EXECUTIVE FUNCTION AND SELF-INJURIOUS BEHAVIOUR IN AUTISTIC CHILDREN WITH A CO-OCCURRING INTELLECTUAL DISABILITY

By

CLAIRE LOUISE WRIGHT

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF CLINICAL PSYCHOLOGY

Centre for Applied Psychology

School of Psychology

The University of Birmingham

June 2022

Overview

This thesis is submitted in partial fulfilment for the degree of Doctor of Clinical Psychology at the University of Birmingham and contains three chapters. The first chapter is a meta-analytic review of the literature on the executive function abilities of autistic children aged six and under in comparison to neurotypical peers. The second chapter is a quantitative empirical study of executive function and self-injurious behaviours in autistic children with a co-occurring intellectual disability. The final chapters present a ‘press release’ for both the literature review and the empirical research, which outline the main findings of the papers in a manner suitable for public dissemination.

Dedication

For my parents, Clive and Christine

& Kat, Ben and Bump

Acknowledgements

Firstly, I am very grateful to all the children and families who gave so much of their time and effort in taking part in the research, and for caregivers' enthusiasm in acting as co-researchers for the study.

I would like to thank my research supervisors Caroline Richards and Rory Devine, for their invaluable expertise and kind patience. Thank you for trusting me with this project, and for all your support throughout this long process. And to the whole team working on the Sleep-Impulsivity-Behaviour Study, it has been a pleasure working alongside such dedicated, competent, and compassionate people.

To my wonderful family and friends, who have been an endless source of love and understanding. Thank you for taking care of me, sending brownies, and for putting up with me cancelling so many plans over recent months. I appreciate you all so much and look forward to spending more time with you now. A special thank you goes to my course-mate and library body double Suchi, I could not have done it without you.

Contents

Chapter 1 Meta-analysis : A meta-analysis of executive function in preschool-aged autistic children

1.1 Abstract

Background: Reduced executive function (EF) abilities are thought to underpin some of the behavioural characteristics of autism. However, inconsistent findings in autistic[1] preschoolers have raised questions about the developmental primacy of EF differences. This meta-analytic review synthesises the empirical evidence on EF performance in autistic preschoolers and neurotypical comparisons, to describe and evaluate putative differences.

Methods: Empirical papers reporting EF outcomes for young autistic children and matched comparisons were identified through searches of MEDLINE, EMBASE, PsycINFO and Web of Science databases. Papers were screened against inclusion criteria and evaluated using a bespoke quality rating tool. A pooled weighted effect size was estimated using a multilevel model to account for dependencies among effect sizes and moderation analyses explored the observed heterogeneity.

Results: The analysis of forty-three papers involving 2591 children (1110 autistic children, mean age = 62.16, range = 27 - 82 months, 1481 neurotypical children, mean age = 53.81, range = 20 - 82 months) revealed a significant moderate effect (g = -0.78) with autistic children scoring lower on EF measures than neurotypical comparisons. Similar effect sizes were observed irrespective of the domain measured, but differences were more pronounced

[1] The term autistic, rather than persons with autism, is used throughout to reflect the preferences for identity-first rather than person-first language, reported by a large sample of autistic individuals in the UK (Kenny et al., 2016).

when informant-based measures were used, among older preschoolers, and when children were matched on age without controlling for ability.

Interpretation: These findings mirror the broad EF differences reported in school-aged autistic children and adults, with true differences in EF apparent even in young autistic preschoolers. The lack of convergence of informant-based and direct measures is discussed, along with the clinical implications of the findings.

1.2 Background

Autism is a common neurodevelopmental condition with an estimated prevalence of one in every 100 children (Zeidan et al., 2022). There are a range of cognitive accounts which together seek to explain the differences in social interaction and communication, and the restricted and repetitive patterns of behaviour observed in autism (American Psychiatric Association [APA], 2013). Three influential hypotheses propose that the core characteristics of autism can be explained by differences in either theory of mind (i.e., the ability to attribute mental states to others; Baron-Cohen et al., 1985), perceptual processing (i.e., the ability to integrate information to construct meaning often termed central coherence theory; Frith, 1989; Frith & Happé, 1994; Van der Hallen et al., 2015) or executive function (i.e., the ability to control thought and action; Ozonoff et al., 1991; Russell, 1998). While prior work has revealed mean differences between autistic[2] and neurotypical individuals on a range of executive function (EF) measures (see previous meta-analyses: Demetriou et al., 2018; Lai et al., 2017), a similarly robust synthesised effect has not been reported in preschool-aged autistic children. This raises questions about the developmental primacy of EF differences in autism (Hill, 2004). If the autistic behavioural phenotype is indeed underpinned by reduced EF abilities, then differences in EF should be apparent in early childhood. Therefore, the present review extended prior meta-analyses by focussing exclusively on EF in autistic children aged six and younger, to evaluate the EF profile in this age group and the potential moderators of any effect.

[2] The term autistic, rather than persons with autism, is used throughout to reflect the preferences for identity-first rather than person-first language, reported by a large sample of autistic individuals in the UK (Kenny et al., 2016).

2.1 Executive function

Executive function (EF) is an umbrella term for a collection of top-down neurocognitive processes which support self-control and flexible, goal-directed behaviour (Zelazo & Carlson, 2012). EF is commonly fractionated into three theoretical core domains: inhibition, working memory and cognitive flexibility (Miyake et al., 2000). Inhibitory control is required to withhold a prepotent or habitual response, and to attend or respond selectively to relevant stimuli while ignoring distracting information (e.g., Munakata et al., 2011). Working memory refers to an ability to hold, update and manipulate information in mind as required (Baddeley, 1996). Cognitive flexibility or set-shifting, refers to an ability to switch attention or action in response to changing rules, goals, or environments (Miller & Cohen, 2001). Aside from these core EF skills, planning and verbal fluency are also commonly studied indices of EF ability. Like other EF processes, verbal fluency is mediated by the prefrontal cortex and requires the coordination of language, working memory and inhibition (Kuwabara et al., 2006). Planning skills involve the ability to think through and carry out a strategic set of actions with a particular end goal in mind (Zelazo et al., 1997). Tasks assessing planning, require participants to conceptualise a problem and discern the most effective sequence of steps to attain this. As such, planning can be seen as a more complex, compound EF skill which recruits several of the core EF skills (Diamond, 2013).

Whether EF domains are in fact dissociable has long been debated, and both unitary (e.g., Baddeley, 1996) and fractionated models of EF (e.g., Diamond, 2013) have been proposed. Evidence from lesion and neuroimaging studies in neurotypical participants have identified both common and preferentially-recruited functional neural networks according to task-type (Kim et al., 2017; Tsuchida & Fellows, 2013). Overall, current consensus points towards EF processes as being both correlated and separable (Best & Miller, 2010; Friedman & Miyake, 2017). Although research has traditionally focused on EF in adults, innovations in

test development over the past 25 years have given rise to a growing body of work focused on the emergence and correlates of EF in the preschool years (Carlson, 2005; Espy, 1997; Garon et al., 2008; Hughes 1998; Hughes & Ensor, 2005).

2.2 EF in autism

Although differing EF abilities have been linked to several neurodevelopmental conditions, autism is posited to have a unique EF profile (e.g., Ozonoff & Jensen, 1999). In an effort to challenge the dominant theory-of-mind account of autism, Russell (1996) first speculated that difficulties with understanding others' thoughts and feelings might stem from difficulties with EF. This gave rise to the so-called executive dysfunction account of autism. The putative associations between EF abilities and autism are well documented and scores on a range of EF measures have been linked to the core characteristics of autism (e.g., Hill & Bird, 2006; Kenworthy et al., 2009). For example, several studies have reported a link between performance on cognitive flexibility tasks, and the degree of repetitive behaviours in autistic individuals (e.g., Lopez et al., 2005; South et al., 2007) and there is growing evidence for the predictive significance of early EF skills on autistic children's emerging behaviours (Pellicano et al., 2010; 2013). In a small 12-year prospective study, EF was shown to predict variance in autism characteristics over and above age and other cognitive abilities (Kenny et al., 2019). Thus, EF is purported to explain both social-communicative characteristics and restricted, repetitive behaviours and interests among autistic children.

While a substantial body of research has demonstrated that autistic children perform worse than their neurotypical peers on tasks demanding greater executive control (e.g., Hughes et al., 1994; Kenworthy et al., 2005; Olde Dubbelink & Geurts, 2017; Ozonoff et al., 1991; Tonizzi et al., 2021), there have been inconsistent conclusions regarding domain-specific differences. A recent meta-analytic review of 235 studies of EF abilities in individuals aged between 6 and 64 years with and without autism, revealed a broad moderate

effect of EF, with autistic individuals scoring lower on EF measures on average when compared to neurotypical controls (Demetriou et al., 2018). Effect sizes were comparable across theoretical EF domains and stable into adulthood. However, studies with autistic participants under the age of six were not included due to the qualitatively different measures used to assess EF in preschool-aged children. Autism is commonly diagnosed in the preschool period (Matson et al., 2008), and can be reliably detected in children as young as two (e.g., Lord et al., 2006), therefore, evaluating the EF abilities of autistic preschoolers is critical to understanding the developmental course of these differences. If the autistic behavioural phenotype is underpinned by early emerging differences in EF, then differences should be apparent in early childhood.

2.3 Measuring EF in early childhood

Executive function emerges in early childhood and develops rapidly during the preschool years (Hughes, 1998). In neurotypical children, early working memory and inhibitory control skills develop expeditiously between the ages of two and six (Carlson et al., 2013; Diamond & Goldman-Rakic, 1989). More complex set-shifting abilities are thought to be underpinned by these early skills (Diamond, 2013). These changes are accompanied by a significant increase in the number of synapses in the prefrontal cortex during early childhood (Twardosz, 2012). However, whether EF emerges as dissociable domains, or a shared capacity is debated. While differentiated EF profiles have been reported in older children (e.g., St. Clair-Thompson & Gathercole, 2006), Wiebe and colleagues (2008) concluded that tasks designed to measure working memory and inhibition were likely measuring a single cognitive ability in neurotypical preschool children.

Emerging EF abilities have been associated with functional outcomes, such as social cognition (Devine & Hughes, 2014), school readiness and occupational achievement in later years (Pellicano et al., 2017). In neurotypical and neurodevelopmental populations, reduced

EF abilities in early childhood are predictive of poorer academic and adaptive outcomes (Blair et al., 2007; Fitzpatrick et al., 2014; Gardiner & Iarocci, 2017; Kenny et al., 2019; Viterbori et al., 2015). Considering the potentially disabling impact of EF difficulties, early assessment of these skills is critical to inform transdiagnostic proactive interventions and service planning. A range of developmentally sensitive direct and informant-based EF measures have been developed to assess EF in this younger population (for a summary of common tasks used to measure EF in early childhood, see Carlson 2005; Garon et al., 2008).

Studies investigating EF in preschool-aged autistic participants have reported inconsistent findings (Garon et al., 2018). Papers have reported moderate to large effect sizes favouring neurotypical groups on tasks designed to test inhibition (e.g., Tan, 2018), response and attention shifting (e.g., Faja & Dawson, 2014; McEvoy et al., 1993) and working memory (e.g., Dawson et al., 1998). However, some studies have found no group differences on similar tasks when matching groups on ability (e.g., Dawson, 2002; Lam & Yeung, 2012; Stahl & Pry, 2002; Yerys et al., 2007). Where some researchers report poorer performance of young autistic children compared to matched neurotypical peers on a broad range of EF tasks (e.g., Pellicano et al., 2017), others report increased difficulties on measures of inhibition and switching but comparable abilities in other domains such as working memory (e.g., Gardiner et al., 2017; Valeri et al., 2020). Several papers report no significant group differences on EF measures between autistic children under the age of five and mental-age-matched neurotypical children (e.g., Dawson et al., 2002; Griffith et al., 1999; Yerys et al., 2007), whereas others reported moderate to large effect sizes with similar-aged participants and comparison groups (e.g., Garon et al., 2018; Tan et al., 2018). This variability poses real issues for executive accounts of autism. Specifically, it is unclear whether (1) differences in EF are independent of group differences in general cognitive or verbal abilities, (2) EF

differences are limited to specific aspects of EF or encompass multiple domains of EF, (3) and whether EF differences are early emerging.

2.4 Sources of heterogeneity

Different methodological approaches, small sample sizes and heterogeneity within participants (i.e., age, verbal and non-verbal ability and autism characteristics), may in part account for the inconsistent findings reported for autistic preschoolers. Jarrold and Brock (2004) highlight the challenge of recruiting an adequate clinical sample, alongside issues matching groups in autism-related research. Larger effect sizes are commonly reported by studies which have recruited same-aged neurotypical comparisons to evaluate a hypothesised delay (e.g., Coldren et al., 2003). Smaller and null effects are more often reported by researchers who have controlled for between-group differences in general ability through age and ability-matched (e.g., Gardiner, 2017) or mental-age-matched neurotypical comparisons (e.g., Zacharov et al., 2021). This has led to conflicting conclusions about whether there are specific reductions in EF ability over and above more general differences in cognitive ability. Other design factors, such as variation in the EF tasks or measures chosen might also account for the mixed findings. Reported effect sizes also may vary according to study quality, for example, whether or not the neurotypical group were screened for autism characteristics. The present meta-analysis aimed to overcome some of these challenges to interpretation by exploring potential moderators through subgroup analyses and meta-regression techniques (Borenstein, 2021).

Informant-based and direct measures are often presumed to target a shared underlying EF construct. However, when used together these measures often only capture a small amount of overlapping variance (e.g., Camerota et al., 2016). Some possible reasons are that informant-based measures may capture behavioural characteristics in addition to EF, and are vulnerable to bias (Denckl, 2002). However, Friedman and Gustavson (2022) propose this

lack of convergence might indicate that direct and informant-based measures capture different but equally valid aspects of self-control. The authors identify several dimensions on which these differ. For example, direct measures assess EF in time-limited, controlled situations supported by goal rehearsal and feedback. Conversely, informant-based measures capture the outcome of goal-directed behaviours as they present in everyday activities, across contexts and over time. Therefore, both direct and informant-based measures may capture supplementary information about an individual's EF abilities across settings. A meta-analytic approach affords an opportunity to evaluate the contribution of measurement approaches to effect size heterogeneity (Ellis, 2010).

2.5 Summary and aims

In summary, recent meta-analytic data based on school-aged children, adolescents and adults revealed marked differences in EF between neurotypical and autistic participants, with reduced abilities observed across EF domains which remain relatively stable across development. In contrast, the data in autistic preschoolers is highly heterogeneous and it is therefore unclear whether these differences are developmentally consistent. Focusing on early childhood will provide insight about the developmental primacy of executive function differences in autism. Therefore, the overarching aim of the current study was to synthesise the empirical evidence on EF performance in autistic children aged six and under in comparison to neurotypical peers. In addition, the review sought to examine whether differences in EF were general (i.e., affecting multiple domains of EF) or specific (i.e., limited to a single domain of EF), whether EF differences were independent of group differences in general cognitive or verbal ability, and methodological sources of heterogeneity between studies.

1.3 Method

The meta-analytic review was pre-registered through Prospero (Available at: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020213276). Primary papers were identified and analysed in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021).

3.1 Identifying primary studies

A systematic search of MEDLINE, EMBASE, PsycINFO and Web of Science electronic databases from inception, was conducted in January 2021. The aim of the search was to obtain a comprehensive overview of the literature on executive function in young autistic children. The search strategy included terms and keywords generated with reference to the literature, as well as in consultation with experts in the field of neurodevelopment conditions and executive function in young children. An example search strategy for PsycINFO is outlined in Table 1. The search terms were adapted for use with other databases as appropriate. Additionally, a hand search of the reference lists of included articles was conducted prior to analyses.

Inclusion Criteria

Searches were restricted to papers available in English but were not limited to published articles to reduce effects driven by publication bias. Eligible studies measured an aspect of EF in children aged six years or younger with either a diagnosed or highly suspected autistic spectrum condition[3], as well as neurotypical peers. Both direct and indirect measures

[3]Papers reporting very young children with 'highly-suspected' autism were considered only if this was corroborated by scores on a well-established screening tool (e.g., The Childhood Autism Rating Scale–Second Edition [CARS-2], Schopler et al., 2010), or later clinically confirmed.

Table 1

Example PsycINFO search strategy

Construct	#	Search Terms
Early Childhood	1	exp Preschool Students/ OR Kindergarten Students/ OR exp Childhood Development
	2	child* OR preschool* OR pre-school* OR kindergarten* OR infant* OR toddler* OR "early childhood"
	3	1 OR 2
Autism	4	exp Autism Spectrum Disorders/
	5	autis* OR "autis* spectrum disorder" OR ASD OR Asperger* OR "pervasive development*" OR PDD OR "PDD NOS" OR PDDNOS
	6	4 OR 5
EF	7	exp Executive Function/ OR exp Attention/
	8	(Executive ADJ (function* OR dysfunction OR control)) OR ((mental OR cognitive) ADJ flexibility) OR ((cognitive OR self- OR impulse OR effortful OR inhibitory OR interference) ADJ control) OR ((response OR conflict) ADJ inhibition) OR (Set ADJ (shifting OR switching)) OR inhibit* OR planning OR fluency OR generativity OR "problem solving" OR "working memory" OR "self-regulation"
	9	7 OR 8
Measure	10	exp Cognitive Assessment/ OR exp Neuropsychological Assessment/
	11	((battery OR test* OR assess* OR measure* OR task* OR rating OR scale) ADJ6 (Executive function* OR EF OR Dysexecutive OR Executive Dysfunction*)) OR BRIEF-P OR BADS-C OR "card sort*" or WCST OR DCCS OR "Childrens Category Test" OR "Childrens Behavio?r Questionnaire" OR CBQ OR "developmental neuropsychological assessment" OR NEPSY OR "Working Memory Test Battery for Children" OR WMTB-C OR ((Gratification OR snack OR Gift) ADJ2 delay) OR ((Corsi OR Digit* OR Word) ADJ2 span) OR (Delayed ADJ (response task OR non-match)) OR (Spatial ADJ (reversal OR conflict)) OR (Object ADJ (reversal OR retrieval)) OR (Tower ADJ3 task) OR (Fluency ADJ4 task) OR Stroop OR A-not-B OR "go no-go" OR "Self-ordered pointing" OR "Invisible displacement" OR "Spin the pots" OR "Detour-reaching" OR "Pattern making test" OR Antisaccade OR "Dont paradigm" OR "Flexible Item Selection" OR "Prohibition Task" OR "Multilocation search" OR "Reverse categorization" OR "Hand game" OR "Lurias hand" OR "Whisper task" OR Knock-tap OR "Noisy book" OR Bear-Dragon OR grass-snow OR (day-night ADJ3 task) OR (black-white ADJ3 task) OR (count-label ADJ3 task) OR (Teddy-bear ADJ3 (shift* OR test OR task)) OR (Truck* ADJ2 task) OR ((Simon OR puppet) ADJ says)
	12	10 OR 11
Combined	13	3 AND 6 AND 9 AND 12

of EF, such as informant questionnaires, were included providing these yielded sufficient quantitative data. Papers evaluating commonly studied EF domains such as inhibition, working memory and cognitive flexibility were selected. The full inclusion and exclusion criteria are summarised in Table 2.

The results of the systematic search are presented in Figure 1. The search yielded a total of 3584 articles, which reduced to 2211 once duplicates were removed. Papers were then screened by title and abstract using the exclusion criteria. The four most common reasons for exclusion were: the study was not focussed on autistic individuals (i.e., other neurodevelopmental conditions, psychiatric conditions, or the experience of family members, N= 475), the study was not focused on children (N = 377), the study did not include a measure of EF (N = 271), and the paper did not report an empirical study (N = 245). A second reviewer independently screened a random sample of 10% (N = 221) of these papers, which yielded good inter-rater reliability (κ = 0.84). The remaining 527 full text articles were then reviewed against the inclusion criteria and 40 articles met all criteria. A second reviewer independently screened a random sample of 10% (N = 53) of these papers, which yielded good inter-rater reliability (κ = 0.81). In total, 43 studies satisfied the criteria for inclusion in the meta-analysis. One additional article was identified from the references of the included studies, and a further two articles were identified through electronic alerts received before the analysis was run.

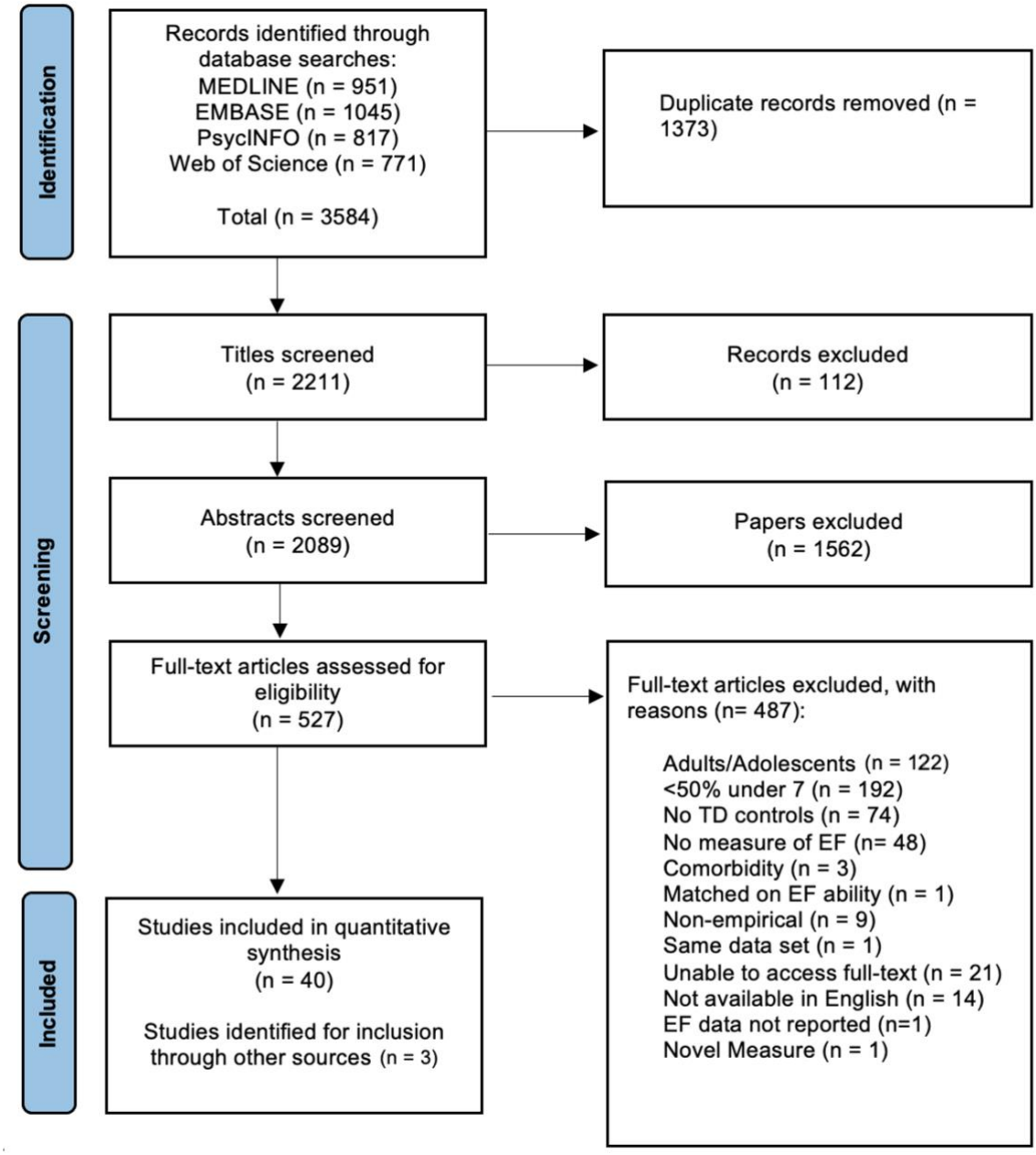
Table 2.

Inclusion and exclusion criteria

Inclusion	Exclusion
Population	
Participants with a mean age of six years or below and either a diagnosed, or highly suspected autistic spectrum condition (AUT) established through recognised diagnostic criteria or use of standardised screening or diagnostic measures	No autistic participants or autism of a known genetic aetiology, participant groups with a mean age over six years old
Comparison group	
A control sample of neurotypical (NT) children, matched by age or ability, without a diagnosis of an autistic spectrum condition or other neurodevelopmental conditions	Comparison groups of children with other neurodevelopmental conditions, developmental delay, or intellectual disability
Executive Function Measures	
At least one measure of EF evaluating one of the following theoretical EF domains: inhibition, shifting, working memory, and/or measures of complex EF e.g., planning and fluency.	No quantitative assessment of EF abilities
Outcome data	
Papers reporting either Means and Standard Deviations, or numbers passing and failing tasks for both AUT and NT groups, or t/F-Test statistics, Cohen's d/Hedges' g or an r effect size.	Papers not reporting sufficient quantitative data where authors have not responded to requests for this
Type of article	
Both published peer-reviewed, and unpublished literature were included	meta-analyses/theoretical papers/ reviews/commentaries/ clinical guidance/non-empirical papers/ validation of psychometric scales/qualitative papers
Design	
Cross-sectional studies, as well as intervention or longitudinal studies which reported baseline data will be included	Single case studies

Figure 1.

PRISMA Flow Diagram illustrating the outcome of search strategy and the application of the inclusion criteria against identified records.



3.2 Quality Criteria

The quality of identified papers was evaluated using a bespoke appraisal tool, which combined criteria from established tools designed to evaluate case-control and cross-sectional studies (Downes et al., 2016), along with criteria which targeted threats to validity and quality specific to the measurement of EF in children. The resulting seven criteria and operationalised scoring are summarised in Table 3. The strengths of each of the primary papers were summed to provide an overall summary score as an indicator of quality. The quality outcome for each paper is presented in Appendix L.

3.3 Data extraction

All data were extracted by the author. It was anticipated that EF performance would be reported as a mean or mean difference and standard deviations for both the autistic and neurotypical comparison groups. Where several outcome variables were reported for an EF task, for example, hit rate and commission errors, all attempts were made to extract the outcome most frequently reported across the papers. Where this was not possible, the mean number or proportion of correct responses was used. The extracted characteristics included sample size, participant demographics, performance on standardised ability measures, approach to confirming diagnostic status, method of matching, EF measure, EF domain, variable extracted and publication status. Experimental tasks and informant-based measures of EF were coded into EF domains in line with Garon, Bryson and Smith's (2008) review of EF in preschoolers. This resulted in five theoretical EF domains of inhibition, fluency, shifting, planning, and working memory.

Table 3

Quality criteria for defining autistic (AUT) and neurotypical (NT) groups, confounders, statistical power, data collection and attrition

	Quality Rating		
	0	1	2
Was an established measure used to assess autism?	Clinical diagnosis only	Clinical Diagnosis + established screening tool e.g., SCQ	Clinical Dx + diagnostic measure e.g., ADOS, ADI-R
Were NT participants screened for autism with an established tool?	No	Yes e.g., SCQ	N/A
Was a measure of verbal or non-verbal ability used to control for this potential confound?	not measured or considered	ability measured & considered e.g., used as covariate	AUT vs NT matched on a measure of ability
Was the study sufficiently powered to find moderate (0.5) effect?	Power <50%	Power >50%,	Power >80%,
Was a battery of EF tests used? 1 = >1 EF task, 0 = single EF task	singular task	>1 task	NA
Where an EF battery or EF questionnaire was used, was the reliability e.g., internal consistency reported?	No	Yes	NA
Is attrition explicitly reported?	Differences in N or df not accounted for	Yes OR number recruited = number analysed	NA

3.4 Data Analysis

Hedges' *g* was calculated as the standardised mean difference between autistic children and matched comparison groups on each measure of EF. Like Cohen's *d*, Hedges' *g* effect sizes are considered small (0.20), moderate (0.50), or large (0.80) and is the preferred unit of analysis when sample sizes are small (Hedges et al., 2009). Using Hedges' *g* also afforded a direct comparison to Demetriou and colleagues' (2018) meta-analysis of school-aged children and adults. Where higher scores on EF variables represented poorer performance, for example, where error scores are reported, the direction of the effect was reversed such that all negative effect sizes reflect a lower average performance of the autistic group in comparison with the neurotypical group.

If means and standard deviations were not reported then t or F statistics were transformed into estimates of Hedges' *g* if the sample sizes were reported. If neither summary statistics nor test statistics were reported, then effect sizes as reported by the primary studies were extracted[4]. Finally, if a paper reported categorical data, such as the number of participants who passed a task in each group, a reported non-parametric test statistic was converted into Hedges' *g*. In instances where it was not possible to convert a chi-squared statistic, for example, because the analysis included an additional and irrelevant comparison group, an appropriate statistical test was calculated using the frequency data before converting this into Hedges' g. Where data were given for multiple conditions of the same task, such as each post-switch condition of the DCCS, an effect size was calculated for each relevant condition and the average effect size was entered into the analysis.

[4] It should be noted that effect sizes as reported in primary studies are frequently calculated from data that has been adjusted for the association with one or more covariates. Such adjustments emphasise the idiosyncratic character of the reported effect and may result in dissimilarity with the effects reported within the other primary studies.

Selecting the appropriate meta-analytic model

As several eligible papers reported multiple measures designed to assess the same or different EF skills in the same group of participants, a three-level random-effects model was chosen to account for the potential interdependency of effect sizes (Assink & Wibbelink, 2016). Although relatively rare in the literature, the 'multilevel' method maximises statistical power in these circumstances, assigning a random effect for each effect size in addition to each study, reducing the likelihood of artificially reducing heterogeneity (Riley, 2009). This adds an additional parameter when compared to traditional random-effects models. This approach has also recently been used in a similar meta-analytic review of EF components in people with intellectual disability (Spaniol & Danielsson, 2022). Given that one of the aims of this analysis was to compare performance across different EF domains, the multilevel approach allowed for an exploration of both differences in outcomes between and within studies, while retaining data on all reported EF measures (Assink & Wibbelink, 2016).

For the selected papers, the estimated percentage of total variation attributed to sampling, within-study and between-study variation was 15.44%, 17.59%, and 66.96% respectively. The likelihood ratio tests indicated significantly better fit of the three-level model, when compared to models without an estimation of within or between studies variance ($AIC_{\text{full model}} = 232.52$, $AIC_{\text{without within-studies variance}} = 261.19$, $AIC_{\text{without between-studies variance}} = 307.78$, $p < .0001$), demonstrating the appropriateness of an additional parameter in explaining more variation than both a fixed and two-level model. Therefore, a three-level random-effects model of pooled standardised mean difference was calculated using the generic inverse variance method, with 95% confidence intervals (CI) using R Studio.

Next, analyses of sensitivity and publication bias were undertaken at the level of the paper, with only the largest effect size in papers reporting multiple measures. The impact of

disproportionately influential studies was assessed using a "leave-one-out" analysis, in which the random effects model was calculated with each of the primary studies removed in turn and the subsequent changes in heterogeneity and weighted average effect size were recorded. Highly discrepant and influential papers were omitted from subsequent analyses. Publication bias is caused by the tendency for statistically significant results to be published over non-significant results (Forstmeier et al., 2017). Small study bias is the tendency for studies with smaller sample sizes to show greater variability in their measurement of EF. These biases were explored with a funnel plot, which plots the magnitude of the study's EF effect against the square root of the study's sampling variances (standard error = $\sigma/\sqrt{N}$).

Sources of heterogeneity were explored using moderator analyses. Borenstein (2020) states that, although commonly misinterpreted, the index I^2 describes the proportion of the variance in the measured effects that can be attributed to variance in the true effect, rather than sampling error. Larger I^2 values (>75%) might indicate inconsistency across included studies (Higgins et al., 2003) and further moderator analyses are required to explain the within-study and between-study heterogeneity (Assink & Wibbelink, 2016). However, it should be noted that high I^2 values were expected in the present review as EF was not measured with a single task type. Potential moderators (EF task, matching strategy, study quality, age and ability) of the overall effect of EF were explored using subgroup analyses and meta-regression techniques.

1.4 Results

4.1 Descriptive statistics

Table 4 presents the characteristics of all included studies. Across the 43 papers reporting EF outcome data for 2591 young children (autism = 1110, neurotypical = 1481), 126 effect sizes were calculated. Sample sizes ranged from 7 to 72, and 7 to 216 for autistic and neurotypical samples respectively. The mean age of autistic participants was 62.16 months (SD = 15.53, range = 27 - 82 months). The mean age of neurotypical participants was 53.81 months (SD = 18.37, range = 20 - 82 months).

4.2 Omnibus test

Do preschool-aged autistic children score differently on measures of executive function than neurotypical comparison groups?

The three-level random effects model yielded a weighted mean effect size of g = -0.78, 95% CI [-.57, -.98]. This represented a moderate effect that was significantly greater than zero, z = -7.35, p <.001. Young autistic children scored lower on measures of executive function than neurotypical comparison groups. Forest plots[5] are presented in Figures 2 – 6. Observed effect sizes ranged from -4.32 to 0.44 and there was significant unexplained heterogeneity, Q_{125} = 738.95, p <.001, I^2 = 83.08%. The calculated I^2 exceeded Higgins' (2003) proposed threshold. Therefore, subsequent analyses explored potential sources of heterogeneity between the estimates of EF differences.

[5] Although the overall effect was calculated across all domains, forest plots are presented according to EF domains for clarity.

