cognitive assessments failing to assess this construct, and those that did often differed in the processes and skills they measured (V. Anderson, 1998, 2001). The previously held belief that executive functions did not develop until early adulthood resulted in few executive function measures for use with children; even as measures of executive function for adults became established. Our understanding of executive function development in childhood and in those who are most commonly exposed to measures of executive function, including those with an ASC, has been limited as a result. The majority of measures that do exist are not suitable across childhood, with a focus on adolescence and preschool-aged children, with younger school-aged children being neglected (Best & Miller, 2010; Henry & Bettenay, 2010).

Acknowledging the importance of executive function processes in childhood development and their associations with positive outcomes, there is a need for appropriate measures of executive function for children, especially children aged 6 to 8 years. Having suitable measures for this age group is particularly important, given that 6 to 8 years has been considered to be a distinct executive function growth period, and the greatest period of development by some (V. Anderson, 1998, 2001; Hudspeth & Pribram, 1990; Romine & Reynolds, 2005). To gain a better understanding of the executive function measures available for children and the associated challenges, a review of the current literature concerning assessment of executive function in young school-aged children and children with an ASC was conducted.

1.5.2. Method

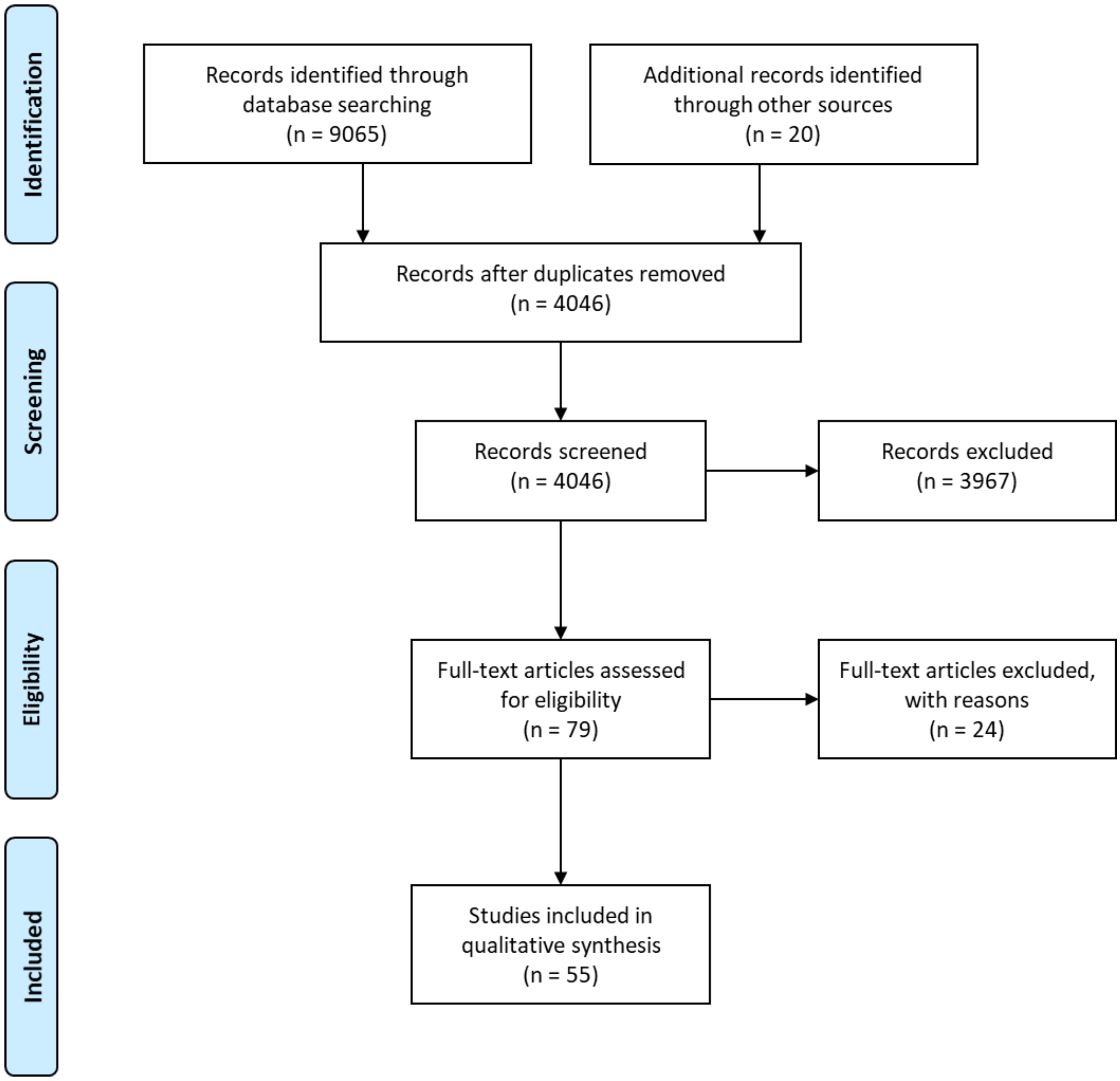
Literature searches were completed using the databases Academic Search Complete, Child Development & Adolescent Studies, PsychINFO, and SCOPUS. Combinations of the following terms were used to search through titles, keywords, and abstracts: “Execution Function*”, child*, (measure* OR assess* OR test). An additional search was completed using the various combinations, plus the search terms (Autism OR ASC OR ASD). From the initial searches, it was observed that game-like formats of executive function measures are being increasingly developed and recommended. Therefore, the term “game” was included in addition to the initial search terms.

1.5.3. Results

The initial searches produced a total of 9065 results, plus 20 additional articles were identified through other sources. These were then filtered to only include articles from academic journals and dissertations in the English language, looking at a school-aged population, and duplicates were removed, which identified 4046 results. These articles were then screened by manually going through the titles, and abstracts where appropriate, to assess whether the article was eligible, along with ensuring the article was accessible. Articles were required to investigate performance-based measures of executive function for neurotypical school-aged children, 6 to 8 years old, or school-aged children with a diagnosis of an ASC. 79 articles were found to be eligible. The full texts of the eligible articles were read to further determine their eligibility to be included in the literature review; 55 articles were included. Figure 1 presents a PRISMA Flow Diagram (Moher et al., 2009), summarising the article selection process.

Figure 1

PRISMA Flow Diagram of Article Selection Process



Note. Adapted from “Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement” (Moher et al., 2009)

1.6. Literature Summary

1.6.1. Measures of Executive Function for Children

A variety of approaches have been used to operationalise executive function in children, including performance-based measures, informant reports and observational tools. The articles included in this literature review focus on performance-based measures, sometimes referred to as direct or behavioural assessments.

Performance-based measures of executive function involve standardised procedures, often directly administered by clinicians or researchers to children on a one-to-one basis and in a laboratory setting. Such measures usually assess the accuracy and/or response time of a given task, with decreased accuracy and increased response time attributed to reduced executive function abilities. Performance-based measures require carefully controlled stimulus presentation to ensure each participant's experience is exactly the same (Toplak et al., 2013). Therefore, they are considered to be objective, and conceptually precise, together with the best form of measure for isolating specific executive function skills (McCoy, 2019). Previous measures of executive function were designed specifically for use in a laboratory setting, however, these measures are being increasingly used in the field and adapted for clinical use (McCoy, 2019). Some of the measures available for children aged 6 to 8 years are summarised below.

1.6.1.1. Adapted Measures

Several measures of executive function for use with children were adapted from adult measures, reflecting the assumption that they measured the same localised dysfunction in both groups (P. Anderson, 2002; V. Anderson, 1998, 2001). An example of an adult measure adapted for younger children is the Modified Wisconsin Card Sorting Test proposed by Nelson (1976). The original Wisconsin Card Sorting Test (WCST; Grant & Berg, 1948) was made simpler, shorter and less stressful by reducing the number of cards and changing the administration procedure. This was in the hope that the measure would be more suitable for children and, therefore, reduce the risk of refusal (Cianchetti et al., 2007). Other examples include the modification of the Tower of London (ToL;

Shallice, 1982) to enhance its utility with children as a measure of executive planning and problem-solving (Culbertson & Zillmer, 1998), and the Iowa Gambling Task (Bechara et al., 1994), which has been administered in its original version to children as young as eight years, and simplified for use with younger children (Children's Gambling Task; Kerr & Zelazo, 2004; Lehto & Elorinne, 2003).

However, these adapted adult measures did not take into account the child's reading ability or their limited attention span. This might have led to inappropriately interpreting deficits to be a result of injury rather than related to the young person's appropriate level of development (V. Anderson, 1998, 2001). The measures were also considered to be inappropriate as they were not of interest or relevance to young children (P. Anderson, 2002). As a result of these challenges, V. Anderson (1998, 2001) spoke of a need for "valid and well-standardized" assessment measures of executive function, specifically designed for children (p.124). Salimpoor and Desrocher (2006) suggested that in order to successfully assess executive function, tests must: be developmentally appropriate, separate components of executive function, vary in levels of difficulty and complexity, use different modalities for presentation and response, and be ecologically valid.

1.6.1.2. Established Measures

Carefully selected batteries measuring executive function in children were reviewed by Henry & Bettenay (2010). When selecting the measures, they considered the developmental appropriateness and the requirement of non-executive cognitive processes. They found that the measures they investigated showed some promise in successfully assessing executive function in children. Two executive function batteries they identified were the Delis-Kaplan Executive Function System (D-KEFS, Delis et al., 2001) and the Behavioural Assessment of the Dysexecutive Syndrome in Children (BADS-C, Emslie et al., 2003). Both batteries are considered suitable for children, however, only those aged 8 years and over. The D-KEFS has been criticised for requiring increased investigator input, due to close monitoring being required throughout administration, along with complex instructions and scoring (Homack et al., 2005). The BADS-C claims to assess executive function in an ecologically valid manner, however, there is little research on precisely which functions the different tasks measure (Henry & Bettenay, 2010).

Henry & Bettenay (2010) also identified the Developmental NEuroPSYchological Assessment-II (NEPSY II; Korkman et al., 2007), a battery of neuropsychological tests for children aged over 3 years; however, not all tests cover the full age range. Measures of executive function include Auditory Attention and Response, which is suitable for children aged 7 years and above, and Inhibition which is suitable for children aged 5 years and above and has been found to have modest to good test-retest reliabilities. The NEPSY-II has been criticised for the lack of standardised data on urban/rural residences and discrepancies in parent education level among certain ethnicities within the standardised data. However, the NEPSY-II has been praised for the inclusion of participants with additional educational needs in the standardised data, including those with a diagnosis of an ASC (Davis & Matthews, 2010).

Some batteries assess specific executive function components, such as the Working Memory Test Battery for Children (WMTB-C; Pickering & Gathercole, 2001). The WMTB-C is a UK normed and standardised battery, which includes the 'listening span task' and a backwards digit span task, both of which are suitable for children aged 5 years and above. The WMTB-C has been praised for being easy to administer and score, however, criticised for variable test-retest reliabilities (Henry & Bettenay, 2010).

A performance-based behavioural assessment used with younger children is the Delay Gratification task (Wilson et al., 2017). The Delay Gratification task, considered to assess hot executive function, is typically used with children aged 5 years and under. However, Wilson et al. (2017) explored the tasks utility in older children aged 5 to 12 years, by making age related changes to the original task, such as offering a selection of gifts as a reward. They found age-related differences between the ages of 5 and 7 years, however, ceiling effects were observed from then on. Researchers suggest that the task could be more suitable for older children, if the delay period was extended.

Other tests suitable for the age range of 6 to 8 years yet to be standardised include Lagattuta et al.'s (2011) and Gerstadt et al.'s (1994) Stroop-like tasks, called the Happy-Sad task and the Day/Night task, respectively. The tasks require children to inhibit a prepotent response, for example, by saying 'happy' when viewing a sad face and 'sad' when viewing a happy face, or 'night' when viewing the day card and 'day' when viewing the night card. The Happy-Sad task was developed to measure executive function across childhood and early adulthood. Challenges associated with these measures will be discussed.

1.6.2. Challenges of Measuring Executive Function in Children

The literature has identified various challenges in measuring executive function in children. Such challenges include, lack of construct and ecological validity, task impurity, differing levels of difficulty and complexity, task dependant variables, lack of sensitivity to individual differences, the need to be administered by highly trained researchers or clinicians in laboratory settings and their use in small, self-selected samples of children whose parents are motivated and willing to participate in such studies (Salimpoor & Desrocher, 2006; Wilbourn et al., 2012). The challenges that most frequently came up in the literature will be discussed in more detail.

1.6.2.1. Heterogeneity Between Measures of Executive Function

There is generally little consensus on how human cognition, including executive function, is operationally defined in research (Kramer et al., 2014), which has led to variability in how it is measured. Such heterogeneity is present in task-dependent variables, which include the measures type of presentation and response required, this is often either auditory or visual and verbal or non-verbal, respectively.

Performance-based measures of executive function for children have been argued to differ greatly in their presentation, particularly in how child-friendly they are, the amount and form of feedback given and the inclusion of game elements (Johann & Karbach, 2018). There has been little research investigating whether the different ways of assessing executive function in children impacts their performance. However, the various forms of presentation require different skills, for example, auditory presentation requires language and as previously discussed, this can impact a young child's performance.

1.6.2.2. *Task Impurity*

There is much uncertainty in the literature around the executive functions different measures tap into, along with to what extent they capture single pure processes (Ballhausen et al., 2017). Task impurity refers to a test performance requiring the use of other cognitive components in addition to the one it is measuring, and a tests inability to differentiate between different cognitive components. This applies to differentiating executive function from other non-executive cognitive skills and differentiating between different components of executive function. Measures of executive function are typically confounded by other non-executive skills and processes required for performance (V. Anderson, 1998, 2001). For example, Ballhausen et al. (2017) found that prospection predicted planning performance on the Zoo Map test in the BADS-C over the traditional marker of executive function, inhibition, measured by the Go/NoGo task. Like much of the research in this area, there are limits to this studies generalisability, due to the participants being from middle-class backgrounds and predominantly Caucasian. Task impurity can result in measures being unable to separate specific cognitive deficits (P. Anderson, 2002), subsequently impacting accurate hypothesis testing and theory building (van der Sluis et al., 2007).

Task impurity is particularly relevant for younger children, who are still developing basic skills needed to understand and complete the task, along with particular sub-groups of children, such as those that struggle with social interaction and communication, and those whose first language is not the test language (McCoy, 2019). Task impurity in these populations raises difficulties in concluding whether improved performance with age is attributable to the development of executive function or non-executive abilities (van der Sluis et al., 2007).

An example of task impurity is demonstrated by Lagattuta et al. (2011), who found differences in children's performance on measures considered to assess the same process. They found that children of all ages performed better on the

Day/Night task than the Happy-Sad task, for both number of errors and speed. However, ceiling effects were demonstrated by 8 years of age on the Day/Night task. Lagattuta et al. (2011) predicted and investigated whether this was a result of the happy-sad cards only being distinguishable by the shape of the mouth, whereas the Day/Night task was distinguishable by multiple features, such as the colour of the background and different shapes (moon versus crescent). They found that even after adapting the cards, so the day/night cards were more similar and the happy-sad cards had more distinguishing factors, children aged 4 to 11 years still found the Happy-Sad task more difficult. This research could not conclude why this task was found to be more difficult across all ages. One explanation could be that the Day/Night task involves arbitrary response mapping in addition to inhibition. Simpson & Riggs (2005) suggested this because of children differentiating between more than the opposite response, day versus night rather than sun versus moon.

Due to the interrelated nature of executive function components, gaining an accurate assessment of a single component can be difficult and often requires the interpretation or triangulation of multiple measures (Salimpoor & Desrocher, 2006). Similarly, to fully account for non-executive abilities, additional measures would need to be carried out.

1.6.2.3. Ecological Validity

Performance-based measures have been argued to underestimate the challenges faced in daily life, due to being highly structured (Annotti and Teglasi, 2017). Requiring a measure of executive function to be novel can impact its ecological validity, due to it not being relatable to a real world situation (P. Anderson, 2002). Performance-based measures have traditionally been developed for administration in a quiet environment to reduce distractions. As such, they do not capture the conditions under which children must regulate their attention, behaviour and emotions (McCoy, 2019). Along with the environment in which executive function measures are typically carried out in, the role of the clinician has also been suggested to impede ecological validity and optimal performance, through investigator and participant bias (P. Anderson, 2002; Berg et al., 2020).

Low ecological validity of executive function measures has been argued to cause incorrect predictions of children's day-to-day executive function behaviour. This has been supported by research examining the distinction between hot and cool executive function, with performance differing based on whether the task is intended to be hot or cool (McCoy, 2019; Zelazo & Carlson, 2012). Traditionally, executive function research has explored the various executive function components in decontextualised, cool settings. McCoy (2019) suggested that hot executive function assessments are more likely to be ecologically valid, compared to cool executive function assessments, as they "require children to recruit a number of regulatory processes in the context of external rewards and motivators" in order to meet the demands of the task (p. 68). For example, Kenworthy et al. (2020) found the "hot" Executive Function Challenge Task, but not the D-KEFS, to be moderately correlated with parent-reported executive function measured by the Behaviour Rating Inventory of Executive Function (BRIEF; Gioia et al., 2002). However, there is uncertainty around the degree to which performance-based assessments could truly replicate hot situations experienced by children (McCoy, 2019).

Informant-response measures were developed to capture observed executive function in real-life settings and quantify qualitative data in a standardised format (P. Anderson, 2002). Such measures were designed to provide an ecologically valid indicator of executive function in everyday settings (Toplak et al., 2013), and include measures such as: the BRIEF (Gioia et al., 2002); and the Childhood Executive Function Inventory (CHEXI; Thorell & Nyberg, 2008). However, informant-response measures have been found to capture different constructs of executive function compared to behavioural measures, as they share little to modest levels of agreement (V. Anderson et al., 2002; Annotti & Teglasi, 2017; McAuley et al., 2010; Ten Eycke & Dewey, 2016; Toplak et al., 2013). Informant-response measures have been considered to more accurately capture how executive function manifests in real-world settings, such as the home or school, compared to the behavioural measures (McAuley et al., 2010). Therefore, informant-response measures are rarely used as a replacement for behavioural measures, rather they are used alongside them to assess whether a child's performance on the behavioural measures translates into real-world settings, thus, demonstrating ecological validity.

Specific performance-based measures have been developed to address the challenge of ecological validity, such as the Multiple Errands Test (Shallice & Burgess, 1991), which is intended to assess how one navigates the complex demands of real life (Gardiner et al., 2017). Obradović et al. (2018) also suggested that changing the environment in which a measure is completed could impact its ecological validity. For example, they found that administration of a performance-based measure (the Hearts and Flowers task), in a group classroom environment was predictive of gains in students' academic achievement, whereas administration in a one-to-one laboratory setting was not.

There remains an increasing demand for measures of executive function that can be used in real-word everyday contexts and evidence ecological validity (McCoy, 2019). Welsh & Peterson (2014) argued that executive function research must adapt its methodologies "to better approximate natural contexts" (p. 152).

1.6.2.4. Use Across Childhood

Due to the rate at which executive function develops throughout childhood, measures of executive function are often suitable for one age group but not another (P. Anderson, 2002; Lagattuta et al., 2011). Measures of executive function suitable in childhood are either too easy for older children or inappropriate for younger children through, for example, requiring children to be literate, and can result in ceiling or floor effects (Lagattuta et al., 2011; Zelazo et al., 2013). As a result, tasks measuring executive function are often designed for a specific age group. Lagattuta et al. (2011) attempted to address this problem through their development of the Happy-Sad task, which was concluded to be a suitable measure of executive function across childhood. However, this measure has not been standardised and there are few standardised executive function batteries suitable across childhood.

1.6.2.5. Cultural Generalisability

Measures of executive function are often, if not always, standardised against White Western populations, which has implications for their validity in populations with differing demographics, culture and language. In line with this, performance-based measures of executive function have been criticised for their lack of transferability to different nations and cultures, such as low- and middle-income Countries (LMICs). It has been argued that the majority of executive function measures have not been adapted or normed for use in LMICs settings, and would not be appropriate, as they require representational symbolic and shared knowledge, along with shared language between administrator and participant (Ford et al., 2019).

To address this challenge, Ford et al. (2019) suggested the use of single, un-normed tasks, where executive function ability is determined by comparing performance on executive function trials to baseline trials. They developed the Rapid Assessment of Cognitive and Emotional Regulation (RACER), designed to be administered by and to those with no literacy or numeracy and in a variety of assessment environments. It includes measures of inhibition and working memory. RACER was designed to not require content knowledge and not be impacted by distraction, through short tasks. The administration is automated by the application itself, including instructional videos, timekeeping and response recording. Ford et al. (2019) found that RACER was appropriate for assessing executive function in children aged 5 to 13 years in LMIC settings and could be applicable to a range of environments and populations. However, some difficulties were experienced in settings with limited electricity. They did find differences in performance across two settings, which they attributed to differences in the countries educational and development levels. This is important to consider when interpreting results, given the diverse populations who complete measures of executive function.

1.6.2.6. Summary of Challenges

The reviewed literature identified a lack of sensitive and specific measures of executive function, especially for use with younger school-aged children (Henry & Bettenay, 2010). Due to measures covering different age ranges, varying use of verbal and non-verbal assessments and inconsistencies between performance on established tests and real-life behaviour, the use of a combination of tests and the use of additional qualitative data is often necessary to comprehensively assess executive function in school-aged children (Henry & Bettenay, 2010). In line with this, P. Anderson (2002) highlights the importance of using qualitative and cognitive-process methodologies alongside the traditional quantitative data to enhance the utility of the measure.

1.6.3. Computerised Measures of Executive Function

In line with the growing dominance of technology, researchers and clinicians are increasingly utilising technology in the development of executive function measures for children. Researchers have turned to using computerised tests as a way of addressing some of the challenges identified.

Computerised neuropsychological tests for children have been considered to have a number of advantages compared to classical methods of examination. Such advantages include improved data collection, such as more precise data on response time and accuracy, than pen-and-paper measures (McCoy, 2019). This is a result of standardised implementation and objective scoring procedures, thus reducing examiner effects, as well as being more easily available in a number of a languages, and increased participant motivation (Forns et al., 2014; McCoy, 2019; Piper et al., 2012). In some cases, established performance-based measure of executive function have been computerised; in others, novel computerised measures of executive function have been developed, examples of both will be discussed.

The n-Back paradigm developed by Nelson et al. (2000) was considered the most widely used computerised measure of working memory and has been demonstrated to show good criterion validity and internal consistency. It has since been used in computerised executive function batteries, such as the NIH Toolbox for Assessment of Neurological and Behavioural Function Cognitive Battery (NTCB; Kramer et al., 2014). The NTCB was designed with the aim of providing a single set of measures that could be used with young children through to adolescence, predominantly in research, through an accessible format for both clinician and participant (Akshoomoff et al., 2014; Bauer & Zelazo, 2014; Weintraub, Bauer, et al., 2013). The battery is made up of measures of cognitive flexibility, inhibitory control and working memory, in accordance with Miyake et al's. (2000) conceptualisation of executive function. However, it also includes measures of verbal fluency, planning and insight, alongside non-cognitive measures of executive function, such as social cognition and behaviour (Kramer et al., 2014). The measures included in the battery were adapted from established measures, including the Dimensional Change Card Sort Test (DCCST; Zelazo, 2006) and the Erikson Flanker task (Eriksen & Eriksen, 1974), to measure cognitive flexibility and inhibition, respectively, along with novel measures, such as the List Sorting Working Memory task. When developing the NTCB, the authors noted the importance of measures using simple instructions and having practice trials (Bauer & Zelazo, 2014).

In a validation study of the NTCB, the measures demonstrated strong test-retest reliability in 3- to 6- and 8- to 15-year-olds (Bauer & Zelazo, 2014); and good convergent validity when compared to established tests in children aged 8 years and above (Weintraub, Dikmen, et al., 2013). Due to the limited measures available for younger children, convergent validity was difficult to assess in those under 8 years. However, Tulsky et al. (2013) demonstrated moderate convergent validity of the working memory measure in the younger age group with the NEPSY-II Sentence Repetition (Korkman et al., 2007). Zelazo et al. (2013) assessed convergent validity of the flexibility and inhibition measures in the younger age group, using the Block Design subtest of the Wechsler Preschool and Primary Scale of Intelligence, 3rd Edition (Wechsler, 2002) and found a strong positive correlation. Akshoomoff et al. (2014), on the other hand, found the NTCB version of the DCCST to be "limited in its utility for measuring cognitive flexibility in children under the age of seven" in a sample including neurotypical children and children with suspected or a diagnosis of a learning disability or Attention-Deficit Hyperactivity Disorder (ADHD; p. 10).

The Cambridge Neuropsychological Test Automated Battery (CANTAB; Cambridge Cognition Ltd, 2006) is a computerised assessment of executive function with normative data available for children aged 4 years and above. The majority of tasks are nonverbal, in both instructions and response, and it has consequently been suggested to be more 'culture-free' than other measures (Henry & Bettenay, 2010). The battery has demonstrated construct and discriminant validity in child populations, along with high internal consistency (Henry & Bettenay, 2010). The battery has been argued to be limited by its price and reduced ecological validity due to the purity of the tasks and assessment format (Luciana, 2003).

The Psychology Experiment Building Language (PEBL) was developed to overcome limitations of previously developed computerised batteries used across the lifespan, such as the cost of use and inaccessible computer codes that underlie the commercial tests, through the software being freely available to download and modify (Piper et al., 2012). PEBL computerised established tests of executive function, such as the Trail Making Test (TMT), the WCST and the ToL. However, similarly to other established measures, the younger age limit is 7 years.

1.6.4. Game-Based Measures of Executive Functions

Performance-based measures of executive function are frequently presented to children as if they are games. It has been suggested that using game-like measures of cognitive function with children could increase motivation to engage in the task and, therefore, improve task performance (Johann & Karbach, 2018). Another benefit of game-based measures is potentially reduced cultural and language bias. Providing a narrative with clear goals in a game environment has been considered to heighten involvement with tasks and increase ecological validity, as it resembles real life more closely (Berg et al., 2020). Game-based measures of executive function have been developed with this in mind.

Józsa et al. (2017) designed a tablet application, made-up of game-like tasks. The application, designed as a school readiness assessment, is appropriate for educators who do not have formal training to administer the test and interpret the results and does not require participant reading ability. The application has been found to have good reliability and criterion validity.

Johann & Karbach (2018) designed a new battery of executive function tasks for children, using game elements based on Ryan & Deci's (2000) self-determination theory (relatedness, competence and autonomy), to enhance intrinsic motivation. They found that interest, perceived competence and relatedness were higher after completing the game-based tasks compared to established measures of executive function, however, no significant differences were found between effort and perceived autonomy. Johann & Karbach (2018) found that their new game-based measure significantly correlated with the established measures, apart from the stroop-like task of inhibition. They did not find performance differences between the game-based and established measures, except for the Flanker task, where performance was improved on the game-based version. They could not establish whether the increase in motivation was a result of the addition of game-elements or due to the inclusion of several motivational features. Problems with this, and other game versions, is that they are 2D and, therefore, resemble established tests.

Berg et al. (2020) evaluated two newly developed child friendly apps aimed at assessing executive function in the classroom, with the intention of them being fun and enjoyable activities, subsequently eliciting greater task performance. Early Years Toolbox (EYT; Howard & Melhuish, 2017) was designed to measure executive function though short, easy to understand tasks suitable for younger children, through the use of publicly available 2D executive function apps. eFun (Berg et al., 2019) is a self-assessed executive function measure involving game-based tasks that are based on established executive function measures, with the inclusion of dynamic elements to sustain attention. Berg et al. (2020) compared the EYT and eFun apps, also assessing how enjoyable, fun, exciting, easy, hard, boring, and frustrating participants found the tasks. Results suggested that children enjoyed playing both sets of tasks. However, children rated the tasks they found easiest as most enjoyable.

Bell et al. (2021) developed the Active Test of Embodied Cognition (ATEC), a computerised measure of motor speed, balance, rhythm, bilateral coordination, attention, memory, response inhibition and self-regulation using body movement. The ATEC was developed on the premise that our sensorimotor experiences and higher cognitive processes are related, referred to as 'embodied cognition'. Bell et al. (2021) suggested that such a measure may be "more sensitive to neurodevelopmental processes which are atypical in children at risk for neurodevelopmental disorders" (p. 3). In order for the game to be more engaging, Bell et al. (2021) designed it in the form of a TV gameshow whereby the participants are contestants who take part in 'The Activate Games'.

The game allows for comparison between sensory modalities through having both verbal and visual trials.

ATEC demonstrated concurrent validity with traditional neuropsychological and parent-report measures, including a measure of real-world functioning. ATEC also demonstrated discriminant validity between children at risk of executive function-related impairments and those not, and high test-retest reliability. Bell et al. (2021) also found that the ATEC total score accounted for a significant amount of variance on the parent measure of real-world functioning and that established measures did not contribute significantly. However, the measure has not yet been trialled in large populations or longitudinal studies, nor have age-related norms been established. The researchers aim to have automated administration and scoring systems. An adult measure is also in development with the hope ATEC can be utilised across the lifespan.

