EXPLORING THE IMPACT OF MOOD ON HIGHER-LEVEL COGNITIVE FUNCTIONS

By

Bryony Jessica Fenton

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

May 2024

Thesis Overview

This thesis contains one volume consisting of five chapters and is submitted by Bryony Fenton for the Clinical Psychology Doctorate. The first chapter comprises a meta-analysis reviewing the executive function performance of children and adolescents with current depression, relative to healthy controls. The key finding in this review was that executive function was significantly worse in children and adolescents with depression. This effect was present for both overall executive function and the three investigated executive functioning subdomains (inhibition, working memory and shifting), with working memory showing the most significant impairment. Methodological limitations were considered, highlighting the need for more rigorous and robust research in the field. Clinical implications related to the assessment, formulation and intervention of childhood/adolescent depression were also explored.

The second chapter details an empirical research project investigating the effect of happiness on the theory of mind ability of 3-4 ½-year-old children. Consistent with research focused on adults, it was found that children induced to feel happiness were more likely to make errors on a classic false belief task than children induced to feel neutral in mood. This paper offers support for happiness in children increasing reliance on more easily accessible egocentric knowledge, when mental state reasoning. However, it is recognised that additional research is required to explore this topic further, with potential avenues discussed.

The third and fourth chapters contain press releases for both the meta-analysis and the empirical research paper respectively, for the purpose of public dissemination.

Acknowledgements

Firstly, thank you to my research supervisor, Dr Andy Surtees. I have appreciated your patience, reassurance, guidance, and continued feedback. I feel very grateful to have had you as my supervisor throughout this process. I would also like to thank Dr Chris Jones for all your support with the complex world of meta-analysis and Jessica La for taking the time to support with various aspects of the meta-analysis. A further thank you goes out to the school from which I recruited and all the children who took part in this research.

Thank you to the wider course team, my appraisal tutor Dr Gary Law and all my clinical supervisors (Dr Louise Cooke, Dr Ben Hobson, Dr Mike Ridley-Dash and Dr Abi Gallivan) for all of your support and advice over the last three years of training.

Finally, thank you to all my friends and family for your love and support. Your unwavering encouragement and belief in me have helped get me through this doctorate and reminded me of what I am capable of.

List of Contents

List of Illustrations

List of Tables

CHAPTER I: DIFFERENCES IN EXECUTIVE FUNCTION PERFORMANCE BETWEEN CHILDREN/ADOLESCENTS WITH DEPRESSION AND HEALTHY CONTROLS: A SYSTEMATIC REVIEW AND META-ANALYSIS

Abstract

Background: Childhood/adolescent depression is increasing in prevalence and can lead to long-lasting impacts. Executive function is an important higher-level cognitive process, suggested to include three subdomains (inhibition, shifting and working memory). There is growing evidence that executive function may be impaired in children/adolescents with depression, although the current literature is inconsistent. This meta-analysis explored differences in executive function performance between children/adolescents with depression and healthy controls. Potentially relevant moderators were also explored.

Method: Three electronic databases were systematically searched to identify any relevant studies. Search terms focused on the constructs of depression, executive function, and children/adolescents. Studies were screened in line with a developed inclusion/exclusion eligibility criteria. Included studies underwent an additional quality assessment. Standardised mean differences were calculated using a random effects meta-analytic model.

Results: Twenty-eight articles were included in the final analysis. Overall, it was found that children/adolescents with depression had significantly worse executive function, particularly in working memory. Of the moderators explored, only age and intellectual ability were found to have a significant impact: groups of children with lower IQ, a younger age and depression showed more marked executive function impairments.

Conclusion: The findings suggested that executive function appears to be significantly worse in children/adolescents with depression. This may lead to important clinical implications regarding the assessment and treatment of young-onset depression. However, additional research is needed to better infer causality and to address the various methodological quality issues identified.

Introduction

Mood disorders, such as depression are growing in children/adolescents worldwide (Erskine et al., 2017). Difficulties in higher-level cognitive processes, such as executive function can lead to a range of difficulties including academic challenges and impaired social functioning (Snyder, 2013). There is increasing evidence that difficulties in executive function may be more prominent in children/adolescents with depression and may exacerbate depressive symptoms (Wagner et al., 2015), although there remains substantial inconsistency within the literature. This paper provides a comprehensive meta-analysis to explore the executive function performance of children/adolescents with current depression.

Depression in Children and Adolescents

Unipolar depression is predominantly characterised by prolonged low mood, loss of interest and pleasure in day-to-day activities, and diminished energy (Gruenberg et al., 2005). Depression covers a spectrum which can range from sub-clinical milder symptoms and more moderate chronic forms (e.g., dysthymia), to a clinical diagnosis of Major Depressive Disorder (MDD) (Schramm et al., 2020; Weavers et al., 2023). Childhood and adolescence are periods particularly associated with emotional, biological, and psychosocial changes which can increase an individual's vulnerability to the development of depression (Baune et al., 2014). Depression is considered one of the most prevalent mental health difficulties for children and adolescents, with an estimated average global prevalence rate of 6.2% in 5–17-year-olds (Erskine et al., 2017). Similarly, a recent study by Shorey et al. (2022) reported a global point prevalence rate of 8% for MDD and 4% for dysthymia in adolescents (age 10-19 years). Indeed, it is estimated that 1 in 5 children/adolescents will experience a depressive

episode before the age of 18 years old (Goodall et al., 2018). The prevalence of elevated depressive symptoms in adolescents also continues to grow, with an increase from 24% (2001-2010) to 38% (2011-2020) (Shorey et al., 2022). Depression in young people is reported to recur frequently, with approximately 70% of children and adolescents experiencing a recurring episode within two years (Shorey et al., 2022). A younger age of onset can significantly predict the likelihood of relapse and many children and adolescents with depression continue to experience severe symptoms in adulthood (Garber et al., 2009; Shorey et al., 2022).

Depression in childhood and adolescence is reported to have wide-ranging adverse impacts on areas such as academic/vocational performance, interpersonal relationships, risk-taking behaviours, physical health, and general quality of life (Shorey et al., 2022; Wagner et al., 2015). Indeed, Baune et al. (2014) discussed how young people with depression are likely to experience severe and enduring impairments in their overall psychosocial functioning. Globally, depression has been ranked as one of the leading causes of disability and is significantly associated with an increased risk of suicide, reported to be the second leading cause of death in adolescents (Shorey et al., 2022). Thus, due to the significant and enduring impacts of young-onset depression, it is vital to understand the factors associated with depression during this critical period. This may contribute to the better identification of young-onset depression and allow for more targeted psychotherapeutic interventions.

Executive Function

Executive function is an “umbrella term”, referring to higher-level cognitive control processes vital for purposeful, goal-directed behaviour and self-regulation (Majeed et

al., 2023; Yang et al., 2022). Executive function is complex and multifaceted in nature, for which several models have been proposed and debated within the literature (Gray-Burrows et al., 2019; Hirst et al., 2017). One particularly influential conceptual framework for understanding executive function is the unity and diversity model outlined by Miyake et al. (2000). In this model executive function is theorised to be underpinned by three separate but related core subdomains: inhibition, shifting and working memory (Khoury et al., 2015). Inhibition refers to the deliberate ability to suppress or avoid prepotent responses to irrelevant stimuli (Majeed et al., 2023). Behavioural tasks typically used to measure inhibition include the Stroop colour-word task and the go/no-go task (Spaniol & Danielsson, 2022). Shifting involves switching flexibly and efficiently between different tasks and mental sets as demands change (Gray-Burrows et al., 2019). The Wisconsin Card Sorting Task (WCST) and Trail Making Test part B (TMT-B) are often cited as two of the most common tasks assessing set shifting (Snyder, 2013). Finally, working memory reflects the ability to store and manipulate information mentally over a short period of time (Lonergan et al., 2019). Typical working memory tasks comprise tests such as spatial span and delayed-match-to-sample (Snyder, 2013).

Within the initial model, Miyake et al. (2000) suggests that, whilst clearly separable, the three subdomains cannot be considered "completely independent" (Miyake et al., 2000, p72). However, Miyake et al's. (2000) rationale for this was based solely on the increased statistical fit of a full three-factor model, in comparison to modelling inhibition, shifting and working memory as independent factors. Resultantly, further exploration of the precise nature of the inter-connections between the three subdomains was limited (Miyake et al., 2000). Therefore, modelling the underlying commonality within the model (i.e. the conceptualisation

of the executive functions as unitary) and its mechanisms has been the source of much debate (Blakey et al., 2016). Some researchers have proposed that the commonality in executive function may strongly reflect inhibitory processes (Hall & Fong, 2015; Valian, 2015). This is linked to the view that most executive function tasks have been noted to require the use of some form of inhibitory control (Friedman & Miyake, 2017). However, this perspective has been critiqued for being both somewhat ambiguous and overly broad in its conceptualisation of inhibition (Friedman & Miyake, 2017). An alternative theory, frequently discussed within the literature, argues that the ability to actively maintain and manage task goals and task-related information may be the central mechanism underpinning the unity of executive functions (Miyake & Friedman, 2012; Friedman & Miyake, 2017). This ability has been hypothesised to be most closely related to working memory (Miyake et al., 2000; Jurado & Rosselli, 2007), which Baddeley (2003) specifically defined as the capacity to temporarily hold and manipulate information necessary for complex cognitive tasks. Consequently, other controlled processes, such as inhibition and shifting are proposed to require the active maintenance and manipulation of task goals and relevant information held within working memory to be executed effectively (Miyake et al., 2000; Fleming et al., 2016). This theory could imply a more hierarchical structure, in which the working memory subdomain plays a more primary role that supports and helps regulate the other two related executive function subdomains. Further research is needed to continue exploring this area in more depth and detail, with falsifiable predictions of competing theories lacking at this stage.

In considering developmental trajectories, general executive function and the specific subdomains are reported to show substantial progression throughout childhood and adolescence (Best & Miller, 2010). More specifically, the ability to hold information in

working memory is described as developing from a very young age, with the subsequent ability to manipulate this information showing a more prolonged period of progression and refinement (Diamond, 2013). Alternatively, inhibitory control is postulated to be particularly difficult for young children, before a rapid improvement occurs across both simple and complex tasks around preschool age (Best & Miller, 2010; Diamond, 2013). Finally, shifting is reported to demonstrate a more gradual and linear improvement throughout childhood (Diamond, 2013).

Several additional skills that build on these three core executive functioning subdomains have also been described in the literature (e.g., planning and verbal fluency; Yang et al., 2022). However, for the scope of this review, due to the various ways executive function can be conceptualised, the three-component unity and diversity model was focused on to ensure clear boundaries in the searching and coding of relevant primary articles. This framework has also been extensively replicated across various age groups and developmental stages (Khoury et al., 2015).

Depression and Executive Function

Depression and Executive Function in Adults

Within the extensive literature focused on adults, moderate but significant neurocognitive impairments, specifically in executive function are one important and well-established factor associated with depression (Goodall et al., 2018; Rock et al., 2014; Wagner et al., 2012). In adults, significant impairments in the inhibition subdomain have been consistently reported (Snyder, 2013; Wagner et al., 2012). Regarding the subdomains of

shifting and working memory, the literature has been more inconsistent, with some studies finding pronounced deficits in shifting and working memory (Christensen et al., 1997; Wagner et al., 2012) and other studies finding no significant impairments (Zakzanis et al., 1998). However, a widely-cited and comprehensive (k = 113) meta-analysis conducted by Snyder (2013), found that patients with MDD were similarly significantly impaired across all executive function tasks requiring shifting, inhibition and working memory, relative to healthy controls.

Executive function is required to carry out most daily tasks, therefore deficits in executive function can contribute to difficulties in academic achievements, family functioning, emotional wellbeing, and social behaviour (Majeed et al., 2023; Snyder, 2013). More specifically within depression, executive function impairments have been reported to have significant detrimental impacts on an individual's coping skills, susceptibility to relapse and treatment response (Wagner et al., 2012). A complex bi-directional relationship may exist between depression and executive function, in which disrupted executive function can contribute to the development/maintenance of depressive symptoms and depressive symptoms can additionally compromise executive function abilities (Ciuhan & Iliescu, 2021). However, it is acknowledged that there is some uncertainty in this regard (Ciuhan & Iliescu, 2021). Some researchers have identified the involvement of executive function in supporting emotion regulation, a central difficulty in depression (Wagner et al., 2012). Research has also discussed how executive function deficits may hinder the management of everyday tasks, alongside the ability to disengage from negative moods/thoughts, thus potentially fostering and perpetuating depressive symptoms (Letkiewicz et al., 2014). Further literature has outlined how depression may interfere with executive function through resulting in both structural and functional

abnormalities in the pre-frontal cortex, a key area of the brain supporting executive function (Snyder, 2013). An alternative 'cognitive effort' hypothesis has been proposed, suggesting that primary depressive symptoms (e.g., negative thinking and rumination) are cognitively demanding and thus lead to a subsequent increased difficulty in allocating cognitive resources to more effortful tasks, such as tasks requiring executive function (Nuño et al., 2021).

Despite the considerable evidence to date regarding increased impairments in executive function for adults with depression, difficulties arise in generalising these findings to children/adolescents (Wagner et al., 2015). Both neurological and cognitive development continue well into adulthood (Goodall et al., 2018). Hence, there is potential for neurocognitive functioning in children/adolescents to experience unique vulnerability to a range of influences, including stressful events and subsequent psychopathology (Nyvold et al., 2022; Yang et al., 2022). Therefore, further clarifying the link between executive function and young-onset depression may lead to important implications for early identification, management, and treatment targets, as a means of improving clinical outcomes and overall wellbeing in this population (Majeed et al., 2023).

Current Understanding of Executive Functioning in Children and Adolescents With Depression

Studies comparing the executive function of children/adolescents with depression to healthy children/adolescents are increasingly emerging. However, the number of studies remains small compared to the literature in adults, and the results reported are often inconsistent (Wagner et al., 2015). Several systematic reviews have examined executive function in children/adolescents with depression. Vilgis et al. (2015) conducted a literature

review summarising the findings across 33 papers investigating executive function and attention in children/adolescents with depressive disorders. It was found that, despite mixed results, most studies did not report significant impairments in the executive functioning subdomains of inhibition, shifting and working memory for children/adolescents with depression compared to healthy controls (Vilgis et al., 2015). The researchers acknowledged that their depression group included participants with both current and remitted depression which may have impacted the ability to detect a relationship between depression and executive function (Vilgis et al., 2015). Baune et al. (2014) calculated the effect sizes from seven papers exploring neuropsychological functioning in adolescents with depression. Again, mixed results were observed in which three of the studies found impairments in executive function of varying effect sizes (small to large) in adolescents with depression compared to healthy controls (Baune et al., 2014). However, the remaining four studies found no significant differences in executive function between the two groups (Baune et al., 2014). Within Baune et al's. (2014) review, the researchers failed to delineate executive functioning into its different subdomains, which could be considered a weakness (Majeed et al., 2023).

Both Wagner et al. (2015) and Goodall et al. (2018) instead conducted meta-analytic reviews. Wagner et al. (2015) identified 17 papers, in which they found children/adolescents with MDD showed significantly worse working memory performance, inhibitory control, shifting ability, and verbal fluency compared to healthy children/adolescents. Within the meta-analysis conducted by Wagner et al. (2015), children/adolescents were only included in the depression group if they had a formal diagnosis of MDD, which raises the question regarding links between executive function and milder forms of depression in childhood/adolescence. In contrast, a meta-analysis of 23 studies by Goodall et al. (2018) found no significant

differences in working memory, response inhibition, planning and set shifting between young people with depression and healthy controls. It is important to note that Goodall et al. (2018) used participants between the ages of 12-25 years in their sample which covers diverse neurodevelopmental stages, potentially leading to age-related neurocognitive differences impacting results. Similarly, small-scale research findings in this area have also been mixed with some studies finding children/adolescents with depression to be impaired on executive function measures of inhibition (Cataldo et al., 2005), shifting ability (Günther et al., 2011) and working memory (Fisk et al., 2019). Conversely, other studies have reported minimal to no impairment in inhibitory control (Fisk et al., 2019), set shifting (Bloch et al., 2013) and working memory (Maalouf et al., 2011) for young people with depression in comparison to healthy children/adolescents.