Table 4

Characteristics of included studies

Study (1st Author, Year)	N AUT	N NT	Age AUT (months)	AUT range	Matched	EF measure	Extracted Variable	*g*	SE
Arbelle 1994	28	28	43.00		MA & G	Prohibition task	#passed	-0.67	.27
Bonli 2005	18	16	68.00	61 - 81	CA	Delayed Response	#reversal trials passed	-2.08	.42
Bonli 2005	9	19	52.00	43 - 57	CA	Delayed Response	#reversal trials passed	-0.68	.40
Bonli 2005	18	16	68.00	61 - 81	CA	Luria's tapping test	#correct	-1.28	.37
Bonli 2005	9	19	52.00	43 - 57	CA	Luria's tapping test	#correct	-1.82	.46
Bonli 2005	18	16	68.00	61 - 81	CA	Scrambled Search	efficiency ratio	-1.42	.38
Bonli 2005	9	19	52.00	43 - 57	CA	Scrambled Search	efficiency ratio	-0.68	.40
Bonli 2005	18	16	68.00	61 - 81	CA	Day/Night Stroop	#correct	-0.87	.35
Bonli 2005	9	19	52.00	43 - 57	CA	Day/Night Stroop	#correct	-0.53	.40
Bonli 2005	18	16	68.00	61 - 81	CA	Spatial reversal	#correct	-1.20	.37
Bonli 2005	9	19	52.00	43 - 57	CA	Spatial reversal	#correct	-1.38	.43
Bonli 2005	18	16	68.00	61 - 81	CA	Stationary Search	efficiency ratio	-0.86	.35
Bonli 2005	9	19	52.00	43 - 57	CA	Stationary Search	efficiency ratio	-1.07	.42
Berg 2016	26	40	58.72	38 - 85	CA	BRIEF-P - Plan	total subscale	-1.79	.29

Berg 2016	26	40	58.72	37 - 85	CA	BRIEF-P - Shift	total subscale	-1.95	.30
Berg 2016	26	40	58.72	36 - 85	CA	Tower of Hanoi-R	total points	-0.88	.26
Campbell 2019	15	124	28.00	22 - 34	CA	Prohibition task	%fail	-0.48	.27
Carotenuto 2019	25	25	37.08	32 - 42	CA & G	BRIEF-P - Inhibition	total subscale	-1.44	.31
Carotenuto 2019	25	25	37.08	32 - 42	CA & G	BRIEF-P - Plan	total subscale	-0.24	.28
Carotenuto 2019	25	25	37.08	32 - 42	CA & G	BRIEF-P - Shift	total subscale	-1.18	.30
Carotenuto 2019	25	25	37.08	32 - 42	CA & G	BRIEF-P - WM	total subscale	-0.02	.28
Coldren 2003	7	7	67.32	52 - 93	VA	Spatial reversal	#post-shift errors	-1.24	.55
Coldren 2003	7	7	67.32	52 - 93	CA	Spatial reversal	#post-shift errors	-1.67	.59
Dawson 1998	20	20	64.60		RL	DNMS	#errors	-1.22	.34
Dawson 1998	20	20	64.60		RL	Delayed Response	%correct searches	-0.59	.32
Dawson 2002	72	39	43.50	34 - 52	MA	A not B	%correct on reversals	0.16	.20
Dawson 2002	72	39	43.50	34 - 52	MA	A not B + invisible	%correct on reversals	-0.14	.20
Dawson 2002	72	39	43.50	34 - 52	MA	DNMS - objects	%correct	-0.34	.20
Dawson 2002	72	39	43.50	34 - 52	MA	DNMS - pictures	%correct	-0.19	.20
Dawson 2002	72	39	43.50	34 - 52	MA	Object Reversal	%criteria met	0.20	.20
Dawson 2002	72	39	43.50	34 - 52	MA	Spatial reversal	%correct on reversals	-0.20	.20
Drayer 2008	29	30	68.00	48 - 83	CA	BRIEF-P - Inhibition	total subscale	-2.55	.35
Drayer 2008	29	30	68.00	48 - 83	CA	BRIEF-P - Plan	total subscale	-2.69	.36
Drayer 2008	29	30	68.00	48 - 83	CA	BRIEF-P - Shift	total subscale	-2.49	.34

Drayer 2008	29	30	68.00	48 - 83	CA	BRIEF-P - WM	total subscale	-2.38	.34
Drayer 2008	29	30	68.00	48 - 83	CA	DCCS	#post-switch correct	-1.47	.29
Drayer 2008	29	30	68.00	48 - 83	CA	Prohibition task	latency	-2.49	.34
Drayer 2008	29	30	68.00	48 - 83	CA	Noisy Book	total t-score	-1.67	.30
Drayer 2008	29	30	68.00	48 - 83	CA	Day/Night Stroop	#correct t score	-0.71	.27
Drayer 2008	29	30	68.00	48 - 83	CA	Tower of Hanoi-R	total t-score	-1.79	.31
Dunn 1996	10	10	81.50	52 - 108	MA	Verbal Fluency	%correct	-0.16	.43
Fabio 2020	20	20	42.93	38 - 45	MA	DCCS	#correct	-0.76	.32
Faja 2014	23	20	82.50	72 - 94	CA & MA	BKWD	raw score	-0.37	.30
Faja 2014	23	20	82.50	72 - 94	CA & MA	DCCS	#phases passed	-0.60	.31
Faja 2015	31	28	82.00	72 - 94	CA & IQ	DoG	latency	-0.89	.27
Gardiner 2017	24	19	66.88	42 - 88	CA & IQ	Boxes	error ratio	-0.05	.30
Gardiner 2017	24	19	66.88	42 - 88	CA & IQ	Go/No-Go	#commission errors	-0.21	.30
Gardiner 2017	24	19	66.88	42 - 88	CA & IQ	Monkey Tower	highest level achieved	-0.65	.31
Gardiner 2017	24	19	66.88	42 - 88	CA & IQ	Continuous Performance	#commission errors	-0.46	.31
Gardiner 2017	24	19	66.88	42 - 88	CA & IQ	Boy-Girl Stroop	#commission errors	-0.03	.30
Garon 2018	18	83	63.17	>52	MA	Flap Book	#correct	-0.92	.27
Garon 2018	16	133	45.38	<52	MA	Flap Book	#correct	-0.57	.27
Garon 2018	18	83	63.17	>52	MA	Hide and Seek	#errors	-0.41	.26
Garon 2018	16	133	45.38	<52	MA	Hide and Seek	#errors	-0.66	.27

Garon 2018	18	83	63.17	>52	MA	Tricky Box	#correct	-1.79	.29
Garon 2018	16	133	45.38	<52	MA	Tricky Box	#correct	-0.62	.27
Hanson 2014	25	25	70.36	38 - 99	MA	BKWD	#correct trials	-0.12	.28
Hanson 2014	25	25	70.36	38 - 99	MA	Count/Label	#correct	-0.19	.28
Hanson 2014	25	25	70.36	38 - 99	MA	DCCS	#post-shift correct	-0.28	.28
Hanson 2014	25	25	70.36	38 - 99	MA	Verbal Fluency	#correct	-0.10	.28
Hanson 2014	25	25	70.36	38 - 99	MA	Black/White Stroop	#correct	-0.86	.29
Hanson 2014	25	25	70.36	38 - 99	MA	Truck Loading	#correct trials	0.10	.28
Hanson 2014	25	25	70.36	38 - 99	MA	Tower of Hanoi	highest level achieved	-0.28	.28
Jahromi 2013	20	20	58.95		MA, G & EL	BRIEF: ISCI	total subscale	-1.06	.33
Jahromi 2013	20	20	58.95		MA, G & EL	Day/Night Stroop	proportion correct	-0.34	.31
Jahromi 2019	18	20	57.61	40 - 77	MA, G, RL & EL	DoG	latency	-0.63	.33
Jahromi 2019	18	20	57.61	40 - 77	MA, G, RL & EL	Luria's Hand Game	#correct conflict	-0.89	.33
Jahromi 2019	18	20	57.61	40 - 77	MA, G, RL & EL	Day/Night Stroop	proportion correct	-0.29	.32
Jones 2013	36	22	75.10	68–93	CA	DNMS	%correct	-1.08	.29
Kimhi 2014	29	29	59.45		CA, G, VMA & IQ	FIST	#post-shift errors	-0.62	.27

Kimhi 2014	29	29	59.45		CA, G, VMA & IQ	Tower of London	efficiency score	-0.53	.26
Lam 2012	12	12	73.37		CA, G, NVIQ, VIQ	WCST	categories completed	0.06	.39
McEvoy 1993	17	16	60.65	40 - 80	MA	A not B	#correct	-0.46	.34
McEvoy 1993	17	16	60.65	40 - 80	MA	A not B + delay	#correct	0.02	.34
McEvoy 1993	12	12	60.65	40 - 80	MA	Alternation task	#perseverations	-0.52	.40
McEvoy 1993	13	12	60.65	40 - 80	MA	Spatial reversal	#perseverations	-0.78	.40
Meir 2020	28	58	81.50	54 - 110	CA & IQ	BKWD	longest string	-0.63	.23
Memisevic 2021	32	32	65.30		CA	BRIEF-P - Inhibition	total subscale	-3.76	.41
Memisevic 2021	32	32	65.30		CA	BRIEF-P - Plan	total subscale	-3.45	.39
Memisevic 2021	32	32	65.30		CA	BRIEF-P - Shift	total subscale	-2.88	.35
Memisevic 2021	32	32	65.30		CA	BRIEF-P - WM	total subscale	-4.27	.45
Ostfeld-Etzion 2016	40	40	63.40	36 - 82	MA	DoG	%fail	-0.52	.23
Pastor-Cerezuela 2016	47	53	80.06	60 - 96	CA & IQ	Verbal Fluency	#correct	-0.83	.21
Pastor-Cerezuela 2020	40	40	81.20	60 - 96	CA & NVIQ	Verbal Fluency	#correct	-1.37	.25
Pastor-Cerezuela 2020	40	40	81.20	60 - 96	CA & IQ	Labyrinths	#errors	-0.77	.23
Pastor-Cerezuela 2020	40	40	81.20	60 - 96	CA & IQ	Nepsy Auditory Attention	#errors	-0.55	.23
Pastor-Cerezuela 2020	40	40	81.20	60 - 96	CA & NVIQ	Numbers Stroop	reaction time	-0.97	.23

Pellicano 2007	30	40	67.15	49 - 88	CA & IQ &G	Luria's Hand Game	#correct	-0.54	.24
Pellicano 2007	30	40	67.15	49 - 88	CA & IQ &G	Mazes	raw score	-0.21	.24
Pellicano 2007	30	40	67.15	49 - 88	CA & IQ &G	Teddy Bear Card Sort	#trials to criterion	-0.73	.25
Pellicano 2007	30	40	67.15	49 - 88	CA & IQ &G	Tower of London	#solved in min moves	-0.86	.25
Pellicano 2010	37	31	67.92	49 - 88	CA & IQ	Teddy Bear Card Sort	error proportion	-1.10	.26
Pellicano 2010	36	36	67.92	49 - 88	CA & IQ	Tower of London	#solved in min moves	-1.52	.27
Pellicano 2017	30	30	53.26	36 - 72	CA & IQ	Backwards Corsi Span	#correct trials	-1.68	.30
Pellicano 2017	29	29	53.26	36 - 72	CA & IQ	DCCS	#correct post-switch	-1.36	.29
Pellicano 2017	30	30	53.26	36 - 72	CA & IQ	Less is More	#optimal selections	-0.93	.27
Precenzano 2017	8	15	37.12	32.4 - 42.02	CA	BRIEF-P - Inhibition	total subscale	-1.90	.51
Precenzano 2017	8	15	37.12	32.4 - 42.02	CA	BRIEF-P - Plan	total subscale	-1.00	.45
Precenzano 2017	8	15	37.12	32.4 - 42.02	CA	BRIEF-P - Shift	total subscale	-1.60	.48
Precenzano 2017	8	15	37.12	32.4 - 42.02	CA	BRIEF-P - WM	total subscale	-1.40	.47
Rutherford 2003	28	26	33.93	26 - 41	MA	Spatial reversal	#correct – error	0.43	.27
Schindler 2018	17	25	71.09	48 - 84	CA & IQ	Shape School - complex	#correct	-0.29	.31

Schindler 2018	17	25	71.09	48 - 84	CA & IQ	Shape School - Inhibit	#correct	-0.32	.31
Schindler 2018	17	25	71.09	48 - 84	CA & IQ	Shape School - Switch	#correct	-0.64	.32
Smith 2019	29	30	27.39	19 - 37	CA	Card Sort	%hit rate	0.28	.26
Tan 2018	13	25	72.96		CA	DCCS	#post-shift correct	-0.57	.34
Tan 2018	13	25	72.96		CA	FIST	%correct	-1.36	.37
Tan 2018	13	25	72.96		CA	Day/Night + Happy/Sad Stroop	reaction time	-1.49	.38
Valeri 2020	27	27	61.63	49 - 71	CA & NVIQ	Card Sort	#post-shift correct	-0.67	.28
Valeri 2020	27	27	61.63	49 - 71	CA & NVIQ	Pattern Making Test	#correct	-0.37	.27
Valeri 2020	27	27	61.63	49 - 71	CA & NVIQ	Day/Night Stroop	#correct	-0.69	.28
Valeri 2020	27	27	61.63	49 - 71	CA & NVIQ	Spin the pots	efficiency score	-0.23	.27
Wang 2022	24	24	55.60	36 - 72	CA	Day/Night Stroop+	#correct	-1.48	.32
Yerys 2006	19	20	80.04	60 - 91	VMA	DCCS	#pass	0.03	.31
Yerys 2006	19	20	80.04	60 - 91	VMA	FIST	proportion correct	-1.65	.37
Yerys 2007	16	15	34.80	26 - 41	MA	A not B	#post-shift errors	-0.01	.35
Yerys 2007	16	15	34.80	26 - 41	MA	Spatial reversal	#correct	0.08	.35
Yerys 2007	17	18	34.80	26 - 41	CA	Spatial reversal	#correct	-0.42	.33
Yerys 2007	17	16	34.80	26 - 41	MA	Windows	#pass	0.03	.34

Yerys 2007	16	16	34.80	26 - 41	CA	Windows	#pass	-0.48	.35
Yi 2012	18	31	79.32	45 - 115	VMA	DCCS	#post-shift correct	-0.69	.30
Yi 2013	18	18	80.21	48 - 101	VMA	DCCS	#post-shift correct	-0.13	.33
Yi 2013	18	25	80.21	48 - 101	CA	DCCS	#post-shift correct	-1.34	.34
Zacharov 2021	9	9	57.44	40 - 68	NVMA	DCCS	%pass	0.00	.45
Zacharov 2021	9	9	57.44	40 - 69	NVMA	Reverse Categorisation	%pass	0.00	.45
Zhao 2019	51	47	63.00	36 - 89	NVIQ	DCCS	#correct	-0.56	.20

Note: AUT = autism; NT = neurotypical; CA = chronological age; EL = expressive language; G = gender; IQ = intelligence quotient; MA = mental age; NVIQ = non-verbal IQ; NVMA = non-verbal mental age; RL = receptive language; VA = verbal ability; VMA = verbal mental age; BKWD = backwards digit span; BRIEF = Behavioural Rating Inventory of Executive Function (Gioia et al., 2000); BRIEF-P = Behavioural Rating Inventory of Executive Function - Preschool version (Gioia et al., 2003); WM = Working Memory scale; ISCI = The Inhibitory Self-Control Index; DCCS = Dimensional Change Card Sort task; DNMS = delayed nonmatch-to-sample task; DoG = delay of gratification task; FIST = Flexible Item Selection Task; WCST = Wisconsin Card Sorting Test

Figure 2

Forest plot of difference between autistic and neurotypical preschool-aged children on measures of inhibition

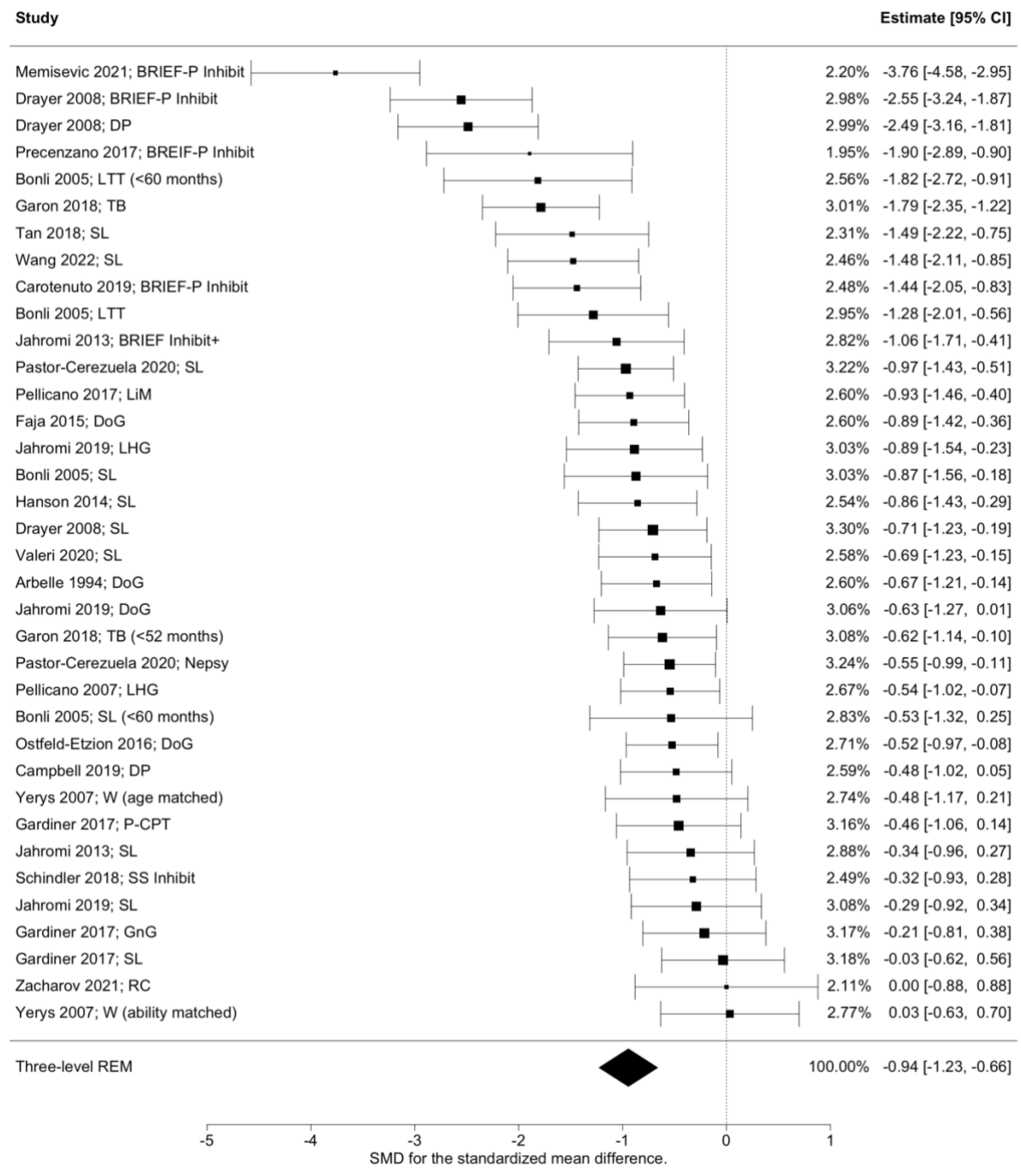
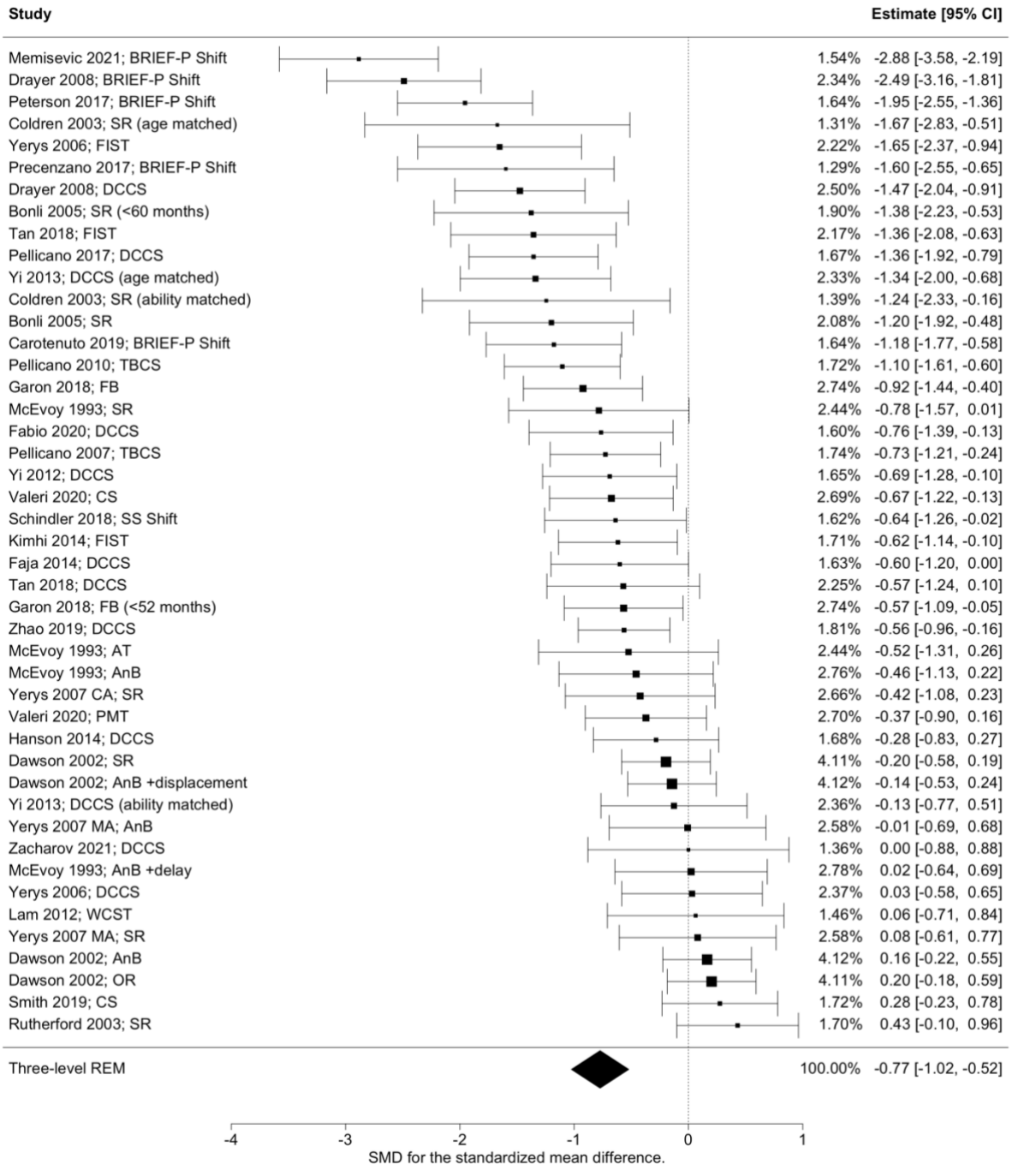


Figure 3

Forest plot of difference between autistic and neurotypical preschool-aged children on measures of cognitive flexibility



.

Figure 4

Forest plot of difference between autistic and neurotypical preschool-aged children on measures of working memory

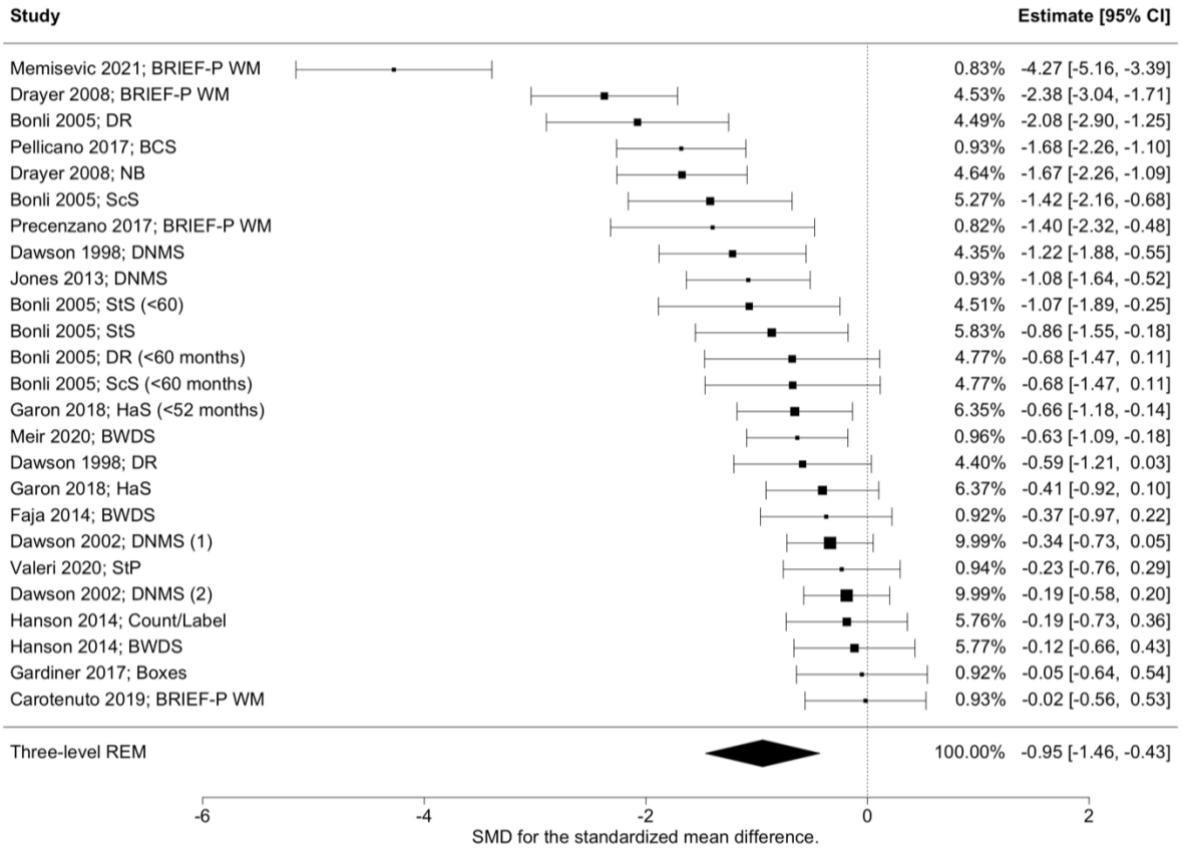


Figure 5

Forest plot of difference between autistic and neurotypical preschool-aged children on measures of verbal fluency

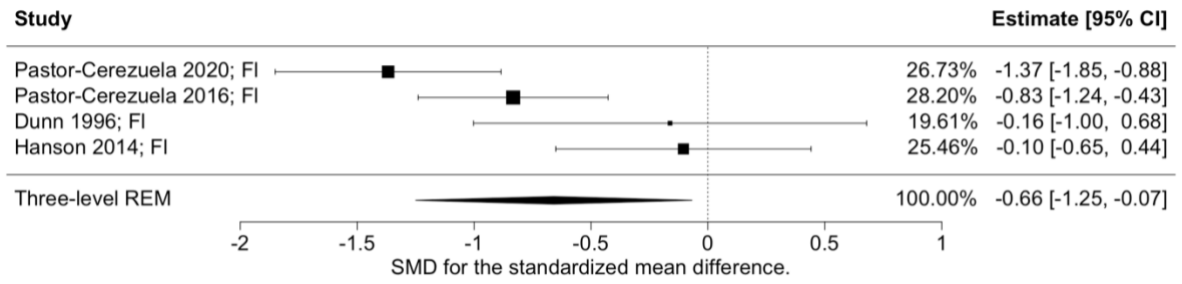
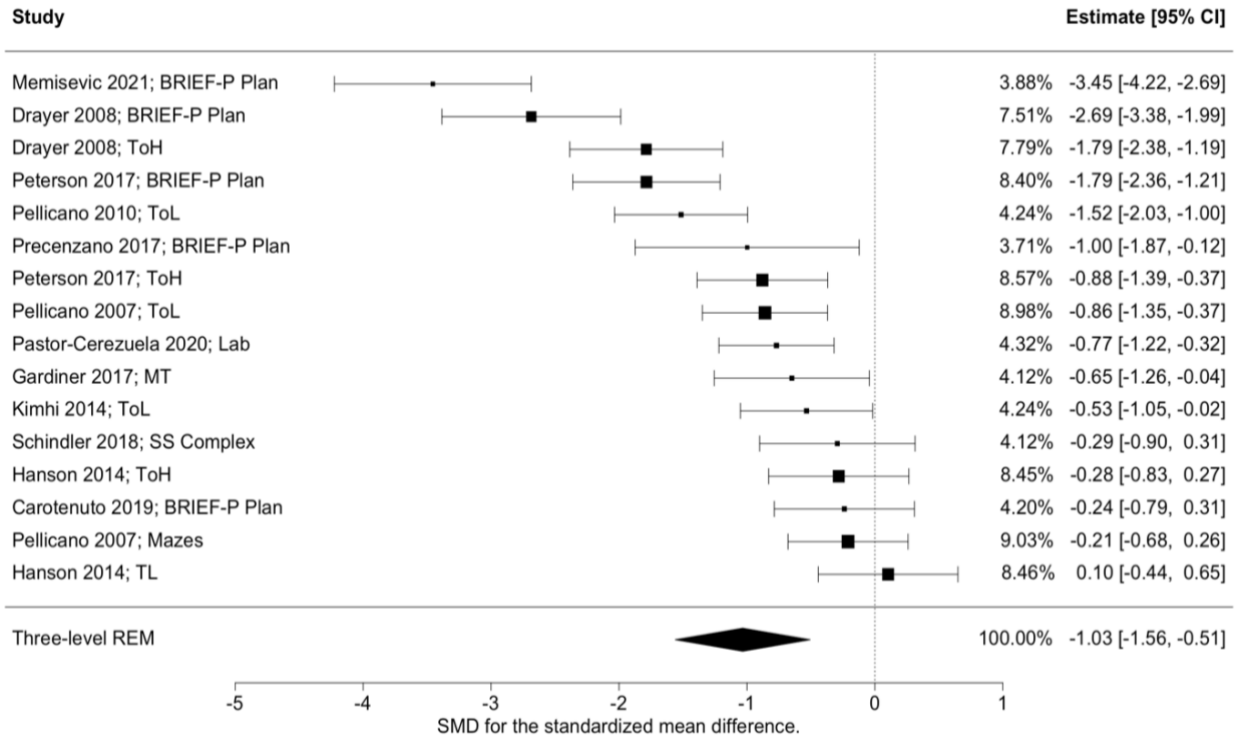


Figure 6

Forest plot of difference between autistic and neurotypical preschool-aged children on measures of complex EF

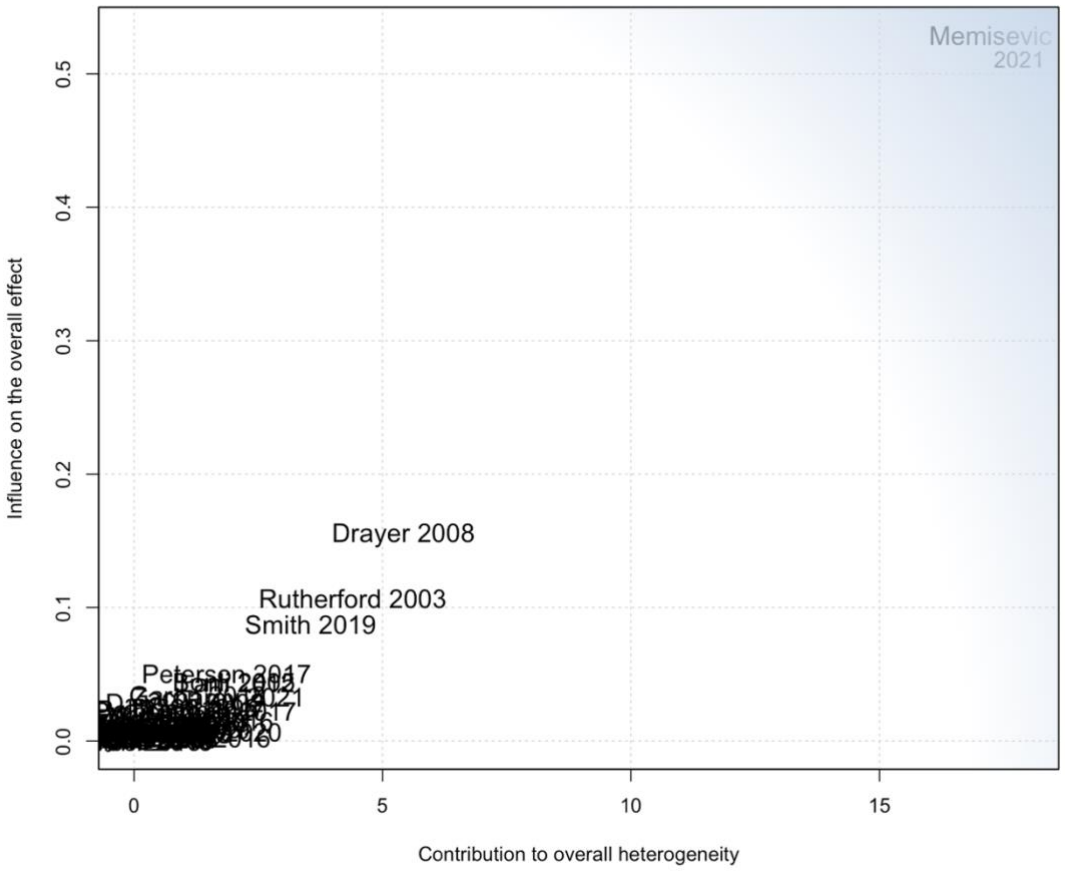


4.3 The impact of influential studies

The result of a "leave-one-out" analysis is presented in the Baujat plot (Baujat, et al., 2002) in Figure 7. On visual inspection of the forest and Baujat diagnostic plots, a clear outlier which may be disproportionately influential can be identified. Therefore, the weighted mean effect size was recalculated excluding this outlier. The effect of EF reduced by 8.97% but remained robust, g = -0.71, 95% CI [-.55, -.87], z = -8.89, p <.001, I^2 = 77.48%.