In an attempt to further address the challenges of executive function measures and taking into consideration the recent developments of computerised and game-based tasks, Davis (2020) developed a 3D computerised game to measure executive function in school-aged children, called Dragon Adventure. Considering the heterogeneity between definitions of executive function, executive function research often requires a narrower focus on models and concepts. As a result of being influential, widely adopted and considered the most frequently evaluated model in the literature, the development of Dragon Adventure was based on Miyake et al.'s (2000) model of executive function and, therefore, targeted flexibility, from now on referred to as switching, inhibition and working memory. Dragon Adventure is comprised of three executive function tasks assessing these components. Davis (2020) tested Dragon Adventure in young people aged 11 to 12 years. The game was found to be moderately to largely correlated with related established measures and a strong association was found with the CHEXI (Thorell & Nyberg, 2008), indicating that Dragon Adventure demonstrated good ecological validity and has the potential to be utilised as a measure of executive function for children.

1.6.5. Challenges of Measuring Executive Function in Children with an Autism Spectrum Condition

Measures of executive function are used to understand and build an executive function profile of those with an ASC (Gómez-Pérez et al., 2016). The contrasting findings surrounding the executive function profile of children with an ASC has been considered to be a result of challenges with the measures used (Mackinlay et al., 2006; Gardiner et al., 2017).

Many of the challenges mentioned above have been reported to be exacerbated when the measures are used with those with an ASC. For example, some have argued that ASC is characterised by deficits in comprehension, processing speed and attention span, all of which have been considered to effect children's performance on executive function tasks adapted from the adult literature (Adams & Jarrold, 2009). Other challenges with executive function measures considered specific to the ASC population will be discussed below.

1.6.5.1. *Task Impurity*

Research has shown that use of executive function measures in the ASC population are more susceptible to the effects of task impurity and this is reflected in the scores (Adams & Jarrold, 2009). For example, differences were not present between children with a diagnosis of an ASC and neurotypical children on the Day/Night task. It was suggested that this was due to the task being underpinned by arbitrary response mapping, rather than inhibition (Russell et al., 1999). However, differences between the populations were present on the Dog/Pig task, which is considered to be a more accurate measure of inhibition. The Dog/Pig task requires children to say 'dog' or 'pig' in response to the sun or moon card, rather than being instructed to "say the opposite" (Ames & Jarrold, 2007; Diamond et al., 2002). The Dog/Pig task is considered a truly opposite response task of inhibition due to 'dog' and 'pig' not being semantically related to the to-be-inhibited response, 'night' or 'day' (Ames & Jarrold, 2007; Diamond et al., 2002).

Russell et al. (1999) proposed that children with ASCs struggle on tasks that require inner speech, such as those that require holding in mind novel, arbitrary information or rules. This has been supported through findings that individuals with ASCs use inner speech less than neurotypical individuals to complete certain executive function tasks (e.g. Wallace et al., 2009). Russell et al. (1999) also found that performance of children with an ASC is impacted by the demands of a task, such as whether the task is verbal or non-verbal.

Children with an ASC have been found to perform similarly to neurotypical children on the Stroop task. Adams & Jarrold (2009) suggested that this was a result of the population being less affected by the semantic demands of this task, due to poor reading comprehension. They evidenced this by comparing performance on the traditional Stroop task to an additional trial, whereby participants were presented with pictures of animals with incongruous heads and bodies and were asked to inhibit the head and name the body, and vice versa. They concluded that the classic Stoop task is not a valid measure of inhibition for children diagnosed with an ASC. However, the authors noted that their findings of impairment in the ASC sample on the new task could have been a result of a "greater tendency to process interfering distractors" (Adams & Jarrold, 2009, p. 1062)., highlighting the exacerbated impact of task impurity in this population.

1.6.5.2. Ecological Validity

Research has demonstrated that structured performance-based measures of executive function do not reflect difficulties children with an ASC face in everyday situations (Chan et al., 2009; Czermainski et al., 2015). Gómez-Pérez et al. (2016) argued that executive function difficulties associated with ASCs may be more accurately observed in real-life situations than clinical contexts. The discrepancy between test performance and real-life situations may be a result of the highly structured nature and clearly defined limits of performance-based measures of executive function, with these masking the real-life difficulties faced by children with an ASC (Czermainski et al., 2015; Mackinlay et al., 2006). Research investigating the executive function profile of those with an ASC has generally been more consistent when informant-rating measures are used, compared to performance-based measures (Gardiner et al., 2017). Multitask tests, such as the Battersea Multitask Paradigm, have been shown to more successfully capture difficulties of everyday life over measures involving a single problem (Mackinlay et al., 2006; Shallice & Burgess, 1991). However, when measuring executive function in children with an ASC through a multitask test, Mackinlay et al. (2006) did not find significant correlations with the BRIEF (Gioia et al., 2002).

The BADS-C (Emslie et al., 2003) was designed to be a more ecologically valid measure of executive function and minimised the requirement for reading, language skills and verbal short-term memory, in order to be suitable for high-functioning populations of children with an ASC (White et al., 2009). However, as previously discussed, this battery is not suitable for children under 8 years and performance has been predicted by other non-executive cognitive abilities (e.g., Ballhausen et al., 2017).

Taking these challenges with ecological validity into account, it has been suggested that performance-based measures of executive function could demonstrate how children with an ASC can succeed (Gardiner et al., 2017).

1.6.5.3. Generalisability

Research investigating measures of executive function in children with an ASC often focuses on a specific part of the population, usually those without language or learning difficulties, who are often described as ‘high-functioning’ within the literature. Due to a proportion of the ASC population having language and learning difficulties, such research cannot be generalised to the wider ASC population (e.g. Gómez-Pérez et al., 2016). This could be a further reason why there is difficulty defining the executive function profile of this population. As a result, it has been suggested that future research should include a greater variability in the ASC sample, in order to increase generalisability (Gardiner et al., 2017; Gómez-Pérez et al., 2016).

1.6.6. Computerised and Game-Based Measures of Executive Function for Children with an Autism Spectrum Condition

The performance of children with an ASC on executive function tasks has been claimed to be negatively impacted when administered by an experimenter. Therefore, it has been suggested that use of computerised executive function measures could aid assessment in this population (Kenworthy et al., 2008). For example, CANTABs presentation and response measurement through touch screen has been considered to be more appropriate for children with an ASC (Luciana, 2003).The inequivalence between established and computerised measures of executive function for this population has been supported by research showing that children with an ASC performed similarly on single and multi-component executive function tasks when computerised, despite this population typically being found to experience increased difficulty with more complex, multi-component executive function tasks (Gardiner et al. (2017). In addition, Ozonoff (1995) found that children with an ASC performed equally to neurotypical individuals on a computerised test of executive function, but not when administered by an experimenter.

Computer-administered tasks may facilitate the performance of children with an ASC through reducing social interaction requirements (Ozonoff, 1995; Ozonoff & Strayer, 2001), though findings are not consistent. Williams and Jarrold (2013) did not find a significant difference between performance on the established and computerised measures in a population of children with an ASC, despite the measures being equivalent apart from the interaction with the experimenter. Gardiner et al. (2017) also found that performance on computerised tasks of executive function differed to parent-report measures, which could indicate that computerised-measures do not address the difficulties of ecological validity.

1.6.7. Summary

Despite a number of challenges with established measures of executive function, recent research advancing these measures through the use of technology and game-based paradigms shows promise towards creating more suitable measures with increased clinical and research utility. These advancements may have additional benefits for those with an ASC.

1.7. The Present Study

1.7.1. Justification

There are currently few standardised measures of executive function suitable for children aged 6 to 8 years. Considering the associations between executive function and positive outcomes in childhood, it is important to have measures appropriate for this population, in order to better understand executive function development. Along with this, the ability to assess executive function is important for a broad range of childhood conditions and behavioural problems, including ADHD, head injury, epilepsy and query learning disabilities. A greater understanding of executive function in childhood could lead to better interventions and rehabilitation for those with executive dysfunction. Better assessments, interventions and rehabilitation could help children to thrive academically and socially (Jurado & Rosselli, 2007).

There is inconclusive evidence around the executive function profile of children diagnosed with ASC, however, this lack of clarity has been attributed to the challenges inherent in established measures of executive function. Having a suitable measure of executive function for this population could further our understanding of their executive function profile. In addition, children with cognitive and behavioural difficulties are the populations most likely to be presented with measures of executive function, and therefore it is important that the measures are suitable for them (Henry & Bettenay, 2010). Often measures are devised with either clinical or non-clinical populations in mind and not both, however it is important that measures of executive function are suitable for both neurotypical and neurodiverse children. More appropriate measures of executive function could improve the identification of executive dysfunction and the implementation of appropriate support. In addition, established measures of executive function often have to be adapted for use with those with communication difficulties, which can be time consuming for clinicians.

My personal interest in the topics of child development and utilising technology in clinical psychology led me to the current area of research, along with the research Davis (2020) had undertaken. Through my Director of Studies, I was able to meet with Davis and discuss the potential future directions of her research, including ideas for the development of Dragon Adventure. From the literature review, I established that there was a gap in measures of executive function for children aged 6 to 8 and that measures were typically not trialled in the populations that are most commonly presented with them. I felt that Dragon Adventure could provide benefits, both as a suitable measure for children aged 6 to 8 and for children with communication difficulties, such as those with an ASC or who do not speak English as a first language.

Building on the research carried out by Davis (2020), this research investigates the use of the game, Dragon Adventure, as a measure of executive function for neurotypical 6- to 8-year-olds and children diagnosed with an ASC aged between 6 and 11 years. The game will be trialled in 6- to 8-year-olds, as an age group that could benefit most from a novel measure. The ASC sample age range is broader in order to include the age group initially assessed in Davis (2020) and enable future comparisons to be made. To address the challenge of studies typically recruiting a high functioning sample of children with an ASC, meaning results are not generalisable to the full population (e.g. Gardiner et al., 2017), this study will aim to recruit a sample with a range of functions and abilities.

Further development of Dragon Adventure could improve engagement and accessibility. For example, additions to Dragon Adventure, such as including non-verbal demonstrations of the tasks, would reduce the reliance on language and literacy and contribute to the game being more culturally fair. Weintraub, Bauer, et al. (2013) suggested that children perform better at a younger age when “simple, easy to follow instructions are used and when task materials are engaging, concrete and familiar” (p. 17). To ensure the instructions are easy to follow and developmentally appropriate, they will be both written and spoken, along with visual demonstrations; this should also limit cognitive demands associated with social interactions. Data will be collected on whether English is the participant’s primary language, to assess whether language affects their performance on Dragon Adventure. Dragon Adventure also holds benefits for the clinician through the ease of administration and automated multidimensional scoring. Dragon Adventure has the potential to be more accurate than established measures and less prone to human error.

In line with Davis (2020), enjoyment levels experienced when completing both Dragon Adventure and established measures will be assessed using a visual analogue scale. Although Davis (2020) did not find any differences in enjoyment ratings between Dragon Adventure and established measures, this could have been as a result of the game not being appropriate for the samples age range (11 to 12 years), but the game arrangement could be more suitable for children of a younger age.

One of the limitations of Davis' (2020) research, was not accounting for other cognitive and non-cognitive factors, such as computer literacy and processing speed. At the ages of 6 to 8 years, it is likely children will have varied exposure to computers, therefore, it is important that this is accounted for; processing speed will also be accounted for. The current study will also attempt to control for fatigue bias and practice effects through randomly allocating participants an order to complete the tasks, either the established measures or Dragon Adventure first. To support children being more accustomed to the test situation, the data collection will be carried out in the participant's school. It has been suggested that executive function abilities might be more appropriately assessed in an environment where the child is expected to be more independent (Mackinlay et al., 2006).

In line with Davis (2020), this research will assess the ecological validity of Dragon Adventure through teacher-report using the CHEXI (Thorell & Nyberg, 2008). Teacher reports have been considered more accurate compared to parent reports. This is because parent reports have been suggested to be impacted by the parents role in day-to-day organising on behalf of their child (Mackinlay et al., 2006).

Evaluation of measures by those completing them is very rarely collected or taken into consideration (Berg et al., 2020). As recommended by Davis (2020), this research will collect formal feedback from participants on how Dragon Adventure could be improved.

1.7.2. Research Aims and Questions

This study aims to establish feasibility, acceptability, concurrent and ecological validity of the newly developed game-based measure of executive function, Dragon Adventure (Davis (2020), in neurotypical children aged between 6 and 8 years and in children with a diagnosis of an ASC aged between 6 and 11 years. This study also aims to establish whether Dragon Adventure improves engagement, compared to established measures of executive function, and gain feedback on how the game could be improved.

The research questions this study will address are:

- Does Dragon Adventure assess inhibition, working memory and switching in line with established measures of executive function and teacher ratings in neurotypical children aged 6 to 8 years?
- Is Dragon Adventure suitable for assessing inhibition, working memory and switching in line with established measures of executive function and teacher ratings for children aged 6 to 12 years with a diagnosis of an ASC?
- Is there a difference in rated enjoyment between the established measures and Dragon Adventure in typically developing children aged 6 to 8 years and children aged 6 to 12 years diagnosed with an ASC?
- How do children and young people feel Dragon Adventure could be improved?

1.7.3. Hypotheses

Taking into account the literature and previous research investigating Dragon Adventure, the following hypotheses are made:

- Dragon Adventure subgames will be moderately to strongly correlated with established measures of executive function in children aged 6 to 8 years.
- Dragon Adventure will be a suitable measure of inhibition, working memory and switching for children diagnosed with an ASC aged 6 to 12 years.
- Dragon Adventure will be moderately to strongly correlated with teacher ratings of executive function in both populations.

- Dragon Adventure will be rated as more enjoyable than established measures of executive function.

2. METHODS

2.1. Epistemology

Epistemology is concerned with the nature of knowledge and the basis for claims to possess knowledge (Willig, 2019). Within research, epistemology is considered to inform the theoretical perspective, methodology and methods (Crotty, 1998). Therefore, it is important that researchers reflect on their own epistemological position.

Two key epistemological positions exist, the realist and relativist. The realist position is concerned with materialism and poses that a reality exists which we can objectively investigate through systematic observation and experimentation, due to structures and objects having cause-effect relationships with each other. The relativist, on the other hand, is concerned with idealism and poses that reality is constructed through the diverse possible interpretations of the world and that societal context informs all epistemic assumptions (Fletcher, 1996). Realism and relativism are dimensions on a spectrum with differing positions. Such positions of epistemological truth or theoretical perspective include positivism, idealism, pragmatism, phenomenology, social constructionism, and critical realism.

2.1.1. A Critical Realist Position

This research is underpinned by a critical realist position (Bhaskar, 1990; Greenwood, 1994). Critical realism proposes that a reality exists, however, we do not have direct contact with it, and therefore, our theories of reality are socially constructed (Pilgrim & Bentall, 1999). Critical realism concerns the perception of knowable and observable phenomena being partly dependent on real processes, along with our beliefs and expectations. Our methods of finding out about the world are understood to be shaped by societal forces and interests, and as a result, are imperfect. The position suggests that we should investigate reality "cautiously and critically", through the consideration of history, social context and power (Pilgrim & Bentall, 1999, p. 262).

Critical realism argues that categorisation, classification and quantification offer the best approach to understanding. Therefore, within psychological research, the focus is on particular constructs, with a preference for self-report questionnaires and experimental studies. This position also holds that causal links can be established between psychological processes and are most likely to be revealed through the study of groups. Therefore, focus is placed on intra-psychic constructs rather than interpersonal ones. Critical realism takes a pragmatic approach to truth, whereby concern lies with methodological issues over conceptual and philosophical issues. Therefore, fundamental concepts are treated as 'real' if they appear to have statistical support.

This research reflects a critical realist epistemological stance as it acknowledges that the concept of executive function only exists due to historical and social context, however, currently represents a socially contextualised way of understanding some aspects of behaviour, that can be measured independently of the researcher. Therefore, this research will follow an experimental quantitative methodology, with a qualitative element to gain feedback on Dragon Adventure. Similarly, this research acknowledges that the classification of ASCs currently represents a helpful approach to understanding a group of people, however, the diagnoses exist as a result of history, social context and power, and can be challenged. Due to 'executive function' and 'ASC' being Western constructs, both have currency and utility in Western clinical, research and academic settings. Problems may arise when such terms are attributed more broadly, and alternative understandings are not considered. Therefore, any interpretations should also be located within the current historical and social context.

2.2. Design

The current study utilises quantitative experimental methodology, adopting both a cross-sectional correlation and within-participants design. A cross-sectional design was adopted for validating Dragon Adventure against the established performance-based and teacher-rating measures for both the neurotypical and ASC sample. Scores from Dragon Adventure were compared to scores on the

Wechsler Nonverbal Scale of Ability Spatial Span (WNV-SS), NEPSY-II Inhibition subtest and the CHEXI. When comparing Dragon Adventure to established measures, computer literacy and processing speed were controlled for using the Children's Technology Use Questionnaire (CTUQ; Horton, 2013) and NEPSY-II Naming Time score, respectively. A within-participants design was adopted to compare enjoyment ratings between Dragon Adventure and the established measures. The type of measure, either established or Dragon Adventure, acted as the independent variable and the enjoyment rating as the dependent variable.

A qualitative methodology was implemented to assess feedback from participants on Dragon Adventure. Feedback responses were analysed through a content analysis.

2.3. Recruitment

The population of 6- to 8-year-old children was chosen based on the results of the literature review, which demonstrated that there is a lack of standardised executive functioning measures for this age group compared to other age groups. The decision to trial Dragon Adventure in a neurodiverse population was made due to such groups typically being presented with performance-based measures, that are often not developed for them. The population of children diagnosed with an ASC was chosen due to personal influence, having previously worked with this population.

Recruiting participants from these populations seemed most appropriately done through schools, both mainstream and specialist. Unfortunately, the data collection took place during the pandemic and most schools were not allowing external visitors. Through personal connections, two schools were willing to host recruitment and data collection. As a result, neurotypical participants were recruited from a mainstream primary school. Participants with a diagnosis of an ASC were also recruited from the mainstream school, along with a specialist primary school for children with a diagnosis of an ASC. Recruiting only from two schools engenders limitations to this research in regard to sample size and generalisability.

The inclusion criteria for neurotypical children were that they were aged 6 to 8 years, able to understand information in English, and did not have a learning disability. The inclusion criteria for children with a diagnosis of an ASC, were that they were aged age 6 to 11 years old, able to understand information in English and had a diagnosis of an ASC. Exclusion of participants with other diagnoses, such as learning disabilities, was not implemented due to previous research being criticised for only representing a specific ‘high-functioning’ sample of the ASC population.

2.4. Measures

2.4.1. Dragon Adventure

Dragon Adventure (Davis, 2020) is a newly developed computer game measuring executive function. It includes three subgames adapted from established measures of inhibition, working memory and switching. Throughout the game, the participant’s character is a dragon. The participant moves the dragon using the laptop arrow keys and space bar, along with using the trackpad to click buttons on screen. The game includes written instructions on screen, presented through short phrases and sentences. The participant must click ‘continue’ to be presented with the next instruction. Due to the instructions not having a back button, which could result in participants missing important information, printed instructions were also provided in this study. To enhance the games usability for those with communication difficulties, instructions were also read out verbally and video demonstrations were shown before each subgame (please see screen shots in Appendix A).

The aim of the game is for the dragon to find its way home by completing three tasks to collect three crystals. One crystal is awarded for each game they complete and opens the ‘gate’ to the next game on to the end of play.

2.4.1.1. Dragon Dash

Dragon Dash measures inhibition, using a Stroop-like paradigm. The dragon must reach the end of the path, avoiding various obstacles either at the sides (left and right) or above or below, to collect the hidden gem. The participant uses the left, right, up and down arrow keys to move the dragon and avoid hitting the obstacles. The participant is instructed, however, that the keys have been reversed and they must use the opposite keys to the direction they want to move in i.e., press the left arrow key to move right. However, when they hear a beep sound, the keys will switch back to normal i.e., press the left key to move left. They are instructed that the keys will switch every time they hear the beep. The beep sounds every 30 seconds, and the game lasts a total of five minutes. If the wrong key is pressed and an obstacle is hit, an error is recorded, and the dragon will not move until the correct key is pressed. Performance was measured by the total number of errors made when the keys are the right way (*Dragon Dash Forward Errors*), reversed (*Dragon Dash Reversed Errors*) and the sum of those to give the overall number of errors (*Dragon Dash Total Errors*), along with distance covered within the five minutes (*Dragon Dash Distance*). Fewer errors and longer distance indicated better performance. Due to there not being a practice condition, errors made in the first minute of the subgame, which included one rule change, were not included when scoring the data.

2.4.1.2. Dragon Sequence

Dragon Sequence measures working memory, based on the WNV-SS backwards condition. Participants are presented with a grid of nine grey squares. The participants are instructed that a sequence of squares will light up in red, and they must reproduce the sequence in the reverse order by clicking on the appropriate squares, using the laptop trackpad, in order to collect the hidden gem. The participant is presented with two consecutive sequences of the same length, if at least one is reversed correctly, the sequence is increased by one square, up to a maximum of a nine-square sequence. Each square lights up red for one second before turning grey. Each sequence reversed incorrectly is marked as an error. If participants get two consecutive sequences wrong, the game concludes, and the participant collects the gem. Performance is measured by the longest correct sequence (*Dragon Sequence Longest Span*) and total number of correct sequences (*Dragon Sequence Total Correct*). More correct spans and longer longest spans indicated better performance. It was not possible to account for there not being a practice condition when scoring, due to the subgame concluding if two consecutive errors were made.

2.4.1.3. Dragon Hunt

Dragon Hunt measures task switching, based on the TMT (Reitan, 1992). The participant is presented with a scene where 40 red and blue eggs and crystals are scattered over the land. The participant is instructed to collect all eggs and crystals to collect the hidden gem. They are instructed, however, that they must collect the eggs and crystals in a specific order, they must start with the blue egg, then switch colour to the red egg and then switch shape to the red crystal and then switch colour to the blue crystal and keep switching shape and colour until they have collected them all. If the participant tries to collect an incorrect item, for example, if the participant walks into a red crystal when the next item is a red egg, this is marked as an error. When the participant collects the correct item, the item disappears accompanied by a sound. If the participant tries to collect the wrong item a different sound is produced, and the item remains on the screen. Performance was measured by total completion time (*Dragon Hunt Time*) and number of errors made (*Dragon Hunt Errors*). Fewer errors and shorter completion time indicated better performance. Due to there not being a practice condition, the number of errors made and the time taken to collect the first 10 items were not included when scoring the data.

2.4.2. Established Measures

2.4.2.1. Wechsler Nonverbal Scale of Ability Spatial Span

The WNV-SS (Wechsler & Naglieri, 2006), was used to measure working memory. The measure consists of a forward and a backward condition. For both conditions, the participant is presented with 10 wooden blocks arranged on an A4 size board. The examiner points to a sequence of blocks, starting with two blocks and gradually increasing, up to a maximum of a nine-block sequence. In the forward condition, the participant is required to repeat the sequence in the same order as the examiner. In the backward condition, the participant is required to repeat the sequence in the reverse order to the examiner.

Participants are given the opportunity to practice before each condition with two-block sequences.

The participant is presented with two sequences of the same number of blocks, if one or both sequences are repeated or reversed correctly, the number of blocks in the sequence is increased by one. This is repeated until the participant makes two consecutive errors within a sequence of the same length, or until they complete sequences of nine blocks. Scores were calculated according to the test instructions. Performance is measured by the total number of correct sequences (*WNV-SS Total Correct*) and the longest correct sequence (*WNV-SS Longest Span*) for both the forwards and backwards conditions. For this study, only the scores from the backwards condition were used, as they directly compare to the Dragon Sequence subgame.

Age-scaled scores were calculated, however, were only available for those aged 8 and above using the total correct score. Age-scaled scores were calculated for those aged 7 and below using norms from Grossi et al. (1979), however, these were only available for longest forward span. Therefore, only the total correct forward span and longest forward span scaled scores are reported below.

2.4.2.2. Developmental NEuroPSYchological Assessment-II Inhibition

The NEPSY-II Inhibition test (Korkman et al., 2007) was used to measure processing speed, inhibition and switching. It involves two similar tasks, both of which are made-up of three subcomponents (Naming, Inhibition and Switching). In the first task, participants are presented with black and white circles and squares. In the second task, participants are presented with black and white arrows, either pointing up or down. For each component of the tasks, participants are presented with 40 shapes or arrows, in five rows of eight.

The first component of the tasks, NEPSY-II Naming, requires participants to name each shape (e.g., say ‘circle’ if it is a circle) and the direction of the arrows they see (e.g., say ‘up’ if the arrow is pointing up), regardless of whether they are black or white, in the order presented on the page.

The second component, NEPSY-II Inhibition, requires participants to say the opposite shape (e.g., to say ‘circle’ if it is a square) and direction (e.g., to say ‘up’ if the arrow is pointing down) to the one presented. Again, regardless of whether the shape or arrow is black or white, and in the order presented.

The third component, NEPSY-II Switching, requires participants to switch between the rules from the first two components, depending on whether the shape or arrow is black or white. If the shape is black, participants are asked to say the shape they see (e.g., to say ‘circle’ if it is a black circle) and the direction the arrow is pointing (e.g., to say ‘up’ if it is a black arrow pointing up). If the shape is white, participants are asked to say the opposite shape (e.g., to say ‘circle’ if it is a white square) or direction (e.g., to say ‘up’ if it is a white arrow pointing down).

The examiner demonstrates each subcomponent on five shapes and arrows, which the participant is then asked to practice with. If the participant makes five uncorrected errors in the practice, they do not complete that component. Scores are calculated according to the test instructions. Performance is measured through the total completion time (*Time*) and errors made (*Errors*) across the two tasks for each component, Naming, Inhibition and Switching, and age-scaled scores are also generated for each component.

The NEPSY-II Inhibition test is recommended for use with children aged between 5 and 16, however, the Switching component of the tasks is not normed for children under 7. Due difficulty finding tests with concurrent validity suitable across the populations age range, it was decided that all components of the NEPSY-II Inhibition test would be used with children aged 6 years and older.

2.4.3. Childhood Executive Functioning Inventory

The CHEXI (Thorell & Nyberg, 2008) is a 24-item informant response-measure of executive function, which can be completed by teachers or caregivers. The measure comprises four subscales measuring inhibition, working memory, planning and regulation. Each item is rated from 1 (definitely not true) to 5 (definitely true). Scores are totalled for the subscales to give an overall score, with higher scores indicating increased executive function difficulties.

The CHEXI was completed by teachers to assess the predictive and ecological validity of Dragon Adventure. It would have been best to have both parental and teacher corroboration of real-world function, but for the current study this was not logistically feasible. Teacher ratings are considered to have good predictive validity due to being objective, functional and school-based (Mackinlay et al., 2006).

Through a factor analysis of the CHEXI, Catale et al. (2015) identified two factors in children aged 8 to 11 years. The two factors were Working Memory (WM), including the working memory and planning subscales, and Inhibition, including the inhibition and regulation subscales. Despite formal normative data not being available for this measure, Catale et al.'s (2015) research provides means, standard deviations (SD) and proposed cut-off scores for 242 neurotypical children aged 8 to 11 years; from which scaled scores were calculated. However, norms are not available for children aged 6 and 7, therefore, scaled scores were only calculated below for those aged 8 and above.

2.4.4. Children’s Technology Use Questionnaire

The CTUQ (Horton, 2013) was used to measure participants’ computer literacy. The CTUQ is a nine-item self-report measure developed to provide evidence of a child’s experience with a specific piece of technology, in this case a laptop or computer. Horton (2013) validated the questionnaire in 39 children aged 7 to 11 years. Good construct validity was found, along with promising test-retest analysis.

The questionnaire was designed to be adaptable, through the inclusion or omission of different questions depending on the research needs. Therefore, there is no standardised scoring for the questionnaire. In this study, the questionnaire was scored by allocating one point or more to each item that indicated further experience with a laptop or computer. A full breakdown of the scoring procedure can be found in Appendix B.

2.4.5. Visual Analogue Scale

The visual analogue scale used in Davis (2020) to measure the acceptability of Dragon Adventure and established measures, was also used in this research (please see Appendix C). The scale states “I enjoyed this task”. Children are asked to rate how much they agree with this statement on a 5-point scale using cartoon faces (sad, neutral and happy face). The scale will be used to measure acceptability and compare enjoyment ratings for Dragon Adventure versus the established measures. Participants will be asked to complete the scale twice, once for completing the established measures and once for completing Dragon Adventure.

2.4.6. Qualitative Question

Taking into consideration the criticism that typically feedback is not gained when developing measures for children (Berg et al., 2020), qualitative feedback was gathered on how the game could be further improved. Participants were asked “how could Dragon Adventure be improved?”, and their responses analysed through a content analysis.