In the eight-ten years since the previous systematic literature searches were conducted, there has been a substantial growth in the available literature exploring executive function ability in children/adolescents with depression. Therefore, alongside the current lack of clarity, a timely review of the existing relevant research is needed to clarify the executive functioning profile of children/adolescents with depression. As one well-recognised source of variability in executive function in depression relates to the executive functioning subdomain assessed (Khoury et al., 2015), exploring both an overall executive function construct (i.e. encompassing the shifting, inhibition and working memory subdomains), alongside each individual subdomain and its relation to childhood/adolescent depression is indicated. This ensures consideration of both the unity and diversity of executive function, alongside helping specify any potential executive function difficulties that may be related to depression in children/adolescents (Majeed et al., 2023; Nyvold et al., 2022).

Possible Moderators Affecting the Association Between Executive Function and Childhood/Adolescent Depression

To understand some of the discrepancies in the literature, it will be important to examine potential moderators (both sample characteristics and methodological factors) that may impact the association between executive function and depression in children/adolescents. Heterogeneity in the current findings may be linked to the significant diversity observed in the samples recruited across different studies, which vary in regard to the presence of comorbidity, use of psychotropic medication, age, sex[1], and intellectual functioning (Majeed et al., 2023).

Comorbidity has been reported in 40-70% of children and adolescents with depression, with co-occurring anxiety disorders, attention deficit hyperactivity disorder and conduct disorders frequently identified (Rao & Chen, 2009). These psychological difficulties are often independently associated with impairments in executive function (Johnson et al., 2015; Majeed et al., 2023; Oosterlaan et al., 2005), potentially compounding or influencing any executive function deficits observed in children/adolescents with depression. It is hypothesised that the use of psychotropic medication may also impact the strength of the relationship between executive function and young-onset depression. Some evidence has indicated that the use of medications (e.g., antidepressants and benzodiazepines) can impair executive function (Majeed et al., 2023; Snyder, 2013), whereas other studies have reported positive effects of medication on cognitive functioning (Prado et al., 2018). Finally, age, sex and intellectual ability may be additional relevant factors to consider when exploring the

[1] [1] The term sex is consistently used in this paper instead of the term gender, due to inconsistency in the included studies reporting and a lack of clear distinction between sex and gender in the primary articles.

relationship between depression in young people and executive function. As previously outlined, differing developmental trajectories of executive function and its specific subdomains have been reported (Khoury et al., 2015). Sex differences may also emerge, as females are more likely to experience depression (Essau et al., 2010), alongside sex differences frequently observed in brain maturation (Vilgis et al., 2015). Whilst, intellectual ability has been found to be generally associated with executive functioning (Arffa, 2007), alongside being reported to be reduced in children/adolescents with depression (Wagner et al., 2015).

When considering methodological factors, various task outcome measures are reported across different studies, in which reaction time (efficiency) and/or performance accuracy (effectiveness) are typically assessed (Majeed et al., 2023). These can reflect differing cognitive processes and may provide further information regarding potential underlying mechanisms (Hoffmann et al., 2017; Majeed et al., 2023). All the above factors may contribute to the nuances found regarding executive function ability in childhood/adolescent depression and thus warrant further, more in-depth review using a meta-analytic approach.

Rationale

The development of depression in childhood and adolescence can have significant and long-lasting impacts. Though executive function deficits have begun to be identified in this population, there is an inconsistency of findings across different studies. To address this, a comprehensive systematic review and meta-analysis was performed to summarise the current literature examining the nature of executive function in children and adolescents with depression. To extend the work conducted by Wagner et al. (2015), depression was

investigated from a broader perspective (i.e. including children and adolescents with an MDD diagnosis, alongside studies including young people with depressive symptoms and/or dysthymia). For the purpose of this meta-analysis, the term child/adolescent refers to participants 18 years old or under.

This meta-analysis firstly aimed to address whether children/adolescents with depression show global and/or more specific differences in executive function compared to healthy children/adolescents. To explore this, a meta-analysis on group-level differences between children/adolescents with and without depression, in their performance on behavioural executive functioning tasks was conducted. Both general executive function and its three individual subdomains (inhibition, shifting and working memory) were analysed. Secondly, using subgroup analysis and meta-regression, this meta-analysis aimed to examine potentially relevant factors (i.e., executive function task outcome measure, age, comorbidities, intellectual ability, sex, and psychotropic medication use) that may moderate the association between executive function and depression in childhood and adolescence.

Method

The study was preregistered using PROSPERO prior to the initiation of study identification (ID number: CRD42023445447). No amendments were made to the protocol. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Page et al., 2021) guidelines were adopted and followed throughout (see Appendix 1 for a PRISMA checklist).

Identifying Primary Studies

Search of Electronic Databases

A systematic search of the current literature was carried out in August 2023 using the APA PsycINFO, MEDLINE, and Embase online databases. The search aimed to provide a comprehensive overview of the existing literature investigating executive function performance in children/adolescents with depression in comparison to healthy controls. The three following constructs were selected *depression*, *executive function,* and *children and/or adolescents*, in which search terms were then developed and categorised accordingly. An abstract, title and keyword search was employed for each construct. The search terms for each construct were guided by relevant existing literature, in which they were particularly adapted from previous related reviews undertaken by Vilgis et al. (2015) and Wagner et al. (2015). To build on the existing literature and capture a broader experience of depression including mild/moderate chronic symptoms, the terms *dysthymia* and *dysthymic disorder* were also included within the *depression* construct. In line with the eligibility criteria outlined further below, limits of English language, human participants, and child/adolescent participants (0-18 years) were applied where possible, as a means of further filtering the search data. For each database, records ranging from the earliest date available to present day were accessed. The full list of search terms used to identify the relevant studies, method of search and applied limits are outlined in Table 1-1.

Table 1-1

Search Criteria Used for the Databases

Construct	Free text search terms	Method of search	Limits
Depression	"major depressive disorder*" "depress*" "childhood depress*" "paediatric depress*" "dysthymi*" "low mood" "affective disorder*" "mood disorder*" "major depress*" "MDD" "Adjustment disorder*" "depression symptom*" "dysthymic disorder*" "depressive disorder*"	Free search terms All search terms within a construct were combined with *OR.* All three constructs were combined with *AND.* Abstract, title and keywords search used for each construct (ab,kw,ti.).	English language Human studies Children (0-18 years) APA PsycINFO: 1967-August 2023 MEDLINE: 1946-August 2023 Embase: 1974-August 2023
Executive function	"executive function*" "executive dysfunction*" "executive control" "cognitive control" "shifting" "set shifting" "switching" "cognitive flexibil*" "inhibition" "inhibitory control" "inhibition capacity" "response inhibition" "working memory" "updating"		
Children and/or adolescents	"child*" "adolescen*" "bab*" "infan*" "Toddler*" "Preschool*" "pre-school" "Kindergar*" "Youth*" "teen*" "school child*" "juvenile*" "Early childhood" "young child*" "boy*"		

Construct	Free text search terms	Method of search	Limits
	"girl*" "pediatric*" "paediatric*" "young people" "young person*" "kid*" "childhood" "preteen*" "primary school" "elementary school" "secondary school" "high school" "school age"		

Inclusion and Exclusion Criteria

The full inclusion and exclusion criteria developed for this meta-analysis are described in Table 1-2. The criterion of *participant characteristics* ensured the clinical group was defined as experiencing a current diagnosis of depression/dysthymia and/or scoring above a cut-off threshold on a depression symptom measure (self-report or clinician-rated). Also required, was a healthy control group in which participants were not diagnosed with any form of depression, nor recruited on the basis of having any other psychopathology, alongside all study participants being classified as children and/or adolescents (either an age range between 0-18 years or if not reported, a mean age of 18 years or under). Studies needed to use standardised behavioural executive function tests assessing at least one of the three core components of executive function investigated in this meta-analysis (inhibition, shifting and working memory) to meet inclusion for the *executive function measurement* criterion. Studies using executive function tests that did not assess any of these three subdomains, that involved a significant affective component or that were from a purely neuroimaging or self-report perspective were not included. The criterion for *outcome data* required studies to report the

appropriate data or statistical information required to calculate a standardised mean difference effect size (e.g., means and standard deviations). Finally, the *study type* criterion was limited to only studies reporting primary data comparing the executive function ability of children/adolescents with depression to healthy controls. Other article types (e.g., systematic reviews) were excluded from the meta-analysis.

Table 1-2

Inclusion and Exclusion Criteria

Criterion	Inclusion criteria	Exclusion criteria	Justification
Participant characteristics	• Studies with a participant group primarily experiencing current depression (determined via either clinical diagnosis of depression or sub-clinical depressive symptoms assessed using cut-off thresholds of a relevant scale/measure)	• Studies with a depression participant group where participants are primarily recruited based on the presence of other co-morbidities (e.g., other mental health, neurological, paediatric or neurodevelopment disorders considered within the wider literature to affect executive functioning) or including participants with remitted depression.	This is to ensure that included studies are primarily investigating the executive functioning ability of the target population of those with depression, as opposed to other co-morbidities known to have an impact on executive functioning. Alongside, broadening the scope of the review to capture those experiencing depressive symptoms but who may not necessarily have been given a diagnosis of MDD.
	• Studies with a healthy control group displaying no evidence of depression and not recruited on the basis of having any other psychopathology.	• Studies that do not include a healthy control group.	Inclusion of a healthy control group allows meaningful comparisons between the two groups as a means of enhancing understanding of the impact of depression on executive function performance, whilst strengthening the

Criterion	Inclusion criteria	Exclusion criteria	Justification
			validity of the study's findings.
	• Studies with participants from a child and adolescent population (up to age 18 years).	• Studies with any participants over the age of 18 years, unless data is separated for participants under and over the age 18.	This is to address the mixed and limited findings in the literature regarding how depression may affect executive function performance in children and adolescents, in comparison to the greater literature focused on adults.
Executive function measurement	• Studies that include executive function tasks assessing at least one of the three core subdomains (shifting, inhibition or working memory)	• Studies that only include executive function tasks not assessing any of the three core subdomains (e.g., instead assessing planning or problem solving.	Executive function is a broad and multifaceted construct. Inhibition, shifting and working memory are considered core components of key executive functioning theoretical models and can be impacted in a range of clinical conditions (including mental health difficulties). Focusing on the three main subdomains allows for clarity within the review and a more focused investigation of executive function and the main cognitive processes involved.
	• Studies using standardised behavioural neuropsychological tests of executive function.	• Studies evaluating executive functioning from a neuroimaging or self-report/questionnaire perspective only or with a significant affective component to the task.	This is to allow for objective and quantifiable results within studies that can measure specific components of executive function (e.g., shifting), allow for comparison of performance between groups and increase methodological rigour. Furthermore, affective manipulation of executive function task stimuli can impact

Criterion	Inclusion criteria	Exclusion criteria	Justification
			participants subsequent performance.
Outcome data	• Studies that have reported or provided on email contact, task performance data as either means and standard deviations, T-test statistics, Cohen's *d* effect size or statistical data that can be converted to one of these statistics. • Studies presenting group level data separately for children/adolescents with depression and healthy children/adolescents.	• Studies that do not present appropriate data for the meta-analysis (e.g., instead reporting median and interquartile range values) and who's authors do not respond to two attempts at contact via email. • Studies that present data in graph form where it is not possible to extract all the data required.	This is to ensure that the outcomes reported in the studies can be calculated into an effect size required for the meta-analysis.
Study characteristics	• English language, studies that report primary data on executive function ability of children and adolescents with depression or depressive symptoms compared to healthy children/adolescents.	The below article types were excluded: • Systematic reviews • Meta-analyses • Intervention studies (unless baseline executive function scores are reported for both groups) • Qualitative papers • Case studies • Poster/conference papers • Theoretical papers • Association studies • Papers not published in English.	This meta-analysis centres on comparing executive function performance between a clinical and control group using primary research studies. The types of articles excluded do not result in the relevant data required for this meta-analysis.

Paper Selection

From the full search, any duplicate records were initially manually removed using features of the Zotero referencing software. The remaining articles were then screened by the researcher using their titles and abstracts, according to the above inclusion and exclusion criteria. For any remaining articles, the full text was accessed to enable a more detailed review

against the inclusion and exclusion criteria. Any articles found to meet the inclusion criteria fully, alongside not meeting any of the exclusion criteria were included in the meta-analysis. To ensure any relevant studies were not missed the reference lists of included articles and of relevant meta-analyses/systematic reviews excluded at the title/abstract stage were additionally screened. In cases where a relevant study did not report the appropriate data required for the meta-analysis, the study's authors were contacted via email to request this data. In any instances of no response, a follow-up email was sent after a two-week window.

The reliability of the screening process was checked by using a second-rater. A 20% pseudo-random sample of studies from the total search were cross-checked against the inclusion/exclusion criteria. Cohen's kappa (κ) was used to calculate inter-rater reliability. Inter-rater reliability was good ($\kappa = 0.64$). Any disagreements were discussed to ensure a consensus was reached.

Risk of Bias Assessment

Specific quality criteria developed for this meta-analysis were used to assess the varying risk of study level bias. As recommended by Higgins et al. (2011), the quality criteria selected for the framework represented the main threats to validity for the question being addressed in this meta-analysis (e.g., problems with measurement of depression), as opposed to the objective overall quality of the studies.

The risk of bias criteria were predominately informed by adapting existing risk of bias frameworks, including The Cochrane Collaboration Risk of Bias Tool (Higgins et al., 2011) and its generalisation to non-randomised studies by Kim et al. (2013). The framework used in this meta-analysis assessed the risk of bias across five domains thought to reflect core threats

to validity: *selection bias*, *depression measurement bias, executive function measurement bias*, *reporting bias,* and *generalisability*. For each of the individual included studies, the risk of bias in the five domains was rated by the researcher as either "low risk", "unclear risk", or "high risk". Table 1-3 outlines the criteria for the above risk categories across each domain. As studies could potentially meet the criteria for more than one level of risk within a domain (e.g., meet both the "high" and "unclear risk" level), a cautious judgement was employed, in which the higher risk of bias level was selected. Due to all primary studies having a homogenous group design, no weighting was placed on study design when considering the overall quality.

Each of the domains within the framework was equally weighted. Scores were applied to each of the three risk categories, such that risk domains ranked as having a "low risk" of bias scored two points, those evaluated as "unclear risk" of bias scored one point and those deemed to have a "high risk" of bias were awarded zero points. Scores across each of the five domains were then summed together. This resulted in a total risk of bias score between 0 – 10 for each included study, in which higher scores indicated a higher level of methodological quality (lower risk of bias). The reliability of the risk of bias ratings was checked by using a second-rater, with a 30% random sample of studies selected. Inter-rater reliability was good ($\kappa = 0.74$). Any discrepancies in ratings were discussed, and final ratings were subsequently agreed.

Table 1-3

Domains of Risk of Bias and the Criteria for Ratings of "Low", "Unclear" or "High Risk"

Domain	Details	Risk of bias
Selection bias	Were efforts made to minimise selection bias in the included studies, such as clearly outlining the sampling method? Was convenience sampling used? Were there efforts to match participant groups on relevant variables?	High Risk- Selection of participants unspecified or a single non-random sample used (e.g., a specialist psychological clinic or single school). No attempts made to match the depression and control group. Unclear Risk- Participants recruited from multiple non-random samples (e.g., multiple specialist clinics or multiple schools). Matching on key relevant variables (e.g., age, education, and IQ). Low Risk- Random sampling. Matching on a range of variables.
Depression measurement bias	Are the depression outcome measures used to classify the group valid and reliable for this population? Have the authors used well validated formal measures to confirm allocation into the depression group?	High Risk- The measure used to identify the depression group is unspecified or a measure that is not validated for use in a child/adolescent population is used. Unclear Risk- Depression classification is based on a self/parent/teacher report that is validated for use in a child/adolescent population. Low Risk- Depression classification is confirmed by formal methods, including validated clinician scales (e.g., the children's depression rating scale) and/or a standardised structured/semi-structured clinical interview (e.g., based on DSM-IV criteria).