Figure 7

Baujat diagnostic plot of sources of heterogeneity. The vertical axis reports the influence of the study on the overall effect and the horizontal axis reports the discrepancy of the study with the rest of the literature. Studies with high values on both the x and y-axis (shaded) can be considered influential



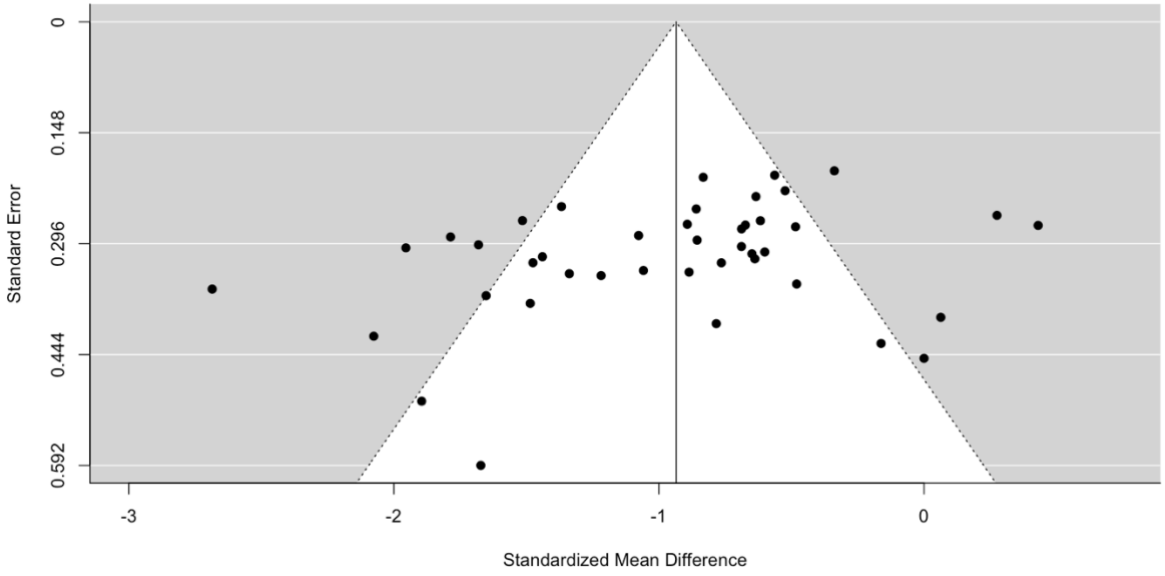
4.4 Publication bias

The funnel plot of Hedges' *g* is presented in Figure 8. Egger's (1997) regression test was significant (Egger's intercept = 0.35, $p = .04$) and there appeared to be a relative dearth of studies with small sample sizes reporting non-significant results, which may suggest that publication bias could be leading to an overestimation of the true effect. However, a trim-and-fill analysis did not result in the imputation of any studies. Rosenthal's (1984) "file drawer"

statistic revealed that 5,962 papers with null findings would be required to decrease the significance of the pooled effect to the just-significant level (p =.05). This exceeded the critical value (i.e., 5k+10) of 220 as recommended by Rosenthal (1984) and indicated that the meta-analytic results were robust to the threat of publication bias (Rosenberg, 2005).

Figure 8

Funnel plot of the Hedges' g. The 95% confidence interval of the expected distribution of Hedges' g is shown as an inverted "funnel"

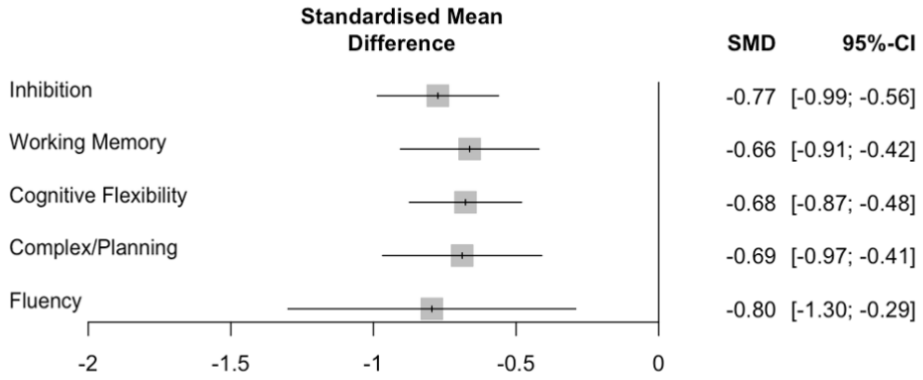


4.5 Are differences in EF general or specific?

EF domain did not significantly moderate the outcome, $Q_4 = 1.47$, $p = .83$ and the analysis revealed comparable, moderate effect estimates for each EF domain (see Figure 9), suggesting that EF differences between the autistic and neurotypical groups were general rather than limited to a specific theoretical EF domain.

Figure 9

Subgroup analysis by EF domain. N.B size of data points do not reflect weighted pooled effect sizes



Are effect sizes consistent across different tasks?

Where an equivalent direct measure was used to assess EF by five or more papers, a moderator analysis was conducted to evaluate whether tasks produced similar effect sizes. Due to the wide range of tasks used across the 43 papers, the only tasks which yielded sufficient data were the DCSS (k = 12), Spatial reversal (k = 9), A not B (k = 5), and Day/Night Stroop (k = 7). Task type did not significantly moderate the outcome, $Q_3 = 3.53$, $p = .32$ and the results of the analysis are presented in Table 5. Comparable effect sizes were obtained for DCCS, spatial reversal and Day/Night stroop-like task, suggesting that these provided similar estimates of EF in autistic preschoolers. No effect of EF was found on A not B tasks, and the performance of autistic preschoolers was comparable to neurotypical peers. Additionally, the type of inhibition task used (delay [k = 6] vs. conflict [k = 25]) did not significantly moderate the effect of EF, $Q_1 = .86$, $p = .35$, although differences in EF were more pronounced in delay paradigms ($g = -.92$) than conflict ($g = -.70$).

Table 5

The effect of EF by task

				95% CI	
Task	*g*	z	*p*	LL	UL
A-not-B	-.17	-0.71	.48	.30	-.64
Day/Night	-.55	-2.83	< .01	-.17	-.93
DCCS	-.67	-4.43	< .001	-.37	-.96
Spatial Reversal	-.55	-2.82	< .01	-.17	-.93

Note: CI = Confidence Interval; LL = lower limit; UL = upper limit; DCCS = Dimensional Change Card Sort task

Is the effect of EF moderated by the measurement approach?

As can be seen through visual inspection of the forest plot (Figures 2 – 6), indirect informant-based measures (k = 14) consistently produced large effect sizes when compared to direct measures (k = 108). This difference was explored in a subgroup analysis which revealed that measurement type significantly moderated the effect of EF, $Q_1 = 24.58$, $p <$ .001, with informant-based measures generating a much larger weighted mean and 95% confidence interval (Hedges' $g = 1.48$, 95% CI [-1.14, -1.81], $I^2 = 84.78\%$), which did not overlap with the overall effect recorded for direct measures (Hedges' $g = .62$, 95% CI [-.48, -.77], $I^2 = 70.27\%$).

4.6 Moderator analyses

Methodological factors

The results of moderation analyses evaluating other methodological factors are presented in Table 6. The primary studies all used a matched sample of autistic and neurotypical children. However, individual studies matched groups on either chronological age, ability, or a combination of both. The subgroup analysis of matching strategy revealed

that the smallest effects were recorded by studies which used mental-age-matching, whereas the largest differences between autistic and neurotypical groups were reported when children were matched on chronological age. The moderated model was a better fit for the data than a three-level model that did not contain a matching method as a moderator (Log-likelihood-ratio = 13.04; $p < .001$).

A meta-regression revealed that the total quality index score did not account for the variation in effect sizes, $Q_1 = 24.58$, $p = .30$. However, a further subgroup analysis was run for each of the quality criteria. When studies matched groups on estimated verbal or non-verbal ability, the effect was reduced yet remained significant. More pronounced differences in EF were observed on measures reported by papers which did not use an established diagnostic measure to validate the diagnostic status of autistic participants. No other index of study quality moderated the effect of EF. A further moderator analysis explored publication status of the included primary studies. The resulting omnibus test was significant, indicating that the overall effect was moderated by publication status. A significantly lower effect size estimate was observed for published papers compared to unpublished studies.

Table 6

The impact of study design factors on the weighted standardised mean EF difference between autistic and neurotypical preschoolers

Moderator	Subgroup	k	*g*	95% CI [LL , UL]	*p*	Q_M	*p*
Matching Strategy						18.54	< .001
	Ability	46	-0.45	[-.25 , -.67]	< .001		
	Age & Ability	33	-0.70	[-.47 , -.93]	< .001		
	Age	43	-1.12	[-.88 , -1.36]	< .001		
Diagnosis Risk						7.86	.005
	Low risk	72	-0.57	[-.40 , -.74]			
	Any risk	50	-1.00	[-.74 , -1.24]			
Ability Risk						4.00	.04
	Low risk	97	-0.64	[-.47 , -.80]			
	Any risk	25	-1.01	[-.68 , -1.33]			
Publication Status						17.45	< .001
	Unpublished	33	-1.25	[-.92 , -1.57]	< .001		
	Published	89	-0.60	[-0.45 , -0.75]	< .001		

Note: CI = confidence interval; LL = lower limit; UL = upper limit

Participant Characteristics

The results of moderation analyses evaluating participant variables are presented in Table 7. A stratified approach was used to separate the reported differences into two age categories: autistic children younger than 52 months and older than 52 months. This threshold was chosen to mirror the age split adopted in some of the included papers (Bonli, 2005; Garon et al., 2018) and resulted in a reasonably even number of effects for 'younger' and 'older' preschoolers. Papers which did not specify the included age range of participants, and those which included an age range which spanned these categories were not included in this moderator analysis. The resulting omnibus test was significant, indicating that the overall effect was moderated by age group, and the effect of EF was reduced for autistic participants younger than 52 but remained significant.

Studies differed in their measurement of ability, with some reporting mental-age equivalent estimates and others reporting standardised intelligence quotient (IQ) scores. Therefore, two separate meta-regressions were run to evaluate whether indices of ability moderated the effect of EF, and if there was a residual effect of EF once this was accounted for. Studies which did not measure or report outcome data for ability measures were excluded from these analyses. The mean estimated IQ score of autistic participants did not account for any of the heterogeneity in effect sizes ($k = 77$, $Q_1 = .08$, $p = .78$), and there were residual EF differences after controlling for ability. Conversely, the mean mental-age estimates of autistic participants did significantly moderate EF effects and an increase in mental age was associated with an increase in effect size. The moderator, standard error, and confidence interval for the mental-age meta-regression model is presented in Table 7.

Table 7

The impact of study design factors on the weighted standardised mean EF difference between autistic and neurotypical preschoolers

Moderator	Subgroup	k	*g*	95% CI [LL , UL]	*p*	Q_M	*p*
Age group						6.56	.01
	< 52	32	-0.47	[-.23 , -.72]	< .001		
	> 52	34	-0.83	[-.62 , -1.04]	< .001		

Meta-regression	k	coefficient	95% CI	Q_M	p
MA	34	-.01	[-.00 , -.02]	6.63	.01

Note: CI = confidence interval; LL = lower limit; UL = upper limit

1.5 Discussion

To the author's knowledge, this was the first meta-analysis to evaluate EF differences exclusively in preschool-aged autistic children, extending the findings from previous meta-analyses which have excluded experimental paradigms (Lai et al., 2017) and autistic participants under the age of six (Demetriou et al., 2018). The meta-analysis synthesised the findings of 43 papers reporting on the performance of preschool-aged autistic children and neurotypical comparisons on a range of direct tasks and questionnaire measures designed to assess theorised components of executive function.

5.1 EF in autistic preschoolers

The overall findings showed a consistent moderate effect, indicating comparatively reduced EF abilities in autistic preschoolers when compared with neurotypical peers. The effect was not domain-specific, but instead moderate effect sizes were observed irrespective of the domain of EF measured. This finding does not support theories of a profile of reduced and spared EF domains which are unique to autism (Hill, 2004), but instead mirror the broad EF differences reported between school-aged children and adults with and without autism (Demetriou et al., 2018). Rosenthal's (1984) "file drawer" statistic far exceeded the calculated critical value, indicating that the meta-analytic results were robust to the threat of publication bias. While certain study and child characteristics such as matching criteria were shown to moderate the effect of EF, all adjusted effect sizes remained significant demonstrating a residual EF difference even after controlling for these variables. Overall, the findings indicated that there were true EF differences between autistic preschoolers and neurotypical peers on EF measures, which could not be explained by differences in group composition or general cognitive or verbal ability. Taken together with Demetriou and colleagues' (2018) findings, the observed effect sizes suggest that EF differences may be present in early

childhood and persist into adulthood. Given the predictive value of EF for education and adaptive outcomes (e.g., Pellicano et al., 2017), intervening to support the development of early EF in autistic preschoolers should be a priority.

5.2 Comparison of direct measures of EF

There was a high degree of heterogeneity in the standardised mean differences of EF. This was anticipated, in part, given the wide range of measures used to assess EF in the available literature. While all papers reporting informant-based mean differences used a version of the Behavioural Rating Inventory of Executive Function (BRIEF: Gioia et al., 2000; BRIEF-P: Gioia et al., 2003), there were 49 different direct EF measures reported across the studies. Analyses evaluating the consistency between direct measures were limited by the relatively low number of papers using equivalent tasks. However, comparable effect sizes were found between DCCS (Zelazo, 2006), Spatial Reversal (Kaufmann et al., 1989) and a Stroop-like (Day/Night: Gerstad et al., 1994) tasks, which all provided similar estimates of EF differences in autistic preschoolers. Spatial reversal and DCCS tasks are both designed to measure set-shifting abilities, whereas Stroop-like tasks are thought to assess more complex inhibition skills (Garon et al., 2008). The strikingly similar effect sizes produced by these tasks could be accounted for by domain-general reductions in EF in young autistic children or could indicate that these tasks recruit the same specific cognitive processes, such as working memory and conflict inhibition. No significant effect was found across A not B tasks (Griffith et al., 1999), where the performance of autistic preschoolers was more comparable to neurotypical peers. It is possible that this task, which is deemed an appropriate measure of response-shifting abilities in children from around six months of age (Garon et al., 2008), was too easy for recruited autistic children whose ages ranged from 34 to 60 months. No significant differences were found between delay and conflict-type inhibition tasks, indicating that young autistic children had less capacity to both withhold a prepotent response

and to preferentially respond to target stimuli while ignoring distractors. This replicated the findings of prior meta-analytic work investigating inhibitory control in autism, which reported similar effect sizes between response inhibition and conflict tasks (Tonizzi et al., 2021). However, there were significantly fewer measures of response inhibition reported in comparison to conflict tasks, which limits the interpretability of this finding in the present analysis.

5.3 Study Characteristics

Moderator analyses revealed four aspects of study design which influenced the magnitude of the EF effect. First, the largest effect sizes came from studies using informant-based measures of EF (Hedges' g = 1.48, 95% CI [-1.14, -1.81]), a more moderate pooled effect size was found for direct measures (Hedges' g = .62, 95% CI [-.48, -.77]). These findings echo the effect sizes according to measure type reported by Demetriou and colleagues (2018) in older autistic children and adults. There are several potential explanations for this common finding. For example, caregivers might not always interpret questionnaire items in the same way as other informants and there is the potential for their estimates to be influenced by the emotional valence of the target behaviours (Denckla, 2002). This appears to be supported by research which has demonstrated low inter-rater reliability between parent and teacher ratings on the BRIEF-P (Gioia et al., 2003), with parents endorsing higher rates of difficulties in their children (Schneider et al., 2020). However, the different yet consistent effect sizes generated by direct and informant-based measures suggest that both may capture valid yet conceptually distinct information pertaining to self-control (Friedman & Gustavson, 2022). While direct measures might measure a child's potential EF capabilities in a scaffolded task (Toplak et al., 2013), it is reasonable to suppose that these abilities might be differentially taxed by real-world settings where responses in unfamiliar circumstances are not always guided by goal rehearsal and corrective feedback. This

interpretation could also account for some of the reported differences between teacher and parent ratings on EF scales (Schneider et al., 2020). Whereas teachers observe a child's behaviour in school, a setting likely structured by predictable routines and age-appropriate tasks, parents see their children across settings which may vary in predictability and, therefore, challenge EF to differing degrees. Informant-based measures also reflect behaviour over several days, whereas direct measures offer a snapshot of performance on a given day (Friedman & Gustavson, 2022). It follows then that current direct EF measures may be too far removed from the real-world settings and, therefore, have less utility in predicting everyday EF abilities. Kenny and colleagues (2022) proposed that this could be addressed through designing carefully controlled measures which are more representative of real-world EF problems and comparing these with commonly-used direct measures. Importantly, despite the possible advantages and limitations to each assessment type, the present analysis has demonstrated that both measures do capture significant EF differences between autistic preschoolers and neurotypical peers.

Second, the matching strategies used by papers significantly moderated the effect of EF. The differences between autistic and neurotypical children were more pronounced when groups were matched on chronological age, without also accounting for differences in general ability. However, even when papers matched children based on mental-age alone, the difference in EF was reduced yet remained significant. This is a key finding as performance on EF measures, particularly those which assess working memory, are often highly correlated with both verbal and nonverbal ability (Friedman et al., 2006; Karbach & Kray, 2007). A residual significant effect in cohorts matched on ability strongly suggests that the reduced EF abilities recorded in autistic preschoolers cannot solely be accounted for by more general delays or differences in cognitive ability. However, an important limitation to this finding was that papers matched based on a range of different indices of ability, including receptive

language, non-verbal and verbal ability, and each of these may be differentially associated with EF across different measures. For example, verbal ability will likely account for more variation in fluency tasks than non-verbal ability.

Third, more conservative effect sizes between young autistic and neurotypical children were observed in papers which used established diagnostic measures to validate clinical diagnoses. One interpretation of this result is that this indicator of experimental quality translates into more rigorously defined and homogenous participant groups, thereby increasing the internal validity of these studies. However, these papers also differed systematically in their EF measurement and group matching strategies, with only one paper using an informant-based measure, and the majority matching autistic and neurotypical groups on a measure of ability. Therefore, correlations with these other moderators might account for the reduced effect of EF observed across these papers. It is worth noting here that although the study criteria allowed for the inclusion of participants with 'highly-suspected' autism, only one of the included papers assessed high-risk toddlers, and only data for the group whose diagnosis was later confirmed were extracted.

Finally, it was initially surprising that a larger effect of EF was found for unpublished studies, rather than published studies. A common finding across empirical research fields, is that studies reporting significant results are more likely to be published than those with null findings (e.g., Song et al., 2000). However, on closer inspection this finding may again be accounted for by differences in other study characteristics between published and unpublished papers. Most of the unpublished papers reported mean differences between chronologically-age-matched participants, and half utilised an informant-based measure of EF. Therefore, measurement type and matching approach appear to account for a large portion of the observed heterogeneity attributable to study design.

5.4 Child characteristics

Interestingly, estimated IQ was not found to be a significant predictor of EF effect. However, this might be due to correlations with other moderating variables such as matching strategy. Many papers reporting standardised indices of ability also matched autistic and neurotypical groups on this basis, thereby controlling for the potential effect of general ability. Furthermore, the limited range in reported mean IQ scores could account for this result. Very few studies recruited autistic children with relatively lower intellectual abilities (IQ < 70), and only one paper reported a mean IQ estimate over one standard deviation above the normative average. Thus, the findings tell us little about EF in preschool-aged autistic children with a co-occurring intellectual disability (ID). Despite ID co-occurring in approximately 50% of the global autistic population (Loomes et al., 2017), autistic children with ID are often excluded from the EF literature to isolate EF from more general differences in cognitive ability (Hill, 2004). However, these children are arguably at an increased risk of the negative outcomes associated with poorer EF. For example, previous research has identified reduced EF as a putative risk marker for self-injurious behaviour in autistic children with ID (Richards et al., 2017). Therefore, future studies should seek to address the underrepresentation of these children in the literature on autism and EF.

Moderator analyses using the mean of participant-level variables are limited in their interpretability, therefore, the effect of age was explored through a stratified subgroup analysis. Interestingly, although the effect of EF remained significant for both younger and older preschoolers, the effect was larger for autistic children older than 52 months. This could suggest diverging developmental trajectories for autistic and neurotypical preschoolers. While consistent EF differences are apparent in autistic children aged four and younger, these comparative differences may be more amplified in older preschoolers due to the rapid development of EF that occurs during the early childhood in neurotypical children (e.g.,

Doebel & Zelazo, 2015). Further research is required with autistic children aged two to three to explore this interpretation, as only a small minority of papers included participants in this age group. A meta-regression conducted with the mean estimated mental-age equivalents where available, captured a similar effect whereby mental age and EF effect sizes varied commensurately. Age and mental-age were strongly related in the studies, with the younger preschoolers (< 52 months) accounting for the lowest mental-age estimates.

Overall, while it is difficult to assess the unique contribution of each moderating variable due to the observed interactions between these, the moderator analyses have accounted for some of the heterogeneity in the effect of EF. The findings suggest that the smallest effect sizes will be observed in studies investigating EF through direct measures, in ability-matched groups comprising very young participants. Conversely, studies gathering questionnaire EF data in age-matched groups of older preschoolers will likely generate large effect sizes.

5.5 Clinical implications

Irrespective of the theoretical implications, these results have important clinical implications. The results revealed that cohorts of young autistic children demonstrate poorer EF skills on average than their neurotypical peers. While there are likely to be individual differences in EF masked by meta-analytic techniques (Geurts et al., 2014), emerging EF and has been shown to predict several functional outcomes, including social cognition, school readiness and adaptive behaviour (Devine & Hughes, 2014; Kenny et al., 2019; Pellicano et al., 2017). Early childhood is, therefore, a critical period in which to identify children at-risk of the suboptimal outcomes associated with reduced EF abilities. Routinely measuring EF in preschool-aged autistic children would help clinicians develop proactive care plans. Research teams have begun to evaluate interventions which target EF, either through direct training (e.g., Kenworthy et al., 2014) or indirectly, for example, through physical activity (e.g., Tse et

al., 2019). Although the evidence for the efficacy of these interventions is limited, randomised controlled trials have demonstrated the feasibility and potential of EF interventions for autistic children (Faja et al., 2022), even over and above well-established social skills interventions (Kenworthy et al., 2014).

Furthermore, EF differences are not unique to autism. Other neurodevelopmental conditions such as attention deficit/hyperactivity disorder (ADHD) are also associated with reduced EF abilities (Barkley, 2012). Reduced EF is also increasingly recognised as a transdiagnostic marker of mental health difficulties (McTeague et al., 2016). Therefore, individual differences in EF within neurodevelopmental groups might usefully explain variation in different functional outcomes. When assessing EF in young children in clinical and research settings, using a combination of measurement techniques is likely to yield both distinct and supplementary information about the individual's EF capabilities. While direct measures have been demonstrated to have superior predictive validity regarding academic attainment (Soto et al., 2020), questionnaire measures, which can be completed by a range of informants, offer additional insight into how reduced EF might manifest differently in real-world contexts (Benallie et al., 2021). Therefore, both will provide valuable data for care and education planning.

5.6 Limitations

The results should be interpreted with the following limitations in mind. A key issue that arose was the vast range of direct tasks used. In total, 51 different measures of EF generated the 126 effect sizes synthesised in this analysis. The DCCS (Zelazo, 2006) was the most frequently used task, despite only accounting for twelve of the included effect sizes. While this highlights the advances in age-appropriate task development (Garon et al., 2008), it poses difficulties for meta-analytic approaches. One suggestion for future research efforts is to reach a consensus regarding the most appropriate tasks for autistic preschoolers. The

development of open-source, standardised EF batteries with appropriately scaled tasks might pave the way to more consistency in the assessment of EF (e.g., Kenny et al., 2022; Zanini et al., 2021). Additionally, the uptake of such tools could result in a large and diverse data set over time. Where equivalent EF tasks were used, papers varied in the performance variables they reported. For example, while some papers reported the mean total of correct responses, others reported commission errors or reaction times. Although all these variables are valid performance indices from which meaningful standardised mean differences can be calculated, this is a potential source of heterogeneity in effect sizes. Future research should evaluate the sensitivity of each of these variables in comparisons between autistic and neurotypical children, to inform investigators about which dependent variables to report as standard. A further caveat for the finding of a broad EF difference in the present analysis, is that only 15 of the 43 studies included tests of three or more EF domains, whereas others reported a single measure, or multiple measures assessing the same domain. Therefore, comparisons between theorised domains are largely between studies, which are known to inflate the standard errors (Moses, 2001).

The meta-analysis also afforded an opportunity to evaluate and compare the research methodologies of studies investigating EF in preschool-aged autistic children. Although studies which screened neurotypical participants for autistic characteristics did not report significantly different EF effects, this reduced the scores on quality criteria for most papers, reducing the internal and external validity of these studies. None of the studies achieved a power of 80% owing to relatively small samples of autistic children, which might account for some of the 'non-significant' effects reported. Where batteries of EF were used, many investigators did not report or appraise the internal consistency of these. For the field to reach a consensus regarding the most appropriate measures to use in this population of children, it is essential that the psychometric properties of unstandardised batteries are reported. A further

concern was that it was not always made explicit by investigators whether a whole sample or a proportion of the sample had been reported previously.

The lowest overall quality scores were recorded for papers reporting informant-based measures only. These studies were less likely to have validated the respective autistic and neurotypical groups using established screening tools, or to have included a measure of general ability to control for this potential confound. It was surprising that while most studies reported administering a diagnostic or screening tool to validate children's clinical diagnoses of an autism spectrum condition, very few provided the subscale or overall scores for these measures. Instead, many papers commented only on whether their autistic sample had scored above a recommended threshold. Consequently, due to insufficient data, the relationship between EF effect and the degree of reported autism characteristics could not be evaluated as a potential source of heterogeneity. Additionally, very few papers screened samples for ADHD, a neurodevelopmental condition which is also associated with reduced EF abilities and shown to co-occur at a high rate in autistic individuals (Hartman et al., 2016). However, prior meta-analyses have reported comparable EF effects for autistic children with and without ADHD (Lai et al., 2017). These are areas which could be improved upon in future exploratory research.

5.7 Summary and Conclusions

The results of the meta-analysis have demonstrated reduced EF abilities in autistic preschoolers when compared with neurotypical peers. Predictions of a differential profile of executive dysfunction based on subdomain performance were not supported and moderate effect sizes were observed across all EF domains. Subgroup and meta-regression analyses revealed more pronounced effect sizes for 'older' preschoolers, those matched on chronological age without accounting for more general cognitive ability, and for informant-based measures of EF. The pooled weighted effect sizes for both direct and informant-based

EF measures were consistent with those from previous reviews (Demetriou et al., 2018; Lai et al., 2017). While the limited data available for autistic children under the age of three makes it difficult to draw conclusions regarding the primacy of EF differences, the findings support the claim that EF plays a part in the developmental trajectory of autism. Prospective longitudinal and EF intervention studies are best placed to further elucidate the putative role of EF. Irrespective of whether differential EF abilities underpin autism phenomenology, the findings indicate that preschool-aged autistic children are at greater risk of the poorer functional outcomes associated with reduced EF abilities than neurotypical peers, which has important clinical implications.

1.6 References

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

Arbelle, S., Sigman, M. D., & Kasari, C. (1994). Compliance with parental prohibition in autistic children. *Journal of Autism and Developmental Disorders*, *24*(6), 693–702. https://doi.org/10.1007/bf02172280

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology*, *12*(3), 154–174. https://doi.org/10.20982/tqmp.12.3.p154

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology Section A*, *49*(1): 5-28.

Barkley, R. A., & Peters, H. (2012). The Earliest Reference to ADHD in the Medical Literature? Melchior Adam Weikard's Description in 1775 of "Attention Deficit" (Mangel der Aufmerksamkeit, Attentio Volubilis). *Journal of Attention Disorders*, *16*(8), 623–630. https://doi.org/10.1177/1087054711432309

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind" ? *Cognition*, *21*(1), 37–46. https://doi.org/10.1016/0010-0277(85)90022-8

Baujat, B., Mahé, C., Pignon, J. P., & Hill, C. (2002). A graphical method for exploring heterogeneity in meta-analyses: application to a meta-analysis of 65 trials. *Statistics in Medicine*, *21*(18), 2641–2652. https://doi.org/10.1002/sim.1221

Benallie, K. J., McClain, M. B., Bakner, K. E., Roanhorse, T., & Ha, J. (2021). Executive functioning in children with ASD + ADHD and ASD + ID: A systematic review. *Research in Autism Spectrum Disorders*, *86*, 101807. https://doi.org/10.1016/j.rasd.2021.101807

Berg, J. L. (2016). *Internalizing Symptoms: Relations to Executive Functions in Young Children with Autism Spectrum Disorder*. Unpublished Doctoral Dissertation, Seattle Pacific University. Retrieved from https://digitalcommons.spu.edu/cpy_etd/14

Best, J. R., & Miller, P. H. (2010). A Developmental Perspective on Executive Function. *Child Development, 81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Blair, C., & Razza, R. P. (2007). Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten. *Child Development, 78*(2), 647–663. https://doi.org/10.1111/j.1467-8624.2007.01019.x

Bonli, R. (2005). *An investigation of the early development of executive functions and theory of mind in autism and in typical development*. Unpublished Doctoral Dissertation, University of Saskatchewan, Saskatoon, SK.

Borenstein, M., Hedges, L., V., Higgins, J. P. T., & Rothstein, H. R. (2021). *Introduction to Meta-Analysis* (2nd ed.). Wiley.

Borenstein, M. (2020). Research Note: In a meta-analysis, the I index does not tell us how much the effect size varies across studies. Journal of Physiotherapy, 66(2), 135–139. https://doi.org/10.1016/j.jphys.2020.02.011

Camerota, M., Willoughby, M. T., Kuhn, L. J., & Blair, C. B. (2016). The Childhood Executive Functioning Inventory (CHEXI): Factor structure, measurement invariance, and correlates in US preschoolers. *Child Neuropsychology*, *24*(3), 322–337. https://doi.org/10.1080/09297049.2016.1247795

Campbell, S. B., Northrup, J. B., & Tavares, A. B. (2019). Resistance to temptation in toddlers at genetic risk for autism spectrum disorder. *Autism*, *23*(4), 1018–1027. https://doi.org/10.1177/1362361318797264

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., Zelazo, P. D., & Faja, S. (2013). Executive function. In P. D. Zelazo (Ed.), *The Oxford handbook of developmental psychology (Vol. 1): Body and mind* (pp. 706–743). Oxford University Press.

Carotenuto, M., Ruberto, M., Fontana, M. L., Catania, A., Misuraca, E., Precenzano, F., Lanzara, V., Messina, G., Roccella, M., Smirni, D. (2019). Executive functioning in

autism spectrum disorders: A case-control study in preschool children. *Current Paediatric Research*, *23*, 112–116.

Coldren, J. T., & Halloran, C. (2003). Spatial Reversal as a Measure of Executive Functioning in Children With Autism. *The Journal of Genetic Psychology*, *164*(1), 29–41. https://doi.org/10.1080/00221320309597501

Dawson, G., Meltzoff, A. N., Osterling, J., & Rinaldi, J. (2008). Neuropsychological Correlates of Early Symptoms of Autism. *Child Development*, *69*(5), 1276–1285. https://doi.org/10.1111/j.1467-8624.1998.tb06211.x

Dawson, G., Munson, J., Estes, A., Osterling, J., McPartland, J., Toth, K., Carver, L., & Abbott, R. (2002). Neurocognitive Function and Joint Attention Ability in Young Children with Autism Spectrum Disorder Versus Developmental Delay. *Child Development*, *73*(2), 345–358. https://doi.org/10.1111/1467-8624.00411

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

Denckla, M. B. (2002). The Behavior Rating Inventory of Executive Function: Commentary. *Child Neuropsychology*, *8*(4), 304–306. https://doi.org/10.1076/chin.8.4.304.13512

Devine, R. T., & Hughes, C. (2014). Relations Between False Belief Understanding and Executive Function in Early Childhood: A Meta-Analysis. *Child Development*, n/a. https://doi.org/10.1111/cdev.12237

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

Diamond, A., & Goldman-Rakic, P. (1989). Comparison of human infants and rhesus monkeys on Piaget's AB task: evidence for dependence on dorsolateral prefrontal cortex. *Experimental Brain Research*, *74*(1). https://doi.org/10.1007/bf00248277

Doebel, S., & Zelazo, P. D. (2015). A meta-analysis of the Dimensional Change Card Sort: Implications for developmental theories and the measurement of executive function in children. *Developmental Review*, *38*, 241–268. https://doi.org/10.1016/j.dr.2015.09.001

Downes, M. J., Brennan, M. L., Williams, H. C., & Dean, R. S. (2016). Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open*, *6*(12), e011458. https://doi.org/10.1136/bmjopen-2016-011458

Drayer, J. D. (2008). *Profiles of executive functioning in preschoolers with autism*. Doctoral dissertation, Northeastern University.

Dunn, M., Gomes, H., & Sebastian, M. J. (1996). Prototypicality of responses of autistic, language disordered, and normal children in a word fluency task. *Child Neuropsychology*, *2*(2), 99–108. https://doi.org/10.1080/09297049608401355

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Ellis, P. D. (2010). *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results* (1st ed.). Cambridge University Press.

Espy, K. A. (1997). The shape school: Assessing executive function in preschool children. *Developmental Neuropsychology*, *13*(4), 495–499. https://doi.org/10.1080/87565649709540690

The essential guide to effect sizes: statistical power, meta-analysis, and the interpretation of research results. (2011). *Choice Reviews Online*, *48*(10), 48–5742. https://doi.org/10.5860/choice.48-5742

Fabio, R. A., Esposito, S., Carrozza, C., Pino, G., & Caprì, T. (2020). Correlations between facial emotion recognition and cognitive flexibility in autism spectrum disorder. *Advances in Autism*, *6*(3), 195–204. https://doi.org/10.1108/aia-02-2019-0005

Faja, S., Clarkson, T., Gilbert, R., Vaidyanathan, A., Greco, G., Rueda, M. R., Combita, L. M., & Driscoll, K. (2022). A preliminary randomized, controlled trial of executive function training for children with autism spectrum disorder. *Autism*, *26*(2), 346–360. https://doi.org/10.1177/13623613211014990

Faja, S., & Dawson, G. (2014). Performance on the dimensional change card sort and backward digit span by young children with autism without intellectual disability. *Child Neuropsychology*, *20*(6), 692–699. https://doi.org/10.1080/09297049.2013.856395

Faja, S., & Dawson, G. (2015). Reduced delay of gratification and effortful control among young children with autism spectrum disorders. *Autism*, *19*(1), 91–101. https://doi.org/10.1177/1362361313512424

Firth, U., & Frith, U. (1989). *Autism: Explaining the Enigma (Cognitive Development)*. Blackwell Publishers.