2.5. Ethics

Ethical approval was gained from the University of East London School of Psychology Research Ethics Committee (Appendix D). The research was conducted in a mainstream primary school in Surrey and a specialist school for children diagnosed with an ASC in London. Members of the senior management teams were approached in both settings and consent was obtained to recruit through the organisations (organisation information sheet and consent form can be found in Appendix E and F). Permission was granted to recruit children aged 6 to 8 years from years one to four in the mainstream primary school, and children aged 6 to 11 years from all classes in the specialist primary school. Information and consent forms were distributed to the appropriate parents and/or carers (Appendix G and H) and returned to the school.

The schools were provided with the researchers Disclosure & Barring Service certificate so the researcher could meet with the children independently. The researcher met with the children of the parents and/or carers who had consented to their participation in the study and provided them with a child-friendly information sheet and consent form (Appendices I and J). Easy-read information sheets and consent forms were developed with the specialist school for children with communication difficulties (Appendices K and L). Children were provided with the opportunity to ask any questions before giving consent. Any children who did not give consent or withdrew consent during data collection did not take part in the remaining data collection and were not included in the research.

Full information was given to parents and/or carers and their children and no deception was used. Risk assessments were completed for settings in which the research was taking place, along with specific covid related risk assessments. The researcher wore a facemask at all times, completed regular sanitising of hands and equipment and took lateral flow tests prior to attending the sites to prevent the spread of infection. Participants were informed that they could take a break or stop at any time and the researcher monitored children for signs of discomfort, distress, and fatigue. Parents and/or carers were also informed that they could withdraw their child from the research before a given time point without having to give a reason.

Participant's data were anonymised through the allocation of a unique number, against which all data was collected. To ensure data could be withdrawn if requested, participant number allocations were stored in an independent, password-protected document, separate from the data collected. Paper consent and scoring forms were scanned and stored onto the researcher's password protected UEL One Drive and then destroyed.

2.6. Materials

Materials used included:

- a laptop to play Dragon Adventure
- the WNV-SS blocks on card
- the NEPSY-II stimulus book
- scoring sheets for the WNV-SS and NEPSY-II Inhibition test
- the CHEXI
- the CTUQ
- information sheets
- consent forms
- visual analogue scale
- a pen and pencil
- a table and chair
- a timer.

2.7. Procedure

Parental informed consent and child assent was obtained prior to participants completing the study. Data collection took place in a quiet room or area in the participant's school, with the researcher and participant sat at the same table at a 90-degree angle from one another. Participants were randomly allocated to an order in which to complete the tasks, either Dragon Adventure or the established measures first. Depending on their allocation, participants were initially presented with either Dragon Adventure or WNV-SS.

Dragon Adventure was presented on the laptop and participants progressed through the subgames until the game was complete, unless participants requested to withdraw. Participants were informed that they could ask questions and the researcher would clarify through repeating the relevant instruction. Following completion of Dragon Adventure, participants were presented with the visual analogue scale to rate how enjoyable they found the game. They were also asked to give verbal or written feedback on how the game could be improved.

The established measures were administered in the order of WNV-SS and then the NEPSY-II Inhibition tests: Naming, Inhibition and then Switching. The instructions for these measures were read to participants from the appropriate instruction manuals. Participants were informed that they could ask questions and the researcher would clarify through repeating the relevant instruction. Following completion of the established measures, participants were presented with the visual analogue scale to rate how enjoyable they found them.

Following completion of Dragon Adventure and the established measures, participants were given the CTUQ, the researcher read out the questions and assisted with completion depending on the participant's ability. On completion, participants were given the opportunity to ask any additional questions they had. Outside of the session, participant's teachers were asked to complete the CHEXI.

2.8. Participants

2.8.1. Demographics

Forty participants participated in this study in total, though not all participants completed every task. Therefore, participants were only included in analyses if they completed at least two of the Dragon Adventure subgames. Participants who were able to complete aspects of Dragon Adventure or the established measure, but not others will be discussed.

All 15 neurotypical children who participated were included in analysis. There were more males than females, 9 males and 6 females, however, a Chi-square test indicated that this difference was not significant, X^2 (1, N = 15) = .60, p = .305. Ages ranged from 6 years to 8 years and 11 months (M = 7.5, SD = 1.00). Twelve participants spoke English as their primary language (EPL) and 3 spoke English as an additional language (EAL); a Chi-square test indicated that this difference was significant, X^2 (1, N = 15) = 5.40, p = .035. Therefore, the difference in group size should be taken into account when examining the effects of being EPL or EAL. The majority of the sample (80%) were of White-

British ethnicity; a breakdown of ethnicities can be found in Table 1. Around half of the sample (53%) were considered by their teachers to be working in the average range of the national curriculum level; a breakdown of nation curriculum level can be found in Table 2.

Nineteen out of 25 children with a diagnosis of an ASC who participated were included in the analysis. There were 14 males and 5 females; a Chi-square test indicated that this difference was not significant, X^2 (1, N = 19) = 4.26, p = .064. Ages ranged from 6 years and 3 months to 11 years (M = 9.07, SD = 1.41). Eighteen participants spoke EPL and one spoke EAL; as a result of this difference, the effects of being EPL or EAL will not be examined in this sample. Just over a third of participants (37%) were of White-British ethnicity; a breakdown of ethnicities can be found in Table 3. The majority of the sample (89%) were considered by their teachers to be working below the average of their national curriculum level; a breakdown of nation curriculum level can be found in Table 4.

Table 1

Neurotypical Sample: Ethnicity

Ethnicity	N
White British	12
Indian	1
Brazilian	1
Slovak and Indian	1

Table 2

Neurotypical Sample: National Curriculum Level

National Curriculum Level	N
Above Average	1
Average – Above Average	1
Average	8
Average – Below Average	2
Below Average	3

Table 3

ASC Sample: Ethnicity

Ethnicity	N
White-British	7
Indian	6
Black-African	2
Black-Caribbean	1
Filipino	1
Other Asian Background	1
White Irish	1

Table 4

ASC Sample: National Curriculum Levels

National Curriculum Level	N
Above Average	1
Average – Above Average	1
Below Average	17

2.8.2. Sample Characteristics

Descriptive statistics of the participants' age-scaled scores on the WNV-SS Forwards Span and NEPSY-II Inhibition tasks are presented in Tables 5 and 7; please note that *WNV-SS Forward Span Total Correct* and *Longest Span* are presented together. Scaled scores are presented for sex and language group, along with overall sample. Descriptive statistics of CHEXI age-scaled scores for participants aged 8 and above are presented in Tables 6 and 8. Raw scores were converted to age-scaled scores using data provided in Catale et al. (2015).

Age-related scaled scores have a mean of 10 and a SD of 3. Scores indicate that neurotypical female participants overall scored above average and neurotypical male participants overall scored below average on the established measures. The neurotypical EPL and EAL groups scored similarly across the NEPSY-II Inhibition tasks, however, on the WNV-SS Forwards, the EPL group scored below average and the EAL group scored above average. Scores of the overall sample confirm that the neurotypical group were typically performing on these measures.

The CHEXI age-scaled scores for the neurotypical sample indicate that males scored much below average and females scored above average. This aligns with performance on the established measures; however, it is unclear why male participants scored so low on the CHEXI WM subscale. Those with EPL and EAL scored similarly on the CHEXI WM, both below average. Those with EPL scored in the average range on the CHEXI Inhibition, whereas those with EAL scored below average. Scores from the overall sample confirm that the neurotypical group were typically scoring on CHEXI Inhibition, however, below average on the CHEXI WM, which has been brought down by male participants' scores. This should be taken into account when examining relationships with the CHEXI.

Participants diagnosed with an ASC consistently scored below average on all established measures, however, male participants were close to average on NEPSY-II Switching. Male participants overall scored higher than female participants on the established measures.

The ASC sample scored much below average on the CHEXI. In line with the age-scaled scores on the established measures, female participants scored lower overall than male participants.

Table 5

Neurotypical Sample: Established Measures Age-scaled Scores

Test	Sex		Language		Overall Sample
M (SD)	Male	Female	EPL	EAL	
WNV-SS Forwards	6.98 (3.36)	12.80 (4.27)	8.44 (3.74)	11.33 (7.51)	9.06 (4.57)
NEPSY-II Naming	9.67 (3.54)	12.67 (3.67)	11.00 (3.44)	10.33 (5.86)	10.87 (3.78)
NEPSY-II Inhibition	7.56 (3.09)	12.50 (2.88)	9.25 (3.52)	10.67 (5.69)	9.53 (3.83)
NEPSY-II Switching	8.13 (2.53)	10.80 (4.44)	9.20 (3.26)	9.00 (5.00)	9.15 (3.48)

Table 6

Neurotypical Sample: CHEXI Age-scaled Scores for Participants Aged 8

		Sex		Language		
		Male	Female	EPL	EAL	Overall Sample
CHEXI	WM	2.89 (3.69)	10.83 (2.64)	5.92 (5.40)	6.67 (4.93)	6.07 (5.15)
	Inhibition	6.89 (2.67)	12.67 (2.25)	9.58 (3.48)	7.67 (5.50)	9.20 (3.80)

Table 7

ASC Sample: Established Measures Age-scaled Scores

Test M (SD)	Sex Male	 Female	Overall Sample
WNV-SSp Longest Forwards	7.91 (4.06)	5.80 (3.87)	7.25 (4.00)
NEPSY-II Naming	7.46 (4.16)	6.50 (2.38)	7.24 (3.77)
NEPSY-II Inhibition	7.55 (3.67)	6.75 (4.19)	7.33 (3.68)
NEPSY-II Switching	9.00 (3.57)	7.67 (4.72)	8.67 (3.70)

Table 8

ASC Sample: CHEXI Age-scaled Scores for Participants Aged 8 and Above

		Sex Male	 Female	Overall Sample
CHEXI	WM	3.18 (4.38)	1.40 (1.14)	2.62 (3.72)
	Inhibition	5.91 (2.66)	4.20 (1.30)	5.38 (2.42)

3. RESULTS

3.1. Methods of Analysis

3.1.1. Quantitative Analysis

The data was analysed using Statistical Package for the Social Sciences Version 26. Exploratory data analysis (EDA) was initially carried out to examine the data for errors, outliers and the normality of distributions, using histograms, skewness (<1), kurtosis (<3) and Shapiro-Wilk's test; a summary is given in Tables 9 and 10.

Due to a proportion of the data significantly deviating from the normal distribution, and the small sample sizes, non-parametric tests were used for analysis. Spearman's rank correlations were used to address the relationships between: Dragon Adventure and age; the relationships within and between the Dragon Adventure subgames; the relationships between Dragon Adventure and the established measures (WNV-SS and NEPSY-II Inhibition), the CHEXI and CTUQ. Spearman's rank partial correlations were used to control for computer literacy and processing speed when analysing the relationships between Dragon Adventure and the established measures. Effect sizes were interpreted according to Cohen's (1988) recommendations: small .10-.29; moderate .30-.49; and large ≥.50. Mann-Whitney U tests were used to compare mean scores on the Dragon Adventure subgames between the sex and language groups. Wilcoxon signed rank tests were used for within group means comparisons of the acceptability ratings on the visual analogue scale for Dragon Adventure and established measures.

When interpreting the results, it is important to take into consideration the small sample sizes, as this will impact the strength of conclusions that can be drawn.

3.1.1.1. Neurotypical Sample Data: Non-Normality

As seen in Table 9, *Dragon Dash Reversed Errors* was found to significantly deviate from the normal distribution by Shapiro-Wilk test. The histogram suggested the data was negatively sewed, with scores clustered around a higher number of errors and a long tail of lower numbers of errors. *Dragon Dash Distance* was also found to be negatively skewed, with scores clustered around longer distances covered.

Dragon Sequence Longest Span was found to significantly deviate from the normal distribution by Shapiro-Wilk test. The histogram suggested the data was slightly negatively skewed with scores clustered around higher longest spans, however, there was a small range.

Dragon Hunt Errors was found to be positively skewed with scores clustered around a lower number of errors. The data was also found to be kurtotic, with most participants making 0 to 20 errors and a flat distribution of the remaining error values.

WNV-SS Longest Span was found to significantly deviate from the normal distribution by Shapiro-Wilk test. The histogram suggested a slight negative skew, with the majority of participants scoring 4 as the longest span.

NEPSY-II Inhibition Errors was found to be positively skewed, with scores clustered around a lower number of errors. *NEPSY-II Switching Time* was found to be positively skewed and significantly deviate from the normal distribution by Shapiro-Wilk test, with scores clustered around a shorter time taken and a tail of longer times.

3.1.1.2. Autism Spectrum Condition Sample Data: Non-Normality

As seen in Table 10, *Dragon Sequence Total Correct* was positively skewed with scores clustered around a lower number of total correct. *Dragon Sequence Longest Correct* was found to significantly deviate from the normal distribution by Shapiro-Wilk test. The histogram suggested the data to be slightly positively skewed with most scores clustered around shorter spans, however, also bimodal with most participants scoring 0 or 2.

Dragon Hunt Errors was found to be positively skewed, with most scores clustered around a lower number of errors, and kurtotic as most participants made between 0 and 20 errors, with a flat distribution of the remaining error values.

WNV-SS Total Correct and *Longest Span* were found to significantly deviate from the normal distribution by Shapiro-Wilk test. The histogram suggested the *WNV-SS Total Correct* data to be slightly negatively skewed, with scores clustered around higher total scores, however, also bimodal with most participants scoring 2 and 5 and 6. The histogram suggested the *WNV-SS Longest Span* to be negatively skewed, with most participants scoring 4.

NEPSY-II Inhibition Errors was found to be positively skewed, with scores clustered around a lower number of errors. *NEPSY-II Inhibition Time* was found to significantly deviate from the normal distribution by Shapiro-Wilk test and be positively skewed, with scores clustered around a shorter time taken. *NEPSY-II Switching Errors* was found to significantly deviate from the normal distribution by Shapiro-Wilk test. The histogram suggested the data was slightly positively skewed, with scores clustered around lower number of errors.

Table 9

Neurotypical Sample: Summary of Exploratory Data Analysis

		Mean	SD	Min.	Max.	Skewness z-score	Kurtosis z-score	Shapiro-Wilk Sig. (2-sided)
Established Measures								
WNV-SS	Total Correct	4.40	1.60	1.00	7.00	-0.52	0.25	.46
	Longest Span	3.67	0.90	2.00	5.00	-0.58	-0.01	**.02**
NEPSY-II Inhibition	Naming Time	63.89	14.35	42.00	95.00	0.59	0.05	.49
	Inhibition Errors	7.73	5.66	1.00	22.00	**1.10**	1.51	.15
	Inhibition Time	96.04	22.68	60.33	137.77	-0.11	-0.65	.70
	Switching Errors	16.00	7.51	3.00	30.00	-0.08	-0.02	.99
	Switching Time	138.31	52.74	97.76	268.44	**1.89**	2.75	**.00**
CHEXI	WM	36.07	15.88	13.00	63.00	0.30	-1.02	.38
	Inhibition	27.67	10.33	13.00	47.00	0.15	-0.75	.62
	Total	63.73	23.86	26.00	93.00	-0.21	-1.49	.89
CTUQ	Total Score	10.13	3.14	6.00	17.00	0.72	0.29	.37
Dragon Adventure								
Dragon Dash	Total Errors	33.73	8.71	17.00	44.00	-0.83	-0.28	.09
	Forward Errors	13.33	4.75	6.00	21.00	0.05	-0.94	.34
	Reversed Errors	20.40	5.18	10.00	26.00	-0.93	-0.28	**.03**
	Distance	3523.63	647.05	1906.50	4340.03	**-1.11**	1.65	.14
Dragon Sequence	Total Correct	2.27	1.83	0.00	6.00	0.51	-0.37	.27
	Longest Span	2.27	1.34	0.00	4.00	-0.77	-0.34	**.01**
Dragon Hunt	Errors	12.73	15.35	0.00	54.00	**1.80**	**3.00**	**.00**
	Time	10:54	4:52	4:16	21:06	0.76	0.12	.38

Note. Statistics that identified non-normality are highlighted in **bold.**

Table 10

ASC Sample: Summary of Exploratory Data Analysis

		Mean	SD	Min.	Max.	Skewness z-score	Kurtosis z-score	Shapiro-Wilk Sig. (2-sided)
Established Measures								
WNV-SS	Total Score	3.88	2.28	0.00	7.00	-0.44	-1.17	**.05**
	Longest Span	3.13	1.59	0.00	5.00	-0.92	-0.15	**.01**
NEPSY-II Inhibition	Naming Time	66.58	17.96	41.99	101.89	0.54	-0.56	.41
	Inhibition Errors	9.20	8.63	1.00	32.00	**1.60**	2.47	**.01**
	Inhibition Time	93.18	23.89	66.86	155.75	**1.38**	2.09	**.05**
	Switching Errors	13.50	9.63	4.00	31.00	0.88	-0.45	**.04**
	Switching Time	115.99	17.59	92.40	153.04	0.60	0.36	.52
CHEXI	WM	45.69	12.12	19.00	62.00	-0.81	0.12	.29
	Inhibition	38.00	6.49	27.00	50.00	-0.17	-0.58	.67
	Total	83.69	15.09	56.00	104.00	-0.63	-0.55	.24
CTUQ	Total Score	11.00	4.24	2.00	17.00	-0.70	0.24	.24
Dragon Adventure								
Dragon Dash	Total Errors	27.37	8.56	10.00	39.00	-0.57	-0.28	.25
	Forward Errors	10.89	4.42	4.00	18.00	0.11	-1.29	.31
	Reversed Errors	16.47	5.26	6.00	25.00	-0.64	-0.34	.23
	Distance	3378.33	944.38	1267.80	4460.21	0.98	-0.17	.29
Dragon Sequence	Total Correct	2.00	2.65	0.00	8.00	**1.49**	1.15	**.00**
	Longest Correct	1.79	1.69	0.00	5.00	0.53	0.57	**.01**
Dragon Hunt	Errors	16.07	22.51	0.00	76.00	**1.89**	**3.21**	**.00**
	Time	7:45	3:32	2:21	14:07	0.23	-0.37	.65

Note. Statistics that identified non-normality are highlighted in **bold.**

3.1.2. Qualitative Analysis

Qualitative data collected from participants answering the question "How could Dragon Adventure be improved?" was analysed through a conventional conceptual content analysis (Hsieh & Shannon, 2005). The responses were collated into one document where the researcher examined and familiarised themselves with the data. From this, coding units were identified and used to analyse the data, through tallying the number of times a coding unit was mentioned in the responses.

3.2. Performance of Neurotypical Children

3.2.1. Performance on Dragon Adventure

3.2.1.1. Age

The relationship between performance on Dragon Adventure and age was analysed using Spearman's rank correlation. A summary of the correlation coefficients can be found in Table 11. A moderate positive correlation was found between *Dragon Dash Distance* and age. Small correlations and no associations were exhibited between the remaining Dragon Dash scores, and Dragon Hunt scores and age.

There is an unexpected moderate negative correlation between *Dragon Sequence Total Correct* and age, indicating that better scores were associated with younger age. Reviewing the scatterplot, seen in Figure 2, suggested that one of the youngest participants scored highest and two of the older participants scored lowest, with little association in-between

Table 11

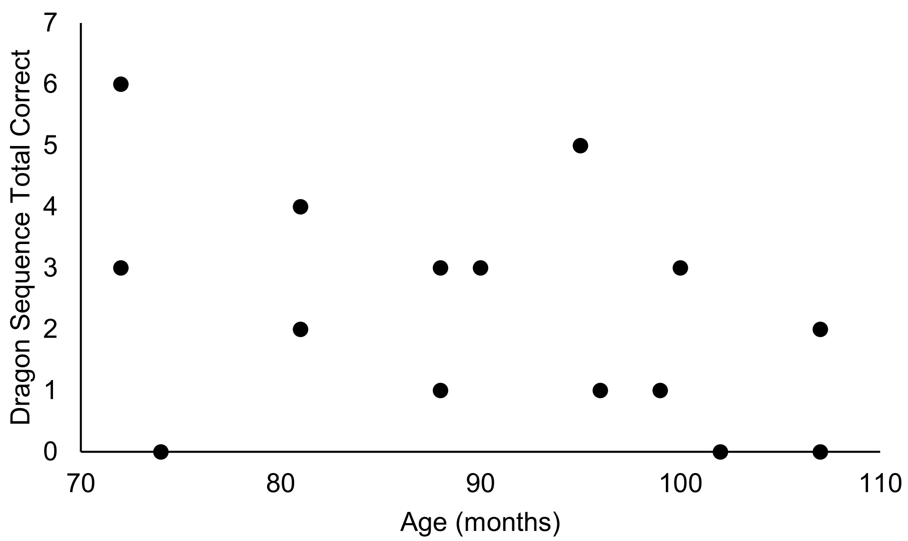
Neurotypical Sample: Spearman's Rank Correlation Coefficients Between Dragon Adventure Subgames and Age

	Dragon Dash				Dragon Sequence		Dragon Hunt	
	Forward Errors	Reversed Errors	Total Errors	Distance	Total Correct	Longest String	Errors	Time
Age	-.09	-.17	-.13	**.34**	**-.43**	-.29	-.07	-.20

Note. Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

Figure 2

Neurotypical Sample: Scatter Plot of Age (months) and Dragon Sequence Total Correct



3.2.1.2. Sex

A summary of the Dragon Adventure subgames means by sex group is presented in Table 12. Female participants performed better on every subgame, apart from *Dragon Dash Distance*, where male participants covered slightly more distance than female participants, and *Dragon Dash Forward Errors*, where male and female participants performed similarly. Mann-Whitney U tests confirmed that the differences in scores between male and female participants were not substantive (all $p > .05$).

3.2.1.3. English Language

A summary of the Dragon Adventure subgame means by language group is presented in Table 13. EAL performed better on Dragon Dash and Dragon Hunt, and EPL performed better on Dragon Sequence. Mann-Whitney U tests confirmed that the differences in scores between EPL and EAL participants were not substantive (all $p > .05$). These results must be interpreted with caution due to the significant difference in sample size between EPL and EAL.

Table 12

Neurotypical Sample: Mean Scores on Dragon Adventure Subgames by Sex

		Male	Female
		Mean (SD)	Mean (SD)
Dragon Dash	Total Errors	34.89 (6.97)	32.00 (11.35)
	Backwards Errors	21.56 (3.71)	18.67 (6.86)
	Forwards Errors	13.33 (4.36)	13.33 (5.72)
	Distance	3537.77 (496.17)	3502.43 (882.75)
Dragon Sequence	Total Correct	1.56 (1.51)	3.33 (1.86)
	Longest Span	1.78 (1.39)	3.00 (0.89)
Dragon Hunt	Errors	18.22 (17.73)	4.50 (4.59)
	Time	11:41 (6:04)	10:24 (2:39)

Table 13

Neurotypical Sample: Mean Scores on Dragon Adventure Subgames by Language Group

		EPL	EAL
		Mean (SD)	Mean (SD)
Dragon Dash	Total Errors	35.50 (7.60)	26.67 (10.97)
	Backwards Errors	21.25 (4.45)	17.00 (7.55)
	Forwards Errors	14.25 (4.63)	9.67 (3.79)
	Distance	3462.43 (678.93)	3768.45 (539.01)
Dragon Sequence	Total Correct	2.50 (1.88)	1.33 (1.52)
	Longest Span	2.42 (1.31)	1.67 (1.53)
Dragon Hunt	Errors	13.08 (16.63)	11.33 (11.15)
	Time	11:38 (4:59)	7:54 (3:34)

3.2.2. Associations Among Dragon Adventure Subgames

The relationships within and between the Dragon Adventure subgames were analysed using Spearman's rank correlations. A summary of the correlation coefficients can be found in Table 14.

Dragon Dash scores exhibited moderate to large correlations with one another. The correlations between Dragon Dash error scores and *Dragon Dash Distance* were negative, suggesting that as the number of errors increased, the distance covered decreased.

Dragon Sequence Total Correct and *Longest Span* exhibited a large positive correlation. With such a large effect, it could be assumed that, in this sample, these two scores are underpinned by the same function.

Dragon Hunt Errors exhibited a large positive correlation with *Dragon Hunt Time*; thus, as the number of errors increased, the taken to complete the task also increased.

3.2.2.1. Associations Between Subgames

Moderate correlations were observed between *Dragon Dash Total Errors* and Dragon Sequence scores, and *Dragon Dash Forward Errors* and *Dragon Sequence Longest Span*. The remaining Dragon Dash scores exhibited small correlations with Dragon Sequence scores, and no association was found between *Dragon Dash Distance* and *Dragon Sequence Total Correct*.

Dragon Dash scores were moderately to largely correlated with Dragon Hunt scores (*Errors and Time*); thus, better scores on Dragon Dash were associated with better scores on Dragon Hunt.

Dragon Sequence scores and Dragon Hunt scores exhibited no associations and small negative correlations with one another.

Table 14

Neurotypical Sample: Spearman's Rank Correlation Coefficients Within and Between Dragon Adventure Subgames

		Dragon Dash				Dragon Sequence		Dragon Hunt	
		FE	RE	TE	D	TC	LS	E	T
DD	Forward Errors	1.00							
	Reversed Errors	**.47**	1.00						
	Total Errors	****.88**	****.79**	1.00					
	Distance	****-.80**	***-.61**	****-.85**	1.00				
DS	Total Correct	-.29	-.19	**-.32**	.09	1.00			
	Longest Span	**-.36**	-.27	**-.42**	.21	****.96**	1.00		
DH	Errors	**.32**	**.45**	**.38**	**-.46**	-.23	-.28	1.00	
	Time	***.63**	***.59**	****.67**	***-.61**	.02	-.12	***.63**	1.00

Note. Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

3.2.3. Associations Between Dragon Adventure and Established Performance-Based Measures

The relationships between the Dragon Adventure subgames and the established measures were analysed using Spearman's rank correlations. A summary of the correlation coefficients can be found in Table 15.

3.2.3.1. Dragon Dash

Dragon Dash scores were moderately correlated with NEPSY-II Inhibition scores; thus, better scores on Dragon Dash were associated with better scores on NEPSY-II Inhibition.

Dragon Dash error scores were moderately to largely correlated with WNV-SS scores; thus, better scores on Dragon Dash were associated with better scores on WNV-SS. However, *Dragon Dash Distance* was only slightly correlated with WNV-SS scores.

Most Dragon Dash scores were moderately to largely correlated with NEPSY-II Switching scores; thus, better scores on Dragon Dash were associated with better scores on NEPSY-II Switching. However, *Dragon Dash Forward Errors*, was only slightly correlated to *NEPSY-II Switching Time*.

3.2.3.2. Dragon Sequence

Dragon Sequence scores were moderately to largely correlated with WNV-SS scores; thus, better scores on Dragon Sequence were associated with better scores on the WNV-SS.

Dragon Sequence Longest Span was moderately correlated with *NEPSY-II Inhibition Errors.* However, the remaining correlations between Dragon Sequence scores and NEPSY-II Inhibition scores were small or exhibited no association.

Dragon Sequence scores showed no association and small correlations with NEPSY-II Switching scores.

3.2.3.3. Dragon Hunt

Dragon Hunt Time was moderately correlated with NEPSY-II Switching scores. However, *Dragon Hunt Errors* and NEPSY-II Switching scores exhibited only small correlations.

Dragon Hunt Errors exhibited moderate negative correlations with WNV-SS scores. However, *Dragon Dash Distance* was only slightly correlated with WNV-SS scores.

Most Dragon Hunt scores were moderately to largely correlated with NEPSY-II Inhibition scores; thus, better performance on Dragon Hunt was associated with better performance on NEPSY-II Inhibition. However, *Dragon Hunt Errors* and *NEPSY-II Inhibition Time* exhibited a small correlation.

Table 15

Neurotypical Sample: Spearman's Rank Correlation Coefficients Between Dragon Adventure Subgames and Established Measures

		WNV-SS		NEPSY-II Inhibition		NEPSY-II Switching	
		Total Correct	Longest Span	Errors	Time	Errors	Time
Dragon Dash	Forward Errors	-.26	**-.32**	**.40**	**.32**	****.82**	.18
	Reversed Errors	***-.61**	***-.55**	**.44**	**.38**	***.67**	**.36**
	Total Errors	**-.50**	**-.50**	**.48**	**.40**	****.80**	**.31**
	Distance	.14	.18	**-.40**	**-.37**	***-.66**	**-.51**
Dragon Sequence	Total Correct	**.5**	**.45**	-.23	.12	.01	-.11
	Longest Span	***.55**	**.44**	**-.39**	.09	-.15	-.13
Dragon Hunt	Errors	**-.39**	**-.34**	****.72**	.26	.16	.28
	Time	-.29	-.27	****.69**	**.36**	**.38**	**.30**

Note. Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

3.2.4. Controlling for Computer Literacy and Processing Speed

Moderate to large correlations found between Dragon Adventure subgames and the established measures were further explored using Spearman's rank partial correlation, to see how the relationship strength was affected when controlling for:

- Computer literacy (measured by the CTUQ)
- Processing speed (measured by the NEPSY-II Naming Time).