Domain	Details	Risk of bias
Executive function measurement bias	Are the executive function tests used valid and reliable for this population?	High Risk – The executive function tests used are unspecified (e.g., an executive function test with no additional description). Major modifications were made to the task. Unclear Risk – The executive function tests used are not validated for use in a child/adolescent population (e.g., only validated in adults) or the validation is unclear. Low Risk – The executive function tests used are validated for use in a child/adolescent population.
Reporting Bias	Are the variables that have been reported for the executive function tests the primary executive functioning variables recommended in the wider literature?	High Risk – Only reported variables not specifically related to executive functioning ability. Unclear Risk – The recommended executive function variable is not reported for the included tests, however a related one that is still eligible to be included in the review is reported. Where both response time and error number are expected only one is reported. Low Risk – The recommended variables identified as the key measure of executive function ability for the included tests are reported. Where expected both response time and error number are reported.
Generalisability	Are there any differences between the study participants and those to whom the review is applicable?	High Risk- Small sample with or without idiosyncratic features (≤20 per group). Unclear Risk- Sufficient sample for generalisation (>20 per group) but

Domain	Details	Risk of bias
	Is there a sufficient participant sample size in the study, allowing it to be meaningful?	with some idiosyncratic features not fully typical of depression (e.g., significant amount of suicide attempters or all one sex). Low Risk- Sufficient sample for generalisation and representative of target population (>20 per group).

Note. Each of the risk categories are colour coded, in which red equates to "high risk" of bias, amber equates to "unclear risk" of bias and green equates to "low risk" of bias.

Data Extraction

The author was responsible for all data extraction, which occurred between October 2023 – November 2023. The data extracted from each primary study included relevant publication details and participant demographic information (e.g., authors, year, age and sex). Details regarding the sample and further methodological information were additionally extracted, such as sample size, the presence of psychiatric comorbidities and psychotropic medication usage in the depression group. Also extracted was the executive function task used, its associated subdomain and clarification of whether it assessed reaction time or accuracy. Finally, executive function task performance on relevant variables in the child/adolescent depression group and the healthy control group were also extracted. Where a broader depression group was reported as smaller subgroups (e.g., a dysthymic disorder and MDD group), performance on the executive function task was combined into a single quantitative outcome using the procedures outlined by Borenstein (2009). Similarly, in instances where studies reported average scores across multiple levels of a task (e.g., spatial

working memory task), these were combined into an overall score using the same method as above.

Table 1-4 outlines the investigated executive functioning subdomains, the tasks used in the primary studies to assess each subdomain and the variables extracted for each task. Within Table 1-4, primary executive function variables for each task are outlined initially, followed by alternative variables also considered acceptable, in studies where the primary variables were not reported. The allocation of executive function tasks to the relevant subdomains and the variables considered appropriate for extraction were based on consultation of the available literature outlining the primary purpose of each task in investigating specific cognitive processes and/or what is most frequently used within neuropsychological practice.

Table 1-4

Extracted Executive Function Tasks and Associated Variables, Separated by Corresponding Executive Function Subdomain

Executive function subdomain	Task	Executive function variables	Supporting reference example
Shifting	Wisconsin card sorting test	Perseverative errors	Dann et al. (2023)
	Trail making test B	Difference in completion time between trial B and trial A. Alternatively, response time and/or error numbers	Christidi et al. (2015)
	Opposite worlds (TEA-Ch)	Opposite worlds response latency	Lakomy (2021)
	Visual set shifting task	Reaction time and/or errors	Schuitema et al. (2015)
	Intra-dimensional–extra-dimensional set-shifting task (CANTAB)	Extra dimensional stage performance. Alternatively, total stage	Jazbec et al. (2007)

Executive function subdomain	Task	Executive function variables	Supporting reference example
		reached and/or adjusted errors	
	Switch task	Switch cost	Kiesel et al. (2010)
	Shifting attention test (CNS vital signs)	Errors and/or reaction time	Brooks and Sherman (2012)
Inhibition	Hayling sentence completion task	Error number and/or reaction time	Robinson et al. (2009)
	Go/No-Go task	False alarms. Alternatively inhibitory accuracy score	Meule (2017)
	Flanker task	Interference score. Alternatively, accuracy and/or reaction time for incongruent trials	Oeri et al. (2019); Paap et al. (2020)
	Stroop task	Interference scores. Alternatively incongruent reaction time and/or errors	Duell et al. (2018)
	Walk-don't walk task (TEA-Ch)	Total number of correct paths	Sutcliffe et al. (2006)
	Stop task	Stop-signal reaction time	Verbruggen and Logan (2008)
	Simon task	Simon effect	de Bruin and Sala (2018)
Working memory	Digit span task (WISC/WAIS)	Total scaled score	Rosenthal et al. (2006)
	Keep track task	Number of words recalled correctly	St Clair-Thompson and Gathercole (2006)
	Spatial working memory (CANTAB)	Between search errors	Aoki et al. (2023)
	Spatial span (CANTAB)	Span length score	Sabahi et al. (2022)

Executive function subdomain	Task	Executive function variables	Supporting reference example
	Sternberg working memory task (verbal and spatial)	Reaction time and/or accuracy	Klabes et al. (2021)
	The California verbal learning test for children	List one items recalled correctly	Graves et al. (2021)
	Delayed match to sample (CANTAB)	Accuracy and/or reaction time	Daniel et al. (2016)
	The Rey Auditory–Verbal Learning Test	Trial one items recalled correctly	Tyagi et al. (2021)
Combined	Combination of WCST, freedom from distractibility errors, Trail Making Test B, Stroop task and verbal fluency tasks	Composite score	Pandina, (2000)

Note. TEA-Ch: Test of Everyday attention for children; CANTAB: Cambridge Neuropsychological Test Automated Battery; WISC: Wechsler Intelligence Scale for Children; WAIS: Wechsler Adult Intelligence Scale

For most studies, executive function test performance was reported as the means, standard deviations, and sample sizes required to calculate the standardised mean difference between children/adolescents with depression and healthy controls. In studies, where standard errors were instead reported they were converted into the standard deviations[2] required for the calculation of the study-level effect. Where needed the direction of the effect size was transformed so that positive effect sizes consistently represented impaired executive function ability in the group with depression compared to the healthy control group.

Numerous included studies reported multiple effects in terms of executive function task variables (e.g., reaction time and error number for the same task). Due to the potential of these effects being meaningfully different, all the different effects were extracted and included

[2] Using the formula $SD \quad SE * \sqrt{n}$

in the meta-analysis, with a label added to clarify the variable of interest. However, it is important to note that the inclusion of multiple effects from the same primary study may lead to some reduction in the size of the confidence intervals around the meta-analytic model, due to the fact that the sample size of that primary study was included more than once (Higgins et al., 2011); consideration of this is noted below.

Data Analysis Strategy

All analyses were undertaken using R. Studio (RStudio, 2020), with significance set at $p < 0.05$. The fixed-effects meta-analytic model and the Random-Effects Model (REM) were initially considered to ascertain which model best fit the extracted data. Quantile-Quantile (QQ) plots of the distribution of primary study effects were generated and reviewed to compare the normality and linearity of both the fixed and REM. Due to several primary studies reporting multiple effects and thus violating the assumption of independence of effect sizes, the standardised mean differences were also calculated for each planned analysis using the three-level meta-analytic model. The three-level model extends the more traditional two-level meta-analytic model by providing separate estimates of sampling variance, within-study variance and between-study variance (Assink et al., 2018). If the three-factor model significantly improved the model fit, this model was accepted and used for the remainder of the analysis. If it did not, the two-factor model was retained for reasons of parsimony (which was ultimately the case). Sensitivity analyses were undertaken, including only one effect size per executive functioning subdomain, per study, as a means of exploring any possible risks related to the overestimation of particular studies in the final meta-analysis. Separate analyses based on executive function tasks reporting accuracy data and tasks reporting reaction time

data were conducted, with the most utilised tasks/outcome variables within this meta-analysis used to guide the selection of the included effect.

For the first stage of the analysis, the standardised mean differences were calculated between children/adolescents with depression and healthy children/adolescents for both overall executive function and each of the three executive functioning subdomains. Within this meta-analysis, a degree of variance is anticipated due to methodological disparities in the primary studies, measurement errors or uncontrolled individual difference factors across the body of literature (Higgins et al., 2003). Higgins I^2 was selected as the measure of heterogeneity, in which higher I^2 values reflect a greater proportion of variability being attributed to heterogeneity. Within this meta-analysis, I^2 values exceeding 75% were defined as problematic heterogeneity. This 75% threshold represents the boundary for high heterogeneity in Higgins original description of the index (Higgins et al., 2003). Additionally, a "leave-one-out" analysis was undertaken, with Baujat plots reviewed, as a means of identifying any significantly influential and discrepant effects within the primary studies. Any identified effects led to the study being re-reviewed to explore any reasons that may account for its discrepancy and/or disproportionate influence and were subsequently removed from the meta-analysis, providing a key reason could be identified. A subgroup analysis on the risk of bias ratings was also calculated to assess the impact of study level risk of bias upon heterogeneity.

For the second stage of the analysis further subgroup analyses were conducted to explore the impact of executive function task outcome measures (reaction time or accuracy), the use of medication in depressed participants and the presence of comorbidities in depressed participants on the effect size. Meta-regressions of depression sample sex proportion, by

study, depression sample average age, by study and depression sample average Full Scale Intelligence Quotient (FSIQ), by study were further included to examine the effects of participant sex, age, and intellectual ability.

Finally, publication bias and the influence of small sample size was assessed using Egger's test (Egger et al., 1997) and visual inspection of created funnel plots.

Results

The results section is divided into three sections. The first section summarises the results of the search, alongside the characteristics and overall quality of the included studies. The remaining sections describe the omnibus analysis and subdomain meta-analytic results, alongside the subsequent subgroup analyses and meta-regressions conducted.

Summary of the Literature Search, Study Content and Quality

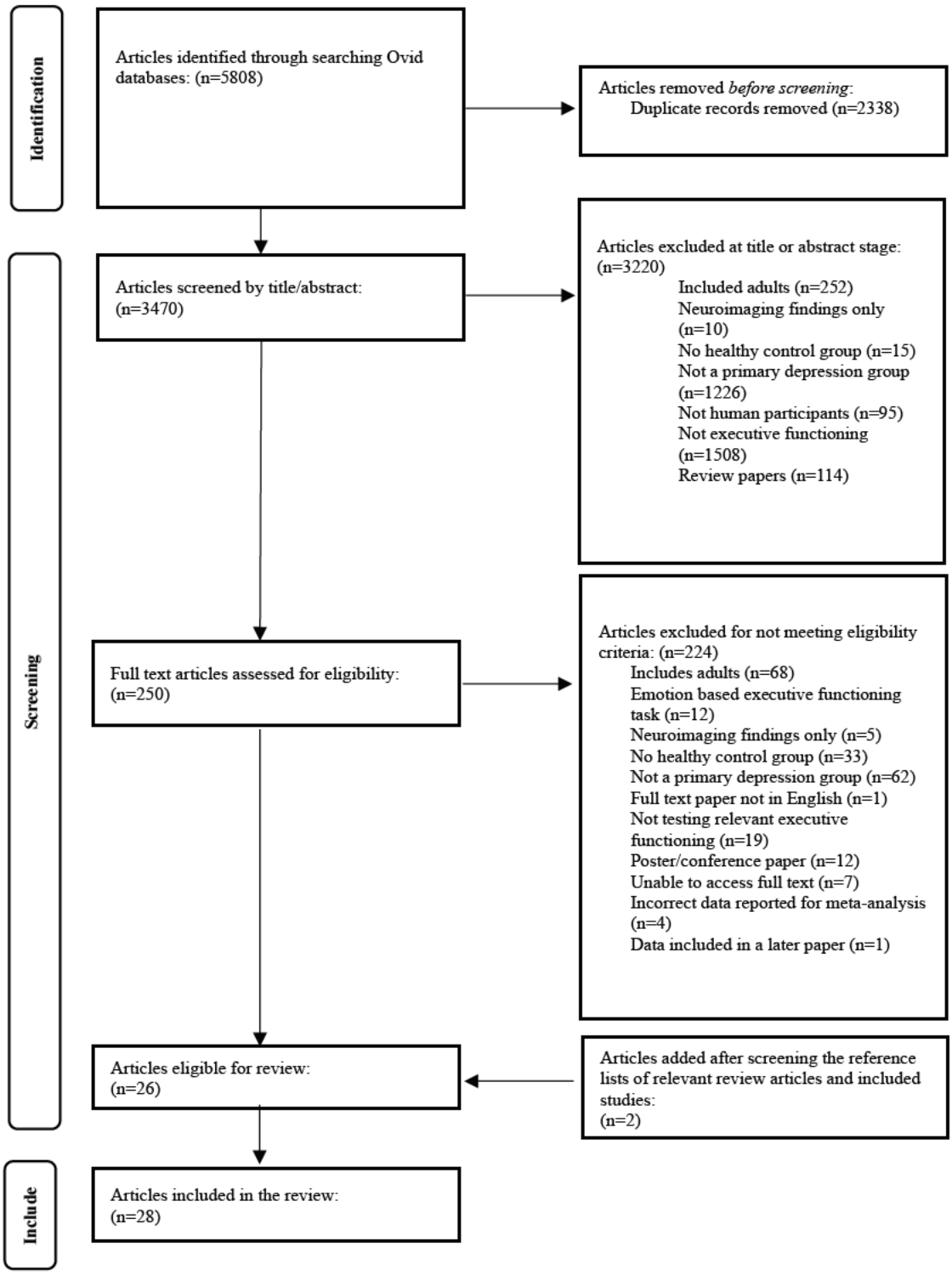
Results of the Search

The search (see Figure 1-1) returned 5808 articles, of which 2338 duplicate records were removed. The remaining 3470 articles were then screened by the researcher using their titles and abstracts, resulting in a further 3220 articles being excluded. The three primary reasons for exclusion at this stage included not testing executive function (k = 1508), not having a primary participant group with depression (k = 1226) and articles including adult participants in their sample (k = 252). For the remaining 250 articles the full text was accessed. Four authors were contacted via email during full-text screening, regarding the

acquisition of relevant data. However, there were no responses to either the initial or follow-up emails, leading to those articles being excluded. In total, 26 articles fully met the inclusion criteria, alongside not meeting any of the exclusion criteria. A further two articles were added in October 2023, after screening the reference lists of included articles and related meta-analyses/systematic reviews. Therefore, a total of 28 articles were deemed eligible and relevant to be included within this meta-analysis.

Figure 1-1

Flow Diagram of the Application of the Inclusion/Exclusion Criteria to the Results of the Electronic Search



Note. Adapted from "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews" by Page et al, 2021, Systematic Reviews, 10(1), p. 8 (https://doi.org/10.1186/s13643-021-01626-4). CC BY

Study and Participant Characteristics

The characteristics of each of the 28 studies included in the meta-analysis are summarised in Table 1-5. Studies compared executive function performance between children/adolescents with depression and healthy children and adolescents across a total of 1958 unique participants (n = 1027 across the depression group and n = 931 across the healthy control group). The mean age of included participants with depression ranged from 9.68 – 16.8 years; the mean age of included healthy controls ranged from 9.68 – 16.51 years. A higher proportion of studies (k = 18) reported over 50% of their sample being female. The presence of comorbidities in the depression group tended to be permitted (k = 20). The findings were more mixed regarding the use of psychotropic medication in participants with depression, in which 14 studies identified it as part of their exclusion criteria. The average FSIQ of the included participants with depression was not consistently reported, with only 12 studies providing a value. Finally, the included studies were undertaken across various worldwide locations both inside and outside of Europe.