Fitzpatrick, C., McKinnon, R. D., Blair, C. B., & Willoughby, M. T. (2014). Do preschool executive function skills explain the school readiness gap between advantaged and disadvantaged children? *Learning and Instruction*, *30*, 25–31. https://doi.org/10.1016/j.learninstruc.2013.11.003

Forstmeier, W., Wagenmakers, E., & Parker, T. H. (2016). Detecting and avoiding likely false-positive findings – a practical guide. *Biological Reviews*, *92*(4), 1941–1968. https://doi.org/10.1111/brv.12315

Friedman, N. P., & Gustavson, D. E. (2022). Do Rating and Task Measures of Control Abilities Assess the Same Thing? *Current Directions in Psychological Science*, *31*(3), 262–271. https://doi.org/10.1177/09637214221091824

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not All Executive Functions Are Related to Intelligence. *Psychological Science*, *17*(2), 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Frith, U., & Happé, F. (1994). Autism: beyond "theory of mind." *Cognition*, *50*(1–3), 115–132. https://doi.org/10.1016/0010-0277(94)90024-8

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. *The Clinical Neuropsychologist*, *31*(8), 1283–1305. https://doi.org/10.1080/13854046.2017.1290139

Gardiner, E., & Iarocci, G. (2017). Everyday executive function predicts adaptive and internalizing behavior among children with and without autism spectrum disorder. *Autism Research*, *11*(2), 284–295. https://doi.org/10.1002/aur.1877

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

Garon, N., Smith, I. M., & Bryson, S. E. (2018). Early executive dysfunction in ASD: Simple versus complex skills. *Autism Research*, *11*(2), 318–330. https://doi.org/10.1002/aur.1893

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: performance of children 3 ½ – 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., 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 GA, Espy KA, Isquith PK. (2003). *Behavior Rating Inventory of Executive Function, Preschool Version (BRIEF-P)*. Psychological Assessment Resources.

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

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

Hanson, L. K., & Atance, C. M. (2014). Brief Report: Episodic Foresight in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *44*(3), 674–684. https://doi.org/10.1007/s10803-013-1896-6

Hartman, C. A., Geurts, H. M., Franke, B., Buitelaar, J. K., & Rommelse, N. N. (2016). Changing ASD-ADHD symptom co-occurrence across the lifespan with adolescence as crucial time window: Illustrating the need to go beyond childhood. *Neuroscience & Biobehavioral Reviews*, *71*, 529–541. https://doi.org/10.1016/j.neubiorev.2016.09.003

Hedges, L., V., Higgins, J. P. T., Rothstein, H. R., & Borenstein, M. (2009). *Introduction to Meta-Analysis* (1st ed.). Wiley.

Higgins, J. P. T. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

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

Hill, E. L., & Bird, C. M. (2006). Executive processes in Asperger syndrome: Patterns of performance in a multiple case series. *Neuropsychologia*, *44*(14), 2822–2835. https://doi.org/10.1016/j.neuropsychologia.2006.06.007

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, *16*(2), 233–253. https://doi.org/10.1111/j.2044-835x.1998.tb00921.x

Hughes, C., & Ensor, R. (2005). Executive Function and Theory of Mind in 2 Year Olds: A Family Affair? *Developmental Neuropsychology*, *28*(2), 645–668. https://doi.org/10.1207/s15326942dn2802_5

Hughes, C., Russell, J., & Robbins, T. W. (1994). Evidence for executive dysfunction in autism. *Neuropsychologia*, *32*(4), 477–492. https://doi.org/10.1016/0028-3932(94)90092-2

Jahromi, L. B., Bryce, C. I., & Swanson, J. (2013). The importance of self-regulation for the school and peer engagement of children with high-functioning autism. *Research in Autism Spectrum Disorders*, *7*(2), 235–246. https://doi.org/10.1016/j.rasd.2012.08.012

Jahromi, L. B., Chen, Y., Dakopolos, A. J., & Chorneau, A. (2019). Delay of gratification in preschoolers with and without autism spectrum disorder: Individual differences and links to executive function, emotion regulation, and joint attention. *Autism*, *23*(7), 1720–1731. https://doi.org/10.1177/1362361319828678

Jarrold, C., & Brock, J. (2004). To Match or Not to Match? Methodological Issues in Autism-Related Research. Journal of Autism and Developmental Disorders, 34(1), 81–86. https://doi.org/10.1023/b:jadd.0000018078.82542.ab

Jones, E. J. H., Webb, S. J., Estes, A., & Dawson, G. (2013). Rule Learning in Autism: The Role of Reward Type and Social Context. *Developmental Neuropsychology*, *38*(1), 58–77. https://doi.org/10.1080/87565641.2012.727049

Karbach, J., & Kray, J. (2007). Developmental Changes In Switching Between Mental Task Sets: The Influence Of Verbal Labeling In Childhood. *Journal of Cognition and Development*, *8*(2), 205–236. https://doi.org/10.1080/15248370701202430

Kaufmann, P. M., Leckman, J. M., & Ort, S. I. (1989). Delayed response performance in males with Fragile-X. *Journal of Clinical and Experimental Neuropsychology*, *12*, 69

Kenny, L., Cribb, S. J., & Pellicano, E. (2019). Childhood Executive Function Predicts Later Autistic Features and Adaptive Behavior in Young Autistic People: a 12-Year Prospective Study. *Journal of Abnormal Child Psychology*, *47*(6), 1089–1099. https://doi.org/10.1007/s10802-018-0493-8

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. Autism, 20(4), 442–462. https://doi.org/10.1177/1362361315588200

Kenny, L., Remington, A., & Pellicano, E. (2022). Comparing the Executive Function Ability of Autistic and Non-autistic Adolescents with a Manualised Battery of Neuropsychological Tasks. *Journal of Autism and Developmental Disorders*, *52*(7), 3169–3181. https://doi.org/10.1007/s10803-021-05034-4

Kenworthy, L., Anthony, L. G., Naiman, D. Q., Cannon, L., Wills, M. C., Luong-Tran, C., Werner, M. A., Alexander, K. C., Strang, J., Bal, E., Sokoloff, J. L., & Wallace, G. L. (2014). Randomized controlled effectiveness trial of executive function intervention for children on the autism spectrum. *Journal of Child Psychology and Psychiatry*, *55*(4), 374–383. https://doi.org/10.1111/jcpp.12161

Kenworthy, L., Black, D. O., Harrison, B., della Rosa, A., & Wallace, G. L. (2009). Are Executive Control Functions Related to Autism Symptoms in High-Functioning Children? *Child Neuropsychology*, *15*(5), 425–440. https://doi.org/10.1080/09297040802646983

Kenworthy, L. E., Black, D. O., Wallace, G. L., Ahluvalia, T., Wagner, A. E., & Sirian, L. M. (2005). Disorganization: The Forgotten Executive Dysfunction in High-Functioning

Autism (HFA) Spectrum Disorders. *Developmental Neuropsychology*, *28*(3), 809–827. https://doi.org/10.1207/s15326942dn2803_4

Kim, N. Y., Wittenberg, E., & Nam, C. S. (2017). Behavioral and Neural Correlates of Executive Function: Interplay between Inhibition and Updating Processes. *Frontiers in Neuroscience*, *11*. https://doi.org/10.3389/fnins.2017.00378

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

Kuwabara, H., Kasai, K., Takizawa, R., Kawakubo, Y., Yamasue, H., Rogers, M. A., Ishijima, M., Watanabe, K., & Kato, N. (2006). Decreased prefrontal activation during letter fluency task in adults with pervasive developmental disorders: A near-infrared spectroscopy study. *Behavioural Brain Research*, *172*(2), 272–277. https://doi.org/10.1016/j.bbr.2006.05.020

Lai, C. L. E., Lau, Z., Lui, S. S. Y., Lok, E., Tam, V., Chan, Q., Cheng, K. M., Lam, S. M., & Cheung, E. F. C. (2017). Meta-analysis of neuropsychological measures of executive functioning in children and adolescents with high-functioning autism spectrum disorder. *Autism Research*, *10*(5), 911–939. https://doi.org/10.1002/aur.1723

Lam, Y. G., & Yeung, S. S. S. (2012). Cognitive deficits and symbolic play in preschoolers with autism. *Research in Autism Spectrum Disorders*, *6*(1), 560–564. https://doi.org/10.1016/j.rasd.2011.07.017

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, *56*(6), 466–474. https://doi.org/10.1016/j.jaac.2017.03.013

Lopez, B. R., Lincoln, A. J., Ozonoff, S., & Lai, Z. (2005). Examining the Relationship between Executive Functions and Restricted, Repetitive Symptoms of Autistic Disorder. *Journal of Autism and Developmental Disorders*, *35*(4), 445–460. https://doi.org/10.1007/s10803-005-5035-x

Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism From 2 to 9 Years of Age. *Archives of General Psychiatry*, *63*(6), 694. https://doi.org/10.1001/archpsyc.63.6.694

Matson, J. L., Wilkins, J., & González, M. (2008). Early identification and diagnosis in autism spectrum disorders in young children and infants: How early is too early? *Research in Autism Spectrum Disorders*, *2*(1), 75–84. https://doi.org/10.1016/j.rasd.2007.03.002

McCarty, N. (2021). Age at autism diagnosis, first intervention drops to under 3 years. *Spectrum*. https://doi.org/10.53053/rfet8671

McEvoy, R. E., Rogers, S. J., & Pennington, B. F. (1993). Executive Function and Social Communication Deficits in Young Autistic Children. *Journal of Child Psychology and Psychiatry*, *34*(4), 563–578. https://doi.org/10.1111/j.1469-7610.1993.tb01036.x

McTeague, L. M., Goodkind, M. S., & Etkin, A. (2016). Transdiagnostic impairment of cognitive control in mental illness. *Journal of Psychiatric Research*, *83*, 37–46. https://doi.org/10.1016/j.jpsychires.2016.08.001

Meir, N., & Novogrodsky, R. (2019). Syntactic abilities and verbal memory in monolingual and bilingual children with High Functioning Autism (HFA). *First Language*, *40*(4), 341–366. https://doi.org/10.1177/0142723719849981

Memisevic, H., & Pasalic, A. (2021). Executive Functions and Developmental Profiles in Preschool Children with Autism Spectrum Disorder [Unpublished manuscript]. Pedagogical Academy, University of Sarajevo. https://doi.org/10.21203/rs.3.rs-438077/v1

Miller, E. K., & Cohen, J. D. (2001). An Integrative Theory of Prefrontal Cortex Function. *Annual Review of Neuroscience*, *24*(1), 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

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

Moses, L. J. (2001). Executive Accounts of Theory-of-Mind Development. *Child Development*, *72*(3), 688–690. https://doi.org/10.1111/1467-8624.00306

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O'Reilly, R. C. (2011). A unified framework for inhibitory control. *Trends in Cognitive Sciences*, *15*(10), 453–459. https://doi.org/10.1016/j.tics.2011.07.011

Olde Dubbelink, L. M. E., & Geurts, H. M. (2017). Planning Skills in Autism Spectrum Disorder Across the Lifespan: A Meta-analysis and Meta-regression. *Journal of Autism and Developmental Disorders*, *47*(4), 1148–1165. https://doi.org/10.1007/s10803-016-3013-0

Ostfeld-Etzion, S., Feldman, R., Hirschler-Guttenberg, Y., Laor, N., & Golan, O. (2016). Self-regulated compliance in preschoolers with autism spectrum disorder: The role of temperament and parental disciplinary style. *Autism*, *20*(7), 868–878. https://doi.org/10.1177/1362361315615467

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

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive Function Deficits in High-Functioning Autistic Individuals: Relationship to Theory of Mind. *Journal of Child Psychology and Psychiatry*, *32*(7), 1081–1105. https://doi.org/10.1111/j.1469-7610.1991.tb00351.x

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., & Moher, D. (2021). Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. *Journal of Clinical Epidemiology*, *134*, 103–112. https://doi.org/10.1016/j.jclinepi.2021.02.003

Pastor-Cerezuela, G., Fernández-Andrés, M. I., Feo-Álvarez, M., & González-Sala, F. (2016). Semantic Verbal Fluency in Children with and without Autism Spectrum Disorder: Relationship with Chronological Age and IQ. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00921

Pastor-Cerezuela, G., Fernández-Andrés, M. I., Sanz-Cervera, P., & Marín-Suelves, D. (2020). The impact of sensory processing on executive and cognitive functions in

children with autism spectrum disorder in the school context. *Research in Developmental Disabilities*, *96*, 103540. https://doi.org/10.1016/j.ridd.2019.103540

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental Psychology*, *43*(4), 974–990. https://doi.org/10.1037/0012-1649.43.4.974

Pellicano, E. (2010a). 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

Pellicano, E. (2010b). The Development of Core Cognitive Skills in Autism: A 3-Year Prospective Study. *Child Development*, *81*(5), 1400–1416. https://doi.org/10.1111/j.1467-8624.2010.01481.x

Pellicano, E. (2013). Testing the Predictive Power of Cognitive Atypicalities in Autistic Children: Evidence from a 3-Year Follow-Up Study. *Autism Research*, *6*(4), 258–267. https://doi.org/10.1002/aur.1286

Pellicano, E., Kenny, L., Brede, J., Klaric, E., Lichwa, H., & McMillin, R. (2017a). Executive function predicts school readiness in autistic and typical preschool children. *Cognitive Development*, *43*, 1–13. https://doi.org/10.1016/j.cogdev.2017.02.003

Pellicano, E., Kenny, L., Brede, J., Klaric, E., Lichwa, H., & McMillin, R. (2017b). Executive function predicts school readiness in autistic and typical preschool children. *Cognitive Development*, *43*, 1–13. https://doi.org/10.1016/j.cogdev.2017.02.003

Precenzano, F., Ruberto, M., Parisi, L. (2017). Executive functioning in preschool children affected by autism spectrum disorder: a pilot study. *Acta Medica Mediterranea, 33*, 35–39.

Richards, C., Davies, L., & Oliver, C. (2017). Predictors of Self-Injurious Behavior and Self-Restraint in Autism Spectrum Disorder: Towards a Hypothesis of Impaired Behavioral Control. *Journal of Autism and Developmental Disorders*, *47*(3), 701–713. https://doi.org/10.1007/s10803-016-3000-5

Riley, R. D. (2009). Multivariate meta-analysis: the effect of ignoring within-study correlation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *172*(4), 789–811. https://doi.org/10.1111/j.1467-985x.2008.00593.x

Rosenberg, M. S. (2005). THE FILE-DRAWER PROBLEM REVISITED: A GENERAL WEIGHTED METHOD FOR CALCULATING FAIL-SAFE NUMBERS IN META-ANALYSIS. *Evolution*, *59*(2), 464. https://doi.org/10.1554/04-602

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Rosenthal, R. (1984). *Meta-Analytic Procedures for Social Research (Applied Social Research Methods)* (Revised ed.). SAGE Publications, Inc.

Russell, J. (1996). Agency: Its Role in Mental Development. Hove, UK: Erlbaum. https://doi.org/10.4324/9780203775691

Russell, J. (1998). *Autism As an Executive Disorder* (1st ed.). Oxford University Press.

Rutherford, M. D., & Rogers, S. J. (2003). Cognitive Underpinnings of Pretend Play in Autism. *Journal of Autism and Developmental Disorders*, *33*(3), 289–302. https://doi.org/10.1023/a:1024406601334

Schindler, R. D. (2018). Individual differences in executive functioning skills in preschoolers: a comparison of traditional and supplemental scoring methods on the shape school. Unpublished Doctoral Dissertation, Case Western Reserve University. Retrieved from https://www.proquest.com/dissertations-theses/individual-differences-executive-functioning/docview/2439997404/se-2?accountid=8630

Schneider, H., Ryan, M., & Mahone, E. M. (2020). Parent versus teacher ratings on the BRIEF-preschool version in children with and without ADHD. *Child Neuropsychology*, *26*(1), 113–128. https://doi.org/10.1080/09297049.2019.1617262

Schopler, E., Reichler, R. J., & Renner, B. R. (2010). *The childhood autism rating scale (CARS)*. WPS.

Sinzig, J., Vinzelberg, I., Evers, D., & Lehmkuhl, G. (2014). Executive function and attention profiles in preschool and elementary school children with autism spectrum disorders or ADHD. *International Journal of Developmental Disabilities*, *60*(3), 144–154. https://doi.org/10.1179/2047387714y.0000000040

Smith, H., Carter, A. S., Blaser, E., & Kaldy, Z. (2019). Successful attentional set-shifting in 2-year-olds with and without Autism Spectrum Disorder. *PLOS ONE*, *14*(3), e0213903. https://doi.org/10.1371/journal.pone.0213903

Song, Eastwood, Gilbody, Duley, & Sutton. (2000). Publication and related biases. *Health Technology Assessment*, *4*(10). https://doi.org/10.3310/hta4100

Soto, E. F., Kofler, M. J., Singh, L. J., Wells, E. L., Irwin, L. N., Groves, N. B., & Miller, C. E. (2020). Executive functioning rating scales: Ecologically valid or construct invalid? *Neuropsychology*, *34*(6), 605–619. https://doi.org/10.1037/neu0000681

South, M., Ozonoff, S., & Mcmahon, W. M. (2007). The relationship between executive functioning, central coherence, and repetitive behaviors in the high-functioning autism spectrum. *Autism*, *11*(5), 437–451. https://doi.org/10.1177/1362361307079606

Spaniol, M., & Danielsson, H. (2021). A meta-analysis of the executive function components inhibition, shifting, and attention in intellectual disabilities. *Journal of Intellectual Disability Research*, *66*(1–2), 9–31. https://doi.org/10.1111/jir.12878

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

Stahl, L., & Pry, R. (2002). Joint Attention and Set-Shifting in Young Children with Autism. *Autism*, *6*(4), 383–396. https://doi.org/10.1177/1362361302006004005

Tan, E., Wu, X., Nishida, T., Huang, D., Chen, Z., & Yi, L. (2018). Analogical Reasoning in Children With Autism Spectrum Disorder: Evidence From an Eye-Tracking Approach. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00847

Tonizzi, I., Giofrè, D., & Usai, M. C. (2021). Inhibitory Control in Autism Spectrum Disorders: Meta-analyses on Indirect and Direct Measures. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-021-05353-6

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*, *54*(2), 131–143. https://doi.org/10.1111/jcpp.12001

Tse, C. Y. A., Lee, H. P., Chan, K. S. K., Edgar, V. B., Wilkinson-Smith, A., & Lai, W. H. E. (2019). Examining the impact of physical activity on sleep quality and executive

functions in children with autism spectrum disorder: A randomized controlled trial. *Autism*, *23*(7), 1699–1710. https://doi.org/10.1177/1362361318823910

Tsuchida, A., & Fellows, L. K. (2013). Are core component processes of executive function dissociable within the frontal lobes? Evidence from humans with focal prefrontal damage. *Cortex*, *49*(7), 1790–1800. https://doi.org/10.1016/j.cortex.2012.10.014

Twardosz, S. (2012). Effects of Experience on the Brain: The Role of Neuroscience in Early Development and Education. *Early Education & Development*, *23*(1), 96–119. https://doi.org/10.1080/10409289.2011.613735

Valeri, G., Casula, L., Napoli, E., Stievano, P., Trimarco, B., Vicari, S., & Scalisi, T. G. (2020). Executive Functions and Symptom Severity in an Italian Sample of Intellectually Able Preschoolers with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *50*(9), 3207–3215. https://doi.org/10.1007/s10803-019-04102-0

Van der Hallen, R., Evers, K., Brewaeys, K., van den Noortgate, W., & Wagemans, J. (2015). Global processing takes time: A meta-analysis on local–global visual processing in ASD. *Psychological Bulletin*, *141*(3), 549–573. https://doi.org/10.1037/bul0000004

Viterbori, P., Usai, M. C., Traverso, L., & de Franchis, V. (2015). How preschool executive functioning predicts several aspects of math achievement in Grades 1 and 3: A longitudinal study. *Journal of Experimental Child Psychology*, *140*, 38–55. https://doi.org/10.1016/j.jecp.2015.06.014

Wang, L., Liang, X., Jiang, B., Wu, Q., & Jiang, L. (2022). What Ability Can Predict Mathematics Performance in Typically Developing Preschoolers and Those with Autism Spectrum Disorder? *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-022-05454-w

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

Yerys, B. E. (2006). Executive functioning in autism: examining the role of verbal labeling and social context. Unpublished Doctoral Dissertation, University of Denver. Retrieved from https://www.proquest.com/dissertations-theses/executive-functioning-autism-examining-role/docview/305325471/se-2?accountid=8630

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

Yi, L., Fan, Y., Zhao, J., Huang, D., Li, Y., & Zou, X. (2013). Atypical understanding of mental terms in Chinese-speaking children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *7*(11), 1411–1417. https://doi.org/10.1016/j.rasd.2013.08.009

Yi, L., Liu, Y., Li, Y., Fan, Y., Huang, D., & Gao, D. (2012). Visual Scanning Patterns during the Dimensional Change Card Sorting Task in Children with Autism Spectrum Disorder. *Autism Research and Treatment*, *2012*, 1–11. https://doi.org/10.1155/2012/123053

Zacharov, O., Huster, R. J., & Kaale, A. (2021). Investigating Cognitive Flexibility in Preschool Children With Autism Spectrum Disorder. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.737631

Zanini, G. A. V., Miranda, M. C., Cogo-Moreira, H., Nouri, A., Fernández, A. L., & Pompéia, S. (2021). An Adaptable, Open-Access Test Battery to Study the Fractionation of Executive-Functions in Diverse Populations. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.627219

Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M. S., Saxena, S., Yusuf, A., Shih, A., & Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. *Autism Research*, *15*(5), 778–790. https://doi.org/10.1002/aur.2696

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., & Carlson, S. M. (2012). Hot and Cool Executive Function in Childhood and Adolescence: Development and Plasticity. *Child Development Perspectives*, n/a. https://doi.org/10.1111/j.1750-8606.2012.00246.x

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early Development of Executive Function: A Problem-Solving Framework. *Review of General Psychology*, *1*(2), 198–226. https://doi.org/10.1037/1089-2680.1.2.198

Zhao, J., Chen, S., Tong, X., & Yi, L. (2019). Advantage in Character Recognition Among Chinese Preschool Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *49*(12), 4929–4940. https://doi.org/10.1007/s10803-019-04202-x

Chapter 2 Empirical Paper: Executive Function and Self-Injurious Behaviour in Autistic Children with a Co-occurring Intellectual Disability

2.1 Abstract

Background: Self-injurious behaviours (SIB) occur at a high rate in autistic[6] children with a co-occurring intellectual disability (ID). Delineating the individual characteristics which contribute to the observed differences in SIB emergence and outcomes will extend the operant model and inform proactive intervention strategies. The present study extended prior work, which has identified a role for reduced behavioural inhibition, through the direct measurement of executive function (EF) abilities in this population.

Method: A developmentally-appropriate EF battery was administered remotely to 60 autistic children with ID (mean age = 10.22, range 5-15, 70% SIB, 14 girls). Questionnaire data on SIB, EF, and adaptive behaviour were collected. Analyses evaluated the relationship between the EF outcome measures and SIB, as well as the unique contribution of EF to the presence and severity of SIB.

Results: SIB was moderately associated with a direct measure of updating and inhibition ($r = -.28$, $p < .05$), and caregiver-rated impulsivity ($r = .41$, $p = .001$) and overactivity ($r = .33$, $p < .05$). Impulsivity and overactivity were significant unique predictors of SIB ($p < .05$), over and above general cognitive ability. In an age-and-ability matched subset, children who displayed SIB were rated as significantly more impulsive on average than those who did not ($g = .78$).

[6] The term autistic, rather than persons with autism, is used throughout to reflect the preferences for identity-first rather than person-first language, reported by a large sample of autistic individuals in the UK (Kenny et al., 2016).

Conclusions: The results replicated previous findings of a robust relationship between caregiver-rated impulsivity and SIB. While no direct measure of EF significantly predicted SIB in this small pilot sample, the direction and magnitude of effects offer tentative support for a role of behavioural inhibition in the developmental course of SIB. Overall, the study supports the feasibility of remote EF assessment in this under-represented group.

2.2 Background

Self-injurious behaviours (SIB), such as head banging and biting, occur at high rates in autistic[7] children and constitute a significant clinical concern (Steenfeldt-Kristensen et al., 2020). The presence and severity of co-occurring intellectual disability (ID) increases the likelihood of SIB, and as many as one in two autistic children with ID will display SIB (Richards et al., 2012). The negative impact of SIB extends beyond physical injury, to poor psychosocial outcomes for individuals (Minshawi et al., 2014) and high levels of caregiver stress (Lecavalier, et al., 2006). While traditional behavioural models are used to understand the maintenance and escalation of self-injury, they are unable to account fully for the differences in SIB emergence and outcomes within and across neurodevelopmental conditions. Recent research efforts have aimed to extend the dominant operant model by delineating individual characteristics associated with SIB in at-risk populations (Oliver & Richards, 2015). The identification of robust behavioural risk factors could reveal causal mechanisms which explain the observed individual differences in prevalence, severity and persistence of SIB. In turn this could improve clinical outcomes through the development of proactive intervention strategies targeting causal mechanisms in at-risk children before SIB becomes established in a child's behavioural repertoire. One promising finding is an association between the presence of SIB and caregiver-reported impulsivity and overactivity in autistic samples (Laverty et al., 2020; Richards et al., 2012; Richman at al., 2013). Increased impulsivity is underpinned by reduced behavioural inhibition (Bari & Robbins, 2013), a core component of executive function (Diamond, 2013). Therefore, the present study

[7] The term autistic, rather than persons with autism, is used throughout to reflect the preferences for identity-first rather than person-first language, reported by a large sample of autistic individuals in the UK (Kenny et al., 2016).

aimed to investigate the putative relationship between both direct and informant-based measures of executive function (EF) abilities, and the presence and severity of SIB in autistic children with co-occurring ID, to evaluate the value of increased impulsivity as a causal mechanism for SIB.

Unlike forms of self-injury such as cutting or burning seen in older or neurotypical populations, SIB describes the repeated and often rhythmic self-directed behaviours observed in individuals with ID and other neurodevelopmental conditions (Rojahn et al., 2007). Common SIB topographies include "head punching, hitting or banging against hard objects, self-biting, hitting other body parts and skin picking/piercing" (Cooper et al., 2009, p. 201). While some definitions stipulate an outcome of actual physical harm (e.g., Matson & Turygin, 2012), SIB is defined here as any self-directed action with the potential to cause harm (Rojahn et al., 2007). Although prevalence estimates of SIB in autism vary, high rates ranging from 30 to 50 percent are well documented (Baghdadli et al., 2003; McClintock et al., 2003; Soke et al., 2016) and the presence of a co-occurring ID has been demonstrated to further increase the likelihood of SIB (Richards et al., 2012). Commonly emerging in early childhood, once manifested SIB can become chronic and has been shown to persist over ten years for a significant minority (Laverty et al., 2020). Within wider ID populations, SIB has been reported to persist into adulthood for most (Taylor et al., 2011).

The physical consequences of SIB can range from tissue damage to, in extreme cases, death (Nissen & Haveman, 1997). While the definition of SIB adopted here requires only the presence of one or more topographies of SIB, the impact of these behaviours in the absence of actual physical harm can be nonetheless pervasive and detrimental. The presence of severe SIB is associated with increased distress and has been shown to negatively impact on the quality of life of both individuals and their caregivers (Rattaz et al., 2015). Alongside other behaviours which challenge, SIB is associated with an increased use of psychotropic

medications (Matson & Neal, 2009) and restrictive practices, including restraint (Heyvaert et al., 2015). Children showing SIB are at higher risk of inpatient admissions (Mandell, 2008), placement breakdown (Emerson et al., 2001) and exclusion from local services (Knapp et al., 2005). As such, understanding the mechanisms underpinning SIB is of critical clinical importance.

SIB in neurodevelopmental and ID populations has predominantly been conceptualised as an operant, functional behaviour mediated by environmental or sensory contingencies (Emerson, 2012). This is reflected in the NICE guidelines (2015), which recommend interventions focussed on managing antecedents, increasing the non-contingent availability of reinforcers, and promoting the replacement of risky or unwanted behaviours with more adaptive, functionally-equivalent skills. While the efficacy of behavioural interventions is well supported (Matson & LoVullo, 2008), there are several limitations to a purely operant account of SIB. First, interventions are contingent upon detrimental behaviours having already emerged and developed a reinforcement history, placing children and families at greater risk of experiencing the associated negative consequences (Richards et al., 2017). Second, SIB can be highly treatment resistant, or re-emerge soon after the completion of clinical interventions (Kelley et al., 2018). Finally, operant models do not fully account for empirical data demonstrating individual differences in the emergence, prevalence, and severity of SIB within high-risk populations (Richards et al., 2012).

Several demographic and personal characteristics have been robustly associated with SIB (Oliver & Richards, 2015). Commonly reported risk makers for SIB are younger age (Esbensen et al., 2009), a greater degree of ID (Richards et al., 2012), and higher levels of autism characteristics (Baghdadli et al., 2003). These demographic risk markers may provide a context in which SIB is more likely to be selected and reinforced by the environment, as all arguably reduce a child's available behavioural repertoire. However, demographic risk

markers are limited in their ability to account for individual differences in SIB presentation in children with autism and a co-occurring ID (MacLean et al., 2010).

The presence of SIB has been associated with significantly higher levels of impulsivity and overactivity (Richards et al., 2012). These behavioural traits have been consistently identified as correlates and predictors of SIB in autism, such that children who are reported as more impulsive or overactive are significantly more likely to engage in SIB at a single time point, and to persist in SIB both three and ten years later (Laverty et al., 2020; Richards et al., 2017; Richman at al., 2013). Overactivity and impulsivity are considered downstream behaviours underpinned by cognitive EF difficulties. Reduced behavioural inhibition has been proposed to influence SIB through difficulties withholding prepotent self-injurious responses to evoking stimuli or antecedents, and/or through a reduced capacity to modify or disengage from self-injurious behaviours once reinforcing consequences are presented (Sonuga-Barke, 2002). To date, studies investigating the relationship between behavioural inhibition and SIB have assessed impulsivity using informant-based measures such as The Activity Questionnaire (TAQ; Burbidge et al., 2010). While these measures are advantageous in gathering data in large samples, they have several limitations. For example, questionnaires may be biased by caregiver recall (Toplak, 2013), and they do not afford the separate evaluation of different aspects of response inhibition (withholding a prepotent response versus discontinuing behaviour once reinforced).The use of direct measures would allow for a more detailed exploration of the EF profiles associated with SIB.

Despite ID co-occurring in approximately 50% of the global autistic population (Loomes et al., 2017), autistic individuals with a co-occurring ID have often been excluded from psychological and EF research. A recent meta-analysis estimated that in 301 primary papers published in autism-specific journals in 2016, 94% of the autistic participants did not have a co-occurring ID (Russell et al., 2019). This lack of representation of the full range of

autistic presentations in the EF literature, can be traced back to Hill's (2004) seminal review. The author highlighted a confound of ability in studies seeking to evaluate the contribution of EF difficulties to the autistic behavioural phenotype. Consequently, studies increasingly recruited only more able participants to control for general ability levels. In papers which have assessed EF in a more representative sample of autistic children, predominantly informant-based EF measures such as the Behavior Rating Inventory of Executive Function (BRIEF: Gioia et al., 2000) have been used (McClain et al., 2022; Tsermentseli, et al., 2018). A very limited number of studies have used direct methods of EF measurement in this population, and those which have, have acknowledged that the chosen tasks were too complex or were confounded by verbal demands (Panerai et al., 2014). Therefore, to explore the relative contribution of EF to the presence and severity of SIB in autism, there is a pressing need to identify appropriately robust direct assessments of EF that are tailored to the needs of autistic children with ID.

2.1 Aims

In summary, SIB is highly prevalent in autistic children with a co-occurring ID. SIB persists beyond childhood and is associated with negative outcomes for individuals and families. Delineating the behavioural characteristics which are reliably associated with SIB is an important step in identifying the risk factors driving SIB. The present study investigates the role of EF in SIB in this under-researched group with the following aims; (1) to evaluate the feasibility of remotely administering a novel battery of direct EF measures in a sample of autistic children with a co-occurring ID; (2) to examine the relationships between different measures of EF across the sample; (3) to examine the relationship between EF and SIB in the full sample, with the hypothesis that the presence and severity of SIB would be associated with greater EF difficulties; (4) to examine the independent effect of EF in an age and ability matched subsample of participants with and without current SIB, with a hypothesis that

children with current SIB would indicate reduced EF skills in comparison to children without current SIB.

2.3 Methodology

The present study was nested within a wider programme of research, the Sleep-Impulsivity-Behaviour study, which was pre-registered on the OSF where the full protocol can be found (https://osf.io/k5qpx, DOI: 10.17605/OSF.IO/K5QPX).

Restrictions due to the COVID-19 pandemic resulted in a significant delay to data collection. Consequently, the data gathered from the first 60 children are reported here as a pilot analysis. Additionally, social distancing precautions meant that in-person cognitive assessments were not possible during the time frame for data collection and the protocol was re-designed to be conducted entirely remotely. For clarity, only the procedure and measures relevant to the author's research questions are described here.

3.1 Recruitment

Autistic children with co-occurring ID were recruited as part of the wider research programme. Eligible children were aged between four and 16 years old, with a diagnosis of autism and ID without a known genetic syndrome. Children who lived outside of the UK, fell outside of the target age range, did not have a diagnosis of both autism and ID, and those with a diagnosed genetic syndrome, were not included in the study. Additionally, each child required a parent or caregiver (hereafter caregiver) able to provide informed consent, complete questionnaires, and that was willing to administer direct cognitive measures in a recorded online session.

Children were recruited through special education schools and in partnership with National Health Service (NHS) trusts. Additionally, caregivers were able to self-select for the study by responding to advertisements published on relevant social media pages and digital platforms, or to physical advertisements displayed in participating schools or clinical settings (see Appendix B for an example advertisement). Following an expression of interest,

caregivers were contacted directly by a member of the research team to confirm eligibility, provide further information, and answer any questions before written consent was sought.

3.2 Participants

A priori power calculations were conducted using G*Power 3.1 (Faul et al., 2009) to determine the sample sizes required to detect a medium effect with 80% power. Effect sizes for power analyses were drawn from published studies demonstrating differences in impulsivity between children with and without SIB (e.g., Licence et al., 2020). The first calculation determined that 62 children were required to observe a significant ($p < .05$) medium effect ($r = 0.3$) in one-tailed Spearman's correlation tests (research question 2 & 3). Next, a sample size of 55 was required to detect a medium effect (Cohen's $f^2 = .15$) at the .05 level of significance in regression analyses evaluating R^2 change with up to three predictors (research question 3). Finally, for the between-group test (research question 4), the calculation determined that 158 children, half of whom engaged in current self-injury, would be required to detect a medium effect (d = 0.4) at the .05 level of significance.