A summary of the correlation coefficients when controlling for the CTUQ and NEPSY-II Naming Time can be found in Tables 16 and 17, respectively.

3.2.4.1. Computer Literacy

As seen in Table 16, most correlations between scores on the Dragon Adventure subgames and scores on the established measures were moderate to large, indicating that the original correlations held up when controlling for CTUQ score. Only the correlation between *Dragon Sequence Longest Span* and *NEPSY-II Inhibition Errors* was small, indicating that CTUQ score partially explained the original correlation.

3.2.4.2. Processing Speed

As seen in Table 17, the majority of correlations between Dragon Dash scores and NEPSY-II Inhibition scores were small or exhibited no association, indicating that NEPSY-II Naming Time partially or fully explained the original correlations. The correlations between Dragon Dash scores and *WNV-SS Total Correct* were modest, however, correlations between Dragon Dash scores and *WNV-SS Longest Span* were small to modest. Most correlations between Dragon Dash scores and NEPSY-II Switching scores were moderate to large, indicating that the original correlations held up when controlling for NEPSY-II Naming Time.

The correlations between Dragon Sequence scores and scores on the established measures were modest to large, indicating that the original correlations held up well when controlling for NEPSY-II Naming Time.

Moderate correlations were observed between *Dragon Hunt Errors* and *WNV-SS Total Correct*, and Dragon Hunt scores and NEPSY-II Inhibition Errors, indicating that the original correlations held up when controlling for NEPSY-II Naming Time. However, the remaining correlations between Dragon Hunt scores and scores on the established measures were small or exhibited no association, indicating that NEPSY-II Naming Time partially or fully explained the original correlations.

Tables summarising the associations between the Dragon Adventure subgames and CTUQ score, and NEPSY-II Naming Time can be found in Appendix M and N, respectively.

Table 16

Neurotypical Sample: Spearman's Rank Partial Correlation Coefficients Between Dragon Adventure Subgames and Established Measures when Controlling for Computer Literacy

		WNV-SS		NEPSY-II Inhibition		NEPSY-II Switching	
		Total Correct	Longest Span	Errors	Time	Errors	Time
Dragon Dash	Forward Errors	-	**-.39**	**.39**	**.51**	****.84**	-
	Reversed Errors	***-.56**	**-.49**	**.41**	**.31**	**.42**	**.43**
	Total Errors	**-.43**	**-.44**	**.48**	**.39**	***.69**	**.44**
	Distance	-	-	**-.45**	***-.56**	**-.54**	****-.73**
Dragon Sequence	Total Correct	***.54**	**.48**	-	-	-	-
	Longest Span	**.48**	**.41**	-.27	-	-	-
Dragon Hunt	Errors	**-.48**	**-.44**	**.48**	-	-	-
	Time	-	-	***.60**	***.56**	**.40**	-

Note. Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

Table 17

Neurotypical Sample: Spearman's Rank Partial Correlation Coefficients Between Dragon Adventure Subgames and Established Measures when Controlling for Processing Speed

		WNV-SS		NEPSY-II Inhibition		NEPSY-II Switching	
		Total Correct	Longest Span	Errors	Time	Errors	Time
Dragon Dash	Forward Errors	-	-.15	.08	-.10	****.83**	-
	Reversed Errors	**-.48**	**-.36**	**.31**	-.05	**.32**	.22
	Total Errors	**-.34**	-.22	.29	-.27	***.58**	.03
	Distance	-	-	-.19	.28	**-.34**	**-.33**
Dragon Sequence	Total Correct	***.58**	***.56**	-	-	-	-
	Longest Span	**.50**	**.45**	**-.32**	-	-	-
Dragon Hunt	Errors	**-.39**	-.29	**.39**	-	-	-
	Time	-	-	**.43**	.01	.15	-

Note. Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

3.2.5. Associations with the CHEXI

The relationships between the Dragon Adventure subgames and the CHEXI factors were analysed using Spearman's rank. A summary of the correlation coefficients can be found in Table 18.

Dragon Dash Reversed Errors was moderately correlated with CHEXI Inhibition; however, the remaining Dragon Dash scores were only slightly correlated or exhibited no association with the CHEXI Inhibition and WM scores.

Dragon Sequence scores were moderately correlated with CHEXI Inhibition and WM scores; thus, better scores on Dragon Sequence were associated with lower scores on the CHEXI (lower scores indicating decreased executive function difficulties).

Dragon Hunt Errors was moderately to largely correlated with CHEXI Inhibition and WM score. Whereas *Dragon Hunt Time* was only slightly correlated with CHEXI inhibition and WM.

Table 18

Neurotypical Sample: Spearman's Rank Correlation Coefficients Between Dragon Adventure Subgames and the CHEXI

	Dragon Dash				Dragon Sequence		Dragon Hunt	
	Forwards Errors	Reversed Errors	Total Errors	Distance	Total Correct	Longest Span	Errors	Time
CHEXI Inhibition	.13	**.33**	.27	-.24	**-.37**	**-.47**	****.65**	.29
CHEXI WM	.19	.29	.28	.02	**-.40**	**-.31**	**.47**	.23

Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

3.2.6. Acceptability

The descriptive statistics for the visual analogue scale can be seen in Table 19. Both established measures and Dragon Adventure were found to be acceptable, with Dragon Adventure being rated very slightly higher. Wilcoxon signed rank test confirmed that the difference between the mean ratings was not substantive (N = 15, Z = -1.58, p = .188).

Table 19

Neurotypical Sample: Descriptive Statistics for Acceptability Ratings of Established Measures and Dragon Adventure

	Mean	SD
Established Measures	4.47	0.74
Dragon Adventure	4.77	0.42

3.3. Performance of Children with an Autism Spectrum Condition

3.3.1. Performance on Dragon Adventure

3.3.1.1. Age

The relationship between performance on Dragon Adventure and age was analysed using Spearman's rank correlation. A summary of the correlation coefficients can be found in Table 20. Dragon Sequence scores and most Dragon Dash scores exhibited moderate to large correlations with age; thus, better performance on Dragon Sequence and Dragon Dash was associated with older age. However, only a small correlation was found between *Dragon Dash Reversed Errors* and age. Small correlations were exhibited between Dragon Hunt scores and age.

Table 20

ASC Sample: Spearman's Rank Correlation Coefficient Between Dragon Adventure Subgames and Age

	Dragon Dash				Dragon Sequence		Dragon Hunt	
	Forward Errors	Reversed Errors	Total Errors	Distance	Total Correct	Longest String	Errors	Time
Age	**-.60	-.22	**-.44**	****.67**	***.49**	***.46**	.24	.15

Note. Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

3.3.1.2. Sex

A summary of the Dragon Adventure subgame means by sex group is presented in Table 21. Male participants performed better on all subgames, apart from Dragon Hunt, where female participants completed the game in less time and with less errors. Mann-Whitney U tests confirmed that the differences in scores between male and female participants were not substantive (all p > .05).

Table 21

ASC Sample: Mean Scores on Dragon Adventure Subgames by Sex

		Male Mean (SD)	Female Mean (SD)
Dragon Dash	Total Errors	26.00 (9.14)	31.20 (5.76)
	Backwards Errors	15.86 (5.93)	18.20 (2.28)
	Forwards Errors	10.14 (4.45)	13.00 (4.00)
	Distance	3373.99 (1053.17)	3390.50 (638.96)
Dragon Sequence	Total Correct	2.36 (2.87)	1.00 (1.73)
	Longest Span	2.07 (0.27)	1.00 (1.41)
Dragon Hunt	Errors	19.36 (24.54)	4.00 (2.65)
	Time	8:02 (3:37)	6:43 (3:40)

3.3.2. Associations Among Dragon Adventure Subgames

The relationships within and between the Dragon Adventure subgames were analysed using Spearman's rank. A summary of the correlation coefficients can be found in Table 22.

Dragon Dash scores (*Forward, Reversed and Total Errors,* and *Distance*) mostly exhibited large correlations with one another, apart from *Reversed Errors* and *Distance*, which exhibited a small correlation. Due to the size of the effect between *Dragon Dash Reversed Errors* and *Total Errors,* they could be considered to measure the same function in this sample. The correlations between Dragon Dash error scores and *Dragon Dash Distance* were negative; thus, as the number of errors increased, the distance covered decreased.

Dragon Sequence Total Correct and *Longest Span* exhibited a large positive correlation with one another, which due to the size of the effect could be considered to measure the same function in this sample.

Dragon Hunt Errors and *Time* exhibited a large positive correlation between one another; thus, as the number of errors increased the time taken to complete the task increased.

3.3.2.1. Associations Between Subgames

Dragon Dash scores and Dragon Sequence scores all exhibited large correlations with one another; thus, better scores on Dragon Dash were associated with better scores on Dragon Sequence.

Dragon Dash Reversed Errors and *Dragon Dash Total Errors* both exhibited large positive correlations with *Dragon Hunt Time*. Small correlations were observed between the remaining Dragon Dash and Dragon Hunt scores, apart from *Dragon Dash Distance* which exhibited no association with Dragon Hunt scores.

No associations and small correlations were found between Dragon Sequence and Dragon Hunt scores.

Table 22

ASC Sample: Spearman's Rank Correlation Coefficients Within and Between Dragon Adventure Subgames

		Dragon Dash				Dragon Sequence		Dragon Hunt	
		FE	RE	TE	D	TC	LS	E	T
DD	Forward Errors	1.00							
	Reversed Errors	***.60**	1.00						
	Total Errors	****.86**	****.90**	1.00					
	Distance	****-.74**	-.25	***-.54**	1.00				
DS	Total Correct	****-.70**	****-.71**	****-.80**	***.55**	1.00			
	Longest Span	****-.63**	***-.52**	****-.64**	***.47**	****.93**	1.00		
DH	Errors	.16	**.30**	.28	-.05	.06	-.05	1.00	
	Time	.23	***.58**	**.51**	-.03	-.22	-.16	****.80**	1.00

Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

3.3.3. Associations Between Dragon Adventure and Established Performance-Based Measures

The relationships between the Dragon Adventure subgames and the established measures were analysed using Spearman's rank correlation. A summary of the correlation coefficients can be found in Table 23.

3.3.3.1. Dragon Dash

Dragon Dash scores were moderately to largely correlated with both NEPSY-II Inhibition and WNV-SS scores; thus, better scores on Dragon Dash were associated with better scores on NEPSY-II Inhibition and WNV-SS.

Most Dragon Dash scores were moderately to largely correlated with *NEPSY-II Switching Time*, however, exhibited only small to moderate correlations with *NEPSY-II Switching Errors*.

3.3.3.2. Dragon Sequence

Dragon Sequence scores were slightly to moderately correlated with both WNV-SS scores and NEPSY-II Inhibition scores.

No associations and small correlations were exhibited between Dragon Sequence scores and NEPSY-II Switching scores.

3.3.3.3. Dragon Hunt

Dragon Hunt scores were slightly to moderately correlated with *NEPSY-II Switching Time*, and no associations were found between Dragon Hunt scores and *NEPSY-II Switching Errors*.

Dragon Hunt scores were largely correlated with WNV-SS scores; thus, better scores on Dragon Hunt were associated with better scores on the WNV-SS.

Dragon Hunt Time was moderately correlated with NEPSY-II Inhibition scores; however, no associations were found between *Dragon Hunt Errors* and NEPSY-II Inhibition scores.

Table 23

ASC Sample: Spearman's Rank Correlation Coefficients Between Dragon Adventure and Established Measures

		WNV-SS		NEPSY-II Inhibition		NEPSY-II Switching	
		Total Correct	Longest Span	Errors	Time	Errors	Time
Dragon Dash	Forward Errors	**-.49**	***-.52**	**.39**	.26	.12	.10
	Reversed Errors	***-.54**	**-.42**	***.58**	***.52**	.21	**.43**
	Total Errors	***-.57**	***-.51**	***.54**	**.50**	.20	**.34**
	Distance	****.69**	****.65**	-.**42**	***-.57**	**-.35**	**-.56**
Dragon Sequence	Total Correct	**.37**	**.40**	**-.47**	-.24	-.08	-.07
	Longest Span	.19	.20	**-.45**	-.10	-.12	.09
Dragon Hunt	Errors	****-.82**	****-.84**	.08	.07	-.02	.25
	Time	****-.70**	***-.60**	**.42**	**.35**	.01	**.30**

Note. Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

3.3.4. Controlling for Computer Literacy and Processing Speed

Moderate to large correlations found between scores on the Dragon Adventure subgames and on the established measures were further explored using Spearman's rank partial correlation, to see how the relationship strength was affected when controlling for:

- Computer literacy (measured by the CTUQ)
- Processing speed (measured by the NEPSY-II Naming Time).

A summary of the correlation coefficients when controlling for CTUQ score and NEPSY-II Naming time can be found in Tables 24 and 25, respectively.

3.3.4.1. Computer Literacy

As seen in Table 24, the majority of correlations between Dragon Dash scores and scores on the established measures were modest to large, indicating that the original correlations held up when controlling for CTUQ score. The correlations between *Dragon Dash Forward Errors* and *NEPSY-II Inhibition*

Errors, and *Dragon Dash Total Errors* and *NEPSY-II Inhibition Time,* were small, indicating that CTUQ score partially explained the original correlations.

Correlations between Dragon Sequence scores and scores on the established measures were small, with no association found between *Dragon Sequence Total Correct* and *WNV-SS Total Correct*, indicating the original correlations were partially or fully explained by CTUQ score.

Correlations between Dragon Hunt scores and WNV-SS scores were moderate to large, indicating that the original correlations held up when controlling for CTUQ score. The correlations between Dragon Hunt scores and both NEPSY-II Inhibition and NEPSY-II Switching scores were small, indicating that CTUQ score partially explained the original correlations.

3.3.4.2. Processing Speed

As seen in Table 25, the majority of correlations between Dragon Dash scores and scores on established measures were moderate to large, indicating that the original correlations help up when controlling for NEPSY-II Naming Time. Correlations between *NEPSY-II Switching Time* and both *Dragon Dash Reversed Errors,* and *Total Errors* were small, indicating that NEPSY-II Naming Time partially explained the original correlations.

The correlations between Dragon Sequence scores and scores on established measures exhibited no association and small correlations, suggesting that NEPSY-II Naming Time partially or fully explained the original correlations.

The majority of correlations between Dragon Hunt scores and scores on established measures were moderate to large, indicating that the original correlations held up when controlling for NEPSY-II Naming Time. However, the correlation between *Dragon Hunt Time* and *WNV-SS Longest Span* was small, and no association was found between *Dragon Hunt Time* and *NEPSY-II Switching Time,* indicating that NEPSY-II Naming Time partially and fully explained the original correlations.

Table 24

ASC Sample: Spearman's Rank Partial Correlation Coefficients Between Dragon Adventure and Established Measures Controlling for Computer Literacy

		WNV-SS		NEPSY-II Inhibition		NEPSY-II Switching	
		Total Correct	Longest Span	Errors	Time	Errors	Time
Dragon Dash	Forward Errors	-.29	**-.44**	.17	**-**	-	-
	Reversed Errors	**-.38**	**-.41**	**.48**	**.30**	-	**.40**
	Total Errors	**-.37**	**-.50**	**.38**	.23	-	**.33**
	Distance	****.67**	****.71**	**-.50**	**-.51**	**-.46**	**-.51**
Dragon Sequence	Total Correct	-.00	.17	-.24	-	-	-
	Longest Span	-	-	-.26	-	-	-
Dragon Hunt	Errors	****-.76**	***-.61**	**-**	-	-	-
	Time	**-.52**	**-.31**	.15	.13	-	.11

Note. Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

Table 25

ASC Sample: Spearman's Rank Partial Correlation Coefficients Between Dragon Adventure and Established Measures Controlling for Processing Speed

		WNV-SS		NEPSY-II Inhibition		NEPSY-II Switching	
		Total Correct	Longest Span	Errors	Time	Errors	Time
Dragon Dash	Forward Errors	**-.36**	**-.42**	**.45**	**-**	**-**	-
	Reversed Errors	**-.45**	**-.32**	****.77**	**.49**	**-**	.24
	Total Errors	**-.43**	**-.43**	****.67**	**.32**	**-**	.20
	Distance	****.66**	***.62**	****-.78**	**-.51**	****-.88**	**-.31**
Dragon Sequence	Total Correct	.13	.22	**-.40**	-	-	-
	Longest Span	**-**	**-**	**-.39**	-	-	-
Dragon Hunt	Errors	****-.76**	***-.58**	**-**	-	-	-
	Time	***-.58**	-.29	**.50**	**.31**	**-**	.04

Note. Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

Tables summarising the associations between the Dragon Adventure subgames and CTUQ score, and NEPSY-II Naming Time can be found in Appendix M and N, respectively.

3.3.5. Associations with the CHEXI

The relationships between the Dragon Adventure subgames and the CHEXI factors were analysed using Spearman's rank correlation. A summary of the correlation coefficients can be found in Table 26.

Dragon Dash Reversed Errors was moderately correlated with CHEXI Inhibition. Whereas *Dragon Dash Forward Errors* and *Total Errors* were moderately correlated with CHEXI WM. The remaining correlations were small, or no association was found.

Unexpectedly, Dragon Sequence scores were moderately correlated with CHEXI Inhibition, and exhibited a small correlation and no association with CHEXI WM.

Dragon Hunt Time was moderately correlated with CHEXI Inhibition and WM, however, *Dragon Hunt Errors* exhibited small and no association with CHEXI scores.

Table 26

ASC Sample: Spearman's Rank Correlation Coefficients Between Dragon Adventure and the CHEXI the CHEXI

	Dragon Dash				Dragon Sequence		Dragon Hunt	
	Forwards Errors	Reversed Errors	Total Errors	Distance	Total Correct	Longest Span	Errors	Time
CHEXI Inhibition	.16	**.35**	.29	-.09	**-.39**	**-.37**	.27	**-.35**
CHEXI WM	**.32**	.26	**.31**	-.27	-.11	-.06	.06	**.41**

Note. Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

3.3.6. Acceptability

The descriptive statistics for the visual analogue scale can be seen in Table 27. Both established measures and Dragon Adventure were found to be acceptable, with Dragon Adventure being rated very slightly higher. Wilcoxon signed rank test confirmed that the difference was not substantive ($N = 16$, $Z = -1.19$, $p = .344$).

Table 27

ASC Sample: Descriptive Statistics for Acceptability Ratings for Established Measures and Dragon Adventure

	Mean	SD
Established Measures	4.25	1.18
Dragon Adventure	4.56	1.04

3.4. Content Analysis

Twenty-eight out of the 34 participants who completed at least two of the Dragon Adventure subgames (82%) answered the question “how could Dragon Adventure be improved?”. Ten (35%) of the respondents stated “nothing”, and 18 (64%) gave suggestions for how Dragon Adventure could be improved. These suggestions were analysed using a conventional conceptual content analysis. A summary of the variable codes and how often they were mentioned in the responses is presented in Table 28.

Half of the respondents who made suggestions (50%) mentioned improving the controls. A third of respondents (33%) mentioned being able to see or engage with additional characters. Of the 27 percent who mentioned the length of the game, four respondents suggested the game be longer or have additional games and one respondent suggested the game be shorter. Of the 27 percent of respondents who mentioned including a specific additional game, two mentioned a flying game, two a chase game and one a maths game. Just less than a quarter of respondents (22%) suggested improvements to the gems, either through suggesting the use of different types of gems or different ways to use the gems in the game. Twenty-two percent also suggested changes to the dragon character, including having the dragon able to fly, breathe fire and fight. Another 22 percent suggested to changes to the game environment or programming. A small number of respondents (11%) suggested having more obstacle courses (referring to Dragon Dash) and making amendments to the game that would make it easier, such as showing which control to press on Dragon Dash and having the eggs and crystals in the correct order on Dragon Hunt. Descriptions and examples of the coding units can be found in Appendix O.

Table 28

Summary of Content Analysis Variable Codes and Frequency Mentioned in Responses

	Frequency	
Codes	No. of participants	Percentage
Game Controls	9	50%
Additional characters	6	33%
Length	5	27%
Additional game	5	27%
Gems	4	22%
Character functions	4	22%
Environment	4	22%
Obstacle courses	2	11%
Ease	2	11%

4. DISCUSSION

The final chapter will summarise the study's findings according to the research aims and questions and in relation to previous research. The strengths and limitations of the current study will be discussed. Both the theoretical and clinical implications will be presented, followed by suggestions for future research. The chapter will conclude with reflections on the professional and ethical issues related to this study, personal reflections on the process and a final summary.

4.1. Aims, Research Questions and Hypotheses

The aim of the current study was to establish the feasibility, acceptability, concurrent and ecological validity of the newly developed game-based measure of executive function, Dragon Adventure (Davis, 2020), in neurotypical children aged 6 to 8 years and in children with a diagnosis of an ASC aged between 6 and 11 years. The study also aimed to establish whether Dragon Adventure improves engagement, compared to established measures of executive function; and gain feedback on how the game could be improved. The research questions were as follows:

- Does Dragon Adventure assess inhibition, working memory and switching in line with established measures of executive function and teacher ratings in neurotypical children aged 6 to 8 years?
- Is Dragon Adventure suitable for assessing inhibition, working memory and switching in line with established measures of executive function and teacher ratings for children aged 6 to 12 years with a diagnosis of an ASC?
- Is there a difference in rated enjoyment between the established measures and Dragon Adventure in neurotypical children aged 6 to 8 years and children aged 6 to 12 years diagnosed with an ASC?
- How do children feel Dragon Adventure could be improved?

Based on previous research findings, it was hypothesised that the Dragon Adventure subgames would be moderately to largely correlated with established measures and teacher ratings of executive function in both the neurotypical and ASC sample, and that Dragon Adventure would be rated as more enjoyable than the established measures of executive function.

4.2. Summary and Interpretation of Results

4.2.1. Does Dragon Adventure Assess Inhibition, Working Memory and Switching in line with Established Measures of Executive Function?

The feasibility and concurrent validity of Dragon Adventure was assessed by analysing the relationship between participants scores on the subgames and their scores on the established measures, through Spearman's rank correlations. The results partially supported the hypothesis that Dragon Adventure subgames would be moderately to largely associated with established measures, in both samples.

Task impurity was partially assessed through controlling for computer literacy and processing speed when analysing the relationships between Dragon Adventure and the established measures. Moderate to large correlations found between the Dragon Adventure subgame scores and scores on the established measures were analysed through Spearman's rank partial correlations to assess whether CTUQ score and NEPSY-II Naming Time impacted the associations.

Dragon Dash was based on the Stroop paradigm to measure inhibition. Performance on this subgame was compared to the NEPSY-II Inhibition task. Dragon Sequence was developed to measure working memory and was compared to the WNV-SS backwards; the established measure it was based on. Dragon Hunt was based on the TMT to assess task switching, and performance was compared to the NEPSY-II Switching task. Associations between all Dragon Adventure subgames and established measures were analysed.

When interpreting the results in real-world and clinical practice, it is important to consider the impact of non-cognitive and additional factors, which are not always possible to account for, such as motivation, attention and emotional status. This is especially important to consider when working with those who are less able to communicate their needs. These factors were monitored throughout data collection and appropriate actions were taken, such as allowing for a break or arranging another time to complete the data collection. However, it is not always possible to fully account for these factors.

4.2.1.1. Neurotypical Children Aged 6 to 8

In the neurotypical sample, Dragon Dash was moderately associated with NEPSY-II Inhibition, however, the majority of the correlations were partially or fully explained by processing speed. Only the association between *Dragon Dash Reversed Errors* and *NEPSY-II Inhibition Errors* remained modest, indicating that *Reversed Errors* could be the most valid measure of inhibition from Dragon Dash. Moderate to large associations were found between Dragon Sequence and WNV-SS, and all of these remained modest to large when controlling for computer literacy and processing speed, suggesting Dragon Sequence is a valid measure of working memory. Dragon Hunt, on the other hand, shared small to moderate associations with NEPSY-II Switching, all of which were partially or fully explained by processing speed. The moderate to large associations found between both Dragon Dash and Dragon Sequence and their related established measures indicate that Dragon Adventure could be a valid measure of inhibition and working memory for children aged 6 to 8 years. However, Dragon Hunt does not show evidence of being a valid measure of switching.

Interestingly, Dragon Dash was strongly associated with NEPSY-II Switching and Dragon Hunt was strongly associated with NEPSY-II Inhibition, with most associations remaining modest to large after controlling for computer literacy and processing speed. Dragon Dash and Dragon Hunt were also found to be strongly associated with one another, suggesting these measures could all be underpinned by a related function.

From these results, it could be argued that Dragon Dash is underpinned by both inhibition and switching. Dragon Dash does require the participant to switch between two rules (using the keys the right way round versus reversing them), similarly to the NEPSY-II Switching task (say the correct shape if it is black versus the opposite shape if it is white). The NEPSY-II Switching is part of the NEPSY-II Inhibition battery and, therefore, each subtask does include elements of inhibition. Both Dragon Dash and NEPSY-II Switching require the participants to inhibit prepotent responses (not using the keys the right way round and not saying the shape they see when it is white), this was evident on Dragon Dash as participants made more errors when the keys were reversed compared to the usual way round.

In addition, it could be argued that Dragon Hunt is underpinned by inhibition, rather than switching. The crystals and eggs in the Dragon Hunt subgame are scattered in a random order and, therefore, it is necessary that participants inhibit the prepotent response of collecting the item they see next. Previously, it has been suggested that switching abilities rely on inhibition, and this could be being captured in the results (Barkley, 1997; Bissonette et al., 2013). It is also possible that the lack of association between Dragon Hunt and NEPSY-II Switching is a result of the NEPSY-II switching task being too difficult for the younger participants.

4.2.1.2. Children with an Autism Spectrum Condition Aged 6 to 11

In the ASC sample, Dragon Dash was moderately to largely associated with the NEPSY-II Inhibition. Small to moderate associations were found between Dragon Sequence and WNV-SS, all of which were small or not associated when controlling for computer literacy. Dragon Sequence data was positively skewed around lower numbers of total correct, whereas WNV-SS data was negatively skewed around higher numbers of total correct. This implies that Dragon Sequence was more difficult that WNV-SS for this sample and could be because Dragon Sequence did not have a practice trial. Dragon Hunt exhibited small associations with NEPSY-II Switching, suggesting the measures are unlikely to be underpinned by a related function. The moderate associations found between Dragon Dash and NEPSY-II Inhibition indicate that Dragon Dash could be a valid measure of inhibition for children with a diagnosis of an ASC aged 6 to 11 years.

Dragon Dash shared large correlations with most of the established measures, which remained moderate to large after controlling for computer literacy and processing speed. Large correlations might not be expected between measures of different executive function components, as according to Miyake et al. (2000), different measures should only be loosely related. Therefore, these results could suggest that executive functions are less distinct in the ASC sample, or that Dragon Dash captures inhibition, working memory and switching abilities; and could be considered a general measure of executive function.

Unexpectedly, Dragon Dash and Dragon Hunt were both largely associated with the WNV-SS. This suggests that Dragon Dash and Dragon Hunt could be underpinned by working memory in this sample. Dragon Dash requires participants to remember what they need to do when they hear the 'beep' and Dragon Hunt requires participants to hold in mind the sequence they are instructed to follow. For Dragon Hunt, if participants were only instructed to start with the blue egg and then keep switching between colour and shape, rather than being told the full sequence to hold in mind, this could make it a purer measure of switching.

It is possible that Dragon Dash and Dragon Hunt are capturing working memory in this sample due to there being a deficit in inhibition and switching, and intact working memory. This supports previous research finding intact working memory in those with an ASC (Ozonoff & Strayer, 2001), however, contradicts the teacher ratings, which reported the most difficulties in working memory on the CHEXI. These results highlight the importance of having measures which capture pure processes in order to make definite conclusions around deficit and intact executive function profiles. Going forward, measures used to assess functions other than working memory in this population should not be underpinned or rely on working memory.

Some of the participants did not complete every measure for various reasons, such as finding it too difficult, not being able to understand the instructions or not wanting to participate. Some participants included in the analysis completed

Dragon Dash and Dragon Sequence, however, did not complete Dragon Hunt; this was largely to do with difficulties using the controls. There were also several participants who completed the WNV-SS but not the NEPSY-II Inhibition battery, or only part of the NEPSY-II Inhibition, this was usually the Naming and Inhibition subtasks. It is possible that this is a result of attention or fatigue, as Dragon Hunt and NEPSY-II Switching are the last segments of Dragon Adventure and the established measures to be completed, respectively. In which case, future research might look to counterbalance the order in which the Dragon Adventure subgames and the individual established measures are completed.