Table 1-5

Table Summarising the Characteristics of All Included Studies

Study	Study location	*N* participants (% male)	Mean age (range)	Measure of depression	Measure(s) of executive function (variable; outcome type)	Comorbidities in the depression group	Medication use in the depression group	Mean FSIQ in the depression group (SD)
Bloch et al., (2013)	Israel	20 (25%) DG 20 (30%) HC	15.42 (DNR) DG 15.75 (DNR) HC	Diagnostic interview, CDRS-R, BDI youths	SWM task (between search errors; accuracy) SSP (span length score; accuracy) ID/ED set-shifting task (stage reached; accuracy and error number; accuracy)	Present	Excluded	DNR
Bloch et al., (2015)	DNR	13 (21%) DG 19 (DNR) HC	15.5 (13-18 years) DG Matched to the DG	CDRS-R, BDI for youths	SWM task (between search errors; accuracy) SSP (span length score; accuracy)	Present	Included	DNR
Brooks et al., (2010)	America	30 (26.7%) DG 30 (matched to DG for sex) HC	14.6 (9-17 years) DG Matched to DG on age	Interview based on DSM-IV criteria	Stroop task (incongruent reaction time) Shifting attention test (errors; accuracy and reaction time)	Present	Included	DNR
Cataldo et al., (2005)	Italy	21 (52%) DG 21 (52%) HC	12 (9-17 years) DG 11.7 (9-17 years) HC	Italian version of the Diagnostic Interview for Children and Adolescents-Revised	Walk-don't walk test (total number of correct paths; accuracy) Stroop task (interference score; RT)	Present	Excluded	DNR
Constantinidou et al., (2011)	America	11 (DNR) DG 13 (DNR) HC	10.39 (8-12 years) DG 10.02 (7-11 years) HC	K-SADS	CLVT for Children (List one items recalled correctly; accuracy	DNR	DNR	105.54 (10.71)

Study	Study location	*N* participants (% male)	Mean age (range)	Measure of depression	Measure(s) of executive function (variable; outcome type)	Comorbidities in the depression group	Medication use in the depression group	Mean FSIQ in the depression group (SD)
Diller et al., (2014)	America	10 (20%) DG 10 (20%) HC	15.9 (12-17 years) DG 15.6 (12-17 years) HC	K-SADS-PL, CDRS-R	Go-no/go task (false alarms; accuracy)	Present	Included	DNR
Favre et al., (2009)	America	39 (51%) DG 24 (46%) HC	12.79 (8-17 years) DG 13.08 (9-17 years) HC	Interview based on DSM-IV criteria, CDRS-R	WCST (Perseverative errors; accuracy) TMT-B (response time and error score; accuracy) Stroop task (interference score; accuracy)	Present	Excluded	105.29 (14.28)
Fisk et al., (2019)	United Kingdom	29 (17.2%) DG 29 (17.2%) HC	14.83 (12-18 years) DG 14.65 (12-18 years) HC	MFQ	HSC task (error number; accuracy and reaction time) KTT (total words recalled correctly; accuracy)	DNR	DNR	92 (8.83)
Franklin et al., (2010)	Australia	26 (77%) DG 28 (61%) HC	9.96 (6-12 years) DG 10.11(6-12 years) HC	ADIS-C, CBCL, CDI	SWM task (between search errors; accuracy) SSP task (span length score; accuracy)	Not present	Excluded	DNR
Gunther et al., (2004)	Germany	31(55%) DG 33 (45%) HC	13.5 (DNR) DG 12.8 (DNR) HC	K-SADS, CDI	RAVLT(Trial one items recalled correctly; accuracy)	Present	Excluded	101 (13)

Study	Study location	*N* participants (% male)	Mean age (range)	Measure of depression	Measure(s) of executive function (variable; outcome type)	Comorbidities in the depression group	Medication use in the depression group	Mean FSIQ in the depression group (SD)
Gunther et al., (2011)	Germany	61 (48%) DG 64 (58%) HC	13.8 (10-15 years) DG 13.6 (10-15 years) HC	Interview based on DSM-IV criteria	Go/no-go task (false alarms; accuracy) Visual set-shifting task (reaction time and errors; accuracy)	Present	Excluded	97 (14.7)
Halari et al., (2009)	United Kingdom	21 (48%) DG 21(matched to DG for sex) HC	16.2 (14-17 years) DG 16.3 (14-17 years) HC	K-SADS	Switch task (switch cost; RT) Simon task (Simon effect; RT) Stop task (stop signal reaction time; RT)	Not present	Excluded	DNR
Hanna et al., (2018)	America	36 (19%) DG 89 (22%) HC	16.8 (13-18 years) DG 16.2 (13-18 years) HC	Interview based on DSM-IV criteria	Flanker task (accuracy and reaction time for incongruent trials)	Present	Included	DNR
Klimkeit et al., (2011)	Australia	34 (32%) DG 33 (27%) HC	15.41 (12.08-17.83 years) DG 15.8 (12.5-17.83 years) HC	K-SADS-PL	WISC-IV digit span (total score; accuracy)	Present	Excluded	DNR
Maalouf et al., (2011)	America	20 (15%) DG 17 (47%) HC	15.3 (DNR) DG 15.2 (DNR) HC	K-SADS, CDRS-R	DMTS task (accuracy)	Present	Included	105 (11)

Study	Study location	*N* participants (% male)	Mean age (range)	Measure of depression	Measure(s) of executive function (variable; outcome type)	Comorbidities in the depression group	Medication use in the depression group	Mean FSIQ in the depression group (SD)
Matthew et al., (2008)	United Kingdom	14 (0%) DG 14 (0%) HC	14.49 (DNR) DG 14.36 (DNR) HC	Semi-structured CAPA-C, MFQ	SWM task (between search errors; accuracy) SSP (span length score; accuracy) DMTS task (accuracy) ID/ED set-shifting task (stage reached; accuracy)	Present	Excluded	DNR
Onat et al., (2019)	Turkey	62 (11%) DG 30 (30%) HC	15.89 (13-18 years) DG 15.11(13-18 years) HC	K-SADS, CDRS-R	Stroop task (incongruent reaction time and errors; accuracy)	Not present	Included	91.78 (DNR)
Pan et al., (2011)	America	30 (37%) DG 14 (47%) HC	16.04 (DNR) DG 15.21 (DNR) HC	K-SADS, BDI for youths	Go-no/go task (false alarms; accuracy)	Present	Included	DNR
Pan et al., (2020)	China	28 (39%) DG 24 (33%) HC	16 (13-17 years) DG 16 (13-17 years) HC	Interview based on DSM-IV criteria	WCST (Perseverative errors; accuracy) TMT-B (response time; RT)	Not present	Excluded	DNR
Pandina, (1999)	America	19 (79%) DG 19 (74%) HC	9.68 (6-13 years) DG 9.68 (6-13 years) HC	K-SADS, CBCL	WCST WISC-III TMT-B Stroop task Verbal fluency (composite score)	Not present	Included	99.83 (10.79)
Peters et al., (2019)	America	18 (44.4%) DG	14.61 (12-17 years) DG	K-SADS, CDRS-R	Go/no-go task (inhibitory accuracy score)	Present	Excluded	DNR

Study	Study location	*N* participants (% male)	Mean age (range)	Measure of depression	Measure(s) of executive function (variable; outcome type)	Comorbidities in the depression group	Medication use in the depression group	Mean FSIQ in the depression group (SD)
		30 (36.7%) HC	14.67 (12-17 years) HC					
Shebab et al., (2016)	Lebanon	24 (33%) DG 24 (42%) HC	14.83 (12-18 years) DG 14.25 (12-18 years) HC	DAWBA, CDRS-R	DMTS task (accuracy)	Present	Excluded	DNR
Shin et al., (2008)	Korea	20 (20%) DG 23 (87%) HC	11.37 (6-16 years) DG 10.13 (6-16 years) HC	Interview based on DSM-IV criteria	WCST (Perseverative errors; accuracy) TMT-B (response time and error score; accuracy) WISC-R digit span (total score; accuracy)	Not present	Included	106.2 (9.53)
Vance & Winther (2021)	Australia	313 (63%) DG 107 (43%) HC	12.2 (6-16 years) DG 10.62 (6-16 years) HC	ADIS-C, CDI, CDS	SWM task (between search errors; accuracy) SSP (span length score; accuracy)	Present	Excluded	91.53 (DNR)
Vijayakumar et al., (2016)	Australia	23 (26%) DG 122 (52%) HC	16.48 (DNR) DG 16.51 (DNR) HC	K-SADS	Stroop task (reaction time and accuracy interference scores)	Present	DNR	107.35 (11)
Vilgis et al., (2014)	Australia	19 (100%) DG 16 (100%) HC	11.2 (8-13.5 years) DG 10.5 (8-13.5 years) HC	ADIS-C, CBCL, CDI	DMTS task (accuracy and reaction time)	Present	Excluded	95.6 (13.3)

Study	Study location	*N* participants (% male)	Mean age (range)	Measure of depression	Measure(s) of executive function (variable; outcome type)	Comorbidities in the depression group	Medication use in the depression group	Mean FSIQ in the depression group (SD)
Vilgis et al., (2022)	DNR	16 (100%) DG 19 (100%) HC	12.94 (DNR) DG 13.04 (DNR) HC	Interview based on DSM-IV criteria, CBCL, CDI	Verbal and spatial Sternberg working memory task (reaction time and accuracy)	Present	Included	DNR
Wilkinson & Goodyer, (2006)	United Kingdom	39 (25%) DG 38 (29%) HC	15 (11 years – 17 years 11 months) DG 14.8 (11 years – 17 years 11 months) HC	Current DSM-IV diagnosis of MDD	Opposite worlds task (opposite worlds response latency; RT)	Present	Included	DNR

Note: DG: Depression Group; HC: Healthy Controls; DSM-IV*: Diagnostic and Statistical Manual of Mental Disorders (4th ed.)*; CDRS-R: Children's Depression Rating Scale-Revised; BDI: Beck Depression Inventory; SWM: Spatial Working Memory; SSP: Spatial Span; ID/ED: Intra-Dimensional–Extra- Dimensional; K-SADS: Kiddie Schedule for Affective Disorders and Schizophrenia; CLVT: California Verbal Learning Test; WCST: Wisconsin Card Sorting Test; TMT-B: Trail Making Test-Part B; MFQ: Mood and Feelings Questionnaire; HSC: Hayling Sentence Completion; KTT: Keep Track Task; RAVLT: Rey Auditory–Verbal Learning Test; DMTS: Delayed Match To Sample; CAPA-C: Child and Adolescent Psychiatric Assessment; DAWBA: Development and Well-Being Assessment; CDS: Children's Depression Scale; ADIS-C: Anxiety Disorders Interview Schedule for Children; CBCL: Child Behaviour Checklist; CDI: Children's Depression Inventory; RT: Reaction Time; DNR: Did Not Report; SD: Standard Deviation; WISC-R: Wechsler Intelligence Scale for Children-Revised; WISC-III: Wechsler Intelligence Scale for Children-Third Edition; WISC-IV: Wechsler Intelligence Scale for Children-Fourth Edition.

Risk of Bias Review

Table 1-6 presents a visual matrix of the application of the risk of bias framework for each of the primary studies, alongside their overall risk of bias score. Further descriptions outlining how each of the included studies met each element of the quality criteria framework are also provided.

Table 1-6

Ratings of Risk of Bias Across the Primary Studies

Study name	Selection bias	Depression measurement bias	Executive function measurement bias	Reporting bias	Generalisability	Risk of bias score
Bloch et al. (2013)	High risk	Low risk	Low risk	Unclear risk	High risk	5
Bloch et al. (2015)	High risk	High risk	Low risk	Low risk	High risk	4
Brooks et al. (2010)	Unclear risk	Low risk	Low risk	Unclear risk	Low risk	8
Cataldo et al. (2005)	High risk	Low risk	Low risk	Low risk	Unclear risk	7
Constantinidou et al. (2011)	High risk	Low risk	Low risk	Low risk	High risk	6
Diler et al. (2014)	High risk	Low risk	Low risk	Low risk	High risk	6
Favre et al. (2009)	High risk	Low risk	Low risk	Unclear risk	Low risk	7
Fisk et al. (2019)	High risk	Unclear risk	Low risk	Low risk	Low risk	7
Franklin et al. (2010)	High risk	Low risk	Low risk	Low risk	Unclear risk	7
Günther et al. (2004)	High risk	Low risk	Low risk	Low risk	Low risk	8
Günther et al. (2011)	Unclear risk	Low risk	Unclear risk	Low risk	Unclear risk	7

Study name	Selection bias	Depression measurement bias	Executive function measurement bias	Reporting bias	Generalisability	Risk of bias score
Halari et al. (2009)	Unclear risk	Low risk	Unclear risk	Low risk	Unclear risk	7
Hanna et al. (2018)	Unclear risk	Low risk	Low risk	Unclear risk	Low risk	8
Klimkeit et al. (2011)	High risk	Low risk	Low risk	Low risk	Low risk	8
Maalouf et al. (2011)	Unclear risk	Low risk	Low risk	Low risk	High risk	7
Matthews et al. (2008)	High risk	Low risk	Low risk	Unclear risk	High risk	5
Onat et al. (2019)	High risk	Low risk	Unclear risk	Unclear risk	Unclear risk	5
Pan et al. (2011)	Unclear risk	Low risk	Low risk	Low risk	High risk	7
Pan et al. (2020)	High risk	Low risk	Low risk	Unclear risk	Unclear risk	6
Pandina (2000)	Unclear risk	Low risk	Low risk	Low risk	High risk	7
Peters et al. (2019)	Unclear risk	Low risk	Low risk	Unclear risk	High risk	6
Shehab et al. (2016)	High risk	Low risk	Low risk	Low risk	Low risk	8
Shin et al. (2008)	High risk	Low risk	Low risk	Unclear risk	High risk	5
Vance and Winther (2021)	High risk	Low risk	Low risk	Low risk	Low risk	8
Vijayakumar et al. (2016)	High risk	Low risk	Low risk	Low risk	Low risk	8
Vilgis et al. (2014)	High risk	Low risk	Low risk	Low risk	High risk	6
Vilgis et al. (2022)	High risk	Low risk	Unclear risk	Low risk	High risk	5
Wilkinson and Goodyer (2006)	Unclear risk	Low risk	Low risk	Low risk	Low risk	9

Note. Red equates to "high risk" of bias, amber equates to "unclear risk" of bias and green equates to "low risk" of bias.

Selection Bias. Regarding selection bias, nine studies were rated as "unclear risk" of bias and 19 were rated as "high risk" of bias. Studies with "unclear risk" of bias mainly used convenience sampling, however, recruited from multiple non-random samples (e.g., multiple clinics) whilst matching the two participant groups on key variables (e.g., age). The "high risk" studies either did not specify their participant selection methods, used single non-random sampling for one or both groups and/or did not employ any matching attempts between the two participant groups. No studies were rated as "low risk" of bias, due to being unable to recruit a truly random sample.

Depression Measurement Bias. Almost all studies (k = 26) were rated as "low risk" of depression measurement bias due to using standardised measures that are well validated in a child/adolescent population to confirm allocation into the depression group. One study was rated as "unclear risk" of bias as the authors used a self-report measure to classify the two groups (Fisk et al., 2019). One study did not clearly specify how the depression group was confirmed and was therefore rated as "high risk" of bias (Bloch et al., 2015).

Executive Function Measurement Bias. Generally, the executive function measurement(s) used within the included studies was/were considered to be good, with 24 being classified as "low risk" of bias. These studies used well-validated tests of executive function (e.g., Stroop interference test). Four studies were found to have an "unclear risk" of bias, resulting from their use of executive function tasks that had either only been validated in adult samples or where their validity in a child/adolescent population was unclear (Günther et al., 2011; Halari et al., 2009; Onat et al., 2019; Vilgis et al., 2022).

Reporting Bias. In total, 19 of the included studies reported the key executive function variables recommended in the literature for the tasks that they used and were thus classified as "low risk" of reporting bias (e.g., perseverative errors in the WCST). The remaining nine studies were found to be of "unclear risk" of bias due to reporting related but not the recommended executive function variable for their included tasks or reporting only error number or response time when both would be expected. It is noted that some potentially relevant studies were excluded at the screening stage, due to not reporting the correct data required for this meta-analysis.

Generalisability. Small sample sizes resulted in 12 studies being rated as "high risk" of bias, due to having 20 or fewer participants in the depression and/or healthy control group. This can result in any statistical analysis lacking power and limits the ability to apply the results found within these studies to other children/adolescents with depression. Additionally, in some studies significant idiosyncratic differences were identified within the sample, such as only one sex being focused on (Matthew et al., 2008; Vilgis et al., 2014, 2022), a specific type of depression being explored (e.g., dysthymia; Franklin et al., 2009) or a significant amount of suicide attempters being included in the depression group (Onat et al., 2019), resulting in a rating of "unclear risk" of bias for six studies. There is potential for these participants to differ systematically from the wider child/adolescent population with depression. Finally, 10 studies appeared to have a "low risk" of bias in terms of generalisability, by using sample sizes of greater than 20 in each of the two groups and having a depression participant group with limited idiosyncratic features.