In total, 60 children with autism and co-occurring ID participated in the study. The demographic characteristics of the sample are presented in Table 1.

Table 1

Demographic characteristics of sample (N = 60).

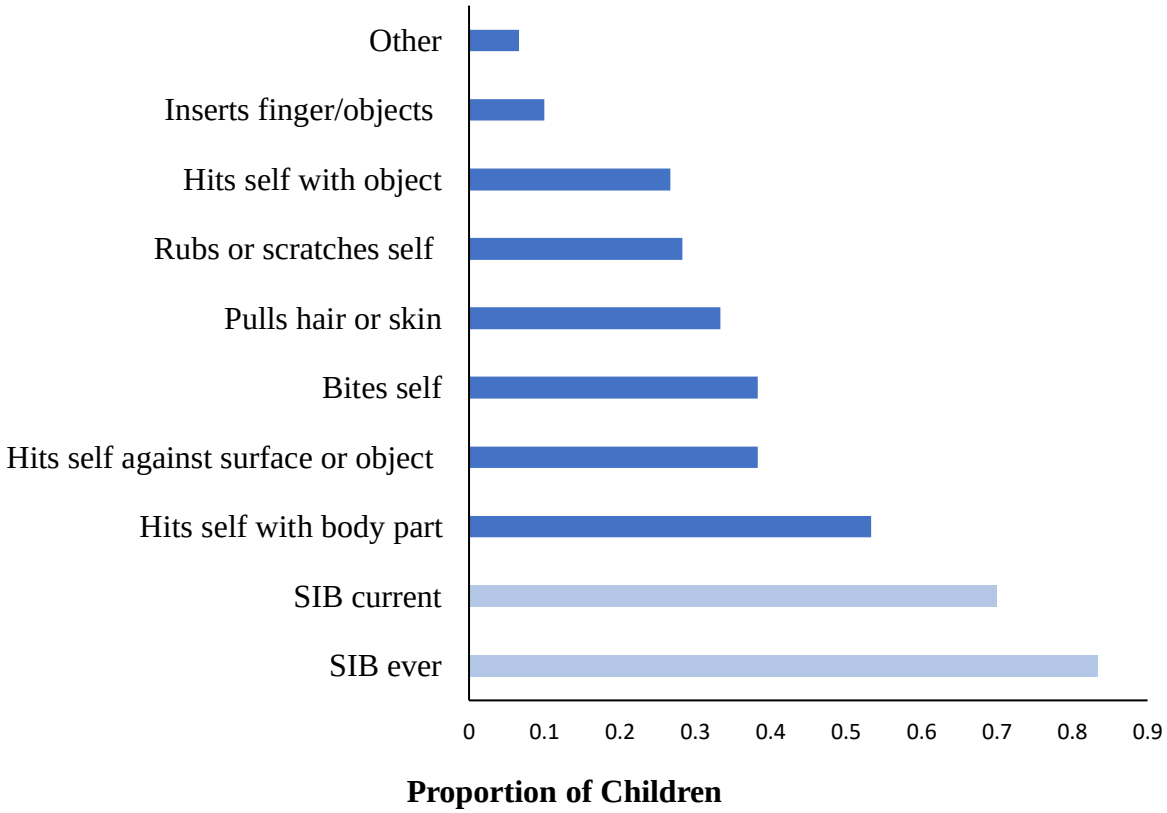
Developmental Variables	M (sd)	Range/%
Age (years)	10.22 (2.68)	5.71 – 15.40
Gender (F/M)	14/46	23% female
Monthly SIB / no current SIB	42/18	70% SIB
Total SIB score [a]	5.77 (4.62)	0 – 13
SCQ total verbal (N = 39)	26.89 (6.68)	11 – 37
SCQ total non-verbal (N = 21)	26.59 (3.66)	20 – 32
SCQ total, verbal items removed	24.30 (5.79)	7 – 33
VABS ABC score	42.45 (14.25)	20 – 80
VABS overall ae	2.79 (1.01)	1 – 7
VABS Receptive Language ae	1.69 (1.63)	0 – 4
BPVS raw score (N = 39)	64.44 (40.38)	2 – 142

Note. SCQ = The Social Communication Questionnaire; VABS = The Vineland Adaptive Behaviour Scales; ABC = Adaptive Behaviour Composite score; ae = age equivalent in years; BPVS = British Picture Vocabulary Scale.
[a] calculated from the Challenging Behaviour Questionnaire.

Of these 60 participants, 83% were reported to have ever displayed a form of self-injury, 70% were reported to have shown SIB within the last month, and 58% were reported to display two or more types of SIB. No current SIB was reported for 18 of the children. Figure 1 presents the prevalence data for the eight SIB topographies within the sample.

Figure 1

The prevalence of historical and current SIB and the relative frequency of SIB topographies in the sample (N = 60), as reported on the Challenging Behaviour Questionnaire (CBQ)



All caregivers provided evidence of their child's clinical diagnosis of autism and ID based on either the DSM (American Psychiatric Association [APA], 2000; 2013) or ICD-10 (World Health Organisation [WHO], 1992) classification systems. As part of the intake questionnaires, caregivers also completed an established autism screening tool, the Social Communication Questionnaire (Rutter et al., 2003). All but one child scored above the recommended clinical cut-off of 15. In lieu of a formal diagnosis of ID, a current Education Health and Care Plan, attendance at special education school, or a requirement of 1:1 support

from a teaching assistant, were recognised as highly suggestive of an ID. Therefore, autistic children meeting one or more of these criteria were included in the present study.

3.3 Informant Measures

The following questionnaire measures (see Appendix E for full copies) were completed by a caregiver:

Background Questionnaire. The demographic information of participants such as gender, age, and verbal ability, was collected by the Background Questionnaire.

The *Challenging Behaviour Questionnaire* (CBQ; Hyman et al., 2002) was included to capture information regarding SIB. Caregivers were asked to indicate whether their child had shown any SIB over their lifetime *and* in the last month. The authors report good test-retest reliability, 0.91 and inter-rater reliability 0.71 for self-injury questions collected by interview (Oliver et al., 2003). However, no psychometric data are reported on the psychometric properties of the questionnaire version of the measure. Caregivers' ratings (0-5) on three items pertaining to duration, severity, and frequency of current SIB, were summed to derive a total SIB score out of 15.

The Childhood Executive Functioning Inventory (CHEXI; Thorell & Nyberg, 2008) is a 26-item measure of everyday EF abilities in children aged 4-12 and was included to complement direct EF measures. The CHEXI has been demonstrated to have adequate test-retest reliability (Thorell & Nyberg, 2008) and internal consistency (Romero-Ayuso et al., 2018). Items which probe core EF domains and were rated from "definitely not true" to "definitely true" (0-5), with higher scores indicating increased EF difficulties. Item scores were summed to calculate the overall inhibition (13 items, $\alpha = .87$) and working memory (11 items, $\alpha = .89$) subscales.

The lifetime version of *The Social Communication Questionnaire* (SCQ; Berument et al., 1999) was included to capture autism characteristics, as observed by the caregiver. The 40-item SCQ was developed as a screening tool for autism in children and adults and has been demonstrated to have good concurrent validity with the Autism Diagnostic Interview (ADI-R; Lord et al., 1994) and the Autism Diagnostic Observation Schedule (Chandler et al., 2007). In samples of autistic children, the SCQ was reported to have adequate internal consistency for both subscale and total scores (Wei et al., 2015). The authors suggest a clinical cut-off of 15 as indicative of autism, however, other researchers investigating the SCQ's sensitivity versus specificity in younger age groups have recommended lower cut-off scores (Barnard-Brak et al., 2016; Snow & Lecavalier, 2008). One item regarding self-injury was omitted from analyses as a possible confound. Additionally, to facilitate comparisons between verbal and non-verbal children, scores on the first 6 verbal items of the SCQ were removed from the total score in later analyses, in line with the approach taken by Richards and colleagues (2017).

The Activity Questionnaire (TAQ; Burbidge & Oliver, 2008) was included as a measure of overactivity, impulsive speech, and behavioural impulsivity in persons with more severe ID. The authors report good internal consistency and robust reliability for the subscales (Burbidge & Oliver, 2008). Caregivers rated the 18 items, measuring difficulties remaining still, resting, or waiting, from "never" to "always" (0-5). Scores were totalled to calculate the impulsivity (6 items, $\alpha = .83$), impulsive speech (3 items, $\alpha = .82$) and overactivity (9 items, $\alpha = .88$) subscales.

Adaptive Behaviour

Children's adaptive functioning was assessed using the Comprehensive Interview Form of the Vineland Adaptive Behaviour Scale (3rd Edition [VABS]; Sparrow, et al., 2016). The VABS has been reported to have excellent reliability and internal consistency (Pepperdine & McCrimmon, 2017), and is extensively used in clinical and research settings as a measure of

adaptive ability in both autism and ID populations (Gillham et al., 2000; Perry et al., 2009). Items were scored based on whether the caregiver reported their child "never", "sometimes" or "usually" carries out a certain behaviour without support (0-2). The standardised overall composite score, receptive language and overall age-equivalent scores were used in analyses.

3.4 Direct measures

Receptive Language

The British Picture Vocabulary Scale (3rd edition [BPVS]; Dunn & Dunn, 2009) was administered as a direct measure of children's receptive vocabulary. The authors report adequate internal reliability and concurrent validity with the criterion-standard measure of ability. The BPVS consists of 14 sets of 12 words which increase in difficulty as the examinee progresses. Each word is displayed as a picture, alongside three similar but incorrect options. Target words include objects, such as 'jewellery', animals, such as 'dog', and verbs, such as 'sharing'. Children are asked to select the image which best matches the spoken word. A total score was calculated by subtracting the number of errors from the total items administered.

Executive Function

Given the study's focus on autistic children with a co-occurring ID, mental-age-appropriate EF tasks were identified based on a battery developed by Devine, Ribner and Hughes (2019). Tasks were designed to have minimal verbal demands, which was particularly important given the anticipated heterogeneity of verbal ability in the sample. Footage of the children's performance on direct EF measures and task validity was coded retrospectively using VLC media player (VideoLan, 2006). The validity criteria applied to tasks is described in Appendix I. A sample of recordings (N = 20, 33%) were coded by a second researcher to evaluate inter-rater reliability.

Prohibition Task (Devine et al., 2019; Friedman et al., 2011)

The Prohibition Task was used as a measure of simple inhibition. The caregiver was instructed to draw the child's attention to a desirable toy, a glittery wand, before giving the verbal instruction "*Look [child's name], no, don't touch"* coupled with a visual gesture with their index finger to emphasise the prohibition. As they delivered this instruction caregivers placed the toy down in reach of the child, turned to face away from the child and avoided interacting with them further for the duration of the trial. The dependent variables of interest were the latency to first touch, and whether a child successfully resisted touching the toy over a period of 30 seconds (pass/fail). The latency was recorded using VCL's timestamp function for accuracy. Double-coding of 20 recordings revealed high agreement on waiting duration, ICC = 0.92, $p < .001$, and almost perfect agreement of waiting outcome (i.e., 0 = fail, = pass), $\kappa = .90$, $p < .001$.

A second 'Don't Touch' paradigm was administered later in the battery, with an array of three attractive novel toys: a soft animal keyring, toy car and sensory ball revealed in a fixed order. The task followed the same procedure as the first prohibition task, but with a one-minute waiting time. Latency to first touch and whether the child passed or not were recorded as the dependent variables. Double-coding of 20 recordings revealed high agreement on waiting duration, ICC = 0.95, $p < .001$, and perfect inter-rater agreement for each trial, $\kappa = 1.00$.

Multi Location Search (Devine et al., 2019; Diamond et al., 1994)

In the multi-location search task children were tasked with finding stickers concealed under distinctively coloured opaque cups, one at a time and with a short delay prior to each search. Successful completion of the task required children to hold in mind cups searched, while avoiding returning to an empty cup. The cups remained stationary throughout the task

and, therefore, the colour, size and location of the cups could be used as cues to support spatial working memory. The caregiver was instructed to line up the five cups in a specific order from left to right, and to attract the child's attention as they placed a coloured sticker under a correspondingly coloured cup. The cups were then concealed from the child's view with a large card as the caregiver counted out loud for five seconds. When the paper was removed, the child was encouraged to find a sticker by pointing to a cup. Children were praised when they successfully retrieved a sticker. The sticker was then placed out of sight and caregivers were instructed to show the child the cup was empty before replacing it. After each search, the cups were concealed for a short delay before the child was encouraged to find another sticker. If a child returned to a previously searched location, the parent was instructed to give feedback by showing them the empty cup and saying "*It's not there. Let's have another go*". The game continued until the child had either found all five stickers or made three consecutive incorrect searches. The number of perseverative responses was recorded, and an efficiency score was calculated by dividing the number of stickers retrieved by the total number of searches, with higher values indicating better coordination of working memory and inhibition skills. Double-coding of 20 recordings revealed excellent inter-rater agreement for number of stickers retrieved, ICC = .99, $p < .001$, and good agreement for number of searches, ICC = .79, $p < .001$.

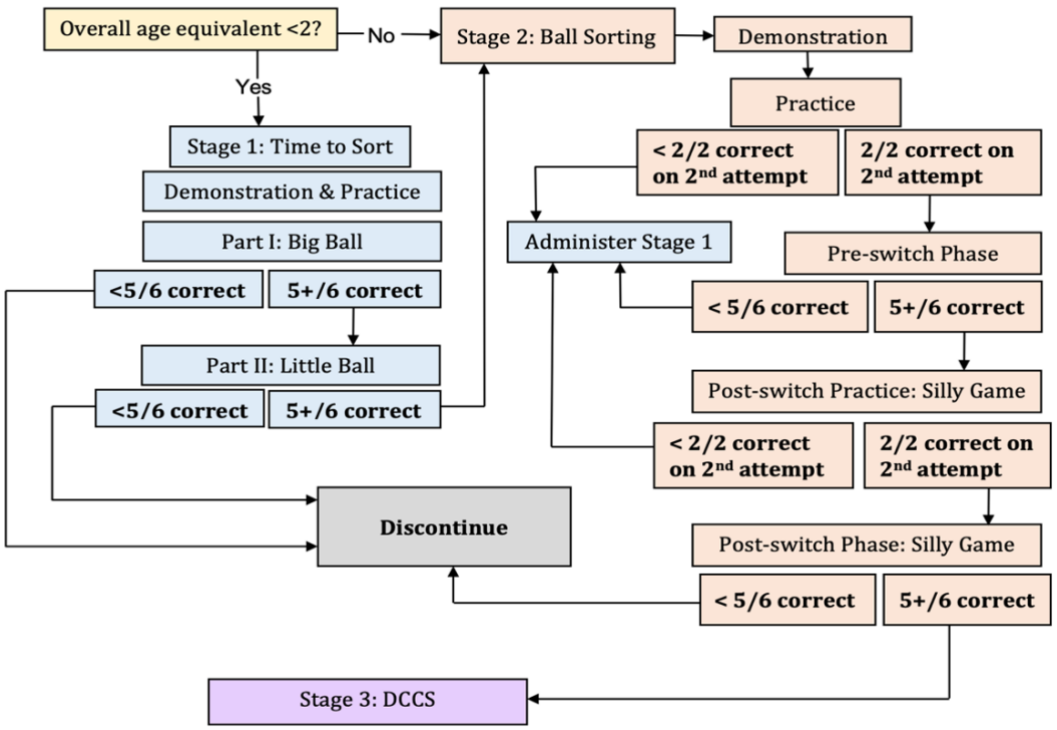
Conflict scales

In an adaptation of Beck and colleagues' (2011) conflict scales, three tasks of increasing complexity were designed to assess aspects of flexibility and conflict across the sample's expected range of abilities. The starting task was determined by a child's estimated adaptive age-equivalent. If a child started on a more advanced game and passed, a pass was inferred for simpler tasks. If a child did not pass this, the order of administration was reversed (see Figure 2). A total 'conflict' score was calculated by summing together the number of

post-switch phases passed across all three tasks out of a possible three. To include as much of the recorded data as possible, missing or incorrectly administered trials were managed by calculating a proportion score for each phase (number correct/valid trials) and scores over 0.80 were considered as a pass for that phase (Zelazo, 2006).

Figure 2

Cognitive flexibility protocol



Sorting task (Devine et al., 2019)

A simple sorting task was used to assess rule learning (set-formation) and reversal. Caregivers were instructed to set up a large box a shoulder-width apart from a smaller box without their lids on. In the rule learning phase, caregivers were asked to gain the child's attention as they demonstrated placing a large ball into the larger box before inviting the child to have a go. Following one practice trial in which feedback was given, children were asked

to sort the same large ball six times consecutively. Caregivers were instructed to pass the ball to the child centrally, and the ball was retrieved from the box after each placement to avoid cuing responses. If a child successfully placed five out of six large balls into the larger box without feedback, they proceeded to the reversal phase where they were tasked with placing a small ball and in the small box instead. Only children who correctly placed five out of six of these balls advanced to the next categorisation task.

Categorization/Reverse Categorization (Carlson et al., 2004)

In the reverse categorisation task, children were required to sort the balls according to a certain rule, which was then changed to assess their ability to learn a new rule and adapt their responses accordingly. The categorisation phase of the task required big balls to be sorted into a larger box, and little balls into a smaller box in a standardised order. Caregivers were again instructed to place the different sized boxes shoulder-width apart with their lids on, before introducing and demonstrate the rule. Children were then invited to practice sorting two big and two little balls alternately. Caregivers were instructed to pass balls centrally and to keep the lids on boxes in-between trials to avoid cueing the child's response. A child could respond by placing a ball into a chosen box, or by pointing to a box for the caregiver to place the ball into. Corrective feedback or praise was given during the practice phase to support learning. If a child sorted both sized balls correctly by their second attempt, the caregiver was instructed to proceed with the task without proving any feedback other than repeating the rule before each trial. If the child correctly sorted five out of six balls, a reversal rule was introduced. In the reverse categorisation phase, children were told that they were going to play a "silly game" in which the big balls now go into the smaller box, and the little balls into the larger box. The caregiver was asked to demonstrate this new rule before a child was asked to practice by sorting two balls of each size according to this new rule. Caregivers were instructed to provide praise or corrective feedback during the practice trials. If a child

correctly sorted both sized balls by their second attempt, the task proceeded as before with caregivers passing balls in a set order, without providing any feedback aside from repeating the reversal rule each time. Children who correctly sorted five out of six balls in the silly game were later given a more advanced set-shifting task. Double-coding of 18 of the sorting and reverse categorisation tasks revealed high inter-rater agreement, ICC = .99, $p < .001$.

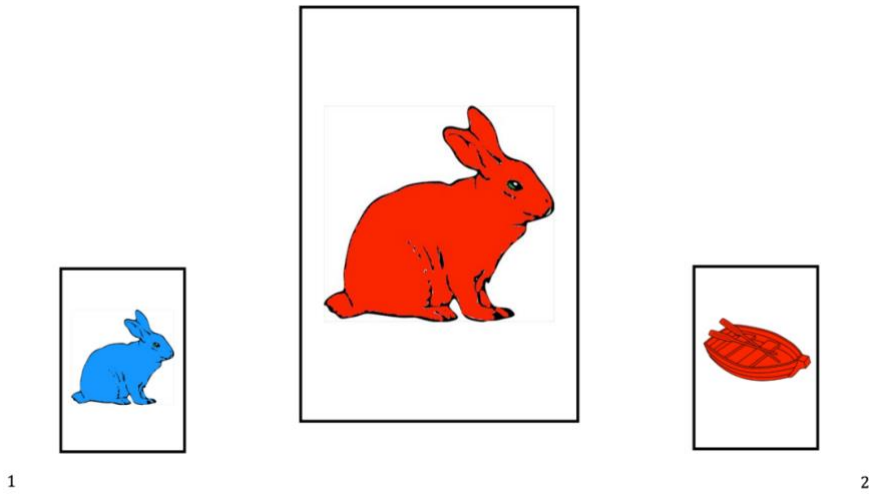
Dimension Change Card Sort (DCCS: Zelazo, 2006)

A slideshow presentation of the DCCS, administered by the researcher, was the most advanced assessment of cognitive flexibility administered. The 'cards', differed on two dimensions: colour (red vs. blue) and shape (rabbit vs. boat). Children were first presented with two cards on the screen, a red rabbit and blue boat, and told that they would be sorting these according to a certain rule. Children were invited to indicate where they thought a card belonged, by either pointing to or naming one of two numbered 'piles' arranged to the left or right of the screen (see Figure 3). If a child opted to point to a pile, then a caregiver was asked to call out the corresponding number for the researcher to record the child's response. In the pre-switch phase, children were asked to sort the cards into the piles according to colour. This was demonstrated before the child was given a practice turn in which they were given explicit feedback. No further feedback was given after this practice trial. Cards were then displayed in a predetermined order. The was rule repeated before each trial and the target feature of each card was verbalised by the researcher: "*here is a blue one, where does it go?"*. If a child successfully sorted five out of six cards according to colour, they progressed onto the post-switch phase in which they were asked to sort cards according to shape instead. A further six cards were shown in a standardised order and the 'shape game' rule was repeated before each trial. The total number of correct trials in each phase was recorded. The key dependent variable was the total number of correct post-switch responses (0-6), with higher scores

indicating better cognitive flexibility. Double-coding of 15 recorded administrations revealed perfect inter-rater agreement, ICC = 1.00, $p < .001$.

Figure 3

Child's view of the online DCCS task



3.5 General Procedure

All data were collected remotely. For a brief review of the evidence on remote testing, see Appendix H. Written information about the study, questionnaires, and a consent form (see Appendices C - E for a copy of the study pack sent to caregivers) were either accessed and completed online through a link emailed to caregivers, or posted as a paper study pack alongside pre-paid return envelopes where preferred. After written consent was obtained from a legal guardian, caregivers attended a video call with a trained researcher during which a risk assessment and VABS interview was completed. Some time was spent familiarising caregivers with the toys and equipment required for their child's 'games session' and demonstrating the EF measures which the caregiver would be supported to deliver. Kits containing the toys sealed in envelopes numbered according to task, and a laminated guide with task instructions and visual prompts were then posted out to families (see Appendix F for

an extract of the instructions sent to caregivers). Wireless headphones were included with the kits if the caregiver did not have access to their own. Caregivers were instructed to keep boxes sealed and toys out of sight of children before the assessment session to retain the novelty of toys across the sample.

Following the caregivers' video call, children attended a remote games session led by a minimum of two researchers based at the University of Birmingham, conducted via the online video conferencing platform 'Zoom'. Verbal consent was sought again immediately prior to the child's participation in the direct measures. The child's performance on all tasks was recorded with their caregiver's consent, and all recordings were stored on the University's secure server. Children completed the session in a quiet room either at home or school, with the support of at least one adult. Most performance measures were administered by a caregiver, guided by a researcher via wireless headphones so that the child could not hear the researcher. The BPVS and the DCCS were both administered remotely by a researcher, with caregivers verbally relaying a child's response to stimuli presented on the screen if required. Aside from the conflict scales, all tasks were attempted, in the same order with all children (See Appendix G for full study flow chart). On average, the online assessment took 75 minutes to complete. The protocol was designed to be flexible, allowing for breaks between tasks if required. Children were motivated through positive feedback and were able to keep all the toys used in the session. Families were further compensated for their time with a feedback report detailing the outcome of all questionnaire and direct measures completed. This was posted alongside an encrypted pen drive with clips of their child completing the tasks, and an electronic copy of the individualised report to allow caregivers to easily share this with involved professionals should they wish to.

Ethical approval for this study was obtained from Health Research England and Bristol Research Ethics Committee (IRAS Project ID 263048), with local research and development approval obtained at all participating NHS trusts (see Appendix A).

Data preparation

Prior to analyses, four variables were extracted from the direct EF data. These were chosen to index the core EF domains of inhibition, updating and set-shifting, while retaining data from as much of the sample as possible. Latency scores were bi-modally distributed for both prohibition tasks, and a Spearman's correlation computed between latency data on the two tasks revealed a strong positive association, $r_s = .69$, $p < .001$. Therefore, a composite prohibition score was calculated for each participant by summing their scores on both tasks (1 = pass, 0 = fail) and dividing this by the number of valid trials (1-2). This resulted in a prohibition score (0-1) for 57 participants. For the multi-location task efficiency scores were retained as an index of working memory ability and a perseveration score, calculated as a proportion of the child's total number of valid searches, was included as a further measure of inhibitory control. Due to data loss on the conflict scale, the conflict score was calculated as a proportion (post-switch phases passed/post-switch phases administered) to retain as much of the recorded data as possible. The inhibition and working memory composite scores from the CHEXI ,and the TAQ's impulsivity, impulsive speech and overactivity subscale scores were included in analyses. Table 2 presents the mean and range of the EF outcome variables.

Table 2

EF Variables Included in Analyses

Measure	N	M (sd)	Range
Direct			
Prohibition	57	.32 (.43)	0 – 1
ML perseveration	53	.12 (.22)	0 - .75
ML efficiency	53	.82 (.26)	.25 – 1
Conflict proportion	45	.80 (.27)	0 – 1
Informant-based			
TAQ Impulsivity	60	20.53 (4.12)	6 – 24
TAQ Impulsive speech [a]	37	6.32 (4.12)	0 – 12
TAQ Overactivity	60	23.53 (9.52)	0 – 36
CHEXI WMF	59	55.22 (6.96)	41 – 65
CHEXI IF	60	47.32 (6.23)	34 – 55

Note. TAQ = The Activity Questionnaire; CHEXI = The Childhood Executive Functioning Inventory; WMF = Working Memory Factor; IF = Inhibition Factor.
[a] verbal participants only

3.6 Data Analysis

Data collection was delayed due to the ongoing COVID-19 pandemic. Consequently, the pilot analysis of an incomplete data set is described here. Analyses aimed to evaluate emerging trends in this small sample to allow for comparison to previous literature. An alpha of .05 rather than a more conservative threshold was chosen for all analyses. This pragmatic decision prioritised identifying emerging effects over an increased risk of Type 1 errors. Effect sizes were calculated to inform discussions on both the scientific and clinical implications of the findings.

All data analyses were conducted using SPSS Statistics, Version 28. Visual inspection of distribution plots and Kolmogorov–Smirnov tests were used to assess normality. Where data were not normally distributed ($p < .05$), appropriate non-parametric tests were used. First, the performance of children on direct EF tasks was summarised. Children's

performance across the two phases of reverse categorisation and DCCS tasks was examined using paired t-tests to evaluate the validity of these tasks in the sample.

Next, the relationships between subscale scores derived from EF questionnaires, composite scores from EF performance measures, demographic variables and total SIB score were evaluated through correlational analyses. Given that much of the data were ordinal and violated the assumptions of parametric tests, Spearman's rank and point-biserial correlations were calculated. Hierarchical regression analyses were then conducted to determine whether EF abilities accounted for variance in SIB scores, over and above differences in adaptive ability. Visual inspection of the data alongside statistical analyses indicated linear relationships between predictor and dependent variables. Given that correlational analyses revealed a large significant relationship between adaptive ability and direct EF measures (see Appendix K), collinearity statistics were consulted. Variance inflation factor (VIF) and tolerance scores did not indicate issues with multicollinearity. Durbin-Watson statistics were within accepted ranges, indicating that the assumption of independence of errors had been met. Scatter and Predicted Probability (P-P) plots suggested that variance of the residuals was constant, and these error terms were normally distributed. Finally, the absence of outliers which may have unduly influenced the outcomes was confirmed by Cook's Distance values of below one. Therefore, data were deemed to have met major assumptions of linear regression including independence, homoscedasticity, and normality of residual terms.

Finally, to test the hypothesis that children presenting with and without SIB would differ on measures of EF, a case-by-case age-matched sample was generated, and independent sample tests were run to evaluate between-group differences in EF abilities.

2.4 Results

4.1 Performance on Direct EF Measures

To explore the first research question regarding the feasibility of remote delivery of EF tasks in this under-represented population, children's performance on each of the administered EF tasks was summarised. Table 3 presents the means and standard deviations of the extracted variables, along with the percentage of data lost and percentage of children scoring at ceiling. Paired t-tests revealed a significant mean difference between children's scores on the reverse categorisation phases (t_{41} = 2.15, p = .04, Hedges' g = .22) and a more pronounced significant difference between average scores on the pre- and post-switch phases of the DCCS (t_{29} = 4.56, p < .001, Hedges' g = .42).

Although all direct EF tasks had non-negligible missing data in contrast to informant-based measures, an analysis of missing values was non-significant, indicating the direct EF data were missing completely at random (MCAR), Little's $\chi^2(135) = 157.65$, p = .09. Table 4 summarises the various reasons for data loss on EF tasks. Data were also missing for 21 children on the receptive language measure (BPVS) due to researcher error (N = 1), disinterest or low ability (N = 13), and frustration or fatigue (N = 7).

Table 3

Descriptive Statistics for Direct Executive Function Measures

EF Measures	N eligible	N completed	% Missing	% ceiling	pass/fail /M(sd)
Inhibition					
Prohibition pass (wand)	60	47	21.67	42.55	20/27
Prohibition pass (array)	60	54	10.00	27.78	15/39
ML #perseverative [a]	60	53	11.67	67.92	0.57 (0.95)
Updating					
ML efficiency	60	53	11.67	58.49	0.82 (0.26)
ML span	60	53	11.67	58.49	3.89 (1.41)
Set-Shifting					
Ball Sort 1 prop. correct	19	7	63.16	57.14	0.84 (0.17)
Ball Sort 2 prop. correct	17	5	70.59	60.00	0.75 (0.50)
RC pre prop. correct	45	43	4.44	97.67	0.97 (0.10)
RC post prop. correct	43	42	2.33	88.10	0.90 (0.24)
DCCS pre prop. correct	38	32	17.95	86.67	0.93 (0.16)
DCCS post prop. correct	32	30	1.33	53.57	0.59 (0.46)
Conflict score	60	36	40.00	47.22	2.28 (0.85)

Note. ceiling performance on the perseveration measure is based on zero perseverative responses.

[a] higher scores indicate poorer performance.

Table 4

Reasons and N for data loss on direct EF measures

Reason	Task	N
Researcher administration error	Wand	2
	Array	2
	Ball sort	3
Caregiver administration error / interference	Wand	2
	Array	2
	Multilocation	2
	Ball sort	3
	Reverse Categorisation	1
	DCCS	2
Child disinterested / did not engage	Wand	2
	Multilocation	2
	Ball sort	5
	Reverse Categorisation	3
	DCCS	3
Child distracted	Wand	3
	Array	1
	Multilocation	2
Child frustration or fatigue	Ball sort	3
	Multilocation	1
	Reverse Categorisation	3
	DCCS	3
External interference	Wand	1
Visibility	Wand	3
	Array	1

4.2 Direct versus informant-based EF measures

To explore the second research question, the correlations between each measure of EF are presented in Table 5. While most direct and informant-based EF variables correlated significantly with one another, there were no significant associations between these measures. Table 5 reports the correlations for the conflict proportion score, however, when the raw conflict scores (generated from complete data) are used, this was robustly correlated with all direct composite scores (r_s = .44 – .55, p < .001) and showed a small but non-significant association with the TAQ impulsivity subscale (r_s = -.21).

Table 5

Spearman's rho correlations between EF variables

		1	2	3	4	5	6	7	8
Direct									
1	Prohibition [a]								
2	ML perseveration	-.36**							
3	ML efficiency [a]	.39**	-.86***						
4	Conflict prop. [a]	.21	-.32*	.36*					
Informant-based									
5	TAQ Impulsivity	-.11	.16	-.27	-.03				
6	TAQ Impulsive Sp.	.14	.02	.11	.20	.40*			
7	TAQ Overactivity	-.12	.17	-.22	.11	.58***	.56***		
8	CHEXI WMF	.21	.03	.03	.30	.41**	.59***	.32*	
9	CHEXI IF	.20	-.03	.05	.12	.51***	.50**	.45**	.68***

[a] higher scores indicate better performance.

*p < .05 **p < .01 ***p < .001 (2-tailed)

SIB and demographic characteristics

There were no significant correlations between children's SIB scores and demographic variables, including age, gender, receptive language skills, adaptive ability, or autism characteristics (see Appendix J).

4.3 Relationship between SIB and EF outcomes

To explore the third research question, Table 6 presents the results correlational analyses between total SIB scores and EF variables. Total SIB score was moderately associated with TAQ Impulsivity and Overactivity subscale scores. A weaker yet significant negative association was revealed between ML efficiency and total SIB score. Additionally, a small positive relationship between the Perseveration and SIB scores approached significance. While no relationship was found between total SIB score and the conflict proportion, when raw conflict scores were used a non-significant trend (r_s = -.23) indicated that the children exhibiting more difficulties with set-shifting were showing more SIB.

Table 6

Relationship between measures of EF and total SIB score

EF measure	r_s	*p*	N	95% CI[a]	
				LL	UL
Direct					
Prohibition score [a]	-.15	.27	57	-.40	.13
ML perseveration	.25	.07	53	-.03	.49
ML efficiency [a]	-.28*	.05	53	-.51	.00
Conflict prop. [a]	.08	.58	44	-.22	.38
Informant-based					
TAQ Impulsivity	.41**	.00	60	.17	.61
TAQ Impulsive Sp.	.12	.47	37	-.22	.44
TAQ Overactivity	.33*	.01	60	.08	.55
CHEXI WMF	.16	.22	59	-.12	.41
CHEXI IF	.10	.46	60	-.17	.35

Note: CI = confidence interval; LL = lower limit; UL = upper limit

[a] higher scores indicate better performance.

[a] based on Fisher's r-to-z transformation.

*$p < .05$ ** $p = .001$ (2-tailed)

Do EF outcome measures predict the total SIB score?