4.2.1.3. Neurotypical and Autism Spectrum Condition Samples

Bringing these results together, it could be concluded that Dragon Dash Reversed Errors is a valid measure of inhibition in both samples and that Dragon Hunt is not a valid measure of switching in either sample. Dragon Sequence was found to be a valid measure of working memory in the neurotypical sample, however, not the ASC sample.

Age was moderately to strongly associated with the Dragon Adventure subgames in the ASC sample, however, this was not the case for the neurotypical sample. This could indicate that Dragon Adventure is more suitable for capturing the development of executive function in those with a diagnosis of an ASC than neurotypical children.

Associations between Dragon Adventure and the established measures were more affected by processing speed in the neurotypical sample, and computer literacy in the ASC sample. Generally, associations were less affected by computer literacy and processing speed in the ASC sample than the neurotypical sample. This contradicts previous research suggesting that task impurity tends to be exacerbated when measures are used with ASC populations (Adams & Jarrold, 2009; Ames & Jarrold, 2007; Diamond et al., 2002; Russell et al. 1999),

4.2.2. Does Dragon Adventure Measure Executive Functions in Line with Teacher Ratings?

The results partially support the hypothesis that Dragon Adventure would be moderately to largely associated with teacher ratings in both samples. In the neurotypical sample, *Dragon Dash Reversed Errors* was moderately associated with CHEXI Inhibition, however, all other associations between Dragon Dash and the CHEXI were small. *Dragon Dash Reversed Errors* could be the most ecologically valid measure of inhibition in Dragon Dash. This makes sense, considering that *Reversed Errors* is the measures of participants ability to inhibit the prepotent response. Dragon Sequence was moderately associated with both CHEXI WM and Inhibition, suggesting Dragon Sequence could be an ecologically valid measure of general executive function abilities. Dragon Hunt was strongly associated with CHEXI Inhibition, further supporting the suggestion that Dragon Hunt could be underpinned by inhibition. The moderate to large associations between the Dragon Adventure subgames and the CHEXI indicates that Dragon Adventure could be an ecological validity measure of executive function for neurotypical children aged 6 to 8 years.

As with the neurotypical sample, *Dragon Dash Reversed Errors* was moderately associated with CHEXI Inhibition in the sample diagnosed with an ASC. Again, this could suggest that *Dragon Dash Reversed errors* is the most ecologically valid measure of inhibition in Dragon Dash. Moderate correlations were also observed between other Dragon Dash scores and CHEXI WM, which aligns with association found between Dragon Dash and WNV-SS. Unexpectedly, Dragon Sequence was moderately associated with CHEXI Inhibition and not CHEXI WM. However, this aligns with results suggesting Dragon Sequence is not a valid measure of working memory in this sample. *Dragon Hunt Time* was moderately associated with both CHEXI Inhibition and WM. This aligns with the associations found between Dragon Hunt and the WNV-SS and could suggest that Dragon Hunt is an ecologically valid measure of general executive function abilities in this sample.

These results indicate that the Dragon Adventure subgames are not consistently associated with the teacher rated executive functions they are considered to measure. However, the results here are mostly consistent with the associations found between the Dragon Adventure subgames and established measures, indicating good ecological validity.

4.2.3. Is there a Difference in Rated Enjoyment Between Dragon Adventure and Established Measures of Executive Function?

The results did not support the hypothesis that Dragon Adventure would be rated as substantially more enjoyable than established measures, in either population. Both Dragon Adventure and the established measures were rated similarly, with Dragon Adventure being rated very slightly higher. There was a small range of acceptability ratings, with most participants rating the game as either 4 or 5 on the visual analogue scale; indicating that the game was found to be acceptable. However, this could be a result of participant bias, through participants wanting to please the researcher. Having a way to rate the measures anonymously may have impacted results, however, in the current study there was not an appropriate way to do this given the age and abilities of the children. As Davis (2020) mentioned, Dragon Adventure is still in an early stage of development and, therefore, lacks some of the more advanced elements of computer games, which could have further increased enjoyment. On the other hand, it is also possible that as children spend more time using computers and other technological devices, the established measures become more novel and enjoyable. Gaining qualitative feedback alongside the ratings for both Dragon Adventure and established measures could provide a deeper understanding of what the participants enjoyed and why.

The neurotypical sample rated both Dragon Adventure and the established measures slightly higher than the ASC sample did. This contradicts research suggesting that computerised measures could be more acceptable to those with a diagnosis of an ASC (Gardiner et al., 2017; Kenworthy et al., 2008; Ozonoff, 1995). However, as a result of the investigator reading out the instructions and helping participants with navigating the keys, Dragon Adventure still required a considerable amount of investigator input, which could have impacted these results.

4.2.4. Comparisons to Davis (2020) Results

To gain a broader understanding of Dragon Adventure as a measure of executive function for children, results from the current study will be compared to the results of Davis (2020). Comparisons will only be made with the results from the neurotypical sample in this study. When comparing the results, it is important to consider that different established measure were used, because the established measures originally used did not cover the age range of the samples in this study, and computer literacy and processing speed were not controlled for in Davis (2020). A summary of the correlation coefficients found between Dragon Adventure and established measures in young people aged 11 to 12 years in Davis (2020) can be seen in Table 29.

Similarly to Davis (2020), Dragon Sequence was found to be moderately to strongly associated with the WNV-SS. This suggests that Dragon Sequence has good concurrent validity in children aged 6 to 8 and 11 to 12 years. Dragon Hunt was only found to have small to moderate associations with the established measures of switching (NEPSY-II Switching and TMT). It could be argued that Dragon Hunt is not an appropriate measure of switching in these age ranges, given that it has not been found to be strongly associated with two different established measures of switching. This also suggests that the results from the current study are less likely to be a result of NEPSY-II Switching being too difficult for the younger ages. Instead, Dragon Hunt appears to be underpinned by inhibition, as both the current study and Davis (2020) found Dragon Hunt to be moderately to strongly correlated with the established measure of inhibition (NEPSY-II Inhibition and Colour-word). Recent research has suggested that adding game elements to a task may result in different processing demands (Johann & Karbach, 2020), and this could be the case for Dragon Hunt.

In contrast to the current study, Davis (2020) found the CHEXI to be moderately to largely associated only with Dragon Sequence, whereas the current study found the CHEXI to be moderately or largely correlated with a score on all Dragon Adventure subgames. These results suggest that Dragon Adventure could be a more ecologically valid measure of executive function in children aged 6 to 8 compared to children aged 11 and 12. Davis (2020) found that male participants performed slightly better on Dragon Adventure subgames in young people aged 11 and 12 years, however, this study found that female participants consistently performed better. Davis (2020) suggested that their findings could be a result of males having increased familiarity with video-game formats, however, the research this suggestion was based on only included those aged 12 years and above. It is possible that children aged 6 to 8 years have less familiarity with video-game formats generally and, therefore, the same effects are not present.

In contrast to Davis (2020), the current study did not find a substantive difference between performance of the EPL and EAL groups on Dragon Adventure subgames. This could be a result of the current study including video demonstrations of each subgame and accounting for the lack of practice trials when scoring Dragon Dash and Dragon Hunt. The latter is supported by the results indicating that the EAL group performed better than the EPL group on the subgames where not having a practice trial was accounted for.

Table 29

Spearman's Rank Correlation Coefficients Between Novel and Existing Measures for Children Aged 11-12 years from Davis (2020)

	Dragon Dash: Total Errors	Backwards spans: Total Correct	Backwards spans: Errors per attempt	Dragon Hunt: Time taken	Flexibility: Total errors
Colour-word: inhibition time	**-.399**	**.470***	**-.509***	**-.787****	**-.602****
Colour-word: inhibition errors	-.190	.276	**-.315**	-.159	.096
Spatial span: backwards total	.046	**.561****	-.297	-.256	-.210
Spatial span: backwards length	.020	**.322**	-.189	**-.420**	**-.348**
Spatial span: combined	-.032	**.501***	**-.326**	-.259	-.194
TMT: combined time	**-.305**	.284	-.008	-.275	**-.342**
TMT: combined errors	-.143	.057	.100	.272	.200
TMT: contrast	-.252	.119	.117	-.104	-.177

Note. Taken from 'Using a Novel Game-Like Computerised Measure to Test Executive Functioning in Children' (Davis, 2020)
*** Correlation is significant at the 0.01 level (2-tailed)*
** Correlation is significant at the 0.05 level (2-tailed)*

4.2.5. How do children feel Dragon Adventure could be improved?

The content analysis indicated that a just over a third of respondents did not feel Dragon Adventure needed to be improved, however, this could be a result of participant bias, due to participants wanting to please the investigator. The majority of respondents did, nevertheless, offer suggestions on how Dragon Adventure could be improved.

The suggestion that came up most was improving the game controls. This was not surprising given that several participants struggled to use the controls throughout the game, and some would ask for help using them. The arrow keys were sensitive, which meant it could take a while for the dragon to go in the desired direction. Visual and verbal frustration was displayed by some participants, and it is possible that this frustration could have impacted the enjoyment ratings of Dragon Adventure.

Most of the suggestions were those that would make Dragon Adventure more alike other computer games, such as having additional characters, more advanced character functions and additional games, such as fighting and racing. Suggestions were also made on more practical elements of the game, such as the length and ease of the game. Of those who commented on the length, the majority mentioned making the game longer through additional games. It is possible that children wanted the game to be longer as they got to take time out of class to participate in the study. A couple of participants made suggestions that would make the game easier. It is possible that the difficulty of the game could have also impacted the enjoyment ratings, as has been found in recent trials of game-based executive function measures (Berg et al., 2020).

Several participants mentioned specifically having more of the obstacle course, referring to Dragon Dash. This was the only specific Dragon Adventure subgame mentioned and could suggest that it was the most enjoyed. This is interesting given that the EDA indicated that this subgame was found to be difficult in the neurotypical sample, as the data was negatively skewed around a higher number of errors.

4.3. Critical Evaluation

4.3.1. Strengths

This study went some way in addressing the limitations of Dragon Adventure identified by Davis (2020). Limitations, such as not being able to go back when reading the instructions and not having a practice trial, were mostly accounted for through additional materials and amending how the subgames were scored. Having the instructions printed out, speaking the instructions through, showing demonstration videos prior to each subgame and accounting for practice trials when scoring, would have made Dragon Adventure more accessible to those with communication difficulties and those with EAL. The results of the current study further highlight the importance of having practice trials, as has been suggested in previous research on measure development (Bauer & Zelazo, 2014). Addressing these limitations could have also increased the resemblance between the Dragon Adventure subgames and the established measures (as established measures typically have a demonstration and practice trials) and, thus, increased their concurrent validity.

The current study also went some way in addressing the challenges of measuring executive function identified in the literature review; for example, controlling for other cognitive and non-cognitive factors to assess task impurity. An attempt was also made to control for practice effects and fatigue bias through counterbalancing the order in which participants completed Dragon Adventure and the established measures.

The current study has increased generalisability in the ASC population compared to previous research. Previous studies have been limited due to only recruiting high-functioning samples (e.g. Gardiner et al., 2017). This study did not exclude children in the ASC sample based on additional diagnoses.

The current study gained formal feedback from participants on how Dragon Adventure could be improved, which can be used in future development of the game-based measure to improve participant experience, motivation, enjoyment and, thus, performance.

4.3.2. Limitations

4.3.2.1. Limitations of Dragon Adventure

It was not possible to account for the lack of a practice trail when scoring Dragon Sequence, unlike the other Dragon Adventure subgames. It is possible that adding a practice trial could strengthen the correlations observed between Dragon Sequence and WNV-SS scores. In addition, Dragon Sequence was scored differently to the WNV-SS, as Dragon Sequence would terminate after two consecutive errors, whereas WNV-SS would stop once a participant has made two consecutive errors in a trial of the same length span. This could have resulted in Dragon Sequence being more difficult to progress through. Amending the Dragon Sequence scoring to mirror that of the WNV-SS could also strengthen the associations.

In order to be accessible for those with communication difficulties, Dragon Adventure only assesses visual executive function. However, measures of executive function have been criticised for not including both visual and verbal modalities (Henry & Bettenay, 2010). Including verbal measures could provide a more detailed understanding of executive function abilities (Salimpoor & Desrocher, 2006). On the other hand, verbal executive function measures often require investigator input for scoring, which has also been criticised (Homack et al., 2005).

A limitation of Dragon Adventure - identified by both researchers and participants - was the game controls being difficult to use, particularly the right, left, forward and backward arrows. This was especially noticeable on Dragon Hunt, as participants would accidentally collide with the wrong egg or crystal due to difficulties controlling the dragon. Using a mouse to click on the items they wish to collect may be easier for and more familiar to participants and, thus, improve accuracy. There were also inconsistencies with the way participants used the control keys during Dragon Dash, which could have impacted scoring. Some participants would hold one key down until they needed to press the next one, whereas others would repeatedly press the same key until they needed to press another. Dragon Dash is scored by recording an error every time the wrong key is pressed. Therefore, more errors were recorded for those who would repeatedly press the same key. Attempts were made to address this when scoring the data, through looking at the times at which errors were recorded.

Dragon Hunt is scored by the time taken to collect all the eggs and crystals. However, from observation this may not have been accurate, as some participants would spend time exploring the game environment between collecting the items. To address this, it could be made clearer in the instructions that participants should work as quickly as possible, or taking into consideration suggestions from the participants, utilise a race format.

4.3.2.2. Study Limitations

Tasks measuring executive function often require multiple abilities, as was the case in this study. Despite a strength of this study being controlling for other cognitive functions (processing speed and computer literacy), other functions and factors could have impacted results. For example, this study did not take into account or control for parental education level, which has consistently been found to account for variance in score differences on measures of cognition (Akshoomoff et al., 2014). Not taking account of this factor could have influenced the validity of this study. Attempts were made to control for fatigue; however, it was not feasible within the current study to consider aspects such as the time of day or what participants had done before. Motivation could have also been impacted depending on which class a participant was missing; however, it was also not feasible to control for this.

The presentation and response modalities of the Dragon Adventure subgames and the established measures did not match. The established measure of working memory was visual, and the Dragon Adventure subgames were all visual. However, the established measures of inhibition and switching both required verbal ability. This could have impacted how comparable the measures were, especially for those who had communication difficulties. However, it was very difficult to find established measures suitable across the age ranges included in the current study.

The questionnaire used to measure computer literacy (the CTUQ) was self-report and, therefore, subjective. A more accurate way of controlling computer literacy could be through performance-based tests, such as motor speed when using the computer and familiarity of using the keys.

Despite trying to account for children not accustomed to the test situation through completing data collection at their school, the classroom used at the specialist school was one that participants had not been in before. It is possible that the unfamiliar room added to the novelty of the situation, as some of the participants were visibly distracted by the new location and wanted to explore. These observations support the suggestion that measures of executive function may be more accurate and ecologically valid when completed in familiar settings and with familiar people, such as at school with teachers (Obradović et al., 2018).

Collecting enjoyability ratings after each established measure and Dragon Adventure subgame, could have allowed for a more detailed breakdown of which Dragon Adventure subgames were considered most acceptable. However, this was considered too much for these samples.

In regard to the sample sizes, larger samples would have been desirable in order to fulfil the normality assumptions and obtain more conclusive and generalisable results. However, this proved difficult in the time restraints. The sample sizes in this study were reasonable for the initial evaluation of Dragon Adventure in the groups recruited.

In regard to the analyses used, while correlations can tell us about associations, they do not demonstrate cause and effect. Therefore, it is not possible to make definite conclusions around the direct relationship between Dragon Adventure and the established measures. In addition to this, running multiple tests to control for additional factors increases the error rate and likelihood of spurious results. However, the current research focused on effect sizes, rather than p-values, whereby the strength of relationships is not affected by multiple testing.

Content analyses have been criticised for being overly reductive and subject to subjective interpretation, which can impact the result and conclusions (Krippendorff, 2018). Given the nature and amount of qualitative data, along with the aims of this research, a content analysis seemed the most appropriate and useful method for summarising the data.

It has been argued that there is a publication bias in favour of Miyake et al's. (2000) three-factor model which has led to researchers accepting this model as the standard within the field (Karr et al., 2018). As a result, researchers have designed studies based on Miyake's model, thus, contributing to the many conceptual replications of this model, despite concerns around the replicability of the model itself. Therefore, this study and Davis (2020) could be considered conceptual replications contributing to the publication bias. It would be important to consider the other executive functions not captured in this model that could be impacting results.

The majority of the neurotypical sample were of White British ethnicity, this reflects the area the school was based in and will impact the results generalisability. Established measures have been criticised for being western and predominantly trialled in white middle-class populations (Ford et al., 2019). Therefore, it is important that Dragon Adventure be trialled in diverse populations, to truly assess whether it could be a more culturally appropriate measure of executive function. Similarly to previous limitations of executive function test research (Wilbourn et al., 2012), this study also used a small self-selected sample of children whose parents are motivated and willing to participate in research.

4.4. Study Implications

4.4.1. Theoretical Implications

The results from the current study support the theories, and other research findings, that executive function develops throughout childhood. Some of the associations found between the Dragon Adventure subgames, and between Dragon Adventure and the established measures, support Miyake et al's. (2000) tripartite model of separable but related executive functions in children from the age of 6. However, the unexpected findings, where Dragon Adventure subgames have been associated with established measures considered to be underpinned by a separable function, could provide evidence of the executive function components being less distinct in younger children and children with an ASC. This is difficult to tease apart when it is possible that these results are also being impacted by task impurity.

The large associations found between the WNV-SS and both Dragon Dash and Dragon Hunt in the ASC sample, could suggest a predominant working memory contribution in those with a diagnosis of an ASC. Both Dragon Dash and Dragon Hunt require working memory in order to inhibit the prepotent response (through remembering that the controls switch when the beep sounds) and engage in task switching (to remember in which order you switch between the objects), respectively. As these results were not as present in the neurotypical sample, it is possible that Dragon Dash and Dragon Hunt are capturing working memory in the ASC sample, as a result of deficits in inhibition and switching.

The results from the current study support the integrated theory of executive function, through evidence of both the unitary and componential view. The intercorrelations between the Dragon Adventure subgames and with the different established measures of executive function supports the unitary view of executive function (Baddeley, 1986, 1992; Dempster, 1992; Norman & Shallice, 1986; Shallice, 1988). The differences in effects of associations between Dragon Adventure and established measures supports the componential view of executive function (Carlson & Moses, 2001; Diamond, 1991; Pennington, 1997; Welsh et al., 1991). Having a suitable measure of executive function across childhood, such as Dragon Adventure, has the potential to provide clarity on whether executive function is a unitary or componential construct in this age group, along with whether and when this changes throughout childhood.

4.4.2. Clinical Implications

The literature review identified a growing need for a suitable measure of executive function appropriate for younger school-aged children. The demonstrated feasibility and acceptability of Dragon Adventure in the current study and Davis (2020), suggests that Dragon Adventure has the potential to be a suitable measure of executive function across the school-aged population and potentially further.

Dragon Adventure improves clinical utility compared to established measures through its benefits for both the clinician and participant. A benefit for the clinician includes Dragon Adventure requiring less resources through reduced administration. Scoring is automatic and, therefore, more accurate than traditional scoring prone to human error. Scoring is also multidimensional and could be further developed to produce scores around inter-trial interactions. Having an inclusive measure suitable those with communication difficulties reduces the time needed to adapt traditional measures for this population. Benefits of Dragon Adventure for participants and clients includes being more engaging, improving motivation and being less stressful than established measures, through familiarity and adopting a game rather than a test format.

Being electronic, Dragon Adventure could be accessed on a range of devices and easily transported to locations. Therefore, Dragon Adventure could be used in a location most suitable for the client and/or clinician; for example, in a setting familiar to the client, which could subsequently improve participant experience and the measures ecological validity (Obradović et al., 2018).

Within child health services, it is important that efficient and accurate tools are available for use when needed, especially within child development services, where cognitive measures are routinely used. As previously stated, further development of Dragon Adventure could result in a measure that reduced demands on clinicians working in clinical settings. In addition, having measures that could potentially be administered in a child's educational setting by a teacher, could further these benefits.

Considering the associations found between executive function and children's behaviour, and positive outcomes (Jurado & Rosselli, 2007), accurately assessing executive function is of paramount importance in order for children to receive the support they need to have the best outcomes academically and socially. Suitable and accurate measures of executive function will further contribute to our understanding of executive function in childhood and have implications for interventions. Such interventions could be informed by findings from research concerning executive function measures, such as utilising computerised technology in a game-based format that is engaging and motivating for the child's age. This could involve future research investigating the cognitive functions required by games and activities currently used by children. Drawing on such information could also have implications for how interventions are administered, for example, moving away from more individualised interventions to interventions administered in group settings, such as in the classroom.

4.5. Future Directions

4.5.1. <u>Current Study</u>

If the current study were to be replicated with the same resources, it would be important to address the limitations mentioned above. A mouse would be used instead of the laptop trackpad, to see if this improves ease of use. To reduce distraction, data collection would take place in an environment familiar to the participant. If possible, parent-rated measures of executive function would also be completed as a comparison, for predictive validity. To reduce investigator input, and potentially bias, verbal instructions for Dragon Adventure would be presented through headphones. If possible, participants would be recruited from more schools in a range of areas, to ensure the sample was representative. This could improve sample size, which would enable more conclusive and generalisable results. A larger sample could also fulfil the data normality assumptions and use of statistical parametric analysis.

4.5.2. Future Research

There are a number of ways future research can be taken forward from the current study, including research to further develop understandings of executive function in children and those with an ASC, and further developing Dragon Adventure as a measure of executive function.

4.5.2.1. Developing Understanding of Executive Functioning in Children

Generally, more research is needed to investigate whether Miyake et al's. (2000) model can be consistently replicated, rather than conceptually reified. This could provide clarity on whether this model is the most appropriate for understanding executive function in children. Such research could include developing measures using alternative models, developing more Dragon Adventure subgames measuring additional executive function components, or comparing Dragon Adventure to established measures of different executive function components.

4.5.2.2. Further Development of Dragon Adventure

To further assess the utility of Dragon Adventure and further develop our understanding around the executive function profile of school-aged children and children with an ASC, the next stage would be to compare the performance of the neurotypical and ASC sample. Future research should continue to trial Dragon Adventure in neurotypical, neurodiverse and clinical samples. It would be important to trial Dragon Adventure in a group with known executive dysfunction, such as those with a Traumatic Brain Injury to the PFC or those with congenital disorders affecting the executive functions, to see if the measure can differentiate these populations. Davis (2020) investigated Dragon Adventure's internal reliability, future studies should also explore Dragon Adventure's test-retest reliability and susceptibility to practice effects.

In order to conduct this future research, the Dragon Adventure subgames should be amended to address the limitations identified, such as improving the controls and scoring output, and developing Dragon Hunt to be a more appropriate measure of switching. In addition, Dragon Adventure could be further developed, to include practice trials and inbuilt measures of motor speed and computer literacy. Future research should continue to assess the relationships between Dragon Adventure and other cognitive and non-cognitive factors, in order to determine the contribution of other skills on performance and to assess whether the game is confounded by the same difficulties as established measures. Acknowledging the result from the ASC sample, adapting Dragon Dash and Dragon Hunt to limit the requirement of working memory could make Dragon Adventure more suitable for those with an ASC.

Taking into consideration the feedback received from participants, Dragon Adventure could be further developed to be more acceptable and engaging, through incorporating more advanced game elements used in standard computer games. Future research should continue to gain feedback and work with those who would be using Dragon Adventure. Gaining feedback from across childhood could inform whether Dragon Adventure needs to be adapted for different age groups. It would be important to gain feedback anonymously or separately from data collection to reduce participant bias, for example, through focus groups.

Dragon Adventure could be adapted to be used on other devices, such as a tablet. Tablets have been suggested to have multiple advantages in their use with children, such as, less attentional demands, not having to reorient attention from the screen to a keyboard or mouse, can be applied in different contexts and can be self-administered (Berg et al., 2020). Using a tablet has been suggested to be a more reliable measurement method of executive function in children (Howard & Okely, 2015). The use of tablets could also be more engaging for children and suit their preferred modality.

4.6. Reflections

4.6.1. Professional and Ethical Issues

An increasing number of measures are using or looking to develop automated administration and scoring systems and it is being increasingly suggested that measures of executive function are conducted by adults familiar to the children in a familiar environment. Supporting these suggestions, in the current study, the established measures took more time to score than Dragon Adventure. Along with this, some of the participants diagnosed with an ASC struggled to or did not engage due to unfamiliarity of the location and the researcher. In accordance with suggestions from the literature, it seems appropriate that measures of executive function are completed by familiar adults in a familiar environment, or automated tests are used under the supervision of a teacher. This poses the question of what a clinician's role will be in the future.

Although the current study has collected formal feedback on Dragon Adventure, this was done after its initial development. Therefore, Dragon Adventure has been developed for the young people, as opposed to with them. It would have also been beneficial to gain feedback from clinicians who use measures of executive function as to what the difficulties are with current measures, how these could be addressed and whether they feel new more appropriate measures of executive function are needed.

Computerised assessments require access to a device, in this case, a laptop or computer, which may not always be accessible, especially in low-income settings. Therefore, Dragon Adventure could be less accessible than established measures to a proportion of the population. Along with this, those who have increased access to a device may perform better due to increased familiarity, hence, it is important to control for computer literacy.

This current research promotes Western narratives of executive function being something children must achieve and be skilled in. Having measures of executive function, whereby children can perform better or worse than one another, benefits the capitalistic social view.

Developing psychological tools based on white Western approaches, in an area where such tools are often normed against the white Western populations, has implications for racially and culturally-minoritised groups. Therefore, it is important to consider the impact of using a tool such as Dragon Adventure on these groups. For example, Dragon Adventure could be used to assess 'behavioural difficulties' that are a result of discrimination and trauma, which could inform a diagnosis and result in the child internalising that there is something ‘wrong’ with them, alongside already present ‘othering’ narratives, which could, in turn, go on to have multiple implications throughout that person’s life.

Taking an anti-racist approach to this research, could involve engaging with a number of different communities to learn about their understanding of the behaviours that have been conceptualised as ‘executive function’ in this research. This could broadly inform thinking around executive function in clinical and research settings. One ambition was for Dragon Adventure to be culturally and linguistically fair through using visual materials, rather than linguistic operations, and basing scores on motor functions, rather than on educational or acquired knowledge.

4.6.2. Personal Reflections

In accordance with the critical realist perspective which underpinned the current research, it is important that a researcher reflect on the context surrounding their study (Flanagan, 1981; Willig, 2019). The realist part of me acknowledges the importance and benefits of being able to understand cognition through concepts and measurement, in order to better human experience. This underpinned the research questions, methodology and design. The critical part of me acknowledges the importance of language and how this shapes experience. This influenced some of the terms used in this study. For example, using the term ‘neurotypical’ instead of ‘typically developing’ and using the phrase ‘participants with an ASC’ rather than ‘autistic participants’. These language choices were driven by a want for this research and future research to be accessible for the communities on which they focus, as ultimately the research is for their benefit. However, this benefit is also dependent on the study being accessible to other researchers. As a result, it felt difficult at times to know how best to balance using the language used in previous research and the language used in communities. For example, I used the term ‘executive dysfunction’ and referred to scores indicating intact executive function abilities as ‘better scores’, as these concepts are commonly used within the literature. However, I acknowledge that such terms lend themselves to a deficit model of understanding executive function and can be medicalising.

Having completed the current study, I feel torn in relation to some of the criticisms of established measures and benefits of newly developed computerised measures. For example, established measures have been criticised for requiring investigator input and their time. To address these, literature has recommended developing measures that do not require investigator input or can be administered by non-clinicians. However, the current study has highlighted the importance of observation while the measures are being completed. Not observing the children while they completed Dragon Adventure, meant I would have missed out on the nuances which have provided context for some of the results. This seems especially important when administering measures in a clinical assessment, as observing, or not observing, could impact clinical decisions. Non-clinicians would need extensive knowledge on the theory behind the measure, the measure itself and the scoring to know what to observe.

Throughout the process, I wondered if I had taken on too much by investigating the use of Dragon Adventure in two different samples in the time available, and whether this would impact the quality of the research. However, it felt important that not only the gap in measures for neurotypical children be addressed, but also the game be trialled in a population that are most likely to be presented with such a measure, and that this be done in the early stages of measure development. I understand that my desire to complete this research with an ASC sample has likely been informed by the work I have done with this population historically.

4.7. Conclusions

The results from the current study have demonstrated that Dragon Adventure shows promise as a measure of executive function for young children aged 6 to 8 and children aged 6 to 11 years with an ASC. In its current state, Dragon Adventure is a valid measure of inhibition in both samples and working memory in the neurotypical sample; Dragon Adventure also demonstrates good ecological validity. Further developments of Dragon Adventure are needed, with the continued input of young people and introducing the input of clinicians that would be using Dragon Adventure in a clinical setting. Future research should continue to trial Dragon Adventure in neurotypical, neurodiverse and clinical populations, recruiting larger sample sizes in order to improve conclusions, and for the development of norms in the future. Comparisons should be made between Dragon Adventure and measures of additional executive functions outside of Miyake's tripartite model, to reduce conceptual replication. Continued research and development of Dragon Adventure could provide benefits for clinicians, services and the lives of children that use them.