Summary of Risk of Bias Assessment. In summary, there was a mixed level of bias across all 28 included studies. Only five studies were not classified as having a "high risk" of bias across any of the quality domains. The risk of bias assessment highlighted that adequate depression assessment and executive function tasks appeared to be used, through studies predominantly using well-validated measures in both domains. There was a notably higher risk of bias across studies in terms of participant selection methods. Although convenience sampling is commonly employed within psychological research studies (Nielsen et al., 2017), it results in a lack of randomisation which may adversely affect the study's precision of measurement. For example, Kunz and Oxman (1998) reported that a "failure to use random allocation and concealment of allocation were associated with relative increases in estimates of effects of 150% or more, relative decreases of up to 90%, [and] inversion of the estimated effect". Therefore, caution may be required when interpreting any findings or conclusions drawn. In the areas of reporting of executive function variables and generalisability, the level of bias was variable. In some cases, samples appeared to be generalisable, and the recommended executive function variable was reported. However, in other cases, it was unclear if the samples enabled generalisation to the wider child/adolescent depression population due to sample size restrictions and/or idiosyncratic differences, alongside if the task variable reported best captured relevant executive function performance.

The overall risk of bias scores ranged from 4-9. Studies assessed as having an "unclear" or "high risk" of bias remained in the meta-analysis due to there being a reduced amount of literature in this field. Therefore, the included studies were considered to be representative of the current evidence-base.

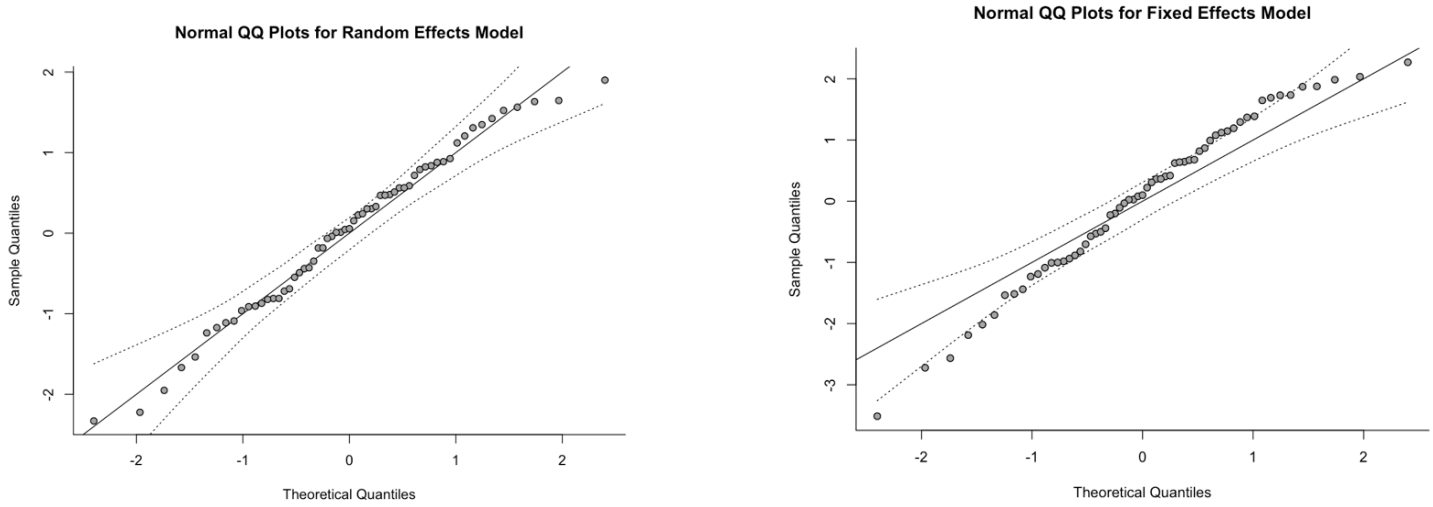
Meta-Analysis Results

Selection of the Meta-Analytic Model

The QQ plots for both the fixed and REM are shown in Figure 1-2. As can be seen from Figure 1-2 there is some evidence of non-linearity and non-normality in the distribution of primary study effects using the fixed effects model, which is absent when the REM is used. Therefore, this indicated that the use of the REM is an appropriate method for the calculation of the weighted average effect. When considering the three-level meta-analytic model, in each instance it did not significantly improve model fit (see Appendix 2). Therefore, for reasons of parsimony, the final analysis used the two-level REM to estimate the effect, in which Cohen's *d* was the selected effect size index and forest plots were produced. The between-studies variance (tau^2) in the random effects model was calculated using the restricted maximum likelihood estimator, which has been shown to be more robust to deviations from normality (Banks et al., 1985). Findings from the sensitivity analyses were consistent with the overall and subdomain effects, impact of influential primary effects and evidence of publication and small study bias. As these analyses suggested inclusion of multiple effects for individual studies made no substantive difference to the central conclusions, the larger analysis (inclusive of all effects) is included here, with the sensitivity analyses in Appendix 3.

Figure 1-2

QQ plots of the Distribution of the Standardised Mean Difference Within the Primary Studies Using Both the Fixed Effects Model (right) and the Random Effects Model (left).



Overall Executive Function

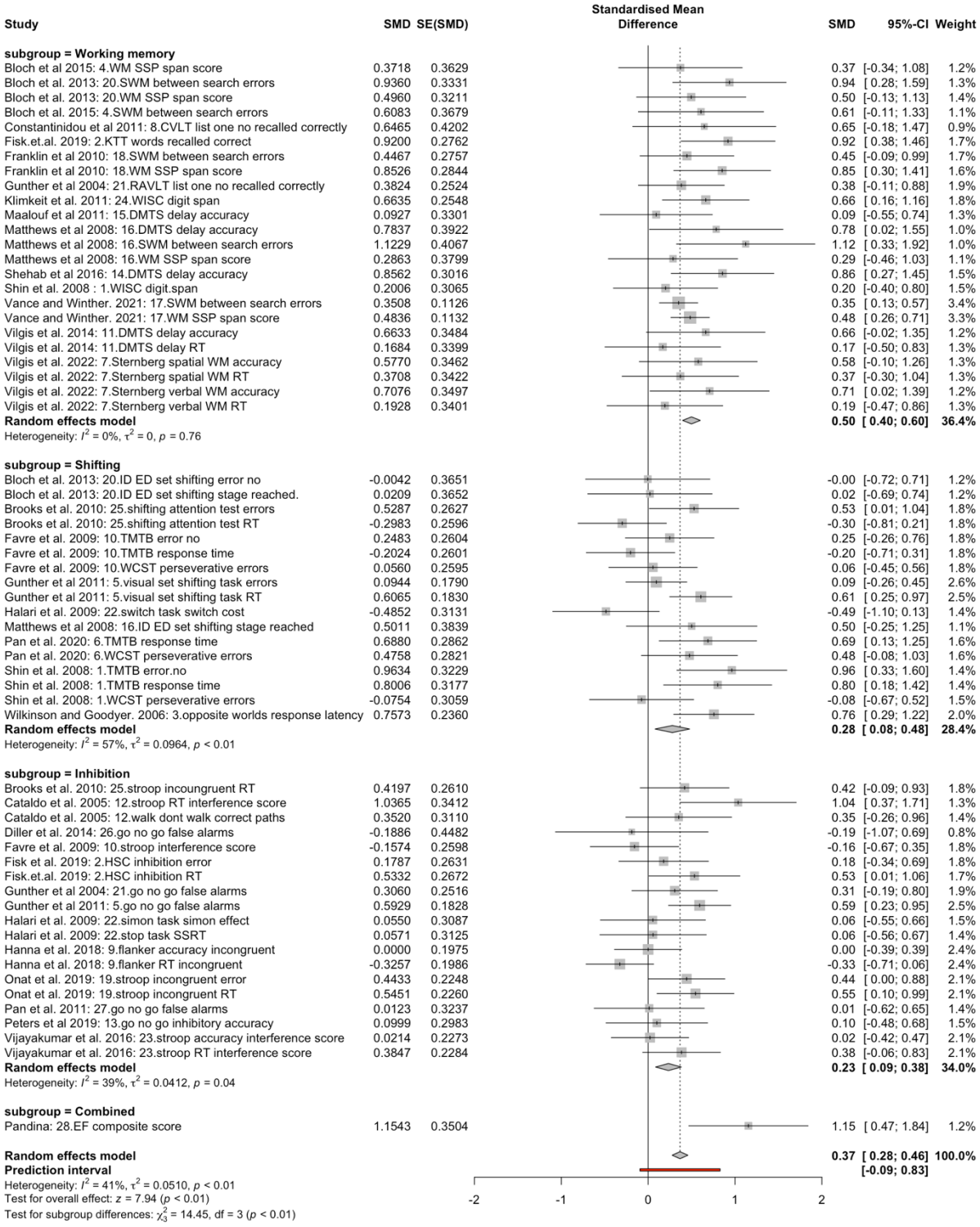
This section reports differences in executive function performance between children/adolescents with depression and healthy controls, using the 61 effects reported across the 28 primary studies. See Appendix 4, for the weighted average standardised mean difference for each of the individual effects reported. The forest plot in Figure 1-3 provides separate estimates for each of the investigated executive function subdomains (shifting, working memory and inhibition), alongside the overall effect across all domains. In total 17 included effects investigated shifting, 24 effects investigated working memory, 19 effects investigated inhibition, and one effect used a composite executive functioning score.

The random effects model suggested a weighted average standardised mean difference for executive function between children/adolescents with depression and healthy

children/adolescents of *SMD* = 0.37 ($z = 7.94$, $p < .0001$) and a 95% confidence interval between 0.28 to 0.46. The positive overall effect reported in this instance indicates worse executive function performance in children/adolescents with depression compared to healthy children/adolescents.

Figure 1-3

Forest Plot of the Weighted Average Standardised Mean Difference of Executive Functioning Ability Between Children and Adolescents with Depression and Healthy Children and Adolescents, Separated by Executive Functioning Subdomain



Note. Diamonds indicate overall meta-analytic effect sizes. Positive effect values reflect better executive function performance for healthy controls. All acronyms are defined in the note for Table 1-5.

Effect of Executive Functioning Subdomain

The random effects model (see Figure 1-3 above) identified a significant difference (X^2 = 14.45, p<.01) between the different subdomains of executive functioning. In each subdomain children/adolescents with depression performed worse on executive function tasks than healthy children and adolescents, with working memory (*SMD* = 0.50, 95%*CI* 0.40 to 0.60) showing the greatest difference between children/adolescents with depression and healthy children/adolescents. Alternatively, the standardised mean differences for inhibition (*SMD* = 0.23, 95%*CI* 0.09 to 0.38) and shifting (*SMD* = 0.28, 95%*CI* 0.08 to 0.48) were observed to be of similar magnitude. Post-hoc subgroup comparisons compared the subgroup differences. The effect size in the working memory subdomain was found to be significantly larger than that in the inhibition subdomain (X^2 = 9.03, p <.01) and showed a trend for being larger than the shifting subdomain (X^2 = 3.72, p = 0.05). There was no significant difference found between the shifting and inhibition subdomain (X^2 = 0.16 p = 0.69). Overall, only one study reported differences in an executive function composite score, in which children/adolescents with depression were also significantly impaired relative to healthy controls.

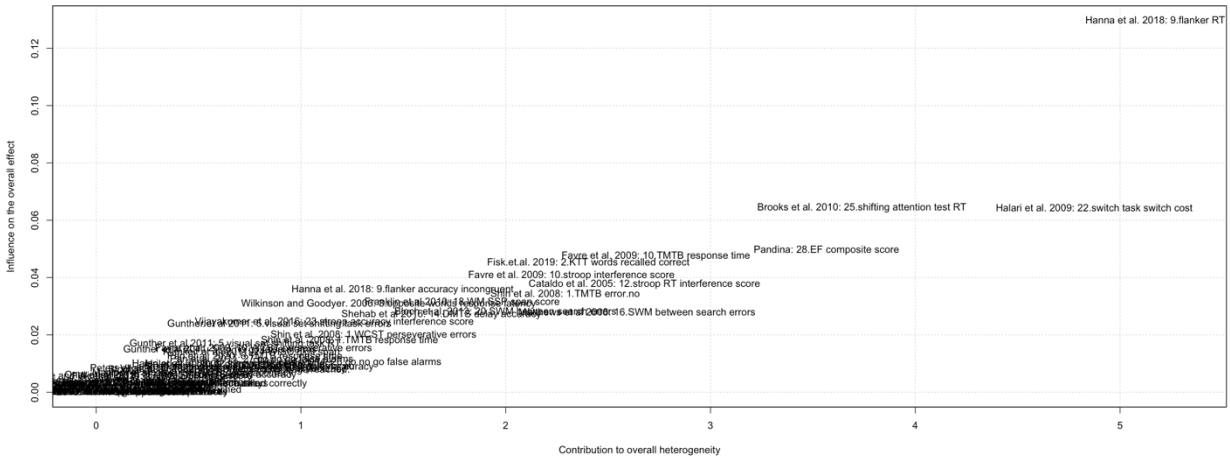
An acceptable level of heterogeneity in the primary studies was observed in the omnibus analysis and across all three executive subdomains investigated: omnibus (Higgins I^2 = 41%), shifting (Higgins I^2 = 57%,), working memory (Higgins I^2 = 0%), and inhibition (Higgins I^2 = 39%). This suggests that the level of variation in the primary studies is consistent with that expected of sampling variation and that these studies are reporting a coherent and consistent effect size for which the weighted average is a valid summary.

The Impact of Influential Primary Effects

Despite an acceptable level of heterogeneity observed across the included effects from the primary studies (defined by Higgins I^2 <75%), a "leave-one-out" analysis was conducted to assess the impact of any discrepant or disproportionality influential primary effects, identified through the construction of a Baujat plot, including each individual effect (see Figure 1-4). The REM was re-calculated with each influential individual effect systematically removed. This allowed for the calculation of the change in weighted average standardised mean difference effect size and the change in heterogeneity.

Figure 1-4

Baujat Plot of Sources of Heterogeneity



Note. The vertical axis represents the effect's influence on the overall effect, whilst the horizontal axis represents the effect's discrepancy with the rest of the included literature.

As highlighted in Figure 1-4, the flanker incongruent reaction time effect in Hanna et al. (2018) was observed to be both influential on the overall effect and discrepant from the

majority of the current literature. Furthermore, both the shifting attention test, reaction time effect in Brooks et al. (2010) and the switch task, switch cost effect in Halari et al. (2009) could also be considered slightly influential and discrepant. Table 1-7 demonstrates the impact of these three effects being removed from the analysis, in which there was a marginal impact observed on the overall effect size (all ≤ 4.31%) and heterogeneity. The studies were also re-reviewed, in which there was limited methodological variation in comparison to other primary studies included within this meta-analysis. Therefore, based on the above all three effects were retained in the final analysis.

Table 1-7

Summary of Influential Effects on Executive Function Performance

	SMD	**95% CI**	**p-value**	**tau^2**	**I^2**
Omitting Hanna et al. (2018) Flanker RT incongruent	0.3828	[0.2970; 0.4686]	< 0.0001	0.0363	34.5%
Omitting Brooks et al. (2010) shifting attention test RT	0.3785	[0.2896; 0.4675]	< 0.0001	0.0447	38.4%
Omitting Halari et al. (2009) switch task switch cost	0.3785	[0.2899; 0.4671]	< 0.0001	0.0442	37.9%

The Effect of Risk of Bias in the Primary Studies

The impact of study level risk of bias on heterogeneity was assessed using a series of subgroup analyses, conducted on the standardised mean differences across the individual studies for the risk of bias ratings of "low risk" and "any risk" (i.e., "unclear risk" and "high

risk" of bias combined) for each of the five types of methodological bias. See Table 1-8 for a summary of the results.

Table 1-8

The Effect of Risk of Bias in the Primary Studies

	Low Risk			Any Risk				
	SMD	**95% CI**	**k**	**SMD**	**95% CI**	**k**	**X^2**	**P**
Selection bias	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Depression measurement	0.4141	[0.2836; 0.5445]	26	0.3241	[-0.0954; 0.7436]	2	0.16	0.69
Executive function measurement	0.4161	[0.2887; 0.5435]	24	0.3341	[-0.1511; 0.8192]	4	0.10	0.75
Reporting bias	0.4725	[0.3317; 0.6133]	19	0.2833	[0.0793; 0.4873]	9	2.24	0.13
Generalisability bias	0.4077	[0.2444; 0.5711]	10	0.4079	[0.2209; 0.5949]	18	0	1.00

Values for selection bias were not included as all studies were rated as "any risk" of bias and the comparison with "low risk" studies was therefore not possible. As can be seen from the table above, none of the remaining areas of risk of bias showed statistically significant differences in the overall weighted average standardised mean differences. These findings therefore suggest that any heterogeneity across the included studies was not primarily the result of methodological quality.