To further explore the third research question, a series of two-step hierarchical regressions were run with the SIB score as the dependent variable. Standardised adaptive behaviour scores were entered at step one of all regression models as a proxy for ability. Each of the EF variables were entered as a second step in separate models. The regression statistics are presented in Table 7. The results indicated that at step one, adaptive ability accounted for only 4.3% of the variation in SIB score, and did not significantly contribute to the model,

$F(1,58) = 2.50$, $p = .12$. Including TAQ Impulsivity resulted in a significant change in R^2, $F(2, 57) = 4.07$, $p = .02$, with a medium effect size, Cohen's $f^2 = 0.15$. TAQ Impulsivity uniquely explained 8.35% (sr^2) of the variance in SIB score. Similarly, introducing TAQ Overactivity as a predictor alongside ability in a separate regression, explained an additional 7.95% of variation in SIB score and resulted in a significant change in R^2, $F(2, 57) = 3.91$, $p = .03$, Cohen's $f^2 = 0.14$. Although no other EF variable explained a significant amount of variance in SIB scores, the direction of all beta coefficients were in line with the hypothesis that increased executive difficulties are associated with higher SIB scores. When raw conflict scores were used instead, the resultant change in R^2 approached significance ($\Delta R^2 = .10$, $p = .06$). Finally, the regressions for prohibition and TAQ Impulsivity scores were re-run to include chronological age and SCQ score respectively at step one, to account for the recorded correlations between these measures (see Appendix K). Prohibition remained non-significant, but with a lower coefficient and TAQ Impulsivity remained a significant predictor over and above ability and autism characteristics ($p = .05$, Cohen's $f^2 = 0.13$).

Table 7

The results of hierarchical regressions with total SIB score as the dependent variable

Variables						
Step 1	Step 2	β	*t*	*R*	R^2	ΔR^2
VABS ABC		-.20	-1.58	.20	.04	.04
	VABS ABC	-.19	-1.37	.21	.05	.00
	Prohibition[c]	-.07	-.49			
	VABS ABC	-.16	-1.08	.23	.06	.01
	ML Perseveration[d]	.12	.84			
	VABS ABC	-.12	-.81	.62	.07	.03
	ML Efficiency[d]	-.18	-1.21			
	VABS ABC	-.19	-1.23	.23	.05	.01
	Conflict Prop.[e]	-.12	-.70			
	VABS ABC	-.09	-.65	.35	.13	.08*
	TAQ Impulsivity	.31	2.33*			
	VABS ABC	.-.21	-1.26	.24	.06	.02
	TAQ Impulsive sp.[a]	.13	.75			
	VABS ABC	-.15	-1.20	.35	.12	.08*
	TAQ Overactivity	.29	2.27*			
	VABS ABC	-.19	-1.47	.25	.06	.02
	CHEXI WMF[b]	.15	1.12			
	VABS ABC	-.20	-1.53	.23	.05	.01
	CHEXI IF	.10	.76			

Note. N = 60 unless otherwise noted. [a]N = 37; [b]N = 59; [c]N = 57; [d]N = 53; [e]N = 44.

CI = confidence interval; LL = lower limit; UL = upper limit

*$p < .05$

4.4 Comparison of a matched sample of children with and without SIB

To explore the fourth research question, children who had not displayed any form of SIB in the preceding month (N = 18), were matched on age (+/- 1 year), adaptive ability (+/- 10 VABS standardised ABC score), and autism characteristics (+/- 4 SCQ score) with children who were reported to display repeated and ongoing SIB (N = 42), resulting in 15 participants in both a SIB and non-SIB group, see Table 8.

Table 8

Demographic characteristics of a subset of SIB and non-SIB children matched on age, adaptive ability and SCQ scores.

	SIB (N = 15)	non-SIB (N = 15)		
	M (sd)	M (sd)	t / χ^2/U	p
CA (years)	10.27 (2.69)	9.94 (2.79)	-0.33	.74
Gender (female, male)	2, 13	3, 12	0.24[a]	.62
BPVS raw [c]	74.13 (41.54)	45.17 (27.50)	-1.89	.08
VABS ABC	39.27 (12.85)	42.73 (12.88)	0.74	.47
VABS RL ae	1.57 (0.66)	1.78 (0.72)	101.00[b]	.65
VABS overall ae	2.54 (0.76)	2.85 (0.91)	89.00[b]	.35
SCQ total (verbal items removed)	24.27 (5.54)	25.27 (4.68)	102.50[b]	.68

Note. CA = chronological age; RL = receptive language.

[a] Chi-square statistic.

[b] Mann-Whitney U statistic.

[c] SIB N = 8; non-SIB N = 12

Although participants were not matched based on gender, independent sample tests indicated no significant group differences in gender, age, autism characteristics, adaptive ability, or receptive language. The groups' scores on measures of EF are displayed in Table 9.

The non-SIB children's scores on all measures indicate better EF on average. While none of the mean differences on direct measures reached significance, effect sizes were largest on inhibition variables. When raw conflict scores were used, no significant group difference was found, although the median was higher for children without SIB.

On average, caregivers of children in the non-SIB group endorsed fewer indicators of EF difficulties on informant-based measures when compared to the SIB group. On the TAQ, children in the non-SIB group were scored as significantly less impulsive (Mdn = 21.00)[8] than children in the SIB group (Mdn = 23.00), U = 64.50, $p < .05$. While other differences were non-significant, the mean difference on the caregiver-rated CHEXI working memory factor approached significance (p = .08).

[8] For outcome data which are not normally distributed, medians (*Mdn*) are reported for dependent variables which have a similarly-shaped distribution across groups.

Table 9

Comparison 15 of children who displayed current SIB with a matched sample of 15 children were not presenting with SIB, on direct and caregiver-rated EF measures.

	SIB			Non-SIB					
EF measure	n	M	SE	n	M	SE	*t* / *U*	*p*	*g*
Direct									
Prohibition [a]	15	0.23	0.11	14	0.32	0.10	87.00 [b]	.41	-0.34
ML perseveration	13	0.17	0.06	15	0.10	0.05	72.50 [b]	.20	0.51
ML efficiency [a]	13	0.77	0.08	15	0.81	0.07	86.50 [b]	.60	-0.19
Conflict prop.[a]	12	0.70	0.10	12	0.81	0.06	60.00 [b]	.51	-0.23
Informant-based									
TAQ Impulsivity	15	22.13	0.65	15	20.00	1.04	64.50 [b]	.04 *	0.78
TAQ Impulsive sp.	8	8.13	1.49	9	6.11	1.42	23.50 [b]	.24	0.58
TAQ Overactivity	15	23.93	2.38	15	22.20	2.28	-0.53	.60	0.19
CHEXI WMF	15	57.40	1.44	15	53.07	1.86	-1.85	.08	0.66
CHEXI IF	15	48.27	1.59	15	46.67	1.67	94.50 [b]	.47	0.27

[a] higher scores indicate better performance.

[b] Mann-Whitney U

*****$p < .05$

2.5 Discussion

The study evaluated differences in the presence and severity of SIB relative to informant-reported and directly assessed EF abilities of sixty autistic children with a co-occurring ID. The results corroborate previous findings of a robust association between informant-reported profiles of impulsive and overactive behaviours and SIB in autistic samples (Laverty et al., 2020; Richards et al. 2017; Richman et al. 2013). The present study extends these results by using well-established measures to control for known correlates of SIB, such as autism characteristics and ability. Although limited conclusions can be drawn from correlational data, the results offer tentative support for a role for behavioural inhibition as a risk factor in SIB, as proposed by Oliver and Richards (2015). The lack of agreement between direct and informant-based measures of EF, along with non-significant findings on direct measures of EF will now be discussed.

Using both informant-based and direct measures of EF allowed for a comprehensive examination of the hypothesised association between SIB and EF abilities. Due to the necessary remote administration of the study, it was not possible to administer a direct measure of intellectual functioning. However, a standardised measure of adaptive functioning, the VABS (Sparrow et al., 2016) was collected for all participants. The VABS has been shown to closely co-vary with estimated intelligence quotients (IQ) in autistic children with lower levels of intellectual ability (e.g., Alvares et al., 2020). Therefore, standardised adaptive behaviour scores were used as a proxy for IQ in analyses. An established autism screening tool was also administered to corroborate diagnoses. Together, these measures ensured the obtained sample met the full inclusion criteria and strengthened the internal and external validity of the study.

5.1 The feasibility of remote measurement of EF in autistic children with ID

Remote assessment methods have not previously been reported for this clinical population. Moreover, autistic individuals with a co-occurring ID are often excluded from psychological and EF research altogether (Russell et al., 2019). Despite the non-negligible data loss, the overall quality of recorded data suggested that the direct assessment of cognitive abilities, even remotely, is both feasible and informative. The appropriateness of the chosen EF battery for this group of children was supported by the variation in recorded performance and the absence of clear floor or ceiling effects. Furthermore, scores on post-switch conditions of the conflict scale were significantly lower than pre-switch scores. This demonstrated that both the reverse categorisation and DCCS tasks were measuring EF abilities in these children, rather than more general difficulties with learning a rule (Carlson, 2005). This echoes the patterns of performance recorded in neurotypical children (Devine et al., 2019), demonstrating the validity of these tasks. While proportion scores were used as a way of reducing data loss, the observed differences between the raw and proportion-scored conflict scale suggested that the latter may have inflated children's scores on this measure. For example, if a child passed the reversal phase on the sorting task but then refused to do the categorisation/reverse categorisation or DCCS tasks, they were awarded a proportion score of 1. This raises questions about how investigators administer and score these tests in the future, as proportion scoring may exert an undue influence on the data, which may be particularly important in samples with co-occurring ID where overall ability level may influence participants' capacity to undertake tasks.

Reviewing the reasons for data loss, a simple set-formation task was the most discontinued in the battery. Given that this was designed for participants with lower ability levels, it is difficult to unpick whether the task was not completed because it was too difficult, or whether it was too simple and therefore not engaging enough. A similar set-formation task

designed by Devine and colleagues (2019) reported far lower levels of attrition in a neurotypical sample of children under the age of two. This version of the task included musical feedback to reinforce game play. The present study was limited by the type of equipment that could be posted to families, and it is possible that this version of the task was not as attractive for the children. Researcher and caregiver administration errors also contributed to data loss. However, this approach to testing, with caregivers administering most tasks was completely novel. A second prohibition task was administered with fewer validity concerns and likely benefited from the caregivers' experience of administering the first task. Future remote research could run additional training sessions with participating caregivers to allow for further practice of novel tasks. Finally, fatigue might have contributed to some data loss, particularly for the BPVS and DCCS which were the final tasks administered in the battery. This might have impacted on the data collected for the conflict scales. Although breaks were offered throughout the session, conducting the data collection in two shorter sessions might have been helpful for some participants and resulted in better uptake of the final tasks. Finally, given the novelty of the methods used, future work should seek to validate the remotely-obtained EF data by repeating the task administration, in-person with a subset of the participants.

5.2 Direct versus informant-based measures of EF

Direct EF measures were robustly intercorrelated, as were informant-based measures. However, none of the trends between these types of measures reached significance. While small to moderate associations between the CHEXI and direct measures of EF have previously been reported in samples of TD and autistic children without a co-occurring ID (Golshan et al., 2019; Thorell & Nyberg, 2008), the non-significant small-to-medium correlations between direct measures and subscales of the CHEXI in the present study were in the opposite direction to that expected, such that higher scores on the CHEXI subscales

(indicating more executive difficulties) were positively correlated with higher scores on direct measures (indicating better executive abilities). Conversely, there were small, non-significant associations between the TAQ and direct measures of EF in the expected direction. Although impulsivity and overactivity are presumed to be underpinned by reduced inhibitory control (e.g., Barkley, 1997), and the TAQ impulsivity subscale was robustly associated with SIB, there was no relationship found between the CHEXI inhibition factor and SIB in the sample. This lack of agreement between EF ratings and performance-based data warrants further exploration.

Although differences between direct and informant-based measures were anticipated, the lack of convergence between informant-measures proposed to capture similar abilities suggested that these may be confounded in some way by other abilities and/or behaviours. All children scored highly on both CHEXI subscales, with caregivers strongly endorsing items relating to EF difficulties, whereas scores on the TAQ and direct measures were more widely dispersed. One possible interpretation of these differences is that the CHEXI may not be suitably calibrated for autistic children with a co-occurring ID and, therefore, may mask children's true EF abilities. The floor on the measure may be higher than that of the TAQ, which was designed for an ID population. Items on the CHEXI, for example, "*Has difficulty understanding verbal instructions unless he/she is also shown how to do something*", may be measuring differences in addition to EF, such as receptive language. Consequently, it might be that these children were scored as very impulsive and with significant working memory difficulties due to generally lower adaptive function and IQ abilities. This interpretation seems likely, as although the recommended age range of the CHEXI encompassed the chronological age range of recruited children, the average estimated age-equivalent adaptive functioning performance of the sample fell below the lower age limit of the measure. Future research could administer the CHEXI across a larger sample of autistic children with ID to evaluate its

factor structure. Alternatively, researchers might consider if there is a way to score the CHEXI excluding items which rely on higher IQ, adaptive or verbal ability.

Informant-based and direct measures of EF often do not correlate as might be expected (Camerota et al., 2016) or produce comparable effect sizes in the same sample (see Chapter 1 meta-analysis). On questionnaire measures, informants are rating the perceived outcome of goal-directed behaviours as they present in everyday activities, across contexts and over time. In contrast, direct measures are designed to capture the efficiency of cognitive processes in a structured and controlled environment (Toplak, 2013). Although informant-based measures may have the advantage of increased ecological validity, they also have certain limitations. For example, respondents might not always interpret questions in the same way, recall behaviour accurately or respond objectively. For example, Denckla (2002) highlights the potential for the emotional experiences of informants, or their relationship with a behaviour of interest, to influence the ratings they provide. Additionally, informant-based measures may capture behavioural characteristics in addition to, or even instead of the targeted EF ability. Upon looking at the items which comprise the TAQ impulsivity subscale (see Table 10), it seems possible that caregiver responses might capture more than impulsive behaviour. For some children, the social communication differences characteristic of autism might include difficulties interpreting social cues around when to interrupt or difficulties with turn-taking (APA, 2013). Therefore, item 18 could be interpreted as an endorsement of these differences rather than purely difficulties with waiting. This could constitute a potential confound in the data, as higher endorsements of autism characteristics have been associated with a higher likelihood of SIB (e.g., Baghdadli et al., 2003). However, although this interpretation of item 18 would unduly inflate the impulsivity score, it is reasonable to assume this may not differ systematically across the relatively homogeneous sample with and without SIB. Moreover, while the impulsivity subscale score does moderately correlate with the adjusted SCQ score,

impulsivity remains a significant predictor of SIB presentation over and above autism characteristics. For some participants, items 11 and 12 may be capturing the challenges associated with keeping a child with SIB safe, and the potential need for increased vigilance. If caregivers of a child with SIB do interpret these items in a systematically different way, this could be confounding the interaction between the impulsivity subscale score and SIB. This should be evaluated through an item by SIB analysis in the full sample, once recruited.

Table 10

Items forming the impulsivity subscale of the TAQ

Item	Question
11	Does the person's behaviour seem difficult to manage/contain whilst out and about?
12	Do you feel that you need to "keep an eye" on the person at all times?
13	Does the person you care for seem to act/do things without stopping to think first?
16	Does the person want things immediately?
17	Does the person find it difficult to wait?
18	Does the person disturb others because they have difficulty waiting for things or waiting their turn?

5.3 SIB prevalence and demographic characteristics

The prevalence of SIB in the recruited sample (70%) was higher than the 20 to 50% commonly reported in other autism samples (Baghdadli et al. 2003; Richards et al. 2012; Soke et al., 2016). It is possible that this was due to a greater number of participants having been recruited through participating NHS sites where children are likely to present with higher levels of behavioural needs. However, a high rate of the behaviour of interest in this small

pilot sample had scientific utility, as it generated a greater range of SIB scores which afforded within-participant analyses. There were no notable relationships between the recorded SIB scores and any demographic variables, including known risk markers such as age, ability, and autism characteristics (Baghdadli et al., 2003; Laverty et al., 2020; Richards et al., 2012, 2017; Soke et al, 2017). This is largely unsurprising given the relatively homogenous high-risk sample recruited, and further illustrates the scientific and clinical value in identifying the personal characteristics which are independently associated with SIB presentations in this population. Interestingly, unlike previous studies in autism and ID populations (e.g., McClintock et al., 2003), there were no significant relationships between reported SIB and measures of the participants' expressive or receptive communication skills. A purely operant account which conceptualises SIB as a functional communication strategy and, therefore, more likely to be reinforced for individuals with lower levels of ability, cannot fully account for this non-significant finding.

5.4 SIB and executive function

Correlational analyses revealed a moderate positive association between SIB and caregiver reported impulsivity and overactivity, with more impulsive and overactive children also displaying more frequent and severe SIB. For direct measures of EF, there was a significant small-to-moderate association between efficiency scores (presumed to index the coordination of working memory and inhibition skills) and SIB, and a small-to-moderate relationship between perseveration and SIB which also approached significance. While all associations between direct EF measures and SIB trended in the expected direction, slightly larger coefficients were recorded for measures of conflict and perseveration than the prohibition score. This suggests that children displaying more frequent and severe SIB, may have relatively more difficulties disengaging from a reinforced response pattern than delaying a prepotent response. Only caregiver-rated impulsivity and overactivity uniquely explained

some of the variance in SIB scores, over and above an index of ability in regression analyses. However, most of the variance in total SIB score was unattributable to the recorded EF or demographic variables. In addition to different functions and reinforcement histories, the unexplained variation in SIB could be associated with factors which were not evaluated in the present study. For example, pain and sensory function, which are associated with SIB in this population (Dimian & Symons, 2022). The robust association between SIB and caregiver-rated impulsivity and overactivity, together with the trends observed in direct EF variables, does support the hypothesis that the presence and severity of SIB is associated with lower scores on direct measures of EF, and higher endorsement of EF difficulties on informant-based measures

A small, matched group comparison controlled for the potential confounds of age, autism characteristics and ability. Scores on all measures indicated better EF abilities in the group without SIB on average. For informant-based measures, only the between-group difference in impulsivity scores was significant, with a large effect size. Although differences on the CHEXI working memory subscale approached significance, an interpretation of this result is limited by the issues discussed. While no significant differences were found on direct measures, the largest effect sizes were recorded for inhibition variables, particularly perseveration, when compared to measures of shifting and working memory. While the results of this analysis do support the hypothesis that the scores of children with current SIB would indicate reduced EF skills in comparison to children without current SIB, these findings should be interpreted with caution. The sample size was very small and multiple comparisons increased the risk of Type I errors. However, it is possible that these trends may reach significance in the full study sample.

The good replication of a robust association between SIB and caregiver-rated impulsivity in this small pilot sample suggests this is a strong effect. Thus, these preliminary

findings do suggest that there might be underlying differences in EF which might predispose children to display SIB in the context of evoking antecedents. Furthermore, the reported effect sizes for direct EF measures also suggest a role for reduced inhibitory control in driving SIB. Reduced abilities to "brake and disengage" might constitute a setting event (Wahler & Fox, 1981), in which SIB is more likely to emerge. More impulsive children might be less able to inhibit SIB in response to an antecedent, or to disengage from behaviour after constraining consequences have elapsed. The results add to the literature describing an emerging behavioural profile associated with SIB in neurodevelopmental populations. The exact nature of this association remains a question for future research. To evaluate the extent to which behavioural inhibition might play a causal role in the developmental trajectory of SIB, future studies could consider whether relatively greater executive difficulties predict the onset of SIB in younger children. Additionally, intervention studies could evaluate the impact of executive function training, or other strategies to reduce motor impulsivity, on SIB frequency and severity. The identification of reliable behavioural risk factors has the potential to extend current operant models. While relatively spared inhibitory control might be necessary to safeguard against SIB, it is unlikely to be sufficient given the heterogeneous nature of SIB. Converging research efforts will be required to corroborate several putative risk factors which may differentially contribute to the emergence and persistence of SIB in autism and ID populations.

5.5 Limitations

Delays due to COVID19-related restrictions ultimately limited the number of participants that could be recruited in the time frame. For correlational analyses, an a priori sample size calculation determined that data from 62 participants were required for sufficient power to detect a moderate effect. While informant-based measures were just shy of this target (N = 60), data loss on direct measures resulted in the analyses of those variables being

underpowered. A similar issue arose in regression analyses, where again data loss on direct measures resulted in some analyses falling short of the 55 participants required to fully power the analysis. For a between-group comparison, 79 participants in each group would be required to detect a medium effect with 80% power. Therefore, a number of the pilot analyses were underpowered. Inspection of the 95% confidence intervals of EF coefficients in regression analyses revealed wide intervals, some spanning zero, which also suggests inadequate sample sizes. For all analyses, an alpha level of .05 was chosen. Although this is a less conservative threshold than the .01 recommended when conducting multiple comparisons (Bender & Lange, 2001), the aim of the analyses was to understand the emerging trends in the pilot sample. As such, the increased risk of Type I errors was deemed preferable to failing to identify an effect which might have both clinical and scientific implications. A between-groups analysis in the full sample will allow for a more scientifically robust and sufficiently powered exploration of the independent contribution of response inhibition to the presence and severity of SIB.

Although the EF tasks were designed to minimise verbal demands, correlational analyses revealed consistent moderate associations between EF performance and receptive language abilities, as measured by the VABS and BPVS. The issue of 'task impurity', whereby performance on a task is impacted by factors other than EF ability, is a widespread challenge for EF research. The present study is strengthened by the inclusion of standardised measures of ability, which allowed for this to be controlled for in analyses.

5.6 Clinical Implications

The findings of this study and suggested avenues for future research have important clinical implications. While risk markers are limited in their clinical utility, delineating the behavioural characteristics associated with SIB will ultimately lead to more clinically informative models of SIB development, persistence and vulnerability. Understanding the

behavioural profiles of children presenting with SIB would inform the early identification of children at higher risk of developing SIB. This in turn could improve outcomes for individuals and families through the development of more proactive treatment approaches, which intervene before SIB becomes embedded in a child's behavioural repertoire (Richards et al., 2017). A further important consideration for clinical interventions is the extent to which consequence-based interventions may be appropriate for SIB. If some children are less able to discontinue behaviour once consequences are presented then, consequence alteration will at best be ineffective, but at worst could be considered unethical.

5.7 Summary and conclusions

The study demonstrated the feasibility of administering a novel battery of direct EF measures remotely in a sample of autistic children with a co-occurring ID. The analyses replicated findings in previous research which has consistently reported caregiver-rated impulsivity and overactivity as reliable correlates and predictors of SIB (Laverty et al., 2020; Richards et al. 2012). Whether this association is driven by reduced behavioural inhibition over and above the potential confounds identified here, requires further exploration. The study extends previous research by administering direct measures of EF to this under-represented sample. While no performance data independently predicted differences in reported SIB, the direction and magnitude of effect sizes in this small pilot sample offer some preliminary and tentative support for an association between reduced behavioural inhibition and increased SIB. Overall, the study supports the feasibility and utility of remote EF assessment in this under-represented group.

2.6 References

Alvares, G. A., Bebbington, K., Cleary, D., Evans, K., Glasson, E. J., Maybery, M. T., Pillar, S., Uljarević, M., Varcin, K., Wray, J., & Whitehouse, A. J. (2020). The misnomer of 'high functioning autism': Intelligence is an imprecise predictor of functional abilities at diagnosis. *Autism*, *24*(1), 221–232. https://doi.org/10.1177/1362361319852831

Aman, M., Singh, N. (1994). *The Aberrant Behavior Checklist-Community*. Slosson Education Publications Incorporated.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision* (4th ed.). American Psychiatric Association.

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders. *(5th Edition)*. https://doi.org/10.1176/appi.books.9780890425596

Baghdadli, A., Pascal, C., Grisi, S., & Aussilloux, C. (2003). Risk factors for self-injurious behaviours among 222 young children with autistic disorders. *Journal of Intellectual Disability Research*, *47*(8), 622–627. https://doi.org/10.1046/j.1365-2788.2003.00507.x

Bari, A., & Robbins, T. W. (2013). Inhibition and impulsivity: Behavioral and neural basis of response control. *Progress in Neurobiology*, *108*, 44–79. https://doi.org/10.1016/j.pneurobio.2013.06.005

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

Barnard-Brak, L., Brewer, A., Chesnut, S., Richman, D., & Schaeffer, A. M. (2016). The sensitivity and specificity of the social communication questionnaire for autism spectrum with respect to age. *Autism Research*, *9*(8), 838–845. https://doi.org/10.1002/aur.1584

Beck, D. M., Schaefer, C., Pang, K., & Carlson, S. M. (2011). Executive Function in Preschool Children: Test–Retest Reliability. *Journal of Cognition and Development*, *12*(2), 169–193. https://doi.org/10.1080/15248372.2011.563485

Bender, R., & Lange, S. (2001). Adjusting for multiple testing—when and how? Journal of Clinical Epidemiology, 54(4), 343–349. https://doi.org/10.1016/s0895-4356(00)00314-0

British Psychological Society. (2015). *Guidance on the assessment and diagnosis of Intellectual Disabilities in adulthood*. Leicester: British Psychological Society.

Burbidge, C., & Oliver, C. (2008). Activity Questionnaire. *PsycTESTS Dataset*. https://doi.org/10.1037/t33334-000

Burbidge, C., Oliver, C., Moss, J., Arron, K., Berg, K., Furniss, F., Hill, L., Trusler, K., & Woodcock, K. (2010). The association between repetitive behaviours, impulsivity and hyperactivity in people with intellectual disability. *Journal of Intellectual Disability Research*, *54*(12), 1078–1092. https://doi.org/10.1111/j.1365-2788.2010.01338.x

Camerota, M., Willoughby, M. T., Kuhn, L. J., & Blair, C. B. (2016). The Childhood Executive Functioning Inventory (CHEXI): Factor structure, measurement invariance, and correlates in US preschoolers. *Child Neuropsychology*, *24*(3), 322–337. https://doi.org/10.1080/09297049.2016.1247795

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., Mandell, D. J., & Williams, L. (2004). Executive Function and Theory of Mind: Stability and Prediction From Ages 2 to 3. *Developmental Psychology*, *40*(6), 1105–1122. https://doi.org/10.1037/0012-1649.40.6.1105

Chandler, S., Charman, T., Baird, G., Simonoff, E., Loucas, T., Meldrum, D., Scott, M., & Pickles, A. (2007). Validation of the Social Communication Questionnaire in a Population Cohort of Children With Autism Spectrum Disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*(10), 1324–1332. https://doi.org/10.1097/chi.0b013e31812f7d8d

Cooper, S. A., Smiley, E., Allan, L. M., Jackson, A., Finlayson, J., Mantry, D., & Morrison, J. (2009). Adults with intellectual disabilities: prevalence, incidence and remission of self-injurious behaviour, and related factors. *Journal of Intellectual Disability Research*, *53*(3), 200–216. https://doi.org/10.1111/j.1365-2788.2008.01060.x

Denckla, M. B. (2002). The Behavior Rating Inventory of Executive Function: Commentary. *Child Neuropsychology*, *8*(4), 304–306. https://doi.org/10.1076/chin.8.4.304.13512

Devine, R. T., Ribner, A., & Hughes, C. (2019). Measuring and Predicting Individual Differences in Executive Functions at 14 Months: A Longitudinal Study. *Child Development*, *90*(5). https://doi.org/10.1111/cdev.13217

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

Diamond, A., Cruttenden, L., & Neiderman, D. (1994). AB with multiple wells: I. Why are multiple wells sometimes easier than two wells? II. Memory or memory+inhibition? *Developmental Psychology*, *30*(2), 192–205. https://doi.org/10.1037/0012-1649.30.2.192

Dimian, A. F., & Symons, F. J. (2022). A systematic review of risk for the development and persistence of self-injurious behavior in intellectual and developmental disabilities. *Clinical Psychology Review*, *94*, 102158. https://doi.org/10.1016/j.cpr.2022.102158

Duerden, E. G., Oatley, H. K., Mak-Fan, K. M., McGrath, P. A., Taylor, M. J., Szatmari, P., & Roberts, S. W. (2012). Risk Factors Associated with Self-Injurious Behaviors in Children and Adolescents with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *42*(11), 2460–2470. https://doi.org/10.1007/s10803-012-1497-9

Dunn, L., & Dunn, D. (2009). *British Picture Vocabulary Scale III*. GL Assessment Limited.

Emerson, E., Hatton, C., Dickson, K., Gone, R., Caine, A., & Bromley, J. (2012). *Clinical Psychology and People with Intellectual Disabilities* (2nd ed.). Wiley.

Emerson, E., Kiernan, C., Alborz, A., Reeves, D., Mason, H., Swarbrick, R., Mason, L., & Hatton, C. (2001). The prevalence of challenging behaviors: a total population study. *Research in Developmental Disabilities*, *22*(1), 77–93. https://doi.org/10.1016/s0891-4222(00)00061-5

Erturk, B., Machalicek, W., & Drew, C. (2017). Self-Injurious Behavior in Children With Developmental Disabilities: A Systematic Review of Behavioral Intervention Literature. *Behavior Modification*, *42*(4), 498–542. https://doi.org/10.1177/0145445517741474

Esbensen, A. J., Seltzer, M. M., Lam, K. S. L., & Bodfish, J. W. (2009). Age-Related Differences in Restricted Repetitive Behaviors in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *39*(1), 57–66. https://doi.org/10.1007/s10803-008-0599-x

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/brm.41.4.1149

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology*, *47*(5), 1410–1430. https://doi.org/10.1037/a0023750

Gillham, J. E., Carter, A. S., Volkmar, F. R., & Sparrow, S. S. (2000). Toward a Developmental Operational Definition of Autism. *Journal of Autism and Developmental Disorders*, *30*(4), 269–278. https://doi.org/10.1023/a:1005571115268

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

Golshan, F., Soltani, A., & Afarinesh, M. R. (2019). The study of executive function domains in children with high-functioning autism. *Learning and Motivation*, *67*, 101578. https://doi.org/10.1016/j.lmot.2019.101578

Heyvaert, M., Saenen, L., Maes, B., & Onghena, P. (2014). Systematic Review of Restraint Interventions for Challenging Behaviour Among Persons with Intellectual Disabilities: Focus on Experiences. *Journal of Applied Research in Intellectual Disabilities*, *28*(2), 61–80. https://doi.org/10.1111/jar.12095

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

Howlin, P., & Karpf, J. (2004). Using the Social Communication Questionnaire to Identify 'Autistic Spectrum' Disorders Associated with Other Genetic Conditions. *Autism*, *8*(2), 175–182. https://doi.org/10.1177/1362361304042721

Hyman, P., Oliver, C., & Hall, S. (2002). Self-Injurious Behavior, Self-Restraint, and Compulsive Behaviors in Cornelia de Lange Syndrome. *American Journal on Mental Retardation*, *107*(2), 146. https://doi.org/10.1352/0895-8017

Kelley, M. E., Jimenez-Gomez, C., Podlesnik, C. A., & Morgan, A. (2018). Evaluation of renewal mitigation of negatively reinforced socially significant operant behavior. *Learning and Motivation*, *63*, 133–141. https://doi.org/10.1016/j.lmot.2018.05.003

Knapp, M., Comas-Herrera, A., Astin, J., Beecham, J., & Pendaries, C. (2005). Intellectual disability, challenging behaviour and cost in care accommodation: what are the links? *Health and Social Care in the Community*, *13*(4), 297–306. https://doi.org/10.1111/j.1365-2524.2005.00539.x

Laverty, C., Oliver, C., Moss, J., Nelson, L., & Richards, C. (2020). Persistence and predictors of self-injurious behaviour in autism: a ten-year prospective cohort study. *Molecular Autism*, *11*(1). https://doi.org/10.1186/s13229-019-0307-z

Lecavalier, L., Leone, S., & Wiltz, J. (2006). The impact of behaviour problems on caregiver stress in young people with autism spectrum disorders. *Journal of Intellectual Disability Research*, *50*(3), 172–183. https://doi.org/10.1111/j.1365-2788.2005.00732.x

Licence, L., Oliver, C., Moss, J., & Richards, C. (2020). Prevalence and Risk-Markers of Self-Harm in Autistic Children and Adults. *Journal of Autism and Developmental Disorders*, *50*(10), 3561–3574. https://doi.org/10.1007/s10803-019-04260-1

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, *56*(6), 466–474. https://doi.org/10.1016/j.jaac.2017.03.013

Lord, C., Rutter, M., & le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*(5), 659–685. https://doi.org/10.1007/bf02172145

MacLean, W. E., Tervo, R. C., Hoch, J., Tervo, M., & Symons, F. J. (2010). Self-Injury among a Community Cohort of Young Children at Risk for Intellectual and

Developmental Disabilities. *The Journal of Pediatrics*, *157*(6), 979–983. https://doi.org/10.1016/j.jpeds.2010.05.052

Mandell, D. S. (2008). Psychiatric Hospitalization Among Children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *38*(6), 1059–1065. https://doi.org/10.1007/s10803-007-0481-2

Matson, J. L., & LoVullo, S. V. (2008). A Review of Behavioral Treatments for Self-Injurious Behaviors of Persons With Autism Spectrum Disorders. *Behavior Modification*, *32*(1), 61–76. https://doi.org/10.1177/0145445507304581

Matson, J. L., & Neal, D. (2009). Psychotropic medication use for challenging behaviors in persons with intellectual disabilities: An overview. *Research in Developmental Disabilities*, *30*(3), 572–586. https://doi.org/10.1016/j.ridd.2008.08.007

McClain, M. B., Golson, M. E., & Murphy, L. E. (2022). Executive functioning skills in early childhood children with autism, intellectual disability, and co-occurring autism and intellectual disability. *Research in Developmental Disabilities*, *122*, 104169. https://doi.org/10.1016/j.ridd.2021.104169

McClintock, K., Hall, S., & Oliver, C. (2003). Risk markers associated with challenging behaviours in people with intellectual disabilities: a meta-analytic study. *Journal of Intellectual Disability Research*, *47*(6), 405–416. https://doi.org/10.1046/j.1365-2788.2003.00517.x

Miller, S. E., & Marcovitch, S. (2015). Examining executive function in the second year of life: Coherence, stability, and relations to joint attention and language. *Developmental Psychology*, *51*(1), 101–114. https://doi.org/10.1037/a0038359

Minshawi, N., Hurwitz, S., Fodstad, J., Biebl, S., Morris, D., & McDougle, C. (2014). The association between self-injurious behaviors and autism spectrum disorders. *Psychology Research and Behavior Management*, 125. https://doi.org/10.2147/prbm.s44635

Nissen, J. M. J. F., & Haveman, M. J. (1997). Mortality and avoidable death in people with severe self-injurious behaviour: results of a Dutch study. *Journal of Intellectual Disability Research*, *41*(3), 252–257. https://doi.org/10.1111/j.1365-2788.1997.tb00704.x

Oliver, C., McClintock, K., Hall, S., Smith, M., Dagnan, D., & Stenfert-Kroese, B. (2003). Assessing the Severity of Challenging Behaviour: Psychometric Properties of the Challenging Behaviour Interview. *Journal of Applied Research in Intellectual Disabilities*, *16*(1), 53–61. https://doi.org/10.1046/j.1468-3148.2003.00145.x

Oliver, C., & Richards, C. (2015). Practitioner Review: Self-injurious behaviour in children with developmental delay. *Journal of Child Psychology and Psychiatry*, *56*(10), 1042–1054. https://doi.org/10.1111/jcpp.12425

Panerai, S., Tasca, D., Ferri, R., Genitori D'Arrigo, V., & Elia, M. (2014). Executive Functions and Adaptive Behaviour in Autism Spectrum Disorders with and without Intellectual Disability. *Psychiatry Journal*, *2014*, 1–11. https://doi.org/10.1155/2014/941809

Pepperdine, C. R., & McCrimmon, A. W. (2017). Test Review: *Vineland Adaptive Behavior Scales, Third Edition (Vineland-3)* by Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. *Canadian Journal of School Psychology*, *33*(2), 157–163. https://doi.org/10.1177/0829573517733845

Perry, A., Flanagan, H. E., Dunn Geier, J., & Freeman, N. L. (2009). Brief Report: The Vineland Adaptive Behavior Scales in Young Children with Autism Spectrum Disorders at Different Cognitive Levels. *Journal of Autism and Developmental Disorders*, *39*(7), 1066–1078. https://doi.org/10.1007/s10803-009-0704-9

Rattaz, C., Michelon, C., & Baghdadli, A. (2015). Symptom severity as a risk factor for self-injurious behaviours in adolescents with autism spectrum disorders. *Journal of Intellectual Disability Research*, *59*(8), 730–741. https://doi.org/10.1111/jir.12177

Richards, C., Davies, L., & Oliver, C. (2017). Predictors of Self-Injurious Behavior and Self-Restraint in Autism Spectrum Disorder: Towards a Hypothesis of Impaired Behavioral Control. *Journal of Autism and Developmental Disorders*, *47*(3), 701–713. https://doi.org/10.1007/s10803-016-3000-5

Richards, C., Moss, J., Nelson, L., & Oliver, C. (2016). Persistence of self-injurious behaviour in autism spectrum disorder over 3 years: a prospective cohort study of risk markers. *Journal of Neurodevelopmental Disorders*, *8*(1). https://doi.org/10.1186/s11689-016-9153-x

Richards, C., Oliver, C., Nelson, L., & Moss, J. (2012). Self-injurious behaviour in individuals with autism spectrum disorder and intellectual disability. *Journal of Intellectual Disability Research*, *56*(5), 476–489. https://doi.org/10.1111/j.1365-2788.2012.01537.x

Richman, D. M., Barnard-Brak, L., Bosch, A., Thompson, S., Grubb, L., & Abby, L. (2013). Predictors of self-injurious behaviour exhibited by individuals with autism spectrum disorder. *Journal of Intellectual Disability Research*, *57*(5), 429–439. https://doi.org/10.1111/j.1365-2788.2012.01628.x

Rojahn, J., Schroeder, S. R., & Hoch, T. A. (2007). *Self-Injurious Behavior in Intellectual Disabilities (Volume 2) (The Assessment and Treatment of Child Psychopathology and Developmental Disabilities, Volume 2)* (1st ed.). Elsevier Science.