REFERENCES

Adams, N. C., & Jarrold, C. (2009). Inhibition and validity of the Stroop task for children with autism. *Journal of Autism and Developmental Disorders*, *39*(8), 1112–1121. https://doi.org/10.1007/s10803-009-0721-8

Akshoomoff, N., Newman, E., Thompson, W. K., McCabe, C., Bloss, C. S., Chang, L., Amaral, D. G., David, G., Casey, B. J., Ernst, T. M., Frazier, J. A., Gruen, J. R., Kaufmann, W. E., Kenet, T., Kennedy, D. N., Libiger, O., Mostofsky, S., Murray, S. S., Sowell, E. R., … Jernigan, T. L. (2014). The NIH toolbox cognition battery: Results from a large normative developmental sample (PING). *Neuropsychology*, *28*(1), 1–10. https://doi.org/10.1037/neu0000001

Allan, D. M., Allan, N. P., Lerner, M. D., Farrington, A. L., & Lonigan, C. J. (2015). Identifying unique components of preschool children's self-regulatory skills using executive function tasks and continuous performance tests. *Early Childhood Research Quarterly*, *32*, 40–50. https://doi.org/10.1016/j.ecresq.2015.02.001

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

Ames, C. S., & Jarrold, C. (2007). The problem with using eye-gaze to infer desire: A deficit of cue inference in children with Autism Spectrum Disorder? *Journal of Autism and Developmental Disorders*, *37*(9), 1761–1775. https://doi.org/10.1007/s10803-006-0309-5

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*(2), 71–82. https://doi.org/10.1076/chin.8.2.71.8724

Anderson, V. (1998). Assessing executive functions in children: Biological, psychological, and developmental considerations. *Neuropsychological Rehabilitation*, *8*(3), 319–349.

Anderson, V. (2001). Assessing executive functions in children: Biological, psychological, and developmental considerations. *Paediatric Rehabilitation*, *4*(3), 119–136. https://doi.org/10.1080/13638490110091347

Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Mikiewicz, O. (2002). Relationships between cognitive and behavioral measures of executive function in children with brain disease. *Child Neuropsychology*, *8*(4), 231–240. https://doi.org/10.1076/chin.8.4.231.13509

Annotti, L. A., & Teglasi, H. (2017). Functioning in the real world: Using storytelling to improve validity in the assessment of executive functions. *Journal of Personality Assessment*, *99*(3), 254–264. https://doi.org/10.1080/00223891.2016.1205075

Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology*, *5*(2), 115–129. https://doi.org/10.1076/chin.5.2.115.3167

Baddeley, A. (1986). *Working memory*. Oxford University Press.

Baddeley, A. (1992). Working memory. *Science*, *255*, 556–559. https://www.science.org/doi/10.1126/science.1736359

Ballhausen, N., Mahy, C. E. V., Hering, A., Voigt, B., Schnitzspahn, K. M., Lagner, P., Ihle, A., & Kliegel, M. (2017). Children's planning performance in the Zoo Map task (BADS-C): Is it driven by general cognitive ability, executive functioning, or prospection? *Applied Neuropsychology: Child*, *6*(2), 138–144. https://doi.org/10.1080/21622965.2015.1124276

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*(1), 65-94. https://doi.org/10.1037/0033-2909.121.1.65

Bauer, P. J., & Zelazo, P. D. (2014). The National Institutes of Health Toolbox for the Assessment of Neurological and Behavioral Function: A tool for developmental science. *Child Development Perspectives*, *8*(3), 119–124. https://doi.org/10.1111/cdep.12080

Baumeister, R. F., & Finkel, E. J. (Eds.). (2010). *Advanced social psychology: The state of the science.* Oxford University Press.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1–3), 7–15. https://doi.org/10.1016/0010-0277(94)90018-3

Bell, M. D., Weinstein, A. J., Pittman, B., Gorman, R. M., & Abujelala, M. (2021). The Activate Test of Embodied Cognition (ATEC): Reliability, concurrent validity and discriminant validity in a community sample of children using cognitively demanding physical tasks related to executive functioning. *Child Neuropsychology*, *27*(7), 973–983. https://doi.org/10.1080/09297049.2021.1923686

Berg, V., McMahon, M., Rogers, S., Garrett, M., & Manley, D. (2019). The eFun App: A new tool to measure executive functions to support learning in a child-friendly, valid and engaging Way. In J. Theo Bastiaens (Ed.), *Proceedings of EdMedia+ Innovate Learning* (pp. 76–85). Association for the Advancement of Computing in Education. *LearnTechLib.* https://www.learntechlib.org/primary/p/209992/.

Berg, V., Rogers, S. L., McMahon, M., Garrett, M., & Manley, D. (2020). A novel approach to measure executive functions in students: An evaluation of two child-friendly apps. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01702

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

Bhaskar, R. (1990). *Reclaiming Reality*. Verso.

Bissonette, G. B., Powell, E. M., & Roesch, M. R. (2013). Neural structures underlying set-shifting: Roles of medial prefrontal cortex and anterior cingulate cortex. *Behavioural Brain Research*, *250*, 91–101. https://doi.org/10.1016/j.bbr.2013.04.037

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*, *57*(2), 111–127. https://doi.org/10.1037/0003-066X.57.2.111

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, *28*(2), 595–616. https://doi.org/10.1207/s15326942dn2802_3

Carlson, S. M., & Moses, L. J. (2001). Individual differences in Inhibitory Control and children's Theory of Mind. *Child Development*, *72*(4), 1032–1053. https://doi.org/10.1111/1467-8624.00333

Catale, C., Meulemans, T., & Thorell, L. B. (2015). The Childhood Executive Function Inventory: Confirmatory factor analyses and cross-cultural clinical validity in a sample of 8- to 11-year-old children. *Journal of*

Attention Disorders, *19*(6), 489–495. https://doi.org/10.1177/1087054712470971

Chan, A. S., Cheung, M., Han, Y. M. Y., Sze, S. L., Leung, W. W., Man, H. S., & To, C. Y. (2009). Executive function deficits and neural discordance in children with Autism Spectrum Disorders. *Clinical Neurophysiology*, *120*(6), 1107–1115. https://doi.org/10.1016/j.clinph.2009.04.002

Christ, S. E., Holt, D. D., White, D. A., & Green, L. (2007). Inhibitory Control in children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *37*(6), 1155–1165. https://doi.org/10.1007/s10803-006-0259-y

Cianchetti, C., Corona, S., Foscoliano, M., Contu, D., & Sannio-Fancello, G. (2007). Modified Wisconsin Card Sorting Test (MCST, MWCST): Normative data in children 4–13 years old, according to classical and new types of scoring. *Clinical Neuropsychologist*, *21*(3), 456–478. https://doi.org/10.1080/13854040600629766

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* (2nd ed.). Lawrence Erlbaum Associates.

Crone, E. A., Richard Ridderinkhof, K., Worm, M., Somsen, R. J. M., & van der Molen, M. W. (2004). Switching between spatial stimulus-response mappings: A developmental study of cognitive flexibility. *Developmental Science*, *7*(4), 443–455. https://doi.org/10.1111/j.1467-7687.2004.00365.x

Crotty, M. (1998). *The Foundations of Social Research: Meaning and Perspective in the Research Process*. Sage.

Culbertson, W. C., & Zillmer, E. A. (1998). The tower of London(DX): A standardized approach to assessing executive functioning in children. *Archives of Clinical Neuropsychology*, *13*(3), 285–301. https://doi.org/10.1016/S0887-6177(97)00033-4

Czermainski, F. R., Bosa, C. A., Mina, C. S., Meimes, M. A., Miranda, M. C., Carim, D., & de Salles, J. F. (2015). Performance of children/adolescents with autism spectrum disorders in executive function: Study of case series. *Psychology and Neuroscience*, *8*(3), 305–320. https://doi.org/10.1037/h0101279

Czermainski, F. R., Riesgo, R. dos S., Guimarães, L. S. P., Salles, J. F. de, & Bosa, C. A. (2014). Executive functions in children and adolescents with

Autism Spectrum Disorder. *Paidéia (Ribeirão Preto)*, *24*(57), 85–94. https://doi.org/10.1590/1982-43272457201411

Damasio, A. R., & Maurer, R. G. (1978). A neurological model for childhood autism. *Archives of Neurology*, *35*(12), 777–786. https://doi.org/10.1001/archneur.1978.00500360001001

Davis, J. (2020). *Using a Novel Game-Like Computerised Measure to Test Executive Functioning in Children* [Doctoral dissertation]. University of East London Research Repository.

Davis, J., & Matthews, R. (2010). NEPSY-II Review. *Journal of Psychoeducational Assessment*, *28*, 175–182. https://doi.org/10.1177/0734282909346716

Dawson, G., Carver, L., Meltzoff, A. N., Panagiotides, H., McPartland, J., & Webb, S. J. (2002). Neural correlates of face and object recognition in young children with Autism Spectrum Disorder, Developmental Delay, and Typical Development. *Child Development*, *73*(3), 700–717. https://doi.org/10.1111/1467-8624.00433

De Vries, M., Prins, P. J. M., Schmand, B. A., & Geurts, H. M. (2015). Working memory and cognitive flexibility-training for children with an autism spectrum disorder: A randomized controlled trial. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *56*(5), 566–576. https://doi.org/10.1111/jcpp.12324

Delis, D. C., Kaplan, E., & Kramer, H. (2001). *Delis Kaplan Executive Function System: Examiner's manual*. Psychological Corporation.

Demetriou, E. A., Lampit, A., Quintana, D. S., Naismith, S. L., Song, Y. J. C., Pye, J. E., Hickie, I., & Guastella, A. J. (2018). Autism spectrum disorders: A meta-analysis of executive function. *Molecular Psychiatry*, *23*(5), 1198–1204. https://doi.org/10.1038/mp.2017.75

Dempster, F. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review*, *12*, 45–75. https://doi.org/10.1016/0273-2297(92)90003-k

Dennis, M. (1991). Frontal lobe function in childhood and adolescence: A heuristic for assessing attention regulation, executive control, and the intentional states important for social discourse. *Developmental Neuropsychology*, *7*(3), 327–358. https://doi.org/10.1080/87565649109540497

Diamond, A. (1991). Guidelines for the study of brain-behavior relationships during development. In H. S. Levin, H. M. Eisenberg, & A. L. Benton (Eds.), *Frontal lobe function and dysfunction.* (pp. 339–378). Oxford University Press.

Diamond, A. (2006). The early development of executive functions. In E. Bialystok & F. I. M. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 70–95). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195169539.003.0006

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Diamond, A., Kirkham, N., & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology*, *38*(3), 352–362. https://doi.org/10.1037/0012-1649.38.3.352

Duan, X., Wei, S., Wang, G., & Shi, J. (2010). The relationship between executive functions and intelligence on 11- to 12-year- old children. *Psychological Test and Assessment Modelling*, *52*(4), 419–431.

Emslie, H., Wilson, F. C., Burden, V., Nimmo-Smith, I., & Wilson, B. A. (2003). *The Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C)*. Thames Valley Test Company.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of clinical and experimental neuropsychology, 26*(7), 874-890. https://doi.org/10.1080/13803390490510680

Flanagan, O. J. (1981). Psychology, progress, and the problem of reflexivity: A study in the epistemological foundations of psychology. *Journal of the History of the Behavioral Sciences*, *17*(3), 375–386. https://doi.org/10.1002/1520-6696(198107)17:3<375::AID-JHBS2300170308>3.0.CO;2-U

Fletcher, G. J. O. (1996). Realism versus Relativism in Psychology. *The American Journal of Psychology*, *109*(3), 409. https://doi.org/10.2307/1423014

Ford, C. B., Kim, H. Y., Brown, L., Aber, J. L., & Sheridan, M. A. (2019). A cognitive assessment tool designed for data collection in the field in low- and middle-income countries. *Research in Comparative and International Education*, *14*(1), 141–157. https://doi.org/10.1177/1745499919829217

Forns, J., Esnaola, M., López-Vicente, M., Suades-González, E., Alvarez-Pedrerol, M., Julvez, J., Grellier, J., Sebastián-Gallés, N., & Sunyer, J. (2014). The n-back Test and the Attentional Network Task as measures of child neuropsychological development in epidemiological studies. *Neuropsychology*, *28*(4), 519–529. https://doi.org/10.1037/neu0000085

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, *133*(1), 101–135. https://doi.org/10.1037/0096-3445.133.1.101

Gardiner, E., Hutchison, S. M., Müller, U., Kerns, K. A., & Iarocci, G. (2017). Assessment of executive function in young children with and without ASD using parent ratings and computerized tasks of executive function. *Clinical Neuropsychologist*, *31*(8), 1283–1305. https://doi.org/10.1080/13854046.2017.1290139

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 31/2-7 years old on a Stroop-like day-night test. *Cognition*, *53*(2), 129–153. https://doi.org/10.1016/0010-0277(94)90068-X

Geurts, H. M., Sinzig, J., Booth, R., & Happé, F. (2014). Neuropsychological heterogeneity in executive functioning in autism spectrum disorders. *International Journal of Developmental Disabilities*, *60*(3), 155–162. https://doi.org/10.1179/2047387714Y.0000000047

Gioia, G. A., Isquith, P. K., Retzlaff, P. D., & Espy, K. A. (2002). Confirmatory factor analysis of the Behavior Rating Inventory of Executive Function (BRIEF) in a clinical sample. *Child Neuropsychology*, *8*(4), 249–257. https://doi.org/10.1076/chin.8.4.249.13513

Gómez-Pérez, M. M., Calero, M. D., Mata, S., & Molinero, C. (2016). Discrepancies between direct and indirect measures of interpersonal and neurocognitive skills in autism spectrum disorder children. *Journal of Clinical & Experimental Neuropsychology*, *38*(8), 875–886. Academic https://doi.org/10.1080/13803395.2016.1170106

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. https://doi.org/10.1037/h0059831

Greenwood, J. D. (1994). *Realism, identity and emotion: Reclaiming social psychology.* Sage.

Griffith, E. M., Pennington, B. F., Wehner, E. A., & Rogers, S. J. (1999). Executive Functions in young children with Autism. *Child Development*, *70*(4), 817–832. https://doi.org/10.1111/1467-8624.00059

Grossi, D., Orsini, A., Monetti, C., & De Michele, G. (1979). Sex differences in children's spatial and verbal memory span. *Cortex*, *15*(4), 667–670. https://doi.org/10.1016/S0010-9452(79)80054-4

Gruber, O., & Goschke, T. (2004). Executive control emerging from dynamic interactions between brain systems mediating language, working memory and attentional processes. *Acta Psychologica*, *115*(2–3), 105–121. https://doi.org/10.1016/j.actpsy.2003.12.003

Henry, L. A., & Bettenay, C. (2010). The assessment of executive functioning in children. *Child and Adolescent Mental Health*, *15*(2), 110–119. https://doi.org/10.1111/j.1475-3588.2010.00557.x

Hill, E. L. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences*, *8*(1), 26–32. https://doi.org/10.1016/j.tics.2003.11.003

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, *16*(3), 174–180. https://doi.org/10.1016/j.tics.2012.01.006

Høie, B., Mykletun, A., Waaler, P. E., Skeidsvoll, H., & Sommerfelt, K. (2006). Executive functions and seizure-related factors in children with epilepsy in western Norway. *Developmental Medicine & Child Neurology*, *48*(6), 519–525. https://doi.org/10.1017/S0012162206001095

Homack, S., Lee, D., & Riccio, C. A. (2005). Test Review: Delis-Kaplan Executive Function System. *Journal of Clinical and Experimental Neuropsychology*, *27*(5), 599–609. https://doi.org/10.1080/13803390490918444

Horton, M. (2013). *Improving Validity and Reliability in Children's Self Reports of Technology Use* [Doctoral dissertation]. University of Central Lancashire.

Howard, S. J., & Melhuish, E. (2017). An Early Years Toolbox for Assessing Early Executive Function, Language, Self-Regulation, and Social Development: Validity, Reliability, and Preliminary Norms. *Journal of Psychoeducational Assessment*, *35*(3), 255–275. https://doi.org/10.1177/0734282916633009

Howard, S. J., & Okely, A. D. (2015). Catching Fish and Avoiding Sharks: Investigating Factors That Influence Developmentally Appropriate Measurement of Preschoolers' Inhibitory Control. *Journal of Psychoeducational Assessment*, *33*(6), 585–596. https://doi.org/10.1177/0734282914562933

Hsieh, H.-F., & Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*, *15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Hudspeth, W. J., & Pribram, K. H. (1990). Stages of Brain and Cognitive Maturation. *Journal of Educational Psychology*, *82*(4), 881–884. https://doi.org/10.1037/0022-0663.82.4.881

Jacob, R., & Parkinson, J. (2015). The Potential for School-Based Interventions That Target Executive Function to Improve Academic Achievement: A Review. *Review of Educational Research*, *85*(4), 512–552. https://doi.org/10.3102/0034654314561338

Johann, V. E., & Karbach, J. (2018). Validation of new online game-based executive function tasks for children. *Journal of Experimental Child Psychology*, *176*, 150–161. https://doi.org/10.1016/j.jecp.2018.07.009

Johann, V. E., & Karbach, J. (2020). Effects of game-based and standard executive control training on cognitive and academic abilities in elementary school children. *Developmental Science*, *23*(4). https://doi.org/10.1111/desc.12866

Józsa, K., Barrett, K. C., & Morgan, G. A. (2017). Game-like tablet assessment of approaches to learning: Assessing mastery motivation and executive functions. *Electronic Journal of Research in Educational Psychology*, *15*(3), 665–695. https://doi.org/10.14204/ejrep.43.17026

Jurado, M. B., & Rosselli, M. (2007). The Elusive Nature of Executive Functions: A Review of our Current Understanding. *Neuropsychology Review*, *17*(3), 213–233. https://doi.org/10.1007/s11065-007-9040-z

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, *144*(11), 1147–1185. https://doi.org/10.1037/bul0000160

Kenworthy, L., Freeman, A., Ratto, A., Dudley, K., Powell, K. K., Pugliese, C. E., Strang, J. F., Verbalis, A., & Anthony, L. G. (2020). Preliminary Psychometrics for the Executive Function Challenge Task: A Novel, “Hot” Flexibility, and Planning Task for Youth. *Journal of the International Neuropsychological Society*, *26*(7), 725–732. https://doi.org/10.1017/S135561772000017X

Kenworthy, L., Yerys, B. E., Anthony, L. G., & Wallace, G. L. (2008). Understanding Executive Control in Autism Spectrum Disorders in the Lab and in the Real World. *Neuropsychology Review*, *18*(4), 320–338. https://doi.org/10.1007/s11065-008-9077-7

Kerr, A., & Zelazo, P. D. (2004). Development of “hot” executive function: The children’s gambling task. *Brain and Cognition*, *55*(1), 148–157. https://doi.org/10.1016/S0278-2626(03)00275-6

Kimhi, Y. (2014). Theory of Mind Abilities and Deficits in Autism Spectrum Disorders. *Topics in Language Disorders*, *34*(4), 329–343. https://doi.org/10.1097/TLD.0000000000000033

Kimhi, Y., Shoam-Kugelmas, D., Agam Ben-Artzi, G., Ben-Moshe, I., & Bauminger-Zviely, N. (2014). Theory of Mind and Executive Function in Preschoolers with Typical Development Versus Intellectually Able Preschoolers with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *44*(9), 2341–2354. https://doi.org/10.1007/s10803-014-2104-z

Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2001). Differential development of attention and executive functions in 3- to 12-year-old Finnish children. *Developmental Neuropsychology*, *20*(1), 407–428. https://doi.org/10.1207/S15326942DN2001_6

Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY II: A developmental neuropsychological assessment manual*. Psychological Corporation.

Kramer, J. H., Mungas, D., Possin, K. L., Rankin, K. P., Boxer, A. L., Rosen, H. J., Bostrom, A., Sinha, L., Berhel, A., & Widmeyer, M. (2014). NIH EXAMINER: Conceptualization and development of an executive function battery. *Journal of the International Neuropsychological Society*, *20*(1), 11–19. https://doi.org/10.1017/S1355617713001094

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology.* Sage Publications.

Lagattuta, K. H., Sayfan, L., & Monsour, M. (2011). A new measure for assessing executive function across a wide age range: Children and adults find happy-sad more difficult than day-night. *Developmental Science*, *14*(3), 481–489. https://doi.org/10.1111/j.1467-7687.2010.00994.x

Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology*, *108*(3), 677–692. https://doi.org/10.1016/j.jecp.2010.11.001

Landa, R. J., & Goldberg, M. C. (2005). Language, Social, and Executive Functions in High Functioning Autism: A Continuum of Performance. *Journal of Autism and Developmental Disorders*, *35*(5), 557–573. https://doi.org/10.1007/s10803-005-0001-1

Lehto, J. E., & Elorinne, E. (2003). Gambling as an executive function task. *Applied Neuropsychology*, *10*(4), 234–238. https://doi.org/10.1207/s15324826an1004_5

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*(1), 59–80. https://doi.org/10.1348/026151003321164627

Liss, M., Fein, D., Allen, D., Dunn, M., Feinstein, C., Morris, R., Waterhouse, L., & Rapin, I. (2001). Executive functioning in high-functioning children with autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 261–270. https://doi.org/10.1017/S0021963001006679

Luciana, M. (2003). Practitioner Review: Computerized assessment of neuropsychological function in children: clinical and research applications of the Cambridge Neuropsychological Testing Automated Battery (CANTAB): Use of the CANTAB in pediatric neuropsychology. *Journal of Child Psychology and Psychiatry*, *44*(5), 649–663. https://doi.org/10.1111/1469-7610.00152

Luciana, M., & Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia*, *36*(3), 273–293. https://doi.org/10.1016/S0028-3932(97)00109-7

Luciana, M., & Nelson, C. A. (2002). Assessment of Neuropsychological Function Through Use of the Cambridge Neuropsychological Testing Automated Battery: Performance in 4- to 12-Year-Old Children. *Developmental Neuropsychology*, *22*(3), 595–624. https://doi.org/10.1207/S15326942DN2203_3

Mackinlay, R., Charman, T., & Karmiloff-Smith, A. (2006). High functioning children with autism spectrum disorder: A novel test of multitasking. *Brain and Cognition*, *61*(1), 14–24. https://doi.org/10.1016/j.bandc.2005.12.006

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in Psychology Research: How Often Do They Really Occur? *Perspectives on Psychological Science*, *7*(6), 537–542. https://doi.org/10.1177/1745691612460688

McAuley, T., Chen, S., Goos, L., Schachar, R., & Crosbie, J. (2010). Is the Behavior Rating Inventory of Executive Function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society*, *16*(3), 495–505. https://doi.org/10.1017/S1355617710000093

McCoy, D. C. (2019). Measuring Young Children’s Executive Function and Self-Regulation in Classrooms and Other Real-World Settings. *Clinical Child and Family Psychology Review*, *22*(1), 63–74. https://doi.org/10.1007/s10567-019-00285-1

Mischel, W., & Ebbesen, E. B. (1970). Attention in delay of gratification. *Journal of Personality and Social Psychology*, *16*(2), 329–337. https://doi.org/10.1037/h0029815

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine*, *151*(4), 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Nelson, C. A., Monk, C. S., Lin, J., Carver, L. J., Thomas, K. M., & Truwit, C. L. (2000). Functional neuroanatomy of spatial working memory in children. *Developmental Psychology*, *36*(1), 109–116. https://doi.org/10.1037/0012-1649.36.1.109

Nelson, H. E. (1976). A Modified Card Sorting Test Sensitive to Frontal Lobe Defects. *Cortex*, *12*(4), 313–324. https://doi.org/10.1016/S0010-9452(76)80035-4

Norman, D., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson., G. E. Schwartz, & D. E. Shapiro (Eds.), *Consciousness and self-regulation* (Vol. 4, pp. 1–18). Plenum Press.

Obradović, J., Sulik, M. J., Finch, J. E., & Tirado-Strayer, N. (2018). Assessing students’ executive functions in the classroom: Validating a scalable group-based procedure. *Journal of Applied Developmental Psychology*, *55*, 4–13. https://doi.org/10.1016/j.appdev.2017.03.003

Ozonoff, S. (1995). Reliability and Validity of the Wisconsin Card Sorting Test in Studies of Autism. *Neuropsychology*, *9*(4), 491–500. https://doi.org/10.1037/0894-4105.9.4.491

Ozonoff, S., & Jensen, J. (1999). Specific executive function profiles in three neurodevelopmental disorders. *Journal of Autism and Developmental Disorders*, *29*(2), 171–177. https://doi.org/10.1023/A:1023052913110

Ozonoff, S., & Strayer, D. L. (2001). Further evidence of intact working memory in autism. *Journal of Autism and Developmental Disorders*, *31*(3), 257–263. https://doi.org/10.1023/A:1010794902139

Pauli-Pott, U., & Becker, K. (2011). Neuropsychological basic deficits in preschoolers at risk for ADHD: A meta-analysis. *Clinical Psychology Review*, *31*(4), 626–637. https://doi.org/10.1016/j.cpr.2011.02.005

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

Pennington, B. F. (1997). Dimensions of executive functions in normal and abnormal development. In N. A. Krasnegor, G. R. Lyon, & P. S. Goldman-Rakic (Eds.), *Development of the prefrontal cortex: Evolution, neurobiology, and behavior.* (pp. 265–281). Paul H Brookes Publishing.

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

Pickering, S. J., & Gathercole, S. E. (2001). *WM Test Battery for Children (WMTB-C) Manual.* The Psychological Corporation.

Pilgrim, D., & Bentall, R. (1999). The medicalisation of misery: A critical realist analysis of the concept of depression. *Journal of Mental Health*, *8*(3), 261–274. https://doi.org/10.1080/09638239917427

Piper, B. J., Li, V., Eiwaz, M. A., Kobel, Y. V., Benice, T. S., Chu, A. M., Olsen, R. H. J., Rice, D. Z., Gray, H. M., & Mueller, S. T. (2012). Executive function on the psychology experiment building language tests. *Behavior Research Methods*, *44*(1), 110–123. https://doi.org/10.3758/s13428-011-0096-6

Reinert, K. R. S., Po'e, E. K., & Barkin, S. L. (2013). The Relationship between Executive Function and Obesity in Children and Adolescents: A Systematic Literature Review. *Journal of Obesity*, *2013*, 1–10. https://doi.org/10.1155/2013/820956

Reitan, R. M. (1992). *Trail Making Test: Manual for administration and scoring.* Reitan Neuropsychology Laboratory.

Riggs, N. R., Blair, C. B., & Greenberg, M. T. (2003). Concurrent and 2-year longitudinal relations between executive function and the behavior of 1st and 2nd grade children. *Child Neuropsychology*, *9*(4), 267–276. https://doi.org/10.1076/chin.9.4.267.23513

Roberts Jr, R. J., & Pennington, B. F. (1996). An interactive framework for examining prefrontal cognitive processes. *Developmental Neuropsychology*, *12*(1), 105–126. https://doi.org/10.1080/87565649609540642

Robinson, S., Goddard, L., Dritschel, B., Wisley, M., & Howlin, P. (2009). Executive functions in children with autism spectrum disorders. *Brain and Cognition*, *71*(3), 362–368. https://doi.org/10.1016/j.bandc.2009.06.007

Romine, C. B., & Reynolds, C. R. (2005). A Model of the Development of Frontal Lobe Functioning: Findings from a Meta-Analysis. *Applied Neuropsychology*, *12*(4), 190–201. https://doi.org/10.1207/s15324826an1204_2

Rosso, I. M., Young, A. D., Femia, L. A., & Yurgelun-Todd, D. A. (2004). Cognitive and Emotional Components of Frontal Lobe Functioning in Childhood and Adolescence. *Annals of the New York Academy of Sciences*, *1021*(1), 355–362. https://doi.org/10.1196/annals.1308.045

Russell, J. E. (1997). *Autism as an executive disorder*. Oxford University Press.

Russell, J., Saltmarsh, R., & Hill, E. (1999). What do executive factors contribute to the failure on false belief tasks by children with autism? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *40*(6), 859–868. https://doi.org/10.1017/S0021963099004229

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Salimpoor, V. N., & Desrocher, M. (2006). Increasing the utility of EF assessment of executive function in children. *Developmental Disabilities Bulletin. 34*(1), 15-42. https://eric.ed.gov/?id=EJ815705

Schoemaker, K., Mulder, H., Deković, M., & Matthys, W. (2013). Executive Functions in Preschool Children with Externalizing Behavior Problems: A Meta-Analysis. *Journal of Abnormal Child Psychology*, *41*(3), 457–471. https://doi.org/10.1007/s10802-012-9684-x

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *298*(1089), 199–209. https://doi.org/10.1098/rstb.1982.0082

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge University Press.