Further Subgroup Analysis and Meta-regression

For this meta-analysis a further three subgroup analyses were completed, alongside three meta-regressions, to test for the impact of various factors on the magnitude of the observed effect.

The Impact of Executive Function Task Outcome Measure

Across the primary studies, 42 included effects reported accuracy data and 18 included effects reported reaction time data. As can be seen from Table 1-9, there was no significant difference in the observed effect between executive function tasks that reported efficiency outcomes and tasks that reported effectiveness outcomes ($X^2 = 0.63$, $p = 0.43$).

Table 1-9

The Impact of Executive Function Task Outcome

Level	Random effects SMD	95% CI	k	X^2	p
Accuracy Data	0.3812	[0.2909; 0.4714]	42	0.63	0.43
Reaction time data	0.2925	[0.0924; 0.4926]	18		

The Impact of Participant-Level Characteristics

As can be seen from Table 1-10 there was no significant difference on executive function performance, for the inclusion of medication ($X^2 = 0.12$, $p = 0.72$) nor for the inclusion of comorbidities ($X^2 = 0.75$, $p = 0.39$) within the child/adolescent group experiencing depression.

Table 1-10

The Impact of Participant-Level Characteristics

	Level	Random effects SMD	95% CI	k	X^2	p
Medication	Included	0.3410	0.1676; 0.5145	22	0.12	0.72
	Excluded	0.3782	0.2663; 0.4900	33		
Comorbidity	Included	0.3292	0.2254; 0.4329	43	0.75	0.39
	Excluded	0.4358	0.2185; 0.6532	14		

Meta-Regression Results

Meta-regressions were carried out to investigate whether study-level, average age, average FSIQ and male % of participants with depression predicted effect size (see Table 1-11). In this context, 'study-level effects' refer to the meta-regressions analysing differences between studies based on aggregated demographic characteristics reported for the entire sample, rather than identifying individual participant data/demographics. Age was a significant predictor of the effect size, (z = -2.48, p = .01), with studies including (on average) older children/adolescents with depression finding smaller effect sizes. FSIQ was also a significant predictor, with studies including children/adolescents with depression with a higher average FSIQ returning smaller effect sizes (z = -2.04, p = .04). The % of male participants with depression was not a significant predictor of the effect size (z = 0.39, p = 0.69).

Table 1-11

Meta-Regression Results for Age, FSIQ and Percentage of Male Participants with Depression

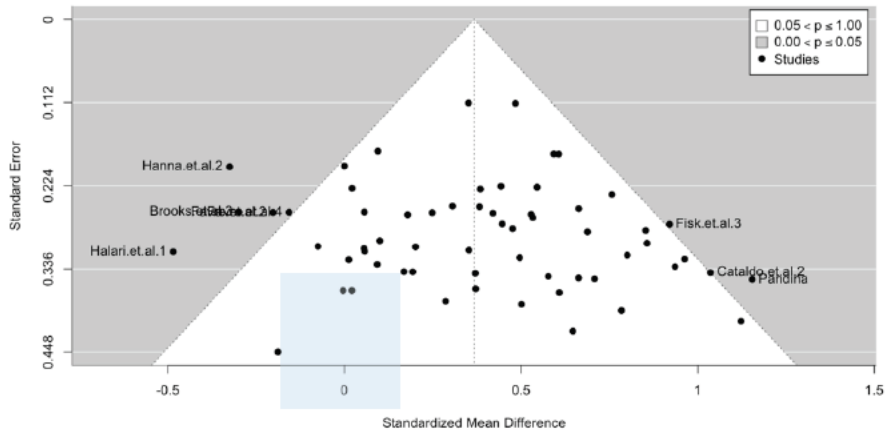
	Coefficient	**SE**	**Z**	**p**
Age	-0.0589	0.0238	-2.4769	0.0133
FSIQ	-0.0169	0.0083	-2.0365	0.0417
% male	0.0007	0.0019	0.3940	0.6935

The Impact of Publication and Small Study Biases

The funnel plot of standardised mean difference across all the included effects is presented in Figure 1-5. There was no clear evidence of publication bias in the distribution of the measures of executive performance. Most studies fall within the range expected of the meta-analytic weighted average effect and the distribution of study level effects appears to be symmetrical. Due to disputes regarding the precision of visual inspection of funnel plots (Terrin et al., 2005), Egger's statistical test was also employed (Egger et al., 1997), and was not significant ($t = 0.58$, $p = 0.56$). Small studies tended to report results both above and below the meta-analytic average, suggesting that there was no systematic bias associated with study precision. Finally, studies with small sample sizes were clearly represented in the shaded area of the graph associated with null effects. Therefore, no simulation of and adjustment for publication bias and small study effects was carried out.

Figure 1-5

Funnel Plot of the Standardised Mean Difference Across All Included Effects



Note. The 95% confidence interval of the expected distribution of executive functioning is shown as an inverted "funnel". The area of the plot shaded in blue is the area associated with publication bias.

Discussion

Due to the inconsistent research findings regarding executive function impairments in children/adolescents with depression, this meta-analysis aimed to examine both general executive function performance and performance across its three specific subdomains (inhibition, shifting and working memory) in children/adolescents with depression. An additional aim was to determine the degree to which variance in the association between executive function and younger-onset depression could be accounted for by factors, including the executive function task outcome reported (accuracy, reaction time) and various participant characteristics (age, comorbidities, intellectual functioning, sex, and psychotropic medication use).

Despite the variability in the methodology, heterogeneity across all the analyses was acceptable, suggestive of a consistent relationship between depression and executive function in children and adolescents. Overall, the results of this meta-analysis suggest that children/adolescents with depression demonstrate worse executive function than their healthy peers. There were significant differences between the three executive functioning subdomains, with working memory showing a larger effect (d = 0.50) in comparison to shifting (d = 0.28) and inhibition (d = 0.23). Whether studies reported executive function tasks that used accuracy or reaction time outcomes did not predict the effect size. There were also no differences observed in executive function outcomes between the two groups based on the presence of comorbidities, sex and psychotropic medication use in the group with depression. Conversely, there were differences observed for age and intellectual ability, with younger children with depression and children with lower FSIQs and depression showing bigger differences from their healthy peers.

Executive Function Impairment in Childhood/Adolescent Depression

Overall, child and adolescent participants with depression demonstrated significant deficits in executive function performance compared to healthy control participants, across all neuropsychological executive function measures. In line with the 'cognitive effort' hypothesis, it is possible that the disrupted motivation and related depressive symptoms (e.g., rumination) associated with childhood/adolescent depression (Roelofs et al., 2009) may substantially occupy an individual's central executive resources, thus leading to a reduced capacity in assigning cognitive resources towards more effortful executive function tasks (Steinberger & Barch, 2023). An additional explanation for the observed findings could relate to the hypothesis that premorbid early executive dysfunction may result in unregulated negative emotionality, due to executive function's important role in moderating the experience of negative mood and thoughts (Nelson et al., 2018). Dysregulated negative emotions and cognitions may lead to the development, intensification and/or perpetuation of depressive symptoms (Nelson et al., 2018). However, due to the cross-sectional nature of the included studies, it is difficult to disentangle whether impaired executive function is a state-related feature of depression, occurring as a consequence of depressive symptoms, or whether executive function impairments are a trait-based feature of the difficulty, representing an early depression risk-marker (Goodall et al., 2018; Steinberger & Barch, 2023). There is likely a complex, interconnected relationship between depression and executive function, and more research employing longitudinal designs and repeated executive function testing would allow for greater insight into the directional links between childhood/adolescent depression and executive function (Snyder, 2013).

Regarding the three executive functioning subdomains, children and adolescents with depression appeared to be most impaired in the working memory subdomain. There is a notable discontinuity between this study's results and the background literature focused on adults with depression, in which the inhibition subdomain tends to be slightly more impaired relative to the other executive function subdomains of shifting and working memory (Snyder, 2013; Wagner et al., 2012). There is currently a lack of clarity regarding what may account for this difference. It is possible that executive function skill development may be differentially impacted by early/first-onset depression, level of treatment exposure, or influenced by neurocognitive developmental stage. Therefore, further research is needed to explore why working memory may be particularly impaired in children/adolescents with depression.

This meta-analysis did not aim to provide a direct test of the conceptual framework and proposed inter-connections between the three subdomains in the unity and diversity model (Miyake et al., 2000). Observing significant impairments across all three executive function subdomains (working memory, inhibition and shifting) is consistent with these components being interrelated and influencing one another. This aligns with the unity aspect of the model, suggesting that while these subdomains are separable, they do not operate in isolation but instead interact dynamically (Miyake et al., 2000). That a subgroup difference between components was identified is consistent with some degree of diversity. Understanding the findings within the alternative proposals (Blakey et al., 2016; Hall & Fong, 2015) could be informative. The hypothesis that actively maintaining and managing task goals and task-related information in working memory serves as a central, foundational process in executive function (Friedman & Miyake, 2017) provides, perhaps, the most parsimonious fit to the current data. An implication of this theory is that factors that show convergence with working

memory skill should also explain variance in inhibition and shifting, but to a lesser degree. This is the pattern we observed linking young-onset depression to the subcomponents of executive function. To our knowledge, however, there is not a clear, a priori, reason to explain why this relationship would subsequently change in adulthood (where inhibition tends to be the most significantly impaired executive function subdomain, Snyder, 2013). In all probability, the bidirectional relationship between depression, executive function and its subdomains is likely multifaceted, with more fine-grained and longitudinal modelling required for a fuller understanding.

It is also important to note that within this meta-analysis working memory was conceptualised as a broad construct. Snyder (2013) found in their related meta-analysis that adult participants with MDD performed worse in working memory tasks requiring manipulation than those requiring maintenance, likely due to manipulation tasks placing an increased demand on cognitive resources. Therefore, it may be helpful for further research to focus on delineating these two aspects of working memory.

Performance was also significantly impaired for children and adolescents with depression across shifting and inhibition. Deficits in the shifting subdomain could reduce children/adolescents' capacity to shift their focus away from negative material, subsequently increasing negative affect and related depressive symptoms (Letkiewicz et al., 2014). Furthermore, the cognitive rigidity and entrenched negative thinking that frequently characterises depression in children/adolescents may also create difficulties in switching responses/strategies according to changing situational demands (Morris & Mansell, 2018). The deficits in inhibitory control observed in children/adolescents with depression are congruent with research proposing that challenges in inhibiting intrusive negative stimuli in

depressed individuals may serve to exacerbate rumination on negative content, intensifying overall negative affect (Joormann et al., 2007).

Moderators

Executive Function Task Outcome

The findings indicated no significant difference between the two executive function task outcome measures explored (reaction time, accuracy), on the executive function performance of children/adolescents with depression relative to healthy controls. Previous research has hypothesised that processing speed deficits may better account for the impairments observed in participants with depression on executive function tasks (Snyder, 2013). However, consistent with Snyder (2013), children/adolescents with depression were significantly impaired on both timed tasks and self-paced accuracy tasks, suggesting that motor slowing cannot fully account for the deficits observed. Furthermore, the results also suggest that unlike findings reported in other mental health conditions, such as anxiety (Majeed et al., 2023), the difficulties observed in young people with depression on reaction time tasks are unlikely to be the result of a speed-accuracy trade-off (Snyder, 2013).

Participant-Level Characteristics

The results suggest that the presence of comorbidities, psychotropic medication use and sex in the group with depression, did not significantly predict the effect size. In considering comorbidity, several studies did not provide clear information regarding the type of comorbidity and/or frequencies, thus preventing more sensitive sub-group analyses from

being undertaken. The findings in relation to medication conflict with previous research, which has reported both positive (Prado et al., 2018) and adverse (Majeed et al., 2023) effects of psychotropic medication on cognitive function. Rogers et al. (2004) proposed a potential theory to account for an apparent lack of effect, in which the complex ways medication may have an impact, lead to the combination of negative and positive effects subsequently cancelling one another out. Therefore, this finding may need to be interpreted with caution. Finally, Wagner et al. (2012) also found no relationship between sex and executive function performance between participants with and without depression. However, their study was based on an adult sample and the impact of sex was only assessed in relation to the verbal fluency aspect of executive functioning.

Conversely, age and FSIQ were found to be significant predictors of the effect size. The greater impact on executive function for younger children with depression and children with lower FSIQs and depression may firstly reflect younger ages being a critical period for the development and refinement of higher-level cognitive skills, such as executive function (Best & Miller, 2010). Therefore, in younger children, the still maturing brain areas required for successful executive function may be more susceptible to the negative effects of depression (Nyvold et al., 2022). Secondly, given that intellectual ability is considered to be generally associated with executive function (Friedman et al., 2006), deficits in intellectual functioning may in part increase children/adolescents' susceptibility to the negative interaction between depression and executive function.

Comparison With Related Reviews

Contrasting with previous meta-analyses/reviews in this area, this review adds definitive data regarding an executive function impairment in children/adolescents with depression. In comparison to previous related meta-analyses, this review identified a larger number of studies/effects (28 studies and 61 effects). Wagner et al. (2015) identified only 15 studies assessing the three executive functioning subdomains focused on in this review, whilst Goodall et al. (2018) identified a total of 20 relevant studies. Alongside increased statistical power and precision in the estimate of effect size, the larger dataset in this review enabled several subgroup analyses to be undertaken, a consistent limitation of previous related reviews. The observed findings were consistent with Wagner et al. (2015) who also reported significant impairments in executive function for children/adolescents with depression, relative to healthy controls. In comparison to Wagner et al. (2015), this meta-analysis broadened its scope to include mild-moderate depressive experiences, alongside an MDD diagnosis. The lack of congruence with Goodall et al. (2018), may relate to the wider age range they used (12-25 years), potentially influencing their observed findings.

In relation to the wider literature focused on executive function ability in adults with depression, the findings of this meta-analysis are also broadly consistent with a number of other meta-analytic reviews. Both Snyder (2013) and Wagner et al. (2012) found significant impairments in adults with depression across multiple executive functioning subdomains. Generally, the similar findings regarding impaired executive function in individuals with depression, across both adult and child/adolescent reviews, could suggest consistency and developmental continuity in the relationship between depression and executive dysfunction. This may have important implications for both future research and clinical practice.

Limitations

It is important to consider the limitations of this meta-analysis when interpreting the findings. Firstly, employing the two-level REM, despite including multiple effects for most of the primary studies could be considered a limitation. However, comparison to the three-level meta-analytic model and additional sensitivity analyses did not significantly improve the model fit. Secondly, a variety of executive function measures were included in this review, in which the 'best' measures were not determined a priori. This enabled the inclusion of more data, which is vital in exploring if an effect is present and has been a common limitation in previous related meta-analyses. However, measurement properties could have influenced the findings, although low levels of heterogeneity suggested this did not have a meaningful impact. Another limitation relates to the focus on the unity and diversity model of executive function, as outlined by Miyake et al. (2000). Use of this well-defined framework enabled clarity and focus within the meta-analysis. However, this also limited the scope of executive functioning explored (e.g., planning), resulting in potentially relevant studies being excluded. Although this meta-analysis was not limited to specific geographical regions and included both peer-reviewed and 'grey literature', the literature search was restricted to only three databases and papers published in English. This may have led to some relevant studies not being included, although the findings observed were predominantly consistent with existing meta-analyses (Snyder, 2013; Wagner et al., 2015).

In considering the limitations related to the field of available literature exploring executive function in children/adolescents with depression, the risk of bias analysis indicated various methodological quality issues within the included studies. Selection bias was consistently rated as "high risk", due to studies using predominantly single non-random

sampling. “High risk” of sampling bias can create substantial challenges to the overall validity and reliability of the meta-analysis. Several studies were also rated as “high risk” of generalisability bias, due to recruiting particularly small sample sizes, limiting the ability to generalise the findings more broadly. Furthermore, the included studies did not consistently report on the severity of depression, preventing further exploration of the specific impact this may have on executive function performance. All included studies were cross-sectional in nature, limiting the ability to infer causality and consider the underlying mechanisms in more depth. Finally, inadequate data were available regarding some of the moderator variables (e.g. FSIQ), reducing statistical power and the potential comprehensiveness of any additional subgroup analyses/meta-regressions.