Romero-Ayuso, D., Jorquera-Cabrera, S., Segura-Fragoso, A., Toledano-González, A., Rodríguez-Martínez, M. C., & Triviño-Juárez, J. M. (2018). Assessment of Sensory Processing and Executive Functions in Childhood: Development, Reliability, and Validity of the EPYFEI. *Frontiers in Pediatrics*, *6*. https://doi.org/10.3389/fped.2018.00071

Russell, G., Mandy, W., Elliott, D., White, R., Pittwood, T., & Ford, T. (2019). Selection bias on intellectual ability in autism research: a cross-sectional review and meta-analysis. *Molecular Autism*, *10*(1). https://doi.org/10.1186/s13229-019-0260-x

Rutter, M., Bailey, A., & Lord, C. (2003). *SCQ*. Western Psychollgical Services.

Snow, A. V., & Lecavalier, L. (2008). Sensitivity and specificity of the Modified Checklist for Autism in Toddlers and the Social Communication Questionnaire in preschoolers suspected of having pervasive developmental disorders. *Autism*, *12*(6), 627–644. https://doi.org/10.1177/1362361308097116

Soke, G. N., Rosenberg, S. A., Hamman, R. F., Fingerlin, T., Robinson, C., Carpenter, L., Giarelli, E., Lee, L. C., Wiggins, L. D., Durkin, M. S., & DiGuiseppi, C. (2016). Brief Report: Prevalence of Self-injurious Behaviors among Children with Autism Spectrum Disorder—A Population-Based Study. *Journal of Autism and Developmental Disorders*, *46*(11), 3607–3614. https://doi.org/10.1007/s10803-016-2879-1

Sonuga-Barke, E. J. (2002). Psychological heterogeneity in AD/HD—a dual pathway model of behaviour and cognition. *Behavioural Brain Research*, *130*(1–2), 29–36. https://doi.org/10.1016/s0166-4328(01)00432-6

Sparrow, S. S., Cicchetti, D. V., Saulnier, C. A. (2016). *Vineland Adaptive Behavior Scales*. PsychCorp.

Tanwar, M., Lloyd, B., & Julies, P. (2016). Challenging behaviour and learning disabilities: prevention and interventions for children with learning disabilities whose behaviour challenges: NICE guideline 2015. *Archives of Disease in Childhood - Education & Practice Edition*, *102*(1), 24–27. https://doi.org/10.1136/archdischild-2015-309575

Taylor, L., Oliver, C., & Murphy, G. (2011). The Chronicity of Self-Injurious Behaviour: A Long-Term Follow-Up of a Total Population Study. *Journal of Applied Research in Intellectual Disabilities*, *24*(2), 105–117. https://doi.org/10.1111/j.1468-3148.2010.00579.x

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*, *54*(2), 131–143. https://doi.org/10.1111/jcpp.12001

Tsermentseli, S., Tabares, J. F., & Kouklari, E. C. (2018). The role of every-day executive function in social impairment and adaptive skills in Autism Spectrum Disorder with intellectual disability. *Research in Autism Spectrum Disorders*, *53*, 1–6. https://doi.org/10.1016/j.rasd.2018.05.006

VideoLan. (2006). VLC media player. Retrieved from https://www.videolan.org/vlc/index.html

Wahler, R. G., & Fox, J. J. (1981). Setting events in applied behavior analysis: Toward a conceptual and methodological expansion. *Journal of Applied Behavior Analysis*, *14*(3), 327–338. https://doi.org/10.1901/jaba.1981.14-327

Wei, T., Chesnut, S. R., Barnard-Brak, L., & Richman, D. (2015). Psychometric Analysis of the Social Communication Questionnaire Using an Item-Response Theory Framework: Implications for the Use of the Lifetime and Current Forms. *Journal of Psychopathology and Behavioral Assessment*, *37*(3), 469–480. https://doi.org/10.1007/s10862-014-9468-4

World Health Organization. (1992). *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines* (1st ed.). World Health Organization.

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

Chapter 3 : Press release for the meta-analysis

Differences in brain-based skills in preschool-aged autistic children support a role for executive function in autism

The results of a recent review have revealed definitive differences in a set of brain-based abilities, termed 'executive function', in young autistic[9] children. This key finding lends support to the much-debated theory that executive function differences could underpin some of the characteristics and experiences of autistic people (Pellicano, 2010). Importantly, the findings highlight how critical it is for clinicians and educators to support the development of these skills in young autistic children.

Executive function (EF) is an umbrella term which describes a set of brain-based skills that support us to resist temptation, respond flexibly to changing contexts, and to hold and manipulate information in mind (Diamond, 2013). Essentially, these skills coordinate our abilities and actions, allowing us to act effectively in response to real-world challenges.

Children aren't born with EF skills, but instead these develop rapidly in early childhood (Hughes, 1998), and some children might have more difficulties developing or learning these skills than others. Importantly, research has shown that the effectiveness of these skills in early childhood predicts later academic, social, and occupational success (Kenny et al., 2019). Early childhood is, therefore, a critical time to intervene and 'scaffold' the development of these skills.

Children with neurodevelopmental conditions like autism have been shown to struggle more with tasks designed to test EF (Benallie et al., 2021). This finding has led researchers to

[9] The term autistic, rather than 'person with autism', is used to reflect the preferences for identity-first rather than person-first language, reported by autistic individuals and advocates in the UK (Kenny et al., 2016).

wonder whether certain differences in EF may explain some of the characteristics of autism, such as repetitive behaviours and special interests (APA, 2013). Although differences in EF are commonly found in school-aged children and adults (Demetriou et al., 2018), investigators have reported conflicting results when testing autistic preschoolers. If early differences in EF do have a role in the development of autism characteristics, then these should be apparent in very young children too. But where some research teams have found similar EF skills to neurotypical children, others have shown that there are EF differences in autistic children as young as two (Campbell et al., 2019).

This new meta-analysis resolves the conflicting findings in younger children by combining and analysing the data from all available studies which compared the EF skills of autistic toddlers and preschoolers to neurotypical peers. The results revealed that, just like in school-aged children and adults, there are indeed true differences in all EF skills between these groups of children from a young age, and young autistic children scored lower than neurotypical children across a range of EF tasks and questionnaires. The analysis also showed that these differences could not be explained by more general differences in intelligence. The inconsistency in previous research in this age group was likely due to differences in the methods that research teams had used to measure EF and select neurotypical children as comparisons.

Although all children, with and without autism, will vary in their EF abilities, the results emphasise that autistic children are at higher risk of the potentially disabling consequences of EF difficulties, and so are likely to need more support to develop these skills than neurotypical children. Routinely assessing EF in early childhood would help identify these children and enable services to intervene early, to promote better outcomes for children and their families. Other research efforts are currently evaluating the potential treatments which might promote the successful attainment of EF skills in autistic children, with some

promising results (Kenworthy et al., 2014). The author emphasises the importance of healthcare services and educational institutions in providing opportunities for very young children to develop effective EF skills, which are so critical to their future success and wellbeing.

3.1 References

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th Edition). https://doi.org/10.1176/appi.books.9780890425596

Benallie, K. J., McClain, M. B., Bakner, K. E., Roanhorse, T., & Ha, J. (2021). Executive functioning in children with ASD + ADHD and ASD + ID: A systematic review. *Research in Autism Spectrum Disorders, 86*, 101807. https://doi.org/10.1016/j.rasd.2021.101807

Campbell, S. B., Northrup, J. B., & Tavares, A. B. (2019). Resistance to temptation in toddlers at genetic risk for autism spectrum disorder. *Autism, 23(4)*, 1018–1027. https://doi.org/10.1177/1362361318797264

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

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

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology, 16(2)*, 233–253. https://doi.org/10.1111/j.2044-835x.1998.tb00921.x

Kenny, L., Cribb, S. J., & Pellicano, E. (2019). Childhood Executive Function Predicts Later Autistic Features and Adaptive Behavior in Young Autistic People: a 12-Year Prospective Study. *Journal of Abnormal Child Psychology, 47(6)*, 1089–1099. https://doi.org/10.1007/s10802-018-0493-8

Kenworthy, L., Anthony, L. G., Naiman, D. Q., Cannon, L., Wills, M. C., Luong-Tran, C., Werner, M. A., Alexander, K. C., Strang, J., Bal, E., Sokoloff, J. L., & Wallace, G. L. (2014). Randomized controlled effectiveness trial of executive function intervention for children on the autism spectrum. *Journal of Child Psychology and Psychiatry, 55(4)*, 374–383. https://doi.org/10.1111/jcpp.12161

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

Chapter 4 : Press release for the empirical paper

A novel online research study reports a strong link between impulsivity and self-injury in autistic children with a learning disability

Early results from a study investigating the individual characteristics associated with self-injurious behaviour (SIB) in autistic children with a learning disability, have shown a strong relationship between SIB and impulsivity. The findings in this relatively small pilot sample points to a role for poorer behavioural inhibition, a so-called 'executive function' skill, which might make it harder for some children to prevent or stop behaviour once started. The outcome of this study will help to extend current understandings of why SIB is seen in some children and not others. This in turn will improve outcomes for children and their families by informing new treatment strategies, which can intervene before this worrying behaviour has become a well-rehearsed habit.

Many young children will show some form of repetitive SIB, such as head banging, biting, and skin scratching. While most will grow out of this, ongoing and repeated SIB is seen in as many as 1 in 2 autistic children with a learning disability (Richards et al., 2012). Once established, SIB can be very difficult to treat and often continues past childhood (Laverty et al., 2020). The consequences of SIB can be devastating for children and their families and can significantly reduce quality of life (Lecavalier et al., 2006).

Despite the high occurrence of SIB in this group, autistic people with a co-occurring learning disability are often not included in research (Russell et al., 2019). The study has reversed this position and built on the existing research showing a link between parent-reported impulsivity and SIB, by directly assessing the executive function skills of this group of at-risk children. 'Executive Function' is an umbrella term which describes a set of brain-based skills, including inhibition, that coordinate our abilities and actions, allowing us to act

effectively in response to real-world challenges (Diamond, 2013). Although difficulties with tasks that require executive function skills are commonly reported in autistic individuals (Demetriou et al., 2018), these skills will vary from person to person.

Due to the COVID-19 pandemic, in-person testing was not possible, and the investigators re-designed the study to be conducted entirely remotely, through questionnaires and an assessment through 'Zoom'. On the video call, parents were guided to deliver the play-based executive function tasks to their children, many of whom showed current SIB. The findings in the first 60 children revealed that those with more frequent and severe SIB also show higher levels of impulsivity and had more difficulties resisting temptation and stopping behaviour once started.

Like other behaviours, SIB can be shaped through subtle learning processes. For example, if ending a preferred activity coincides with a child showing SIB, a parent might decide to resume this activity to prevent the child from coming to any harm. Over time, an association might form between showing SIB and having access to preferred activities, making this more likely to happen when preferred activities are not available. Treatments for SIB aim to understand whether this serves a particular function, reduce potential triggers, and teach children new skills to use instead. Although this can reduce SIB for some children, this type of treatment can only start once a child has started to show SIB, leaving families vulnerable to a whole host of negative consequences.

A learning account of SIB also doesn't fully explain why some children show SIB and others don't. Understanding the individual characteristics of children who show SIB is key to understanding which children are most at-risk. The preliminary results from this ongoing study suggest that behavioural inhibition is one such characteristic that the children with and without SIB seem to differ on. Children's scores on questionnaires and the play-based tasks

showed that for those with SIB, being less able to “brake and disengage” might make it harder for these children to refrain from SIB or to stop this behaviour once started. Although this was a novel approach, the results suggest that testing this group, even online, can work well and investigators are hopeful that this will encourage other researchers to include this under-represented group in future studies.

4.1 References

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

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

Laverty, C., Oliver, C., Moss, J., Nelson, L., & Richards, C. (2020). Persistence and predictors of self-injurious behaviour in autism: a ten-year prospective cohort study. *Molecular Autism*, *11*(1). https://doi.org/10.1186/s13229-019-0307-z

Lecavalier, L., Leone, S., & Wiltz, J. (2006). The impact of behaviour problems on caregiver stress in young people with autism spectrum disorders. *Journal of Intellectual Disability Research*, *50*(3), 172–183. https://doi.org/10.1111/j.1365-2788.2005.00732.x

Richards, C., Oliver, C., Nelson, L., & Moss, J. (2012). Self-injurious behaviour in individuals with autism spectrum disorder and intellectual disability. *Journal of Intellectual Disability Research*, *56*(5), 476–489. https://doi.org/10.1111/j.1365-2788.2012.01537.x

Russell, G., Mandy, W., Elliott, D., White, R., Pittwood, T., & Ford, T. (2019). Selection bias on intellectual ability in autism research: a cross-sectional review and meta-analysis. *Molecular Autism*, *10*(1). https://doi.org/10.1186/s13229-019-0260-x

APPENDICES

1.1 Appendix A: Correspondence granting full ethical approval.

Ymchwil Iechyd a Gofal Cymru
Health and Care Research Wales

NHS
Health Research Authority

Dr Caroline Richards
52 Pritchatts Road, School of Psychology
University of Birmingham, Edgbaston
Birmingham
B15 2TT

Email: hra.approval@nhs.net

19 August 2019

Dear Dr Richards

HRA and Health and Care Research Wales (HCRW) Approval Letter

Study title:	**Sleep-Impulsivity-Behaviour (SIB) Study: examining pathways to self-harm in children with autism and intellectual disability**
IRAS project ID:	**263048**
Protocol number:	**RG_19-087**
REC reference:	**19/SW/0141**
Sponsor	**University of Birmingham**

I am pleased to confirm that **HRA and Health and Care Research Wales (HCRW) Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications received. You should not expect to receive anything further relating to this application.

Please now work with participating NHS organisations to confirm capacity and capability, in line with the instructions provided in the "Information to support study set up" section towards the end of this letter.

How should I work with participating NHS/HSC organisations in Northern Ireland and Scotland?
HRA and HCRW Approval does not apply to NHS/HSC organisations within Northern Ireland and Scotland.

If you indicated in your IRAS form that you do have participating organisations in either of these devolved administrations, the final document set and the study wide governance report (including this letter) have been sent to the coordinating centre of each participating nation. The relevant national coordinating function/s will contact you as appropriate.

Please see IRAS Help for information on working with NHS/HSC organisations in Northern Ireland and Scotland.

How should I work with participating non-NHS organisations?
HRA and HCRW Approval does not apply to non-NHS organisations. You should work with your non-NHS organisations to obtain local agreement in accordance with their procedures.

What are my notification responsibilities during the study?

The document "*After Ethical Review – guidance for sponsors and investigators*", issued with your REC favourable opinion, gives detailed guidance on reporting expectations for studies, including:

- Registration of research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics, and is updated in the light of changes in reporting expectations or procedures.

Who should I contact for further information?
Please do not hesitate to contact me for assistance with this application. My contact details are below.

Your IRAS project ID is **263048**. Please quote this on all correspondence.

Yours sincerely,
Gemma Oakes

Approvals Specialist

Email: hra.approval@nhs.net

Copy to: *Dr Birgit Whitman*

1.2 Appendix B: Sleep-Impulsivity-Behaviour study recruitment flyer

1.3 Appendix C: Information sheet for prospective participants

Participant Information Sheet

We would like to invite your child/person you care for to take part in a research study being conducted in the School of Psychology at the University of Birmingham in collaboration with [insert name of school or NHS Trust].

Please read this information carefully before deciding whether it is in the best interests of your child to take part in the study. If you would like any more information about the study or you have any health and/or language difficulties which make it difficult for you to read this information please contact the research team, Dr Caroline Richards on [redacted] or [redacted] or Dr Georgie Agar on [redacted] or [redacted] or you can ask someone to contact the research team on your behalf.

When you are happy that you have all of the information you need to be able to decide whether or not the person you care for would like to take part in the study, please complete the consent form and post it to the research team at the earliest opportunity in the pre-paid envelope provided.

What is the research study about and why is it important?
This research project, led by Dr Caroline Richards, is based in the School of Psychology at the University of Birmingham. The project aims to better understand behaviour in children with autism and intellectual disability. In the future, the information obtained during this study could help us to identify ways in which we may be able to prevent the onset of particular behaviours that have a negative impact on the quality of life of children with autism and intellectual disability.

This research project is split into three separate studies. We are contacting you about taking part in Study 1 of the project. If your child takes part in Study 1 and you would like to be contacted about participation in future parts of the study, we may contact you a few weeks later with information about taking part in Study 2. This is optional and participation in Study 1 does not mean that your child needs to take part in future studies.

Why have I been invited to take part in the study?
You have been contacted about this study because you are the parent/carer of a child aged between 4 years and 15 years 9 months old who has autism and intellectual disability. You have also expressed an interest in your child taking part in research.

If I decide to participate in the research study, what will I be asked to do?

If you would like your child to participate in the research, you will be asked to sign a consent form and provide us with a few personal details (your name, postal address, telephone number, email address). You can return this to the research team using the pre-paid envelope provided. Alternatively, if you have been given this document by a health professional you can complete the consent form with them, who will pass this onto the research team at the University of Birmingham.

After you have consented to take part in the study, you will be asked to:

- Complete some questionnaires online
- Complete an interview about your child's adaptive functioning via video conferencing
- Take part in some fun play-based games and activities via video conferencing with your child and the research team

We will send a package containing all of toys for you to use in the play session. We will also send instructions for how to play the games and will guide you through each step using in-ear headphones. You can keep the toys as a gift for you and your child. We will thoroughly disinfect all study materials and prepare your package while wearing gloves. We will follow the University of Birmingham's approved risk assessments for the safe preparation and transportation of testing materials at all times.

You may be invited by the research team to complete this study again, so that we can assess how your child's behaviour may have changed over time. There may be the opportunity to take part in other aspects of the study at a later date. This could involve the research team carrying out some more study activities with your child face to face, either in your home, the child's school or at the University of Birmingham. Tasks will all be play-based games and activities, which will be engaging for your child. We will use picture cards, puzzles and role-playing activities to learn about your child's behaviour. We will only invite your child to take part in this part of the study if it is safe for the research team to visit your child. We will continue to follow government guidelines on social distancing at all times.

Where will the study activities take place?
You can complete the questionnaires at your convenience and post these back to the research team at the University of Birmingham. We can also complete the interview via video conferencing at a time that is convenient for you. We will then arrange a time to have a play session with you and your child at your home to carry out the study activities. Alternatively, if you prefer, we can ask your child's school to allow us to conduct the play session there with your child, with a member of staff as the 'co-researcher'.

Who will be involved in collecting the data?
Only suitably trained and experienced researchers will be involved in data collection.

How long will participation in the study take?
Completing the questionnaire and interview is expected to take approximately 1.5 hours of your time. The play-based assessments with you and your child are estimated to take one hour in total. Some children may be able to complete all assessments in one session, but others will require frequent breaks between activities meaning that assessments could be spread over two sessions to suit your child.

If you are invited to complete the assessments again in the future to help us understand how your child's behaviour may have changed, we will contact you at a later date depending on government guidelines. We will likely contact you around 12 months after completing this part of the study and activities will take place again over 1-2 days.

Will assessments/interviews be recorded?
During some of the assessments, your child's behaviour and the behaviour of people in your child's immediate surroundings will be recorded using the webcam of your laptop or tablet device so that we can score the assessments, and check the accuracy of the observations with another researcher. These videos will be captured via Zoom and the recording will be stored on the University's secure Research Data Store.

Are there any risks for individuals taking part in the study?
The questionnaire and interview that you will be asked to complete include questions about your child's behaviour that you may find difficult to talk about. If you feel uncomfortable answering any questions you can discuss this with the research team at any stage of participation. We may be able to support you or signpost you to helpful resources or professionals.

Some children taking part in the study may display challenging behaviours (e.g. aggression, destruction of property) and self-injurious behaviour (e.g. head-banging, biting self, eye poking). These behaviours may occur in the presence of the researcher or when they are interacting with you during the study activities. If your child does display any challenging behaviours during the play session the research team will discuss with you how to proceed or if activities should cease. Before the play session, we will ask you about any challenging behaviours that your child shows and, if appropriate, create a risk-management plan to keep your child and those around them safe during the play session.

What are the potential benefits for individuals taking part in the study?
Following your child's participation in the study, you will receive a personalised feedback report describing the results of the assessments conducted. This report may be useful to you and the health and education professionals involved with your child to highlight strengths and difficulties that your child experiences and identify resources that may be useful for them. We will happily share a copy of this report with any health/education professionals if you request this and provide us with written consent.

If I would like to participate in the project, what should I do now?
Please remember that participation in the project is purely optional and the decision not to participate will not restrict access or affect the right to any education/health services. When you are satisfied that you have all the information you need to be able to decide whether or not you would like your child to take part in the study and if you decide that you would like your child to participate, please complete the consent form and return it to us in the prepaid envelope provided.

You can then complete the questionnaire pack and return this to us using a second prepaid envelope that has been provided. <u>It is important that you return the consent forms and questionnaire pack separately</u> so that your personal information and data cannot be linked in

the unlikely event that it is lost in the post. This is for your safety and to ensure your data is kept confidential.

What if I change my mind about participating after I have provided consent?
Even after you have provided consent for your child's participation in the study, you can request for your child to be withdrawn from the study. This will not restrict access or affect the right to any education/health services. With your permission, personal and research data collected up until the point of withdrawal will be retained by the research team.

If I decide to participate, what will happen to the information I provide?
In order to carry out the research project described above, we will need to collect information about your child, and some of this information will be personal data. Under data protection law, we have to provide you with very specific information about what we do with your data and about your rights. We have set out below the key information you need to know about how we will use your personal data.

More information on how the University processes personal data can be found on the University's website on the page called 'Data Protection - How the University Uses Your Data' (https://www.birmingham.ac.uk/privacy/index.aspx).

Who is the Data Controller?
The University of Birmingham, Edgbaston, Birmingham B15 2TT is the data controller for the personal data that we process in relation to you.

What data are we processing and for what purpose will we use it?
We will collect and process your personal data to conduct the research project, as explained in the Participant Information Sheet.

What is our legal basis for processing your data?
The legal justification we have under data protection law for processing your personal data is that it is necessary for our research, which is a task we carry out in the public interest.

Who will my personal data be shared with?
We will not share your child's data with any third party without your written informed consent.
Sometimes, external organisations assist us with processing your information, for example, in providing IT support. These organisations act on our behalf in accordance with our instructions and do not process your data for any purpose over and above what we have asked them to do. We make sure we have appropriate contracts in place with them to protect and safeguard your data. If your personal data are transferred outside the European Union (for example, if one of our partners is based outside the EU or we use a cloud-based app with servers based outside the EU), we make sure that appropriate safeguards are in place to ensure the confidentiality and security of your personal data.

How will my personal data be kept secure?
The University takes great care to ensure that personal data is handled, stored and disposed of confidentially and securely. Our staff receive regular data protection training, and the University has put in place organisational and technical measures so that personal data is processed in accordance with the data protection principles set out in data protection law.

The University has an Information Security Management System based on ISO27001 with a range of controls covering the protection of personal information. Annual security awareness training is mandatory for staff and the University is accredited under the NHS Information Governance Toolkit, the Payment Card Industry Data Security Standard and is in the process of gaining Cyber Essentials Plus for defined services.

While government regulations are in place advising workers to continue their work remotely from home, the research team will operate as such. During this period, physical data will be stored in a secure lockbox at an approved researcher's home. Electronic data will be password protected and stored on an encrypted online server held by the University of Birmingham (Research Data Store). Only approved members of the research team will have access to this data. When such restrictions end, physical data will be transferred in the lockbox to the University of Birmingham and be stored in a locked filing cabinet in a lockable office. Once government regulations are relaxed normal data storage procedures will ensue, as described below.

In relation to this project, personal identifying information will be stored for the duration of the study on a password protected portable hard drive that is kept in a locked filing cabinet in a lockable office at the University of Birmingham and on the University's secure server Your completed physical data (e.g. questionnaires, scoring sheets) will be stored separately from the personal identifying information. Yours or your child's name will not appear on the physical data. Instead, each participant in the study will be allocated a participant number and this will appear on the data instead of names. A password protected electronic file will be created which links the participant number to the participant name, which will be stored on a password protected portable hard drive and kept in a locked filing cabinet in a lockable office at the University of Birmingham.

The video recordings created during the study are considered to necessarily contain personal identifying information. We will therefore store the recordings of you and your child separately to the other information we have collected about you. These recordings will be labelled with participant numbers and stored on a password protected portable hard drive kept in a locked filing cabinet in a lockable office at the University of Birmingham and on the University's secure server

What will happen to my personal information after the project ends?
If you are not already on an existing participant database held by the researchers involved in this study and have not consented to be added to the database, your contact details will be destroyed at the end of the study and you will not be contacted further. This means that we would no longer be able to trace the results of your child's assessments back to you. If you withdraw from the project, we will keep the information we have already obtained but, to safeguard your rights, we will use the minimum personally-identifiable information possible. The document titled 'Regular Participant Database Information Sheet' gives information about the database that we use to store the personal details of some participants. Please read this document so you can decide if you would like to join the database.

What will happen to the data afterwards?

You will be able to decide whether or not you want to make your child's research data available to any professionals or clinicians working with the person you care for should they wish to see it. Ten years after the end of the study we will destroy all of the data collected during the course of the study. This includes video footage unless you have given us your consent for the recordings to be used for teaching and/or dissemination.

The research team will publish the findings from the study in scientific journals and will present the results at relevant conferences and in newsletters. Any published reports which use the information you have provided us with will be completely confidential and will never use yours or your child's name. Anonymous datasets (with all personal information removed) will be made available to users on an open science platform at the discretion of Dr Caroline Richards. This data will be unidentifiable. Audio/Video recordings will not be shared.

What are my rights in relation to my data?
You may have the following rights in respect of your personal data:

- The right to access to your data (often referred to as a Subject Access Request).
- The right to rectification of inaccuracies in your data.
- The right to erasure of your data (in certain circumstances).
- The right to restrict processing of your data (in certain circumstances).
- The right to object to the processing of your data (in certain circumstances).
- The right to ask for your personal data to be transferred electronically to a third party.
- If the research is being done on the legal basis of your consent, the right to withdraw consent.

However, your rights to access, change or move your information are limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate. If you withdraw from the project, we will keep the information we have already obtained but, to safeguard your rights, we will use the minimum personally-identifiable information possible.

If you would like more information on your rights, would like to exercise any right or have any queries relating to our processing of your personal data, please contact:

The Information Compliance Manager, Legal Services, The University of Birmingham, Edgbaston, Birmingham B15 2TT
Email: dataprotection@contacts.bham.ac.uk Telephone: +44 (0)121 414 3916

If you wish to make a complaint about how your data is being or has been processed, please contact our Data Protection Officer.

Mrs Carolyn Pike, OBE, The Data Protection Officer, Legal Services, The University of Birmingham, Edgbaston, Birmingham B15 2TT

Email: dataprotection@contacts.bham.ac.uk Telephone: +44 (0)121 414 3916

You also have a right to complain to the Information Commissioner's Office (ICO) about the way in which we process your personal data. You can make a complaint using the ICO's website.

Confidentiality
Your confidentiality and the confidentiality of your child will be ensured. If published, any information we collect will be presented without reference to names or any other identifying

information.

In the unlikely event that the research team becomes concerned for your welfare or the welfare of your child, this information will be disclosed to the appropriate authorities to safeguard those who may be at risk. The Chief Investigator, Dr Caroline Richards, will take appropriate action by contacting the University of Birmingham safeguarding team or the relevant authorities.

What can I do if I have any concerns about the research or there is a problem?
If you have any cause for concern about any aspect of the research, in the first instance, please contact Dr Caroline Richards (Chief Investigator) on [redacted] or Dr Georgie Agar (Research Fellow) on [redacted] or [redacted] If you have any concerns related to the research after contacting the Chief Investigator or Research Fellow you can contact the Head of the School of Psychology at the University of Birmingham, Professor Ed Wilding on [redacted]

How do I contact the research team to find out more about the research study?
Please contact Dr Georgie Agar on [redacted] or [redacted] Alternatively, you can post to the following address:

Richards Lab
52 Pritchatts Road
School of Psychology
University of Birmingham
Edgbaston
Birmingham
B15 2TT

Consent
After having read all of the information and having received appropriate responses to any questions that you may have about the study, you will be asked to give your and your child's consent to participate in the study if you decide that you do wish to participate. A consent form has been enclosed for you to complete if you do wish to participate. Please complete the form and post it to the research team in the pre-paid envelope provided at your earliest convenience.

Although you are being asked to consent on behalf of your child, we still seek to take their wishes into account. If at any point during the study your child appears to be overly distressed by the study activities e.g. video recording, then we may discuss withdrawing them from the study. Both parents/carers and members of the research team are able to withdraw a participant from the study if they feel that participation is not in their best interests.

Consent to be included in the 'Regular Participant Database' is separate to consent for the study and you are welcome to participate in the study without being included in the database.

Review

The research project has been approved by the South West Central Bristol Research Ethics Committee. The study reference number is 19/SW/0141.

1.4 Appendix D: Guardian consent form

Site name: _____ Principal Investigator: _____

If you are reading this information on behalf of someone you care for who is a child under the age of 16 years, then we would like to ask you to decide whether or not you think that it is in your child's / person you care for's best interests for them to participate in the study. If you would like your child / person you care for to participate in this study, please complete this consent form. **Please return the consent form to the research team in the prepaid envelope provided.**

Please initial box...

1. I confirm that I have read and understood the information sheet dated [insert date and version number of Study 1 information sheet] for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that the participation of my child / person I care for is voluntary and that we are free to withdraw at any time without giving any reason, without my or that of my child's / person I care for's medical care or legal rights being affected. I agree that data collected prior to withdrawal can still be used in the research analysis. ☐

3. I understand that all information collected during the study will be confidential. Participant names will not be published in any research reports. Information will be treated as strictly confidential and handled in accordance with the provisions of the General Data Protection Regulation 2018 (GDPR) and the Data Protection Act 2018. ☐

4. I understand that as part of the above study, a video recording will be made of me, my child and members of my family who are present at the time of the recording may also be captured on the video and/or audio. This footage will be captured via Zoom and the recording will be stored on the University's secure Research Data Store. I consent for this recording to be made and stored for further review. ☐

5. I understand that the University of Birmingham will hold the copyright of any video/voice recordings collected during the study but that this does not entitle the University of Birmingham to edit, copy or use the videos for teaching purposes without my written permission. ☐

6. I understand that relevant sections of my child's / person I care for medical notes and data collected during the study may be looked at by individuals from regulatory authorities, Sponsors and/or NHS bodies, where it is relevant to my child / person I care for taking part in this research. I give permission for these individuals to have access to my records. ☐

7. I understand that anonymised data collected during the study will be made available on an open science platform and that it will not be possible to identify any participant in the study from this data. I give permission for my child's anonymised data to be made available. ☐

8. I agree to my child's GP being informed of my child's participation in this study. ☐

OFFICE USE To be removed following return of consent form. **ID Number:**

9. I agree to the University of Birmingham research team sharing my research data with any health professionals or clinicians working with the person I care for. ☐

10. I agree that my child can take part in the above study. ☐

11. I am willing to administer the tasks as required in the above study. ☐

Optional clauses: The statements below are optional. Please delete Yes/No as appropriate:

1. I agree to being contacted with information about participating in future parts of the above study using the contact details I provide. Yes/No

2. I agree for my contact information to be kept so that study newsletters and feedback can be sent to me. Yes/No

3. I agree to the University of Birmingham research team using images and video recordings taken as part of the study for non-research purposes e.g. teaching, presentations, websites Yes/No

Please complete the details below

Your Name: _____

Address: _____

_____

Telephone Number 1: _____ Telephone Number 2: _____

Email: _____

Full Name of Your Child / Person You Care for: _____

Relationship to Participant: _____

Signature: _____ Date: _____

To be completed at the play session:

Researcher responsible for confirming consent: _____

Signature: _____ Date: _____

Completed Consent Form Filing
Original completed consent forms will be filed in a secure and safe location at the University of Birmingham. One copy will also be stored in the Investigator Site File and one copy will be given to the parent/carer of the participant.