Shallice, T., & Burgess, P. W. (1991). Deficits in Strategy Application Following Frontal Lobe Damage in Man. *Brain*, *114*(2), 727–741. https://doi.org/10.1093/brain/114.2.727

Simpson, Andrew., & Riggs, K. J. (2005). Inhibitory and working memory demands of the day-night task in children. *British Journal of Developmental Psychology*, *23*(3), 471–486. https://doi.org/10.1348/026151005X28712

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. https://doi.org/10.1080/17470210500162854

Stern, R. A., & Prohaska, M. L. (1996). Neuropsychological evaluation of executive functioning. *American Psychiatric Press Review of Psychiatry*, *15*, 243–266.

Stuss, D. T., & Alexander, M. P. (2000). Executive functions and the frontal lobes: A conceptual view. *Psychological Research*, *63*(3–4), 289–298. https://doi.org/10.1007/s004269900007

Ten Eycke, K. D., & Dewey, D. (2016). Parent-report and performance-based measures of executive function assess different constructs. *Child Neuropsychology*, *22*(8), 889–906. https://doi.org/10.1080/09297049.2015.1065961

Thorell, L. B., & Nyberg, L. (2008). The Childhood Executive Functioning Inventory (CHEXI): A New Rating Instrument for Parents and Teachers. *Developmental Neuropsychology*, *33*(4), 536–552. https://doi.org/10.1080/87565640802101516

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner Review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*(2), 131–143. https://doi.org/10.1111/jcpp.12001

Tulsky, D. S., Carlozzi, N. E., Chevalier, N., Espy, K. A., Beaumont, J. L., & Mungas, D. (2013). V. NIH Toolbox Cognition Battery (CB): Measuring Working Memory. *Monographs of the Society for Research in Child Development*, *78*(4), 70–87. https://doi.org/10.1111/mono.12035 van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence*, *35*(5), 427–449. https://doi.org/10.1016/j.intell.2006.09.001

Vanegas, S. B., & Davidson, D. (2015). Investigating distinct and related contributions of Weak Central Coherence, Executive Dysfunction, and Systemizing theories to the cognitive profiles of children with Autism Spectrum Disorders and typically developing children. *Research in Autism Spectrum Disorders*, *11*, 77–92. https://doi.org/10.1016/j.rasd.2014.12.005

Vaughan, L., & Giovanello, K. (2008). Executive function in daily life: Age-related influences of executive processes on instrumental activities of daily living. *Psychology and Aging 25*(2), 343-355. https://doi.org/10.1037/a0017729

Wallace, G. L., Silvers, J. A., Martin, A., & Kenworthy, L. E. (2009). Brief report: Further evidence for inner speech deficits in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*(12), 1735–1739. https://doi.org/10.1007/s10803-009-0802-8

Wechsler, D. (2002). *Wechsler primary and preschool scale of intelligence.* The Psychological Corporation.

Wechsler, D., & Naglieri, J. A. (2006). *Wechsler Nonverbal Scale of Ability (WNV)*. Harcourt Assessments.

Weintraub, S., Bauer, P. J., Zelazo, P. D., Wallner-Allen, K., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Slotkin, J., Blitz, D. L., Carlozzi, N. E., Havlik, R. J., Beaumont, J. L., Mungas, D., Manly, J. J., Borosh, B. G., Nowinski, C. J., & Gershon, R. C. (2013). National Institutes of Health Toolbox Cognition Battery (NIH Toolbox CB): Validation for children between 3 and 15 years: I NIH Toolbox Cognition Battery (CB): Introduction and pediatric data. *Monographs of the Society for Research in Child Development*, *78*(4), 1–15. https://doi.org/10.1111/mono.12031

Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., Carlozzi, N. E., Slotkin, J., Blitz, D., Wallner-Allen, K., Fox, N. A., Beaumont, J. L., Mungas, D., Nowinski, C. J., Richler, J., Deocampo, J. A., Anderson, J. E., Manly, J. J., Borosh, B., … Gershon, R. C. (2013). Cognition assessment using the NIH Toolbox. *Neurology*, *80*(Issue 11, Supplement 3), S54–S64. https://doi.org/10.1212/WNL.0b013e3182872ded

Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology*, *7*(2), 131–149. https://doi.org/10.1080/87565649109540483

Welsh, M., & Peterson, E. (2014). Issues in the conceptualization and assessment of hot executive functions in childhood. *Journal of the International Neuropsychological Society*, *20*(2), 152–156. https://doi.org/10.1017/S1355617713001379

White, S. J., Burgess, P. W., & Hill, E. L. (2009). Impairments on ‘open-ended’ executive function tests in autism. *Autism Research*, *2*(3), 138–147. https://doi.org/10.1002/aur.78

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, *44*(2), 575–587. https://doi.org/10.1037/0012-1649.44.2.575

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology*, *108*(3), 436–452. https://doi.org/10.1016/j.jecp.2010.08.008

Wilbourn, M. P., Kurtz, L. E., & Kalia, V. (2012). The Lexical Stroop Sort (LSS) picture-word task: A computerized task for assessing the relationship between language and executive functioning in school-aged children. *Behavior Research Methods*, *44*(1), 270–286. https://doi.org/10.3758/s13428-011-0142-4

Williams, D., & Jarrold, C. (2013). Assessing planning and set-shifting abilities in autism: Are experimenter-administered and computerised versions of tasks equivalent? *Autism Research*, *6*(6), 461–467. https://doi.org/10.1002/aur.1311

Williams, D. L., Goldstein, G., & Minshew, N. J. (2006). The profile of memory function in children with autism. *Neuropsychology*, *20*(1), 21–29. https://doi.org/10.1037/0894-4105.20.1.21

Willig, C. (2019). Ontological and epistemological reflexivity: A core skill for therapists. *Counselling and Psychotherapy Research*, *19*(3), 186–194. https://doi.org/10.1002/capr.12204

Wilson, J., Andrews, G., & Shum, D. H. K. (2017). Delay of gratification in middle childhood: Extending the utility and sensitivity of the standard task. *PsyCh Journal*, *6*(1), 8–15. https://doi.org/10.1002/pchj.161

Wing, L., Gould, J., & Gillberg, C. (2011). Autism spectrum disorders in the DSM-V: Better or worse than the DSM-IV? *Research in Developmental Disabilities*, *32*(2), 768–773. https://doi.org/10.1016/j.ridd.2010.11.003

Yerys, B. E., Hepburn, S. L., Pennington, B. F., & Rogers, S. J. (2007). Executive function in preschoolers with autism: Evidence consistent with a secondary deficit. *Journal of Autism and Developmental Disorders*, *37*(6), 1068–1079. https://doi.org/10.1007/s10803-006-0250-7

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, *1*(1), 297–301. https://doi.org/10.1038/nprot.2006.46

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). NIH toolbox cognition battery (CB): Measuring executive function and attention. *Monographs of the Society for Research in Child Development*, *78*(4), 16–33. https://doi.org/10.1111/mono.12032

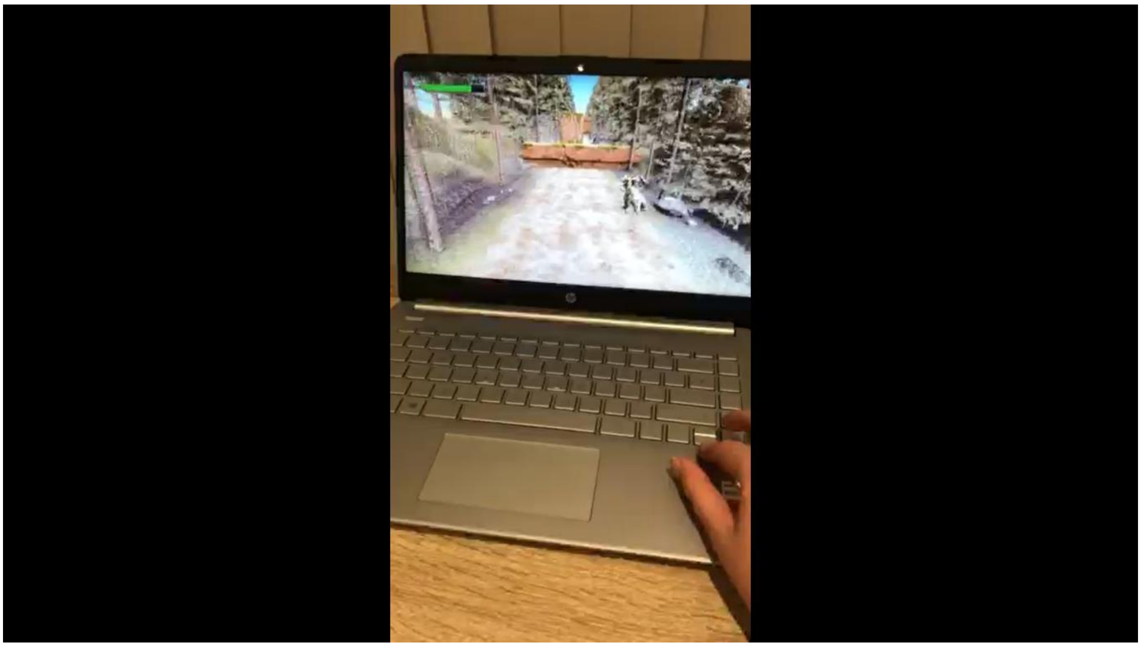
Zelazo, P. D., & Carlson, S. M. (2012). Hot and Cool Executive Function in Childhood and Adolescence: Development and Plasticity. *Child Development Perspectives*, *6*(4), 354-360. https://doi.org/10.1111/j.1750-8606.2012.00246.x

Zelazo, P. D., Carlson, S. M., & Kesek, A. (2008). The development of executive function in childhood. In C. A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 553–574). MIT Press.

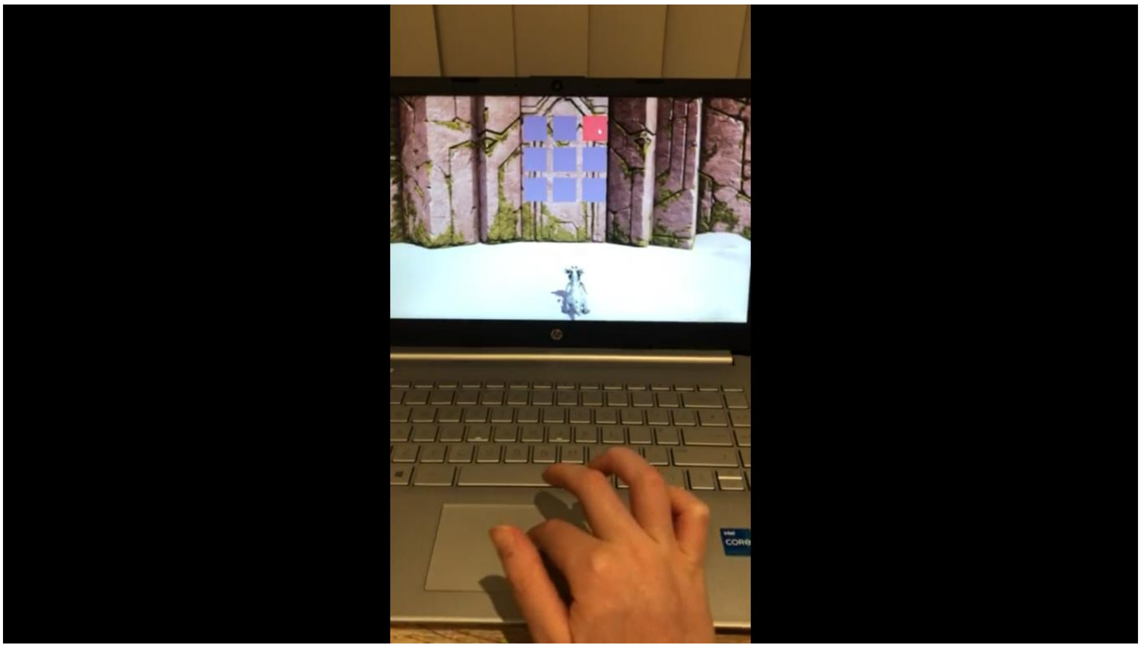
Zelazo, P. D., & Müller, U. (2002). The Balance Beam in the Balance: Reflections on Rules, Relational Complexity, and Developmental Processes. *Journal of Experimental Child Psychology*, *81*(4), 458–465. https://doi.org/10.1006/jecp.2002.2667

APPENDIX A: Video Demonstration Screenshots

Dragon Dash



Dragon Sequence:



Dragon Hunt

APPENDIX B: CTUQ Scoring Procedure

1. Do you know what a [technology] is?
   Yes **(1 point)** No **(0 points)**

2. Do you have your own [technology]?
   Yes **(1 point)** No **(0 points)**

3. Do you use a [technology] at home?
   Yes **(1 point)** No **(0 points)**

4. How many days a week do you use a [technology] (circle your answer)?
   0 days **(0 points)**
   1 or 2 days **(1 point)**
   3 or 4 days **(2 points)**
   5 or 6 days **(3 points)**
   7 days **(4 points)**

5. How good do you think you are at using a [technology] (circle your answer)?
   very good **(4 points)**
   good **(3 points)**
   okay **(2 points)**
   not very good **(1 point)**
   poor **(0 points)**

6. Does your school have [technology] you can use in class?
   Yes **(1 point)** No **(0 points)**

7. 7. Do you use it with your teacher in your class?
   Yes **(1 points)** No **(0 points)**

8. Have you ever [carried out a task] on a [technology]?
   Yes **(1 point)** No **(0 points)**

9. Can you write down things that you use your [technology] for?
   (1 point for each item)

APPENDIX C: Visual Analogue Scale

Please tell us how you found the tasks……………

I enjoyed the task

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

APPENDIX D: University of East London School of Psychology Research Ethics Approval

UNIVERSITY OF EAST LONDON
School of Psychology

APPLICATION FOR RESEARCH ETHICS APPROVAL FOR RESEARCH INVOLVING HUMAN PARTICIPANTS (Updated October 2019)

FOR BSc RESEARCH
FOR MSc/MA RESEARCH
FOR PROFESSIONAL DOCTORATE RESEARCH IN CLINICAL, COUNSELLING & EDUCATIONAL PSYCHOLOGY

1. Completing the application

1.1 Before completing this application please familiarise yourself with the British Psychological Society's Code of Ethics and Conduct (2018) and the UEL Code of Practice for Research Ethics (2015-16). Please tick to confirm that you have read and understood these codes: ✓

1.2 Email your supervisor the completed application and all attachments as ONE WORD DOCUMENT. Your supervisor will then look over your application.

1.3 When your application demonstrates sound ethical protocol, your supervisor will submit it for review. By submitting the application, the supervisor is confirming that they have reviewed all parts of this application, and consider it of sufficient quality for submission to the SREC committee for review. It is the responsibility of students to check that the supervisor has checked the application and sent it for review.

1.4 Your supervisor will let you know the outcome of your application. Recruitment and data collection must NOT commence until your ethics application has been approved, along with other research ethics approvals that may be necessary (see section 8).

1.5 Please tick to confirm that the following appendices have been completed. Note: templates for these are included at the end of the form.

- The participant invitation letter ✓
- The participant consent form ✓

- The participant debrief letter ☐

1.6 The following attachments should be included if appropriate. In each case, please tick to either confirm that you have included the relevant attachment, or confirm that it is not required for this application.

- A participant advert, i.e., any text (e.g., email) or document (e.g., poster) designed to recruit potential participants.

  Included ☐ or

  Not required (because no participation adverts will be used) ☑

- A general risk assessment form for research conducted off campus (see section 6).

  Included ☑ or

  Not required (because the research takes place solely on campus or online) ☐

- A country-specific risk assessment form for research conducted abroad (see section 6).

  Included ☐ or

  Not required (because the researcher will be based solely in the UK) ☑

- A Disclosure and Barring Service (DBS) certificate (see section 7).

  Included ☑ or

  Not required (because the research does not involve children aged 16 or under or vulnerable adults) ☐

- Ethical clearance or permission from an external organisation (see section 8).

  Included ☑ or

  Not required (because no external organisations are involved in the research) ☐

- Original and/or pre-existing questionnaire(s) and test(s) you intend to use.

  Included ☑ or

  Not required (because you are not using pre-existing questionnaires or tests) ☐

- Interview questions for qualitative studies.

  Included ☑ or

Not required (because you are not conducting qualitative interviews) ☐

- Visual material(s) you intend showing participants.
  Included ☑ or

  Not required (because you are not using any visual materials) ☐

2. Your details

2.1 Your name: **Amber Scott**

2.2 Your supervisor's name: **Dr Matthew Jones Chesters**

2.3 Title of your programme: **Professional Doctorate in Clinical Psychology**

2.4 UEL assignment submission date (stating both the initial date and the resit date): **May 2022**

3. Your research

Please give as much detail as necessary for a reviewer to be able to fully understand the nature and details of your proposed research.

3.1 The title of your study:
THE USE OF A NEWLY DEVELOPED COMPUTER GAME TO TEST EXECUTIVE FUNCTIONING IN YOUNG TYPICALLY DEVELOPING CHILDREN AND CHILDREN ON THE AUTISM SPECTRUM

3.2 Your research question:
Does Dragon Adventure assess inhibition, working memory flexibility in line with established tests of Executive Function and caregiver/teacher ratings in typically developing children aged 6 to 8? If so, what are the norms for this age group?
Is Dragon Adventure suitable for assessing inhibition, working memory and flexibility in line with established tests of EF and caregiver/teacher ratings for children ages 6 to 12 on the autism spectrum?
Is there a difference in rated enjoyment between the established tests and Dragon Kingdom in typically developing children aged 6 to 8 and children on the autism spectrum aged 6 to 12?

3.3 Design of the research:
Cross-sectional correlation design for validating Dragon Adventure against the established tests
Within-subjects design for comparison of engagement ratings
Content analysis on responses to how the game could be developed.

3.4 Participants:

30 children aged 6 to 8 years old
30 children aged 6 to 12 on the autism spectrum

3.5 Recruitment:

Recruitment of children aged 6 to 8 will be completed through primary schools. Recruitment of children on the autism spectrum will be completed through a school and a local charity which supports such children. Consent will initially be sought from the parents of both populations, then consent will also be obtained from the children themselves.

3.6 Measures, materials or equipment:

Dragon Adventure – computerised game-based test of executive functions (inhibition, working memory and flexibility)
D-KEFS Colour -word inference test assessing inhibition in children 8+
Day-Night task (Gerstadt et al., 1994) assessing inhibition in children under 8
WNV Spatial span assessing working memory in all children
D-KEFS Trail making assessing flexibility in children aged 8+
Child Color Trials (Williams et al., 1995) to assess flexibility in children under 8
Visual analogue Likert scale of enjoyability
Childhood Executive Functioning Inventory (CHEXI; Thorrell & Nyber, 2008)
Children's Technology Use Questionnaire (CTUQ; Horton, 2013)
Question about how the game could be developed
Equipment – Laptop

3.7 Data collection:

Data will be collected within the primary schools and the charity organisation. Each participant will take part in a session with myself whereby all the measures and materials are completed.

3.8 Data analysis:

Quantitative data will be analysed using SPSS, a correlation will be completed to assess the relationship between scores on Dragon Adventure and the established measures, controlling for age, gender, culture, English as a first language and computer literacy. A t-test will be completed to assess the difference between the enjoyability ratings on the established measures and Dragon Adventure, also controlling for the above.
A content analysis will be used to analyse qualitative feedback on the further development of the game.

4. Confidentiality and security

It is vital that data are handled carefully, particularly the details about participants. For information in this area, please see the UEL guidance on data protection, and also the UK government guide to data protection regulations.

4.1 Will participants data be gathered anonymously?

Yes – participants will be allocated a number by which their data will be gathered and stored.

4.2 If not (e.g., in qualitative interviews), what steps will you take to ensure their anonymity in the subsequent steps (e.g., data analysis and dissemination)?

4.3 How will you ensure participants details will be kept confidential?
Any information which is not anonymous e.g. consent forms, will be scanned and stored securely, then deleted once the research has been completed and assessed. All data will be anonymised through recording against an allocated number.

4.4 How will the data be securely stored?
Folders or documents containing data will be password protected and stored securely on UEL One Drive.

4.5 Who will have access to the data?
The only person who will have access to the data is myself, however, the data may also be looked at by my Director of Studies and could be requested by examiners.

4.6 How long will data be retained for?
The data will be kept for 3 years following the completion of the research.

5. Informing participants

Please confirm that your information letter includes the following details:

5.1 Your research title: ✓

5.2 Your research question: ✓

5.3 The purpose of the research: ✓

5.4 The exact nature of their participation. This includes location, duration, and the tasks etc. involved: ✓

5.5 That participation is strictly voluntary: ✓

5.6 What are the potential risks to taking part: ✓

5.7 What are the potential advantages to taking part: ✓

5.8 Their right to withdraw participation (i.e., to withdraw involvement at any point, no questions asked): ✓

5.9 Their right to withdraw data (usually within a three-week window from the time of their participation): ✓

5.10 How long their data will be retained for: ✓

5.11 How their information will be kept confidential: ☑

5.12 How their data will be securely stored: ☑

5.13 What will happen to the results/analysis: ☑

5.14 Your UEL contact details: ☑

5.15 The UEL contact details of your supervisor: ☑

Please also confirm whether:

5.16 Are you engaging in deception? If so, what will participants be told about the nature of the research, and how will you inform them about its real nature.
NO

5.17 Will the data be gathered anonymously? If NO what steps will be taken to ensure confidentiality and protect the identity of participants?
YES

5.18 Will participants be paid or reimbursed? If so, this must be in the form of redeemable vouchers, not cash. If yes, why is it necessary and how much will it be worth?
NO

6. Risk Assessment

Please note: If you have serious concerns about the safety of a participant, or others, during the course of your research please see your supervisor as soon as possible. If there is any unexpected occurrence while you are collecting your data (e.g. a participant or the researcher injures themselves), please report this to your supervisor as soon as possible.

6.1 Are there any potential physical or psychological risks to participants related to taking part? If so, what are these, and how can they be minimised?
There is a risk of completing the research during a pandemic. To minimise risk of infection for the participant, current guidelines will be followed i.e. masks will be worn, the room will be large enough for social distancing and hands and surfaces will be regularly washed/sanitized. The researcher will be completing lateral flow tests twice a week and will isolate for 10 days if the test is positive. Public transport will be avoided where possible when travelling, if this is not possible, the safest routes will be taken. The researcher will adhere to the school's process for risk assessments.

6.2 Are there any potential physical or psychological risks to you as a researcher? If so, what are these, and how can they be minimised?
There is a risk of completing the research during a pandemic. To minimise risk of infection for the researcher, guidelines will be followed i.e. masks will be worn, the room will be large enough for social distancing and hands and surfaces will be regularly washed/sanitized. The researcher has received both doses of the vaccine and will be completing lateral flow tests twice a week. Public transport will be avoided where possible when travelling, if this is not possible, the safest routes will be taken. The researcher will adhere to the school's process for risk assessments.

6.3 Have appropriate support services been identified in the debrief letter? If so, what are these, and why are they relevant?
N/A

6.4 Does the research take place outside the UEL campus? If so, where?
Yes – the research will take place on primary school campuses and within the charity organisations base.

If so, a 'general risk assessment form' must be completed. This is included below as appendix D. Note: if the research is on campus, or is online only (e.g., a Qualtrix survey), then a risk assessment form is not needed, and this appendix can be deleted. If a general risk assessment form is required for this research, please tick to confirm that this has been completed: ✓

6.5 Does the research take place outside the UK? If so, where?
No

If so, in addition to the 'general risk assessment form', a 'country-specific risk assessment form' must be also completed (available in the Ethics folder in the Psychology Noticeboard), and included as an appendix. [Please note: a country-specific risk assessment form is not needed if the research is online only (e.g., a Qualtrix survey), regardless of the location of the researcher or the participants.] If a 'country-specific risk assessment form' *is* needed, please tick to confirm that this has been included:

However, please also note:

- For assistance in completing the risk assessment, please use the AIG Travel Guard website to ascertain risk levels. Click on 'sign in' and then 'register here' using policy # 0015865161. Please also consult the Foreign Office travel advice website for further guidance.
- For *on campus* students, once the ethics application has been approved by a reviewer, all risk assessments for research abroad must then be signed by the Head of School (who may escalate it up to the Vice Chancellor).

- For *distance learning* students conducting research abroad in the country where they currently reside, a risk assessment must be also carried out. To minimise risk, it is recommended that such students only conduct data collection on-line. If the project is deemed low risk, then it is not necessary for the risk assessments to be signed by the Head of School. However, if not deemed low risk, it must be signed by the Head of School (or potentially the Vice Chancellor).
- Undergraduate and M-level students are not explicitly prohibited from conducting research abroad. However, it is discouraged because of the inexperience of the students and the time constraints they have to complete their degree.

7. Disclosure and Barring Service (DBS) certificates

7.1 Does your research involve working with children (aged 16 or under) or vulnerable adults (*see below for definition)?

YES / ~~NO~~

7.2 If so, you will need a current DBS certificate (i.e., not older than six months), and to include this as an appendix. Please tick to confirm that you have included this: ☑

Alternatively, if necessary for reasons of confidentiality, you may email a copy directly to the Chair of the School Research Ethics Committee. Please tick if you have done this instead: ☐

Also alternatively, if you have an Enhanced DBS clearance (one you pay a monthly fee to maintain) then the number of your Enhanced DBS clearance will suffice. Please tick if you have included this instead: ☐

7.3 If participants are under 16, you need 2 separate information letters, consent form, and debrief form (one for the participant, and one for their parent/guardian). Please tick to confirm that you have included these: ☑

7.4 If participants are under 16, their information letters consent form, and debrief form need to be written in age-appropriate language. Please tick to confirm that you have done this ☑

* You are required to have DBS clearance if your participant group involves (1) children and young people who are 16 years of age or under, and (2) 'vulnerable' people aged 16 and over with psychiatric illnesses, people who receive domestic care, elderly people (particularly those in nursing homes), people in palliative care, and people living in institutions and sheltered accommodation, and people who have been involved in the criminal justice system, for example. Vulnerable people are understood to be persons

who are not necessarily able to freely consent to participating in your research, or who may find it difficult to withhold consent. If in doubt about the extent of the vulnerability of your intended participant group, speak to your supervisor. Methods that maximise the understanding and ability of vulnerable people to give consent should be used whenever possible. For more information about ethical research involving children click here.

8. **Other permissions**

9. Is HRA approval (through IRAS) for research involving the NHS required? Note: HRA/IRAS approval is required for research that involves patients or Service Users of the NHS, their relatives or carers as well as those in receipt of services provided under contract to the NHS.

   ~~YES~~ / NO   If yes, please note:

- You DO NOT need to apply to the School of Psychology for ethical clearance if ethical approval is sought via HRA/IRAS (please see further details here).
- However, the school *strongly discourages* BSc and MSc/MA students from designing research that requires HRA approval for research involving the NHS, as this can be a very demanding and lengthy process.
- If you work for an NHS Trust and plan to recruit colleagues from the Trust, permission from an appropriate manager at the Trust must be sought, and HRA approval will probably be needed (and hence is likewise strongly discouraged). If the manager happens to not require HRA approval, their written letter of approval must be included as an appendix.
- IRAS approval is not required for NHS staff even if they are recruited via the NHS (UEL ethical approval is acceptable). However, an application will still need to be submitted to the HRA in order to obtain R&D approval. This is in addition to a separate approval via the R&D department of the NHS Trust involved in the research.
- IRAS approval is not required for research involving NHS employees when data collection will take place off NHS premises, and when NHS employees are not recruited directly through NHS lines of communication. This means that NHS staff can participate in research without HRA approval when a student recruits via their own social or professional networks or through a professional body like the BPS, for example.

9.1 Will the research involve NHS employees who will not be directly recruited through the NHS, and where data from NHS employees will not be collected on NHS premises?

   ~~YES~~ / NO

9.2 If you work for an NHS Trust and plan to recruit colleagues from the Trust, will permission from an appropriate member of staff at the Trust be sought, and will

HRA be sought, and a copy of this permission (e.g., an email from the Trust) attached to this application?

~~YES / NO~~ - N/A

9.3 Does the research involve other organisations (e.g. a school, charity, workplace, local authority, care home etc.)? If so, please give their details here.
Yes – [SCHOOL NAME] and [SCHOOL NAME] have given permission for the recruitment and data collection to be completed through them. Permission may also be obtained from other schools at a later date.

Furthermore, written permission is needed from such organisations if they are helping you with recruitment and/or data collection, if you are collecting data on their premises, or if you are using any material owned by the institution/organisation. If that is the case, please tick here to confirm that you have included this written permission as an appendix:

☑

In addition, before the research commences, once your ethics application has been approved, please ensure that you provide the organisation with a copy of the final, approved ethics application. Please then prepare a version of the consent form for the organisation themselves to sign. You can adapt it by replacing words such as 'my' or 'I' with 'our organisation,' or with the title of the organisation. This organisational consent form must be signed before the research can commence.