Of particular note, is the ‘task impurity problem’, frequently discussed in the wider executive function literature (Snyder, 2013). This refers to the fact that although neuropsychological tests are predominantly well-established (Nyvold et al., 2022), they were not necessarily developed to specifically assess the three executive functioning subdomains outlined by Miyake et al. (2000). As a result, many of the tasks used within this review can tap various executive function abilities, alongside other cognitive processes, such as visual processing (Nyvold et al., 2022; Snyder, 2013). Attempts were made to mediate this, through pre-defining the three subdomains and associated tests, based on general consensus in the current evidence-base. However, there remains potential for ‘task impurity’ to result in challenges in inferring clear conclusions regarding the impact of childhood/adolescent depression on specific executive functioning subdomains.

Future Research Directions

Additional research is required to clarify the association between executive function and childhood/adolescent depression. Specifically, more longitudinal research will be valuable in exploring the developmental impact of depression on executive function, across different ages, alongside enabling causality to be better inferred. Secondly, within the literature focused on adult depression, the severity of depression has been reported to predict more significant deficits in executive function performance (Snyder, 2013). To enable a greater understanding of the impact depression symptom severity may have on child/adolescent executive function performance, future research could ensure depression severity is more clearly recorded and subsequently combine groups based on symptom severity level. In addition, more consistent reporting in future research, of the moderator variables explored (e.g., FSIQ) would enable better interpretation of their impact. Due to the 'task impurity problem' discussed above and substantial variability in the executive function tasks used across studies, future research may benefit from establishing a more standardised and consistent approach to executive function assessment, with tests clearly outlined and designed in relation to specific executive function subdomains (Goodall et al., 2018). This will increase the comparability across different studies and thus ability to draw firmer conclusions. Finally, in terms of general methodological quality, future studies in the field should focus on increasing sample sizes and improved participant recruitment processes, as a means of increasing the generalisability and representativeness of findings.

Clinical Implications

The findings of this meta-analysis may have important clinical implications. Firstly, to support the earlier identification of depression, clinicians would benefit from including a comprehensive exploration of executive function when assessing and formulating a young person's difficulties with depression. This should include the consideration of specific executive function subdomains. Furthermore, research has discussed how a targeted and person-centred approach to mental health care is vital (Stanhope et al., 2021). Therefore, in considering the development of tailored and effective interventions for children/adolescents with depression, clinicians should remain mindful of targeting any particular executive function difficulties when creating individual treatment plans. Psychological interventions often used to support children/adolescents with depression (e.g., cognitive-behavioural therapy) also rely on executive function skills to facilitate meaningful engagement (Goodall et al., 2018). This may require clinicians to adapt these interventions accordingly as a means of accounting for specific executive function difficulties. Indeed, improved early intervention/prevention for depression in this age group may help reduce the likelihood of depressive symptom progression into adulthood (Hetrick et al., 2008), alongside reducing pressures on the often-stretched health services (Thomas et al., 2016). Finally, continued monitoring of executive function and ongoing cognitive training may have a key role in reducing the risk of relapse, whilst increasing overall quality of life in children and adolescents who have/had depression (Snyder, 2013; Vilgis et al., 2015).

Conclusions

This meta-analysis found executive function to be significantly impaired in children/adolescents with depression relative to healthy controls, with the largest effect size being in the subdomain of working memory. These findings were broadly consistent with literature focused on adult samples, supporting the notion of developmental continuity between depression and executive dysfunction. Additional research is required to further clarify the association between executive function and childhood/adolescent depression. Improvements could be made by the use of longitudinal designs, greater specificity regarding depression severity, use of standardised and consistent executive function tests, alongside more robust sampling methods. However, the findings also highlight the importance of clinicians incorporating an understanding of executive function into the assessment, formulation, and intervention of children/adolescents with depression, to help alleviate distress and improve clinical outcomes.

References

Aoki, S., Nagatani, F., Kagitani-Shimono, K., Ohno, Y., Taniike, M., & Mohri, I. (2023). Examining normative values using the Cambridge neuropsychological test automated battery and developmental traits of executive functions among elementary school-aged children in Japan. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1141628

Arffa, S. (2007). The relationship of intelligence to executive function and non-executive function measures in a sample of average, above average, and gifted youth. *Archives of Clinical Neuropsychology*, *22*(8), 969–978. https://doi.org/10.1016/j.acn.2007.08.001

Assink, M., Spruit, A., Schuts, M., Lindauer, R., van der Put, C. E., & Stams, G.-J. J. M. (2018). The intergenerational transmission of child maltreatment: A three-level meta-analysis. *Child Abuse & Neglect*, *84*, 131–145. https://doi.org/10.1016/j.chiabu.2018.07.037

Baddeley, A, (2003) Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829-839. https://doi.org/10.1038/nrn1201.

Banks, B. D., Mao, I. L., & Walter, J. P. (1985). Robustness of the restricted maximum likelihood estimator derived under normality as applied to data with skewed distributions. *Journal of Dairy Science*, *68*(7), 1785–1792. https://doi.org/10.3168/jds.S0022-0302(85)81028-6

Baune, B. T., Fuhr, M., Air, T., & Hering, C. (2014). Neuropsychological functioning in adolescents and young adults with major depressive disorder: A review. *Psychiatry Research*, *218*(3), 261–271. https://doi.org/10.1016/j.psychres.2014.04.052

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

Blakey, E., Visser, I., & Carroll, D. J. (2016). Different executive functions support different kinds of cognitive flexibility: Evidence from 2-, 3-, and 4-year-olds. *Child Development, 87(*2), 513–526. https://doi.org/10.1111/cdev.12468

Bloch, Y., Aviram, S., Braw, Y., Gvirts, H. Z., Rabany, L., & Walter, G. (2015). Attention improves after clinical improvement in acutely depressed adolescents. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *27*(2), 153–156. https://doi.org/10.1176/appi.neuropsych.13120372

Bloch, Y., Aviram, S., Faibel, N., Govezensky, J., Braw, Y., Rabany, L., & Walter, G. (2013). The correlation between impaired attention and emotional reactivity in depressed adolescent patients. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *25*(3), 233–236. https://doi.org/10.1176/appi.neuropsych.12080194

Borenstein, M. (2009). Subgroup analyses. In M. Borenstein, L. V. Hedges, J. P. T. Higgins, & H. R. Rothstein (Eds.), *Introduction to Meta-Analysis* (pp. 149–186). John Wiley & Sons.

Brooks, B. L., Iverson, G. L., Sherman, E. M. S., & Roberge, M.-C. (2010). Identifying cognitive problems in children and adolescents with depression using computerized neuropsychological testing. *Applied Neuropsychology*, *17*(1), 37–43. https://doi.org/10.1080/09084280903526083

Brooks, B. L., & Sherman, E. M. S. (2012). Computerized neuropsychological testing to rapidly evaluate cognition in pediatric patients with neurologic disorders. *Journal of Child Neurology*, *27*(8), 982–991. https://doi.org/10.1177/0883073811430863

Cataldo, M. G., Nobile, M., Lorusso, M. L., Battaglia, M., & Molteni, M. (2005). Impulsivity in depressed children and adolescents: A comparison between behavioral and neuropsychological data. *Psychiatry Research*, *136*(2–3), 123–133. https://doi.org/10.1016/j.psychres.2004.12.012

Christensen, H., Griffiths, K., Mackinnon, A., & Jacomb, P. (1997). A quantitative review of cognitive deficits in depression and Alzheimer-type dementia. *Journal of the International Neuropsychological Society*, *3*(6), 631–651.

Christidi, F., Kararizou, E., Triantafyllou, N., Anagnostouli, M., & Zalonis, I. (2015). Derived trail making test indices: Demographics and cognitive background variables across the adult life span. *Aging, Neuropsychology, and Cognition*, *22*(6), 667–678. https://doi.org/10.1080/13825585.2015.1027650

Ciuhan, G. C., & Iliescu, D. (2021). Depression and learning problems in children: Executive function impairments and inattention as mediators. *Acta Psychologica*, *220*. https://doi.org/10.1016/j.actpsy.2021.103420

Constantinidou, F., Danos, M. A., Nelson, D., & Baker, S. (2011). Effects of modality presentation on working memory in school-age children: Evidence for the pictorial superiority hypothesis. *Child Neuropsychology*, *17*(2), 173–196. https://doi.org/10.1080/09297049.2010.525503

Daniel, T. A., Katz, J. S., & Robinson, J. L. (2016). Delayed match-to-sample in working memory: A brainmap meta-analysis. *Biological Psychology*, *120*, 10–20. https://doi.org/10.1016/j.biopsycho.2016.07.015

Dann, K. M., Veldre, A., Miles, S., Sumner, P., Hay, P., & Touyz, S. (2023). Measuring cognitive flexibility in anorexia nervosa: Wisconsin card sorting test versus cued task-

switching. *Eating and Weight Disorders*, *28*(1), 60. https://doi.org/10.1007/s40519-023-01589-6

de Bruin, A., & Sala, S. D. (2018). Effects of age on inhibitory control are affected by task-specific features. *Quarterly Journal of Experimental Psychology*, *71*(5), 1219–1233. https://doi.org/10.1080/17470218.2017.1311352

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

Diler, R. S., Pan, L. A., Segreti, A. M., Ladouceur, C. D., Forbes, E., Romero, C. S., Almeida, J. R. C., Birmaher, B., Axelson, D. A., & Phillips, M. L. (2014). Differential anterior cingulate activity during response inhibition in depressed adolescents with bipolar and unipolar major depressive disorder. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *23*(1), 10–19.

Duell, N., Icenogle, G., Silva, K., Chein, J., Steinberg, L., Banich, M. T., Di Guinta, L., Dodge, K. A., Fanti, K. A., Lansford, J. E., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Tapanya, S., Uribe Tirado, L. M., Alampay, L. P., Al-Hassan, S. M., Takash, H. M. S., … Chaudhary, N. (2018). A cross-sectional examination of response inhibition and working memory on the Stroop task. *Cognitive Development*, *47*, 19–31. https://doi.org/10.1016/j.cogdev.2018.02.003

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

Erskine, H. E., Baxter, A. J., Patton, G., Moffitt, T. E., Patel, V., Whiteford, H. A., & Scott, J. G. (2017). The global coverage of prevalence data for mental disorders in children and

adolescents. *Epidemiology and Psychiatric Sciences*, *26*(4), 395–402. https://doi.org/10.1017/S2045796015001158

Essau, C. A., Lewinsohn, P. M., Seeley, J. R., & Sasagawa, S. (2010). Gender differences in the developmental course of depression. *Journal of Affective Disorders*, *127*(1), 185–190. https://doi.org/10.1016/j.jad.2010.05.016

Favre, T., Hughes, C., Emslie, G., Stavinoha, P., Kennard, B., & Carmody, T. (2009). Executive functioning in children and adolescents with major depressive disorder. *Child Neuropsychology*, *15*(1), 85–98. https://doi.org/10.1080/09297040802577311

Fisk, J., Ellis, J. A., & Reynolds, S. A. (2019). A test of the CaR-FA-X mechanisms and depression in adolescents. *Memory*, *27*(4), 455–464. https://doi.org/10.1080/09658211.2018.1518457

Fleming, K. A., Heintzelman, S. J., & Bartholow, B. D. (2016). Specifying associations between conscientiousness and executive functioning: Mental set shifting, not prepotent response inhibition or working memory updating. *Journal of Personality*, *84*(3), 348–360. https://doi.org/10.1111/jopy.12163

Franklin, T., Lee, A., Hall, N., Hetrick, S., Ong, J., Haslam, N., Karsz, F., & Vance, A. (2010). The association of visuospatial working memory with dysthymic disorder in pre-pubertal children. *Psychological Medicine*, *40*(2), 253–261. https://doi.org/10.1017/S0033291709990365

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

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

Garber, J., Gallerani, C., & Frankel, S. (2009). Depression in children. In I. Gotlib & C. Hammen (Eds.), *Handbook of Depression* (2nd ed., pp. 405–443). The Guilford Press.

Goodall, J., Fisher, C., Hetrick, S., Phillips, L., Parrish, E. M., & Allott, K. (2018). Neurocognitive functioning in depressed young people: A systematic review and meta-analysis. *Neuropsychology Review*, *28*(2), 216–231. https://doi.org/10.1007/s11065-018-9373-9

Graves, L. V., Drozdick, L., Courville, T., Farrer, T. J., Gilbert, P. E., & Delis, D. C. (2021). Cohort differences on the CVLT-II and CVLT3: Evidence of a negative flynn effect on the attention/working memory and learning trials. *The Clinical Neuropsychologist*, *35*(3), 615–632. https://doi.org/10.1080/13854046.2019.1699605

Gray-Burrows, K., Taylor, N., O'Connor, D., Sutherland, E., Stoet, G., & Conner, M. (2019). A systematic review and meta-analysis of the executive function-health behaviour relationship. *Health Psychology and Behavioral Medicine*, *7*(1), 253–268. https://doi.org/10.1080/21642850.2019.1637740

Gruenberg, A., Goldstein, R., & Pincus, H. (2005). Classification of depression: Research and diagnostic criteria: DSM-IV and ICD-10. In J. Licinio & M. Wong (Eds.), *Biology of Depression: From Novel Insights to Therapeutic Strategies* (pp. 1–12). Wiley VCH.

Günther, T., Holtkamp, K., Jolles, J., Herpertz-Dahlmann, B., & Konrad, K. (2004). Verbal memory and aspects of attentional control in children and adolescents with anxiety

disorders or depressive disorders. *Journal of Affective Disorders*, *82*(2), 265–269. https://doi.org/10.1016/j.jad.2003.11.004

Günther, T., Konrad, K., De Brito, S. A., Herpertz-Dahlmann, B., & Vloet, T. D. (2011). Attentional functions in children and adolescents with ADHD, depressive disorders, and the comorbid condition. *Journal of Child Psychology and Psychiatry*, *52*(3), 324–331. https://doi.org/10.1111/j.1469-7610.2010.02320.x

Halari, R., Simic, M., Pariante, C. M., Papadopoulos, A., Cleare, A., Brammer, M., Fombonne, E., & Rubia, K. (2009). Reduced activation in lateral prefrontal cortex and anterior cingulate during attention and cognitive control functions in medication-naive adolescents with depression compared to controls. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *50*(3), 307–316. https://doi.org/10.1111/j.1469-7610.2008.01972.x

Hall, P. A., & Fong, G. T. (2015). Temporal self-regulation theory: A neurobiologically informed model for physical activity behavior. *Frontiers in Human Neuroscience*, *9*(117). https://doi.org/10.3389/fnhum.2015.00117

Hanna, G. L., Liu, Y., Isaacs, Y. E., Ayoub, A. M., Brosius, A., Salander, Z., Arnold, P. D., & Gehring, W. J. (2018). Error-related brain activity in adolescents with obsessive-compulsive disorder and major depressive disorder. *Depression and Anxiety*, *35*(8), 752–760. https://doi.org/10.1002/da.22767

Hetrick, S. E., Parker, A. G., Hickie, I. B., Purcell, R., Yung, A. R., & McGorry, P. D. (2008). Early identification and intervention in depressive disorders: Towards a clinical staging model. *Psychotherapy and Psychosomatics*, *77*(5), 263–270. https://doi.org/10.1159/000140085

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savovic, J., Schulz, K. F., Weeks, L., Sterne, J. A. C., Cochrane Bias Methods Group, & Cochrane Statistical Methods Group. (2011). The cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ (Clinical Research Ed.)*, *343*. https://doi.org/10.1136/bmj.d5928

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ (Clinical Research Ed.)*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hirst, R. B., Beard, C. L., Colby, K. A., Quittner, Z., Mills, B. M., & Lavender, J. M. (2017). Anorexia nervosa and bulimia nervosa: A meta-analysis of executive functioning. *Neuroscience and Biobehavioral Reviews*, *83*, 678–690. https://doi.org/10.1016/j.neubiorev.2017.08.011