OFFICE USE To be removed following return of consent form. **ID Number:**

1.5 Appendix E: Questionnaire measures

ID number:
Date:

Background Questionnaire v1.0 05.04.19
SIB Study
IRAS ID: 263048

Background Questionnaire

Please tick or write your response to these questions concerning background details of the person you care for:

1. **Today's date:** _____
2. **Gender:** Male ☐ Female ☐
3. **Date of Birth:** ___/___/____ **Age:**_____
4. **Is the person you care for verbal?** (i.e. more than 30 signs/words in their vocabulary)

   Yes/No (delete as appropriate)

5. **Is the person you care for able to walk unaided?**

   Yes/No (delete as appropriate)

6. **Has the person you care for been diagnosed with autism?** Yes/No (delete as appropriate)

   Please indicate which diagnosis the person you care for has been given:

Autism	☐	Autism Spectrum Disorder	☐
Asperger Syndrome	☐	High Functioning Autism	☐
Autistic Features	☐	Autistic (like) traits	☐
Autistic Continuum	☐	Pervasive Developmental Disorder	☐
Atypical Autism	☐	Autistic Spectrum	☐

7. **Has the person you care for been diagnosed with intellectual disability, learning disability or global developmental delay?** Yes/No (delete as appropriate)

   Please indicate the level of intellectual disability/learning disability of the person you care for:

Mild	☐	Moderate	☐
Severe	☐	Profound	☐
Unknown	☐		

   Other _____

8. **When was the person you care for diagnosed?** _____

9. Who diagnosed the person you care for?

Paediatrician ☐ Clinical Geneticist ☐
GP ☐

Other _____

10. Has the person you care for ever been diagnosed with a biological sleep disorder? Yes/No (delete as appropriate)

If yes, please indicate below which diagnosis the person you care for has been given. If no, please move on to question 11.

Sleep-related breathing disorder (e.g. obstructive sleep apnea syndromes, central sleep apnea syndromes, sleep-related hypoventilation or hypoxemia disorders) ☐
Periodic limb movement disorder or restless leg syndrome ☐
Narcolepsy ☐

Other _____

11. Does the person you care for take any medication relating to sleep. If yes, please give details:

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

12. Has the person you care for ever been diagnosed with epilepsy? Yes/No (delete as appropriate)

If yes, please answer questions 13 to 16b. If no, please move on to question 17.

13. When was the person you care for diagnosed with epilepsy? _____

14. If known, please indicate the cause of the person you care for's epilepsy

_____

15. When did the person you care for last have a seizure? _____

_____

16. Is the person you care for currently taking medication to control their epilepsy?
Yes/No (delete as appropriate)

16a. If no, has the person you care for ever taken medication to control their epilepsy?
Yes/No (delete as appropriate)

16b. If yes to either of these, please give details below and indicate whether the person you care for is currently taking the medication or has stopped and for what reason.

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

17. Has the person you care for had any medical/health difficulties in the last six months? If yes, please give details:

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

18. Does the person you care for take any medication relating to behaviour? If yes, please give details:

In the information sheet and consent form we informed you that we may wish to contact your child's/person you care for's GP (or alternative health professional e.g. Paediatrician) in order to communicate to them information regarding your child's health and diagnostic status (see consent form and information sheet for more information). If you have already indicated on the consent form that you are happy for us to do this, please complete the relevant details below:

19. Name of the person you care for's GP_____

GP Address_____

GP Telephone number_____

ID number:
Date:

CBQ v1.0 20.04.20
SIB Study
IRAS ID: 263048

Challenging Behaviour Questionnaire

Self-injurious behaviour

A) Has the person **ever** shown self-injurious behaviour? (e.g. head banging, head-punching or slapping, removing hair, self-scratching, body hitting, eye poking or pressing).

Yes ☐ No ☐

1) Has the person shown self-injurious behaviour **in the last month**? (e.g. head banging, head-punching or slapping, removing hair, self-scratching, body hitting, eye poking or pressing).

Yes ☐ No ☐

If the behaviour has not occurred, please go to question B.
If the behaviour has occurred during the past month, please answer questions 1.1 to 1.4.:

1.1) Place a tick next to the items in the following list of behaviours, which the person displays in a repetitive manner (repeats the same movement/ behaviour twice or more in succession):

Hits self with body part (e.g. slaps head or face).....	☐
Hits self against surface or object (e.g. bangs head on floor or table).....	☐
Hits self with object.....	☐
Bites self (e.g. bites hand on wrist or arm).....	☐
Pulls (e.g. pulls hair or skin).....	☐
Rubs or scratches self (e.g. rub marks on arm or leg).....	☐
Inserts finger or objects (e.g. eye poking).....	☐
Other form of self-injury, please specify:.....	☐

1.2) In the last month, for how long did the **longest** episode or burst of this behaviour last? *(Please circle one number)*

1	2	3	4	5
Less than a minute	Less than 5 minutes	Less than 15 minutes	Less than an hour	More than an hour

1.3) In the last month, as a result of this behaviour, has physical contact, prevention or restraint by others been necessary e.g. blocking, taking objects from an individual, temporary restraint of an arm? *(Please circle one number)*

0	1	2	3	4
Never	At least once a month	At least once a week	At least once a day	At least once an hour

1.4) Think about how often this behaviour occurred in the last month. If there was no change and you watched the person now, then would you definitely see the behaviour:

1	2	3	4	5
By this time next month	By this time next week	By this time tomorrow	In the next hour	In the next 15 minutes

Total SIB Score:

ID number:
Date:

SCQ
SIB Study
IRAS ID: 263048

Social Communication Questionnaire (Lifetime)

Please circle 'yes' if <u>any</u> one of the following behaviours is present. Although you may be uncertain about whether some behaviours are present or not, please do answer 'yes' or 'no' to every question on the basis of what you think.

1. Is she/he now able to talk using short phrases or sentences? If no, skip to question 8.	**Yes** **No**
2. Can you have a to and fro "conversation" with her/him that involves taking turns or building on what you have said?	**Yes** **No**
3. Has she/he ever used odd phrases or said the same thing over and over in almost exactly the same way (either phrases that she/he has heard other people use or ones that she/he has made up?	**Yes** **No**
4. Has she/he ever used socially inappropriate questions or statements? For example, has she/he ever regularly asked personal questions or made personal comments at awkward times?	**Yes** **No**
5. Has she/he ever got her/his pronouns mixed up (e.g., saying you or she/he for I)?	**Yes** **No**
6. Has she/he ever used words that she/he seemed to have invented or made up her/himself; put things in odd, indirect ways; or used metaphorical ways of saying things (e.g., saying hot rain for steam)?	**Yes** **No**
7. Has she/he ever said the same thing over and over in exactly the same way or insisted that you say the same thing over and over again?	**Yes** **No**
8. Has she/he ever had things that she/he seemed to have to do in a very particular way or order or rituals that she/he insisted that you go through?	**Yes** **No**
9. Has her/his facial expression usually seemed appropriate to the particular situation, as far as you could tell?	**Yes** **No**
10. Has she/he ever used your hand like a tool or as if it were part of her/his own body (e.g., pointing with your finger, putting your hand on a doorknob to get you to open the door)?	**Yes** **No**
11. Has she/he ever had any interests that preoccupy her/him and might seem odd to other people (e.g., traffic lights, drainpipes, or timetables)?	**Yes** **No**
12. Has she/he ever seemed to be more interested in parts of a toy or an object (e.g., spinning the wheels of a car), rather than using the object as it was intended?	**Yes** **No**
13. Has she/he ever had any special interests that were unusual in their intensity but otherwise appropriate for her/his age and peer group (e.g., trains, dinosaurs)?	**Yes** **No**

14. Has she/he ever seemed to be unusually interested in the sight, feel, sound, taste, or smell of things or people?	**Yes** **No**
15. Has she/he ever had any mannerisms or odd ways of moving her/his hands or fingers, such as flapping or moving her/his fingers in front of her/his eyes?	**Yes** **No**
16. Has she/he ever had any complicated movements of her/his whole body, such as spinning or repeatedly bouncing up and down?	**Yes** **No**
17. Has she/he ever injured her/himself deliberately, such as by biting her/his arm or banging her/his head?	**Yes** **No**
18. Has she/he ever had any objects (other than a soft toy or comfort blanket) that she/he had to carry around?	**Yes** **No**
19. Does she/he have any particular friends or a best friend?	**Yes** **No**
20. When she/he was 4 to 5, did she/he ever talk with you just to be friendly (rather than to get something)?	**Yes** **No**
21. When she/he was 4 to 5, did she/he ever spontaneously copy you (or other people) or what you were doing (such as vacuuming, gardening, or mending things)?	**Yes** **No**
22. When she/he was 4 to 5, did she/he ever spontaneously point at things around her/him just to show you things (not because she/he wanted them)?	**Yes** **No**
23. When she/he was 4 to 5, did she/he ever use gestures, other than pointing or pulling your hand, to let you know what she/he wanted	**Yes** **No**
24. When she/he was 4 to 5, did she/he nod her/his head to mean yes?	**Yes** **No**
25. When she/he was 4 to 5, did she/he shake her/his head to mean no?	**Yes** **No**
26. When she/he was 4 to 5, did she/he usually look at you directly in the face when doing things with you or talking with you?	**Yes** **No**
27. When she/he was 4 to 5, did she/he smile back if someone smiled at her/him?	**Yes** **No**
28. When she/he was 4 to 5, did she/he ever show you things that interested her/him to engage your attention?	**Yes** **No**
29. When she/he was 4 to 5, did she/he ever offer to share things other than food with you?	**Yes** **No**
30. When she/he was 4 to 5, did she/he ever seem to want you to join in her/his enjoyment of something?	**Yes** **No**
31. When she/he was 4 to 5, did she/he ever try to comfort you if you were sad or hurt?	**Yes** **No**

32. When she/he was 4 to 5, when she/he wanted something or wanted help, did she/he look at you and use gestures with sounds or words to get your attention?	**Yes** **No**
33. When she/he was 4 to 5, did she/he show a normal range of facial expressions?	**Yes** **No**
34. When she/he was 4 to 5, did she/he ever spontaneously join in and try to copy the actions in social games, such as The Mulberry Bush or London Bridge Is Falling Down?	**Yes** **No**
35. When she/he was 4 to 5, did she/he play any pretend or make-believe games?	**Yes** **No**
36. When she/he was 4 to 5, did she/he seem interested in other children of approximately the same age whom she/he did not know?	**Yes** **No**
37. When she/he was 4 to 5, did she/he respond positively when another child approached her/him?	**Yes** **No**
38. When she/he was 4 to 5, if you came into a room and started talking to her/him without calling her/his name, did she/he usually look up and pay attention to you?	**Yes** **No**
39. When she/he was 4 to 5, did she/he ever play imaginative games with another child in such a way that you could tell that they each understood what the other was pretending?	**Yes** **No**
40. When she/he was 4 to 5, did she/he play cooperatively in games that required joining in with a group of other children, such as hide-and-seek or ball games?	**Yes** **No**

ID number: Date:

TAQ
SIB Study
IRAS ID: 263048

The Activity Questionnaire

Please read each item carefully and consider whether the behaviour has occurred in the last four weeks. Circle the appropriate number on the scale. Please ensure that you indicate a response for every item. If the particular behaviour does not apply, for example, if the person is not verbal, please circle 0 on the scale.

1) Does the person wriggle/squirm about when seated or lying down?

Never/almost never 0	Some of the time 1	Half of the time 2	A lot of the time 3	Always/almost all of the time 4

2) Does the person fidget or play with their hands and/or feet when seated or lying down?

Never/almost never 0	Some of the time 1	Half of the time 2	A lot of the time 3	Always/almost all of the time 4

3) Does the person find it difficult holding still?

Never/almost never 0	Some of the time 1	Half of the time 2	A lot of the time 3	Always/almost all of the time 4

4) Does the person find it difficult to remain in their seat even when in situations where it would be expected?

Never/almost never 0	Some of the time 1	Half of the time 2	A lot of the time 3	Always/almost all of the time 4

5) Does the person prefer to be moving around or becomes frustrated if left in one positions for too long?

Never/almost never 0	Some of the time 1	Half of the time 2	A lot of the time 3	Always/almost all of the time 4

6) When the person is involved in a leisure activity (e.g. watching TV, playing a game etc.) do they make a lot of noise?

Never/almost never	Some of the time	Half of the time	A lot of the time	Always/almost all of the time
0	1	2	3	4

7) When the person is involved in an activity, are they boisterous and/or rough?

Never/almost never	Some of the time	Half of the time	A lot of the time	Always/almost all of the time
0	1	2	3	4

8) Does the person act as if they are "driven by a motor" (i.e. often very active)?

Never/almost never	Some of the time	Half of the time	A lot of the time	Always/almost all of the time
0	1	2	3	4

9) Does the person seem like they need very little rest to recharge their battery?

Never/almost never	Some of the time	Half of the time	A lot of the time	Always/almost all of the time
0	1	2	3	4

10) Does the person often talk excessively?

Never/almost never	Some of the time	Half of the time	A lot of the time	Always/almost all of the time
0	1	2	3	4

11) Does the person's behaviour seem difficult to manage/contain whilst out and about (e.g. in town, in supermarkets etc.)?

Never/almost never	Some of the time	Half of the time	A lot of the time	Always/almost all of the time
0	1	2	3	4

12) Do you feel that you need to "keep an eye" on the person at all times?

Never/almost never	Some of the time	Half of the time	A lot of the time	Always/almost all of the time
0	1	2	3	4

13) Does the person you care for seem to act/do things without stopping to think first?

Never/almost never	Some of the time	Half of the time	A lot of the time	Always/almost all of the time
0	1	2	3	4

14) Does the person blurt out answers before questions have been completed?

Never/almost never	Some of the time	Half of the time	A lot of the time	Always/almost all of the time
0	1	2	3	4

15) Does the person start to respond to instructions before they have been fully given or without seeming to understand them?

Never/almost never	Some of the time	Half of the time	A lot of the time	Always/almost all of the time
0	1	2	3	4

16) Does the person want things immediately?

Never/almost never	Some of the time	Half of the time	A lot of the time	Always/almost all of the time
0	1	2	3	4

17) Does the person find it difficult to wait?

Never/almost never	Some of the time	Half of the time	A lot of the time	Always/almost all of the time
0	1	2	3	4

18) Does the person disturb others because they have difficulty waiting for things or waiting their turn?

Never/almost never	Some of the time	Half of the time	A lot of the time	Always/almost all of the time
0	1	2	3	4

CHILDHOOD EXECUTIVE FUNCTIONING INVENTORY (CHEXI)
FOR PARENTS AND TEACHERS

Below, you will find a number of statements. Please read each statement carefully and thereafter indicate how well that statement is true for the child. You indicate your response by circling one of the numbers (from 1 to 5) after each statement.

Definitely not true	Not true	Partially true	True	Definitely true
1	2	3	4	5

1.	Has difficulty remembering lengthy instructions	1	2	3	4	5
2.	Seldom seems to be able to motivate him-/herself to do something that he/she doesn't want to do	1	2	3	4	5
3.	Has difficulty remembering what he/she is doing, in the middle of an activity	1	2	3	4	5
4.	Has difficulty following through on less appealing tasks unless he/she is promised some type of reward for doing so	1	2	3	4	5
5.	Has a tendency to do things without first thinking about what could happen	1	2	3	4	5
6.	When asked to do several things, he/she only remembers the first or last	1	2	3	4	5
7.	Has difficulty coming up with a different way of solving a problem when he/she gets stuck	1	2	3	4	5
8.	When something needs to be done, he/she is often distracted by something more appealing	1	2	3	4	5
9.	Easily forgets what he/she is asked to fetch	1	2	3	4	5
10.	Gets overly excited when something special is going to happen (e.g., going on a field trip, going to a party)	1	2	3	4	5
11.	Has clear difficulties doing things he/she finds boring	1	2	3	4	5
12.	Has difficulty planning for an activity (e.g., remembering to bring everything necessary for a field trip or things needed for school)	1	2	3	4	5
13.	Has difficulty holding back his/her activity despite being told to do so	1	2	3	4	5
14.	Has difficulty carrying out activities that require several steps (e.g., for younger children, getting completely dressed without reminders; for older children, doing all homework independently)	1	2	3	4	5

www.chexi.se

Definitely not true	Not true	Partially true	True	Definitely true
1	2	3	4	5

15. In order to be able to concentrate, he/she must find the task appealing	1	2	3	4	5
16. Has difficulty refraining from smiling or laughing in situations where it is inappropriate	1	2	3	4	5
17. Has difficulty telling a story about something that has happened so that others may easily understand	1	2	3	4	5
18. Has difficulty stopping an activity immediately upon being told to do so. For example, he/she needs to jump a couple of extra times or play on the computer a little bit longer after being asked to stop	1	2	3	4	5
19. Has difficulty understanding verbal instructions unless he/she is also shown *how* to do something	1	2	3	4	5
20. Has difficulty with tasks or activities that involve several steps	1	2	3	4	5
21. Has difficulty thinking ahead or learning from experience	1	2	3	4	5
22. Acts in a wilder way compared to other children in a group (e.g., at a birthday party or during a group activity)	1	2	3	4	5
23. Has difficulty doing things that require mental effort, such as counting backwards	1	2	3	4	5
24. Has difficulty keeping things in mind while he/she is doing something else	1	2	3	4	5

1.6 Appendix F: Extract from EF task instructions sent to caregivers

Game 1 – The Magic Wand
In this game, we are interested in how children respond to being asked not to do something.

Game 2 – Sticker Hunt
In this game, we are interested in how children remember things.

Game 3 – Time To Sort!
In this game, we are interested in how well children can follow the rules of a game.

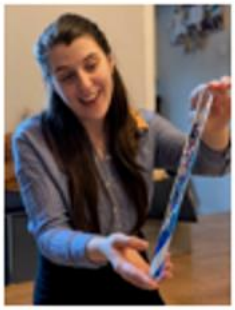
Game 4 – Ball Sorting
In this game, we are interested in how well children can follow the rules of a game, and then adapt when the rules of the game change.

Game 5 – Bag of Toys
In this game, we are interested in how children respond to being asked not to do something.

Game 6 - Playtime
In this game, we are interested in how your child plays without our prompting!

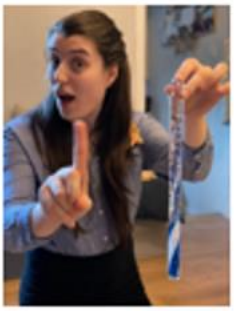
Game 7 – Point To A Picture
In this game, we are interested in your child's understanding of words. We will ask them to point to some pictures on a screen.

How to play



6. Attract your child's attention, then show them the glitter wand.

   Say: **"Oh look [say child's name]"**.



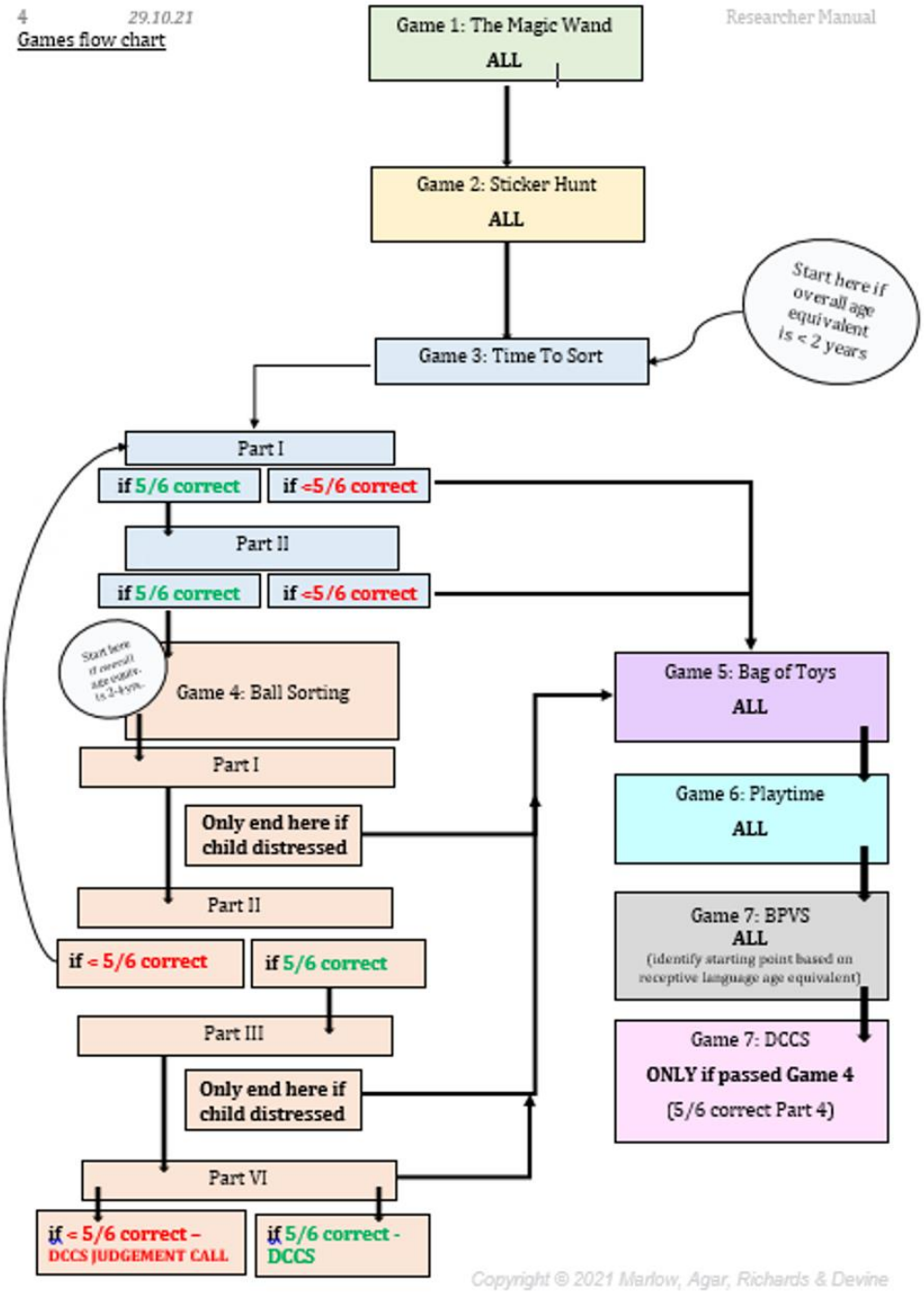
7. Say: **"Look [say child's name]. No, don't touch"**.

   Use your finger to emphasise 'no'.

8. Put the wand down within your child's reach.

9. Turn around, away from your child.

1.7 Appendix G: Full assessment protocol flowchart



1.8 Appendix H: A brief review of the literature on remote testing

There are existing bodies of literature on both the feasibility of remote experimental methods, and the reliability and validity of administering some neuropsychological assessments using telehealth platforms. Online data collection methods have been reported to produce performance equivalent to the same protocol delivered in person (e.g., Oliver & Pike, 2021). Recent research has compared online automated and in-person measurement of executive function in a matched sample of 96 neurotypical children aged 9-15 (Segura & Pompéia, 2020). On a battery which included assessments of working memory, inhibition and switching, the authors found no statistically significant difference between the type of test administration on all but one of the working memory tasks. The authors suggest that education differences between the groups might have contributed to this finding.

While much of the 'telepsychology' research using counterbalanced, within-participants designs has yielded similar performance results for video teleconference-based and face-to-face administration of neuropsychological assessments, including with lower ability cohorts (for a meta-analytic review, see Brearly et al., 2017), most papers focus on older adults (e.g., Hildebrand et al., 2004). Harder and colleagues (2020) were the first to repeat a neuropsychological battery with children, counterbalancing in-person testing and assessment over videoconferencing with the child at home. They concluded that results obtained remotely appeared valid and comparable to in-person testing, and that this method was largely well received by families.

Overall, the evidence supports the use of a remote battery in the context of a global pandemic, while also highlighting the possible utility of this approach to include harder to reach groups in research efforts (Stifel et al., 2020). However, most of the research to date has focussed on fully computerised batteries, or verbally mediated assessments led by an experienced researcher or clinician with the support of a caregiver only to set up tasks. To the

authors' knowledge, the current study is the first to guide caregivers in administering performance-based measures remotely.

References

Brearly, T. W., Shura, R. D., Martindale, S. L., Lazowski, R. A., Luxton, D. D., Shenal, B. V., & Rowland, J. A. (2017). Neuropsychological Test Administration by Videoconference: A Systematic Review and Meta-Analysis. *Neuropsychology Review*, *27*(2), 174–186. https://doi.org/10.1007/s11065-017-9349-1

Harder, L., Hernandez, A., Hague, C., Neumann, J., McCreary, M., Cullum, C. M., & Greenberg, B. (2020). Home-Based Pediatric Teleneuropsychology: A validation study. *Archives of Clinical Neuropsychology*, *35*(8), 1266–1275. https://doi.org/10.1093/arclin/acaa070

Hildebrand, R., Chow, H., Williams, C., Nelson, M., & Wass, P. (2004). Feasibility of neuropsychological testing of older adults via videoconference: implications for assessing the capacity for independent living. *Journal of Telemedicine and Telecare*, *10*(3), 130–134. https://doi.org/10.1258/135763304323070751

Oliver, B. R., & Pike, A. (2019). Introducing a Novel Online Observation of Parenting Behavior: Reliability and Validation. *Parenting*, *21*(2), 168–183. https://doi.org/10.1080/15295192.2019.1694838

Segura, I. A., & Pompéia, S. (2021). Feasibility of Remote Performance Assessment Using the Free Research Executive Evaluation Test Battery in Adolescents. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.723063

Stifel, S. W. F., Feinberg, D. K., Zhang, Y., Chan, M. K., & Wagle, R. (2020). Assessment During the COVID-19 Pandemic: Ethical, Legal, and Safety Considerations Moving Forward. *School Psychology Review*, *49*(4), 438–452. https://doi.org/10.1080/2372966x.2020.1844549

1.9 Appendix I: Validity criteria for direct EF tasks

No.	Description	Examples
1	Was the child in the room when the instruction was given?	One point was awarded if the researcher could see that the child was in the room when the task instructions were given. A point was not given if the child could not be seen on camera when the instructions were delivered, or if they were seen to leave the room at this time.
2	Did the child attend to the instruction?	One point was awarded if the child could be seen to look at either the caregiver or materials as an instruction was delivered, or if they gave verbal or non-verbal feedback to acknowledge an instruction. A point was not given if the child appeared disengaged, distracted or if the caregiver indicated they believed the child had not attended to the instruction.
3	Did the researcher deliver the instructions correctly?	One point was awarded if the researcher directed the caregiver to set up the task appropriately, delivered key instructions according to the approved wording and order detailed in the researcher manual, and prompted the provision or withholding of feedback as required. A point was not given where a researcher did not clearly communicate the correct set up of a task, did not direct a caregiver to use key wording, or did not prompt a caregiver to either give or withhold praise in line with the standardised instructions.
4	Did the caregiver deliver the instructions correctly?	One point was awarded if the caregiver set up the task and repeated instructions as directed, and they did not give any form of feedback or prompts unless explicitly asked to. A point was not awarded if, for example, a caregiver did not repeat the researcher's instructions verbatim where required, gave praise at an inappropriate time, or cued a child to respond in a certain way either verbally or through positioning or gestures.
5	Was the child not holding another toy or object during the task?	One point was awarded if the child was not holding any toy, material, or device unrelated to the task. A point was not given when a child held

		an unrelated item during the task, which was not retrieved by a caregiver.
6	Were the instructions delivered without interference from additional caregivers?	One point was awarded if no other caregivers or children were present, or if others were present but did not interfere verbally or physically with task delivery. A point was not given if persons other than the caregiver delivering the task instructions interfered with the task, for example, by prompting, hinting, distracting, or physically preventing a child from responding as they might otherwise.
7	Did the administering caregiver not interfere when the child was responding to an item?	One point was awarded if the caregiver administering the task allowed a child to respond freely to the instructions and did not interrupt or prevent a child from responding in any way. A point was not given if the lead caregiver stopped a child from responding how they wanted to, or would have otherwise, for example, through redirection of their attention or response.
8	Was the task delivered without any other validity concerns?	A point was awarded if no other validity concerns arose aside from the issues covered by items 1-7. A point was not awarded if another validity concern not better captured by the preceding items was identified. For example, if toys were placed outside of the camera's field of view during inhibition tasks, making it difficult to measure if, and when a child touched these. Any observations deemed to breach this criterion was specified by the researcher coding the task.
	Total	A total for each of the delivered tasks was calculated by summing the outcome of items 1-8. Higher scores indicated fewer concerns about the validity of task administration which might impact performance data. A cut-off score of 6 was applied, such that tasks were omitted from the analyses if they received a total validity score of 5 or lower.

1.10 Appendix J: Spearman's rho and point-biserial correlations between demographic characteristics and SIB.

		1	2	3	4	5	6	7
1	SIB score [a]							
2	Gender	.04						
3	Chronological age	.04	.16					
4	BPVS raw score	.01	.17	.25				
5	VABS ABC	-.19	.20	-.37**	.45**			
6	VABS RL ae	-.23	.19	.23	.40*	.55**		
7	VABS Overall ae	-.21	.27*	.28*	.64**	.68**	.76**	
8	SCQ Total	.13	-.13	.04	-.34*	-.38**	-.42**	-.43**

[a]calculated from the Challenging Behaviour Questionnaire.

*$p < .05$ **$p < .01$

1.11 Appendix K: Spearman's rho and point-biserial correlations between demographic characteristics and EF variables.

	Prohibition	ML perseveration	ML efficiency	Conflict score	TAQ Impulsivity	TAQ Impulsive Sp.	TAQ Overactivity	CHEXI WMF	CHEXI IF
CA	.45**	-.03	.11	.11	.01	.05	-.21	.10	.12
Gender	.23	-.21	.21	.20	-.03	.14	-.09	.08	.05
Verbal Status	.39**	-.46**	.52**	.53**	-.01		-.18	.01	.12
BPVS raw	.42**	-.52**	.54**	.47**	-.17	.38*	-.24	.09	-.07
VABS ABC	.25	-.47**	.47**	.28	-.28*	.06	-.12	-.10	-.05
VABS RL ae	.39**	-.31*	.45**	.29	-.35**	.15	-.29*	-.27	-.05
VABS OA ae	.47**	-.49**	.53**	.41**	-.26*	.25	-.24	-.08	-.01
SCQ total	-.08	.12	-.16	.09	.28*	-.30	.11	.12	.04

*$p < .05$ **$p < .01$

1.12 Appendix L: Quality assessment of papers included in the meta-analysis

Study (1st Author, year)	AUT Measure	TD Screen	Ability Ax	Power	Battery	α	Attrition	Total
Arbelle 1994	2	0	2	0	0	n/a	1	5
Bonli 2005	0	0	2	0	1	0	1	4
Campbell 2019	2	1	1	0	0	n/a	1	5
Coldren 2003	0	0	2	0	0	n/a	1	3
Dawson 1998	1	0	2	0	1	0	0	4
Dawson 2002	2	0	2	1	1	0	0	6
Drayer 2008	0	0	1	0	1	1	1	4
Dunn 1996	0	0	2	0	0	n/a	1	3
Fabio 2020	2	1	2	0	0	n/a	1	6
Faja 2014	2	1	2	0	1	0	1	7
Faja 2015	2	1	2	0	0	n/a	1	6
Gardiner 2017	2	1	2	0	1	1	1	8
Garon 2018	2	0	2	0	1	1	0	6
Hanson 2014	1	1	2	0	1	0	0	5
Jahromi 2013	2	1	2	0	1	1	0	7
Jahromi 2019	2	1	2	0	1	0	0	6
Jones 2013	2	0	1	0	0	n/a	0	3
Kimhi 2014	2	0	2	0	1	0	1	6
Lam 2012	0	0	2	0	0	n/a	0	2
McEvoy 1993	1	0	2	0	1	0	1	5
Meir 2020	2	0	1	1	0	n/a	1	5
Memisevic	1	0	1	0	0	1	1	4
Ostfeld-Etzion 2016	2	1	2	1	0	n/a	1	7

Pastor-Cerezuela 2016	1	0	2	1	0	n/a	1	5
Pastor-Cerezuela 2020	0	0	2	1	1	1	1	6
Pellicano 2007	2	1	2	1	1	0	1	8
Pellicano 2010	2	1	2	1	1	0	1	8
Pellicano 2017	2	1	2	0	1	0	1	7
Berg 2016	1	1	2	1	1	1	1	8
Precenzano 2017	0	0	0	0	0	1	1	2
Rutherford 2003	2	1	2	0	0	n/a	0	5
Schindler 2018;	1	1	2	0	1	0	1	6
Smith 2019	2	1	1	0	0	n/a	1	5
Tan 2018	1	1	1	0	1	0	1	5
Valeri 2020	2	0	2	0	1	1	1	7
Wang 2022	1	0	1	0	0	n/a	1	3
Yerys 2006	1	0	2	0	1	0	1	5
Yerys 2007	2	1	2	0	1	0	1	7
Yi 2012;	1	0	2	0	0	n/a	1	4
Yi 2013	1	0	2	0	0	n/a	1	4
Zacharov 2021	1	0	2	0	1	0	1	5
Zhao 2019	2	1	2	1	0	n/a	1	7