Finally, please note that even if the organisation has their own ethics committee and review process, a School of Psychology SREC application and approval is still required. Ethics approval from SREC can be gained before approval from another research ethics committee is obtained. However, recruitment and data collection are NOT to commence until your research has been approved by the School and other ethics committee/s as may be necessary.

9. Declarations

Declaration by student: I confirm that I have discussed the ethics and feasibility of this research proposal with my supervisor.

Student's name (typed name acts as a signature):
Amber Scott
Student's number: u1945528 Date: 22/04/2021

As a supervisor, by submitting this application, I confirm that I have reviewed all parts of this application, and I consider it of sufficient quality for submission to the SREC committee.

School of Psychology Research Ethics Committee

NOTICE OF ETHICS REVIEW DECISION

For research involving human participants
BSc/MSc/MA/Professional Doctorates in Clinical, Counselling and Educational Psychology

REVIEWER: Matthew Boardman

SUPERVISOR: Matthew Jones Chesters

STUDENT: Amber Scott

Course: Prof Doc Clinical Psychology

DECISION OPTIONS:

1. **APPROVED:** Ethics approval for the above named research study has been granted from the date of approval (see end of this notice) to the date it is submitted for assessment/examination.

2. **APPROVED, BUT MINOR AMENDMENTS ARE REQUIRED BEFORE THE RESEARCH COMMENCES** (see Minor Amendments box below)**:** In this circumstance, re-submission of an ethics application is not required but the student must confirm with their supervisor that all minor amendments have been made before the research commences. Students are to do this by filling in the confirmation box below when all amendments have been attended to and emailing a copy of this decision notice to her/his supervisor for their records. The supervisor will then forward the student's confirmation to the School for its records.

3. **NOT APPROVED, MAJOR AMENDMENTS AND RE-SUBMISSION REQUIRED** (see Major Amendments box below)**:** In this circumstance, a revised ethics application must be submitted and approved before any research takes place. The revised application will be reviewed by the same reviewer. If in doubt, students should ask their supervisor for support in revising their ethics application.

DECISION ON THE ABOVE-NAMED PROPOSED RESEARCH STUDY
(Please indicate the decision according to one of the 3 options above)

Approved

Minor amendments required *(for reviewer)*:

NA

Major amendments required *(for reviewer):*

NA

Confirmation of making the above minor amendments *(for students)*:

I have noted and made all the required minor amendments, as stated above, before starting my research and collecting data.

Student's name *(Typed name to act as signature):*
Student number:
Date:
(Please submit a copy of this decision letter to your supervisor with this box completed, if minor amendments to your ethics application are required)

ASSESSMENT OF RISK TO RESEACHER *(for reviewer)*

Has an adequate risk assessment been offered in the application form?
YES

If the proposed research could expose the researcher to any of kind of emotional, physical or health and safety hazard? Please rate the degree of risk:

☐ HIGH

Please do not approve a high risk application and refer to the Chair of Ethics. Travel to countries/provinces/areas deemed to be high risk should not be permitted and an application not approved on this basis. If unsure please refer to the Chair of Ethics.

☐ MEDIUM (Please approve but with appropriate recommendations)

X ***LOW***

Reviewer comments in relation to researcher risk (if any).

NA

Reviewer *(Typed name to act as signature)*: Dr Matthew Boardman

Date: 27.5.21

This reviewer has assessed the ethics application for the named research study on behalf of the School of Psychology Research Ethics Committee

RESEARCHER PLEASE NOTE:

For the researcher and participants involved in the above named study to be covered by UEL's Insurance, prior ethics approval from the School of Psychology (acting on behalf of the UEL Research Ethics Committee), and confirmation from students where minor amendments were required, must be obtained before any research takes place. For a copy of UELs Personal Accident & Travel Insurance Policy, please see the Ethics Folder in the Psychology Noticeboard

School of Psychology Ethics Committee

REQUEST FOR TITLE CHANGE TO AN ETHICS APPLICATION

For BSc, MSc/MA and taught Professional Doctorate students

Please complete this form if you are requesting approval for a proposed title change to an ethics application that has been approved by the School of Psychology

By applying for a change of title request, you confirm that in doing so, the process by which you have collected your data/conducted your research has not changed or deviated from your original ethics approval. If either of these have changed, then you are required to complete an 'Ethics Application Amendment Form'.

How to complete and submit the request	
1	Complete the request form electronically.
2	Type your name in the 'student's signature' section (page 2).
3	Using your UEL email address, email the completed request form along with associated documents to Dr Jérémy Lemoine (School Ethics Committee Member): j.lemoine@uel.ac.uk
4	Your request form will be returned to you via your UEL email address with the reviewer's decision box completed. Keep a copy of the approval to submit with your dissertation.

Required documents	
A copy of the approval of your initial ethics application.	**YES** ☒

Details	
Name of applicant:	**Amber Scott**
Programme of study:	**Professional Doctorate in Clinical Psychology**
Title of research:	**The Use of a Newly Developed Computer Game to Test Executive Functioning in Young Typically Developing Children and Children on the Autism Spectrum**

Name of supervisor:	**Dr Matthew Jones Chesters**

Proposed title change

Briefly outline the nature of your proposed title change in the boxes below

Old title:	The Use of a Newly Developed Computer Game to Test Executive Functioning in Young Typically Developing Children and Children on the Autism Spectrum
New title:	The Use of a Newly Developed Computer Game to Measure Executive Functioning in Young Neurotypical Children and Children with a Diagnosis of an Autism Spectrum Condition
Rationale:	So the title better reflects the language used in the write-up of the research.

Confirmation

Is your supervisor aware of your proposed change of title and in agreement with it?	**YES** ☒	**NO** ☐
Does your change of title impact the process of how you collected your data/conducted your research?	**YES** ☐	**NO** ☒

Student's signature

Student: (Typed name to act as signature)	**Amber Scott**
Date:	**30/03/2022**

Reviewer's decision

Title change approved:	**YES** ☒	**NO** ☐
Comments:	**The new title reflects better the research study and will not impact the process of how the data are collected or how the research is conducted.**	
Reviewer: (Typed name to act as signature)	**Dr Jérémy Lemoine**	
Date:	**30/03/2022**	

APPENDIX E: Organisation Information Sheet

ORGANISATION INVITATION LETTER

The Use of a Newly Developed Computer Game to Test Executive Functioning in Young Typically Developing Children and Children on the Autism Spectrum

You are being invited to participate in a research study. Before you agree it is important that you understand what participation would involve. Please take time to read the following information carefully.

Who am I?

My name is Amber Scott, I am a Trainee Clinical Psychologist. This study is being conducted as part of my Professional Doctorate in Clinical Psychology at the University of East London.

What is the research?

I am conducting research into improving neuropsychological tests of executive function for children. Executive functions include children's ability to plan, adjust and organise their thinking and behaviour. The aim of this study is to assess whether a newly developed game can successfully test executive function in young children and children on the autism spectrum in a more engaging and accessible manner than tests that are currently available. I am also looking to obtain feedback from children on how the game could be further developed.

Current tests of executive function are often limited by time, cultural norms and language. This newly developed game hopes to address some of these limitations and if children do find it more engaging it could help us measure these skills more accurately.

My research has been approved by the School of Psychology Research Ethics Committee. This means that the Committee's evaluation of this ethics application has been guided by the standards of research ethics set by the British Psychological Society.

Why has your organisation been asked to participate?

Your organisation has been invited to participate in my research as you host the kind of people I am looking for to help me explore my research topic. I am looking to involve children aged 6 to 8 and children aged 6 to 12 on the autism spectrum.

The children who agree to participate will not be judged or personally analysed in any way and will be treated with respect.

You, as an organisation, are quite free to decide whether or not to participate and should not feel coerced.

What will your participation involve?

If you agree to participate, the children you host, and their parents, will be asked whether they would like to participate in this study. Children will attend a session with myself, where they will be asked to complete some pen and paper neuropsychological measures and a new game developed to measure executive function in children, rating their level of enjoyment for each. They will also be asked to complete a questionnaire about their level of computer literacy and for feedback on how the game could be further developed. The session should take about an hour and will take place at your location. We would also ask the child's teacher to fill in a brief questionnaire about the child's ability to plan, adjust and organise their thinking and behaviours. The aim of this is to find out whether the measures are related to real-life strengths and/or difficulties.

I will not be able to pay for children's participation in my research, but their participation would be very valuable in helping to develop knowledge and understanding of my research topic.

Taking part will be safe and confidential

The children's privacy and safety will be respected at all times. Participant's data will be kept anonymous, meaning they will not be able to be identified by the data collected, on any written material or in the write-up of the research. Parent's and children's consent forms will be stored securely and separately from the rest of the data and will be destroyed following completion of the research.

Participants do not have to complete all tasks asked of them and are free to stop their participation at any time.

To ensure the children's and my own safety, social distancing will be maintained at all times, I will wear a mask and sanitizing of hands and equipment will be completed regularly.

What will happen to the information provided?

What I will do with the material children provide will involve anonymously storing all data on a personal drive, only I have access to, which will be password protected. Data will be anonymised through participants being allocated a number which their data will be recorded against; there will be no way of identifying who has been assigned to each number. The anonymised data will be reviewed by myself and my supervisor and may be requested by examiners. Summaries of the data collected will be available in the write-up and may be published in an academic journal, the thesis will also be publicly accessible on UEL's institutional repository. Some broad demographic information may appear in the thesis and works based on it but that this will not be such as to permit the identification of individual participants. Once the research has been completed, the data will be kept for three years, following this, the data will be destroyed. Once the data has been collected children and their parents can withdraw the data up to the end of January 2022.

What if a child or their parent wants to withdraw?

Parents and children are free to withdraw from the research study at any time without explanation, disadvantage or consequence. Separately, parents and children may also request to withdraw their data even after they have participated, provided that this request is made before the end of January 2022 (after which point the data analysis will begin, and withdrawal will not be possible).

Contact Details

If you would like further information about my research or have any questions or concerns, please do not hesitate to contact me. My email address is u1945528@uel.ac.uk.

If you have any questions or concerns about how the research has been conducted please contact the research supervisor Dr Matthew Jones Chesters. School of Psychology, University of East London, Water Lane, London E15 4LZ,
Email: m.h.jones-chesters@uel.ac.uk.

or

Chair of the School of Psychology Research Ethics Sub-committee: Dr Trishna Patel, School of Psychology, University of East London, Water Lane, London E15 4LZ.

(Email: t.patel@uel.ac.uk)

APPENDIX F: Organisation Consent Form

UNIVERSITY OF EAST LONDON

Consent to participate in a research study

The Use of a Newly Developed Computer Game to Test Executive Functioning in Young Typically Developing Children and Children on the Autism Spectrum

I confirm that I have read the information sheet for the above study and that I have been given a copy to keep. ☐

I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

I understand that the organisations participation in the study is voluntary and that we may withdraw at any time, without providing a reason for doing so. ☐

I understand that any personal information and data from the research will be securely stored and remain strictly confidential. Only the research team will have access to this information, to which I give my permission. ☐

It has been explained to me what will happen to the data once the research has been completed. ☐

I understand that the thesis will be publicly accessible in the University of East London's Institutional Repository (ROAR). ☐

I would like to receive a summary of the research findings once the study has been completed and am willing to provide contact details for this to be sent to. ☐

I agree to take part in the above study. ☐

Organisation's Name (BLOCK CAPITALS)

……………………………………………………………………………………….

Identified Persons Name (BLOCK CAPITALS)

……………………………………………………………………………………….

Identified Persons Signature

………………………………………………………………………………………..

Researcher's Name (BLOCK CAPITALS)

………………………………………………………………………………………..

Researcher's Signature

…………………………………………………………………………………………

Date: ……………………..…….

APPENDIX G: Parent Information Sheet

University of East London

PARENT/CARER PARTICIPANT INVITATION LETTER

The Use of a Newly Developed Computer Game to Test Executive Functioning in Young Typically Developing Children and Children on the Autism Spectrum

Your child, or the child you care for, is being invited to participate in a research study. Before you agree it is important that you understand what your child's participation would involve. Please take time to read the following information carefully and if you are happy for your child to participate, please complete the consent form attached. I have also included a children's version of this information sheet and a consent form, which you can go through with your child.

Who am I?

My name is Amber Scott, I am a Trainee Clinical Psychologist. This study is being conducted as part of my Professional Doctorate in Clinical Psychology at the University of East London.

What is the research?

I am conducting research into improving neuropsychological tests of executive function for children. Executive functions include children's ability to plan, adjust and organise their thinking and behaviour. The aim of this study is to assess whether a newly developed game can successfully measure executive function in young children and children on the autism spectrum in a more engaging and accessible manner than tests that are currently available. I am also looking to obtain feedback from children on how the game could be further developed.

Current tests of executive function are often limited by time, cultural norms and language. This newly developed game hopes to address some of these limitations and if children do find it more engaging it could help us measure these skills more accurately.

My research has been approved by the School of Psychology Research Ethics Committee. This means that the Committee's evaluation of this ethics application has been guided by the standards of research ethics set by the British Psychological Society.

Why have you been asked to participate?

Your child has been invited to participate in my research as someone who fits the kind of people I am looking for to help me explore my research topic. I am looking to involve children aged 6 to 8 and children aged 6 to 12 on the autism spectrum.

Your child will not be judged or personally analysed in any way and will be treated with respect.

You and your child are free to decide whether or not to participate and should not feel coerced.

What will your child's participation involve?

If you agree for your child to participate, your child will be asked to attend a session with myself, where they will be asked to complete pen and paper neuropsychological measures and a new game developed to measure the child's ability to plan, adjust and organise their thinking and behaviours, rating their level of enjoyment for each. Your child will be asked to complete a questionnaire about their level of computer literacy and for feedback on how the game could be further developed. The session should take about an hour and will take place in school. We would also ask your child's teacher to fill in a brief questionnaire about the child's ability to plan, adjust and organise their thinking and behaviours. The aim of this is to find out whether the measures are related to real-life strengths and/or difficulties.

I will not be able to pay your child for participating in my research, but their participation would be very valuable in helping to develop knowledge and understanding of my research topic.

Your child's taking part will be safe and confidential

Your child's privacy and safety will be respected at all times. Your child's data will be kept anonymous, meaning they will not be able to be identified by the data collected, on any written material or in the write-up of the research. Your and your child's consent form will be stored securely and separately from the rest of the data and will be destroyed following completion of the research.

Your child does not have to complete all tasks asked of them and are free to stop their participation at any time.

To ensure the children's and my own safety, social distancing will be maintained at all times, I will wear a mask and sanitizing of hands and equipment will be completed regularly.

What will happen to the information that your child provides?

What I will do with the material your child provides will involve anonymously storing all data on a personal drive, only I have access to, which will be password protected. Data will be anonymised through your child being allocated a number which their data will be recorded against; there will be no way of identifying who has been assigned to each number. The anonymised data will be reviewed by myself and my supervisor and may be requested by examiners. Summaries of the data collected will be available in the write-up and may be published in an academic journal, the thesis will also be publicly accessible on UEL's institutional repository. Some broad demographic information may appear in the thesis and works based on it but that this will not be such as to permit the identification of individual participants. Once the research has been completed, the data will be kept for three years, following this, the data will be destroyed. Once the data has been collected children and their parents can withdraw the data up to the end of January 2022.

What if your child wants to withdraw?

Your child is free to withdraw from the research study at any time without explanation, disadvantage or consequence. Separately, you or your child may also request to withdraw their data even after they have participated data, provided that this request is made before the end of January 2021 (after which point the data analysis will begin, and withdrawal will not be possible).

Contact Details

If you would like further information about my research or have any questions or concerns, please do not hesitate to contact me. My email address is u1945528@uel.ac.uk.

If you have any questions or concerns about how the research has been conducted please contact the research supervisor Dr Matthew Jones Chesters. School of Psychology, University of East London, Water Lane, London E15 4LZ,
Email: m.h.jones-chesters@uel.ac.uk.

or

Chair of the School of Psychology Research Ethics Sub-committee: Dr Trishna Patel, School of Psychology, University of East London, Water Lane, London E15 4LZ. (Email: t.patel@uel.ac.uk)

APPENDIX H: Parent Consent Form

UNIVERSITY OF EAST LONDON

Consent for my child, or the child I care for, to participate in a research study:

The Use of a Newly Developed Computer Game to Test Executive Functioning in Young Typically Developing Children and Children on the Autism Spectrum

I confirm that I have read the information sheet for the above study and that I have been given a copy to keep. ☐

I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

I understand that my child's participation in the study is voluntary and that I may withdraw their participation at any time, without providing a reason for doing so. ☐

I understand that if I withdraw my child from the study, their data will not be used. ☐

I understand that I have until the end of January 2022 to withdraw my child's data from the study. ☐

I understand that my child's personal information and data from the research will be securely stored and remain strictly confidential. Only the research team will have access to this information, to which I give my permission. ☐

It has been explained to me what will happen to the data once the research has been completed. ☐

I understand that the thesis will be publicly accessible in the University of East London's Institutional Repository (ROAR). ☐

I would like to receive a summary of the research findings once the study has been completed and am willing to provide contact details for this to be sent to. ☐

I agree for my child to take part in the above study. ☐

Child's Name (BLOCK CAPITALS)

……………………………………………………………………………….

Parent/Carer's Name (BLOCK CAPITALS)

……………………………………………………………………………….

Parent/Carer's Signature

………………………………………………………………………………..

Researcher's Name (BLOCK CAPITALS)

………………………………………………………………………………..

Researcher's Signature

…………………………………………………………………………………

Date: ……………………..…….

APPENDIX I: Child Information Sheet

INVITE TO TAKE PART IN MY RESEARCH

The Use of a Newly Developed Computer Game to Test Executive Functioning in Young Typically Developing Children and Children on the Autistic Spectrum

You are being invited to take part in some research. Before you agree it is important that you understand what it would involve. Please take the time to read this with your parent or caregiver.

Who am I?

My name is Amber Scott, I am a Trainee Clinical Psychologist. I am doing some research as part of my studies at the University of East London.

What is the research?

Sometimes our brains work differently from one another and it can be helpful to measure how our brain is working. We need good measures to assess this so we can learn more about our brains and put things in place to make learning easier.

My research will assess whether a new computer game for children can measure our brain properly and find out what children think of it.

Why you?

I want to see whether the new computer game works properly for children like yourself. That is why I am looking for children just like you to take part.

You and the person who looks after you can decide whether you would like to take part. It is completely up to you!

What will you be asked to do?

You will meet with me and I will ask you to do a set of pen and paper exercises which involve me asking you to do things and then I will ask you to play the computer game, where you will follow the instructions on the screen. I will ask you a question about how much you enjoyed the pen and paper exercises and the computer game, how much you use computers and what you think about the new computer game. I will also ask your teacher some questions about you, which will help us to assess how good the new computer game is.

Your information?

Any information you tell me will be anonymised, which means rather than recording your name I will give you a number, so no one will know it is your information.

The information will be stored in an electronic cloud with a password only I will know. I will look at the information with my supervisor, who I work with. The information will then be put into writing for other psychologists to read.

Want if you change your mind?

If you decide you do not want to take part anymore, that is fine! You can tell me, or you can tell the person who looks after you and they can tell me. You can also change your mind after we have met if it is before January 2022. After January I will have already used your information.

Contact Details

If you have any questions you can ask the person who looks after you to email me. My email address is u1945528@uel.ac.uk.

They can also contact the research supervisor Dr Matthew Jones Chesters. School of Psychology, University of East London, Water Lane, London E15 4LZ, Email: m.h.jones-chesters@uel.ac.uk.

or

Chair of the School of Psychology Research Ethics Sub-committee: Dr Trishna Patel, School of Psychology, University of East London, Water Lane, London E15 4LZ.

(Email: t.patel@uel.ac.uk)

APPENDIX J: Child Consent Form

UNIVERSITY OF EAST LONDON

Consent to participate in a research study:

The Use of a Newly Developed Computer Game to Test Executive Functioning in Young Typically Developing Children and Children on the Autism Spectrum

I have read or been read the information sheet and have been given a copy to keep. ☐

I have been able to ask questions and have them answered. ☐

I know that I can change my mind at any time if I don't want to take part anymore without saying why. ☐

I know that if I no longer want to take part in the study, my answers will not be used. ☐

I know that I have until the end of January 2022 to change my mind. ☐

I know that my information and answers will be stored securely and will only be shared with the research team. ☐

I know what will happen with my information and answers once the research has finished. ☐

I know that other people will be able to read the final report through the researcher's university. ☐

I would like to receive a summary of the research once the study has finished and will ask my parent or caregiver to send contact details for this to be sent to. ☐

I agree to take part in the study. ☐

Your Name (BLOCK CAPITALS)

……………………………………………………………………………………….

Your Signature

……………………………………………………………………………………….

Researcher's Name (BLOCK CAPITALS)

……………………………………………………………………………………….

Researcher's Signature

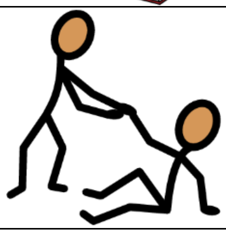
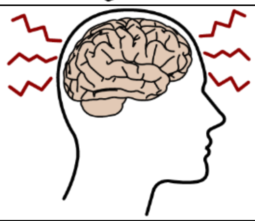
……………………………………………………………………………………….

Date: ……………………………

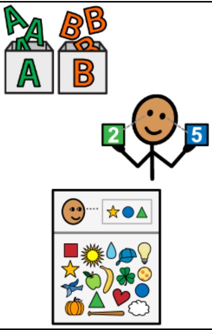
APPENDIX K: Easy-read Information Sheet

INVITE TO TAKE PART IN MY RESEARCH

The Use of a Newly Developed Computer Game to Test Executive Functioning in Young Typically Developing Children and Children on the Autistic Spectrum

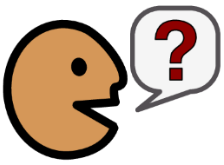
	Hello, this is Amber, she is a trainee psychologist
	Amber is doing some research
	She wants you to help
	The research is about how our brains work

She wants to see if a computer game can show how our brains work

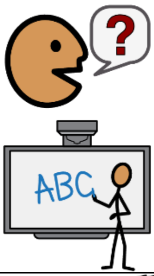


She will ask you to do some activities

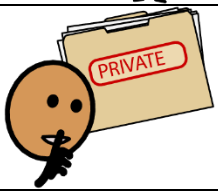
She will also ask you to play a computer game



She will then ask you some questions

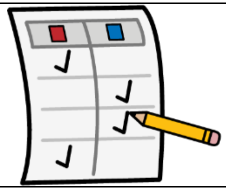
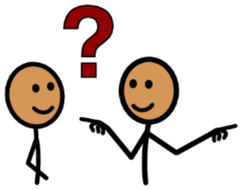


She will also ask your teacher some questions



Your information will be kept safe and no-one else will see it.

You can change your mind until Febuary 1st 2022

	Amber will write a report for others to read. Your name will not be in the report.
	Do you have any questions? Tell your teacher, parent or carer and they can contact Amber.

Contact Details

My email address is u1945528@uel.ac.uk.

They can also contact the research supervisor Dr Matthew Jones Chesters. School of Psychology, University of East London, Water Lane, London E15 4LZ, Email: m.h.jones-chesters@uel.ac.uk.

or

Chair of the School of Psychology Research Ethics Sub-committee: Dr Trishna Patel, School of Psychology, University of East London, Water Lane, London E15 4LZ.

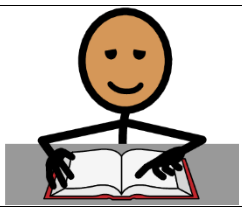
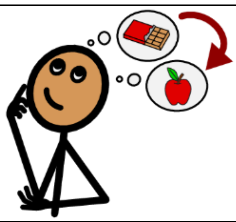
(Email: t.patel@uel.ac.uk)

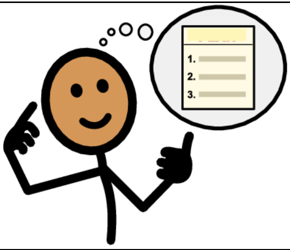
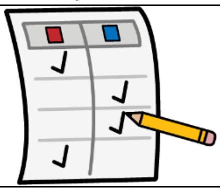
APPENDIX L: Easy-read Consent Form

UNIVERSITY OF EAST LONDON

Consent to participate in a research study:

The Use of a Newly Developed Computer Game to Test Executive Functioning in Young Typically Developing Children and Children on the Autism Spectrum

	I have read the information or someone has read it to me.	
	I have asked questions if I wanted to.	
	I know I can change my mind	

	If I change my mind, it is okay. They will not use my information.	
	I can change my mind until 1st February 2022.	
	I know what will happen with my information	
	Someone will write a report that other people can read.	
	I want to do this.	

Your Name (BLOCK CAPITALS)

……………………………………………………………………………………………

Your Signature

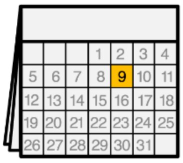
……………………………………………………………………………………………

Researcher's Name (BLOCK CAPITALS)

…………………………………………………………………………………………….

Researcher's Signature

……………………………………………………………………………………………



Date: ……………………………

APPENDIX M: Associations Between the Dragon Adventure Subgames and CTUQ Score

Table 30

Neurotypical Sample: Spearman's Rank Correlation Coefficients Between Dragon Adventure Subgames and CTUQ Score

	Dragon Dash				Dragon Sequence		Dragon Hunt	
	Forwards Errors	Reversed Errors	Total Errors	Distance	Total Correct	Longest Span	Errors	Time
CTUQ	-.18	-.18	-.21	.27	-.22	-.08	**-.32**	-.19

Note. Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

Table 31

ASC Sample: Spearman's Rank Correlation Coefficients Between Dragon Adventure and CTUQ Score

	Dragon Dash				Dragon Sequence		Dragon Hunt	
	Forwards Errors	Reversed Errors	Total Errors	Distance	Total Correct	Longest Span	Errors	Time
CTUQ	**-.42**	****-.64**	***-.59**	****.63**	**.45**	.28	.08	**-.42**

Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

APPENDIX N: Associations Between the Dragon Adventure Subgames and NEPSY-II Naming Time Score

Table 32

Neurotypical Sample: Spearman's Rank Correlation Coefficients Between Dragon Adventure Subgames and NEPSY-II Naming Time

	Dragon Dash				Dragon Sequence		Dragon Hunt	
	Forwards Errors	Reversed Errors	Total Errors	Distance	Total Correct	Longest Span	Errors	Time
NEPSY-II Naming Time	****.67**	**.34**	***.57**	****-.81**	.23	-.01	**-.41**	**.44**

Note. Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

Table 33

ASC Sample: Spearman's Rank Correlation Coefficients Between Dragon Adventure and NEPSY-II Naming Time

	Dragon Dash				Dragon Sequence		Dragon Hunt	
	Forwards Errors	Reversed Errors	Total Errors	Distance	Total Correct	Longest Span	Errors	Time
NEPSY-II Naming Time	**.31**	.26	**.32**	**-.42**	-.08	.03	.26	**.36**

Note. Moderate to large effect sizes are marked in **bold**
** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

APPENDIX O: Coding Units Explained

Table 34

Descriptions and Examples of Coding Units used in the Content Analysis

Coding Unit	Description	Examples
Game Controls	Responses that mentioned how the participants character is controlled in the game and how they would change this.	*"Different type of controls – like an xbox controller"* *"Sensitivity of the keys"* *"Easier controls"*
Additional Characters	Responses that mentioned having additional characters that you could interact with throughout the game or being able to see characters that were mentioned.	*"More dragons around the game"* *"Have a pet through a hatched egg"* *"Seeing Seizure the dragon"*
Length	Responses that mentioned changes that would impact the length of the game.	*"Longer game"* *"More quests"* *"Shorter"*
Additional Games	Responses that mentioned specific additional games they would like included in Dragon Adventure.	*"Boss fight"* *"Chase game"* *"Maths game"*
Gems	Responses that mentioned making changes to the gems that are collected after each subgame or the way they are used.	*"Rainbow gem – multicoloured gems that go into one"* *"Different coloured gems, treasure chest gems"* *"Puzzle pieces instead of gems"*
Character Functions	Responses that mentioned additional advanced functions for the main dragon that they control.	*"Breathes fire"* *"Make him fly"* *"Being able to attack bad guys"*
Environment	Responses that mentioned changed to the game environment or the programming of the game.	*"Doesn't glitch"* *"All flat texture"* *"Bridges lower and easier to get on"*
Obstacle Courses	Responses that mentioned specifically having more obstacle courses, which related to the Dragon Dash game.	*"Go through obstacles for all of it"* *"More obstacle courses"*
Ease	Responses that made suggestions of ways to make the game easier for participants.	*"Pictures on screen to show which controls to touch during dragon dash, instructions on which egg or crystal to collect during dragon hunt"* *"Have eggs and gems in the right order"*