Hoffmann, A., Ettinger, U., Reyes del Paso, G. A., & Duschek, S. (2017). Executive function and cardiac autonomic regulation in depressive disorders. *Brain and Cognition*, *118*, 108–117. https://doi.org/10.1016/j.bandc.2017.08.003

Jazbec, S., Pantelis, C., Robbins, T., Weickert, T., Weinberger, D. R., & Goldberg, T. E. (2007). Intra-dimensional/extra-dimensional set-shifting performance in schizophrenia: Impact of distractors. *Schizophrenia Research*, *89*(1–3), 339–349. https://doi.org/10.1016/j.schres.2006.08.014

Johnson, V. A., Kemp, A. H., Heard, R., Lennings, C. J., & Hickie, I. B. (2015). Childhood- versus adolescent-onset antisocial youth with conduct disorder: Psychiatric illness, neuropsychological and psychosocial function. *PLoS ONE*, *10*(4). https://doi.org/10.1371/journal.pone.0121627

Joormann, J., Yoon, K. L., & Zetsche, U. (2007). Cognitive inhibition in depression. *Applied and Preventive Psychology*, *12*(3), 128–139. https://doi.org/10.1016/j.appsy.2007.09.002

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, *17*(3), 213–233. https://doi.org/10.1007/s11065-007-9040-z

Khoury, J. E., Milligan, K., & Girard, T. A. (2015). Executive functioning in children and adolescents prenatally exposed to alcohol: A meta-analytic review. *Neuropsychology Review*, *25*(2), 149–170. https://doi.org/10.1007/s11065-015-9289-6

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching: A review. *Psychological Bulletin*, *136*(5), 849–874. https://doi.org/10.1037/a0019842

Kim, S. Y., Park, J. E., Lee, Y. J., Seo, H.-J., Sheen, S.-S., Hahn, S., Jang, B.-H., & Son, H.-J. (2013). Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *Journal of Clinical Epidemiology*, *66*(4), 408–414. https://doi.org/10.1016/j.jclinepi.2012.09.016

Klabes, J., Babilon, S., Zandi, B., & Khanh, T. Q. (2021). The sternberg paradigm: Correcting encoding latencies in visual and auditory test designs. *Vision*, *5*(2), 21. https://doi.org/10.3390/vision5020021

Klimkeit, E. I., Tonge, B., Bradshaw, J. L., Melvin, G. A., & Gould, K. (2011). Neuropsychological deficits in adolescent unipolar depression. *Archives of Clinical Neuropsychology*, *26*(7), 662–676. https://doi.org/10.1093/arclin/acr051

Kunz, R., & Oxman, A. D. (1998). The unpredictability paradox: Review of empirical comparisons of randomised and non-randomised clinical trials. *BMJ*, *317*(7167), 1185–1190.

Lakomy, E. (2021). The utility of the Test of Everyday Attention for Children (TEA-Ch) for attention-related referrals. *Dissertations*, *611*. https://digitalcommons.nl.edu/diss/611

Letkiewicz, A. M., Miller, G. A., Crocker, L. D., Warren, S. L., Infantolino, Z. P., Mimnaugh, K. J., & Heller, W. (2014). Executive function deficits in daily life prospectively predict increases in depressive symptoms. *Cognitive Therapy and Research*, *38*(6), 612–620. https://doi.org/10.1007/s10608-014-9629-5

Lonergan, A., Doyle, C., Cassidy, C., MacSweeney Mahon, S., Roche, R. A. P., Boran, L., & Bramham, J. (2019). A meta-analysis of executive functioning in dyslexia with consideration of the impact of comorbid ADHD. *Journal of Cognitive Psychology*, *31*(7), 725–749. https://doi.org/10.1080/20445911.2019.1669609

Maalouf, F. T., Brent, D., Clark, L., Tavitian, L., McHugh, R. M., Sahakian, B. J., & Phillips, M. L. (2011). Neurocognitive impairment in adolescent major depressive disorder: State vs. Trait illness markers. *Journal of Affective Disorders*, *133*(3), 625–632. https://doi.org/10.1016/j.jad.2011.04.041

Majeed, N., Chua, Y., Kothari, M., Kaur, M., Quek, F., Ng, M., Ng, W., & Hartanto, A. (2023). Anxiety disorders and executive functions: A three-level meta-analysis of reaction time and accuracy. *Psychiatry Research Communications*, *3*. https://doi.org/10.1016/j.psycom.2022.100100

Matthews, K., Coghill, D., & Rhodes, S. (2008). Neuropsychological functioning in depressed adolescent girls. *Journal of Affective Disorders*, *111*(1), 113–118. https://doi.org/10.1016/j.jad.2008.02.003

Meule, A. (2017). Reporting and interpreting task performance in go/go-go affective shifting tasks. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00701

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

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Morris, L., & Mansell, W. (2018). A systematic review of the relationship between rigidity/flexibility and transdiagnostic cognitive and behavioral processes that maintain psychopathology. *Journal of Experimental Psychopathology*, *9*(3), 1–40. https://doi.org/10.1177/2043808718779431

Nelson, T. D., Kidwell, K. M., Nelson, J. M., Tomaso, C. C., Hankey, M., & Espy, K. A. (2018). Preschool executive control and internalizing symptoms in elementary school. *Journal of Abnormal Child Psychology*, *46*(7), 1509–1520. https://doi.org/10.1007/s10802-017-0395-1

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, *162*, 31–38. https://doi.org/10.1016/j.jecp.2017.04.017

Nuño, L., Gómez-Benito, J., Carmona, V. R., & Pino, O. (2021). A systematic review of executive function and information processing speed in major depression disorder. *Brain Sciences*, *11*(2), 147. https://doi.org/10.3390/brainsci11020147

Nyvold, O., Nygaard, E., Augusti, E.-M., & Tamnes, C. K. (2022). Unity or diversity of executive functioning in children and adolescents with post-traumatic stress symptoms? A systematic review and meta-analysis. *Child Neuropsychology*, *28*(3), 374–393. https://doi.org/10.1080/09297049.2021.1979950

Oeri, N., Buttelmann, D., Voelke, A. E., & Roebers, C. M. (2019). Feedback enhances preschoolers' performance in an inhibitory control task. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00977

Onat, M., Emiroğlu, N., Baykara, B., Özerdem, A., Özyurt, G., Öztürk, Y., Şahin, U., Ildız, A., Bilgiç, B. K., Ongun, C. H., & Akay, A. P. (2019). Executive functions and impulsivity in suicide attempter adolescents with major depressive disorder. *Psychiatry and Clinical Psychopharmacology*, *29*(3), 332–339. https://doi.org/10.1080/24750573.2018.1541647

Oosterlaan, J., Scheres, A., & Sergeant, J. A. (2005). Which executive functioning deficits are associated with AD/HD, ODD/CD and comorbid AD/HD+ODD/CD? *Journal of Abnormal Child Psychology*, *33*(1), 69–85. https://doi.org/10.1007/s10802-005-0935-y

Paap, K. R., Anders-Jefferson, R., Zimiga, B., Mason, L., & Mikulinsky, R. (2020). Interference scores have inadequate concurrent and convergent validity: Should we stop using the flanker, Simon, and spatial Stroop tasks? *Cognitive Research: Principles and Implications*, *5*(1), 7. https://doi.org/10.1186/s41235-020-0207-y

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ (Clinical Research Ed.)*, *372*. https://doi.org/10.1136/bmj.n71

Pan, F., Xu, Y., Zhou, W., Chen, J., Wei, N., Lu, S., Shang, D., Wang, J., & Huang, M. (2020). Disrupted intrinsic functional connectivity of the cognitive control network underlies disease severity and executive dysfunction in first-episode, treatment-naive adolescent depression. *Journal of Affective Disorders*, *264*, 455–463. https://doi.org/10.1016/j.jad.2019.11.076

Pan, L. A., Batezati-Alves, S. C., Almeida, J. R. C., Segreti, A., Akkal, D., Hassel, S., Lakdawala, S., Brent, D. A., & Phillips, M. L. (2011). Dissociable patterns of neural activity during response inhibition in depressed adolescents with and without suicidal behavior. *Journal of the American Academy of Child and Adolescent Psychiatry*, *50*(6), 602–611. https://doi.org/10.1016/j.jaac.2011.03.018

Pandina, G. J. (2000). Neuropsychological concomitants of bipolar disorder in youth. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, *61*(3-B), 1648.

Peters, A. T., Ren, X., Bessette, K. L., Goldstein, B. I., West, A. E., Langenecker, S. A., & Pandey, G. N. (2019). Interplay between pro-inflammatory cytokines, childhood trauma, and executive function in depressed adolescents. *Journal of Psychiatric Research*, *114*, 1–10. https://doi.org/10.1016/j.jpsychires.2019.03.030

Prado, C. E., Watt, S., & Crowe, S. F. (2018). A meta-analysis of the effects of antidepressants on cognitive functioning in depressed and non-depressed samples. *Neuropsychology Review*, *28*(1), 32–72. https://doi.org/10.1007/s11065-018-9369-5

Rao, U., & Chen, L.-A. (2009). Characteristics, correlates, and outcomes of childhood and adolescent depressive disorders. *Dialogues in Clinical Neuroscience*, *11*(1), 45–62. https://doi.org/10.31887/DCNS.2009.11.1/urao

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

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: A systematic review and meta-analysis. *Psychological Medicine*, *44*(10), 2029–2040. https://doi.org/10.1017/S0033291713002535

Roelofs, J., Rood, L., Meesters, C., te Dorsthorst, V., Bögels, S., Alloy, L. B., & Nolen-Hoeksema, S. (2009). The influence of rumination and distraction on depressed and anxious mood: A prospective examination of the response styles theory in children and adolescents. *European Child & Adolescent Psychiatry*, *18*(10), 635–642. https://doi.org/10.1007/s00787-009-0026-7

Rogers, M. A., Kasai, K., Koji, M., Fukuda, R., Iwanami, A., Nakagome, K., Fukuda, M., & Kato, N. (2004). Executive and prefrontal dysfunction in unipolar depression: A review of neuropsychological and imaging evidence. *Neuroscience Research*, *50*(1), 1–11. https://doi.org/10.1016/j.neures.2004.05.003

Rosenthal, E. N., Riccio, C. A., Gsanger, K. M., & Jarratt, K. P. (2006). Digit Span components as predictors of attention problems and executive functioning in children.

Archives of Clinical Neuropsychology, *21*(2), 131–139. https://doi.org/10.1016/j.acn.2005.08.004

RStudio. (2020). *RStudio Team* [Open source & professional software for data science teams]. https://rstudio.com/

Sabahi, Z., Farhoudi, M., Naseri, A., & Talebi, M. (2022). Working memory assessment using cambridge neuropsychological test automated battery can help in the diagnosis of mild cognitive impairment: A systematic review and meta-analysis. *Dementia & Neuropsychologia*, *16*(4), 444–456. https://doi.org/10.1590/1980-5764-dn-2022-0006

Schramm, E., Klein, D. N., Elsaesser, M., Furukawa, T. A., & Domschke, K. (2020). Review of dysthymia and persistent depressive disorder: History, correlates, and clinical implications. *Lancet Psychiatry*, *7*(9), 801–812. https://doi.org/10.1016/S2215-0366(20)30099-7

Schuitema, I., de Sonneville, L., Kaspers, G., van der Pal, H., Uyttebroeck, A., van den Bos, C., & Veerman, A. (2015). Executive dysfunction 25 Years after treatment with cranial radiotherapy for pediatric lymphoid malignancies. *Journal of the International Neuropsychological Society*, *21*(9), 657–669. https://doi.org/10.1017/S1355617715000788

Shehab, A. A. S., Brent, D., & Maalouf, F. T. (2016). Neurocognitive changes in selective serotonin reuptake inhibitors-Treated adolescents with depression. *Journal of Child and Adolescent Psychopharmacology*, *26*(8), 713–720. https://doi.org/10.1089/cap.2015.0190

Shin, M.-S., Choi, H., Kim, H., Hwang, J.-W., Kim, B.-N., & Cho, S.-C. (2008). A study of neuropsychological deficit in children with obsessive-compulsive disorder. *European Psychiatry*, *23*(7), 512–520. https://doi.org/10.1016/j.eurpsy.2008.03.010

Shorey, S., Ng, E. D., & Wong, C. H. J. (2022). Global prevalence of depression and elevated depressive symptoms among adolescents: A systematic review and meta-analysis. *The British Journal of Clinical Psychology*, *61*(2), 287–305. https://doi.org/10.1111/bjc.12333

Snyder, H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: A meta-analysis and review. *Psychological Bulletin*, *139*(1), 81–132. https://doi.org/10.1037/a0028727

Spaniol, M., & Danielsson, H. (2022). 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

Stanhope, V., Choy-Brown, M., Williams, N., & Marcus, S. C. (2021). Implementing person-centered care planning: A randomized controlled trial. *Psychiatric Services*, *72*(6), 641–646. https://doi.org/10.1176/appi.ps.202000361

Steinberger, D. C., & Barch, D. M. (2023). Investigating the link between depression, cognition, and motivation in late childhood. *Child Psychiatry & Human Development*, *54*(2), 570–581. https://doi.org/10.1007/s10578-021-01267-7

Sutcliffe, P. A., Bishop, D. V. M., & Houghton, S. (2006). Sensitivity of four subtests of the Test of Everyday Attention for Children (TEA-Ch) to stimulant medication in children with ADHD. *Educational Psychology*, *26*(3), 325–337. https://doi.org/10.1080/01443410500341031

Terrin, N., Schmid, C. H., & Lau, J. (2005). In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *Journal of Clinical Epidemiology*, *58*(9), 894–901. https://doi.org/10.1016/j.jclinepi.2005.01.006

Thomas, S., Jenkins, R., Burch, T., Calamos Nasir, L., Fisher, B., Giotaki, G., Gnani, S., Hertel, L., Marks, M., Mathers, N., Millington-Sanders, C., Morris, D., Ruprah-Shah, B., Stange, K., Thomas, P., White, R., & Wright, F. (2016). Promoting mental health and preventing mental illness in general practice. *London Journal of Primary Care*, *8*(1), 3–9. https://doi.org/10.1080/17571472.2015.1135659

Tyagi, R., Arvind, H., Goyal, M., Anand, A., & Mohanty, M. (2021). Working memory alterations plays an essential role in developing global neuropsychological impairment in duchenne muscular dystrophy. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.613242

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, *18*(1), 3–24. https://doi.org/10.1017/S1366728914000522

Vance, A., & Winther, J. (2021). Spatial working memory performance in children and adolescents with major depressive disorder and dysthymic disorder. *Journal of Affective Disorders*, *278*, 470–476. https://doi.org/10.1016/j.jad.2020.09.093

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*, *12*(11), 418–424. https://doi.org/10.1016/j.tics.2008.07.005

Vijayakumar, N., Whittle, S., Yucel, M., Byrne, M. L., Schwartz, O., Simmons, J. G., & Allen, N. B. (2016). Impaired maturation of cognitive control in adolescents who develop major depressive disorder. *Journal of Clinical Child and Adolescent Psychology*, *45*(1), 31–43. https://doi.org/10.1080/15374416.2014.987381

Vilgis, V., Chen, J., Silk, T. J., Cunnington, R., & Vance, A. (2014). Frontoparietal function in young people with dysthymic disorder (DSM-5: Persistent depressive disorder) during spatial working memory. *Journal of Affective Disorders*, *160*, 34–42. https://doi.org/10.1016/j.jad.2014.01.024

Vilgis, V., Silk, T. J., & Vance, A. (2015). Executive function and attention in children and adolescents with depressive disorders: A systematic review. *European Child & Adolescent Psychiatry*, *24*(4), 365–384. https://doi.org/10.1007/s00787-015-0675-7

Vilgis, V., Yee, D., Silk, T. J., & Vance, A. (2022). Distinct neural profiles of frontoparietal networks in boys with ADHD and boys with persistent depressive disorder. *Cognitive, Affective & Behavioral Neuroscience*, *22*(5), 1183–1198. https://doi.org/10.3758/s13415-022-00999-w

Wagner, S., Doering, B., Helmreich, I., Lieb, K., & Tadić, A. (2012). A meta-analysis of executive dysfunctions in unipolar major depressive disorder without psychotic symptoms and their changes during antidepressant treatment. *Acta Psychiatrica Scandinavica*, *125*(4), 281–292. https://doi.org/10.1111/j.1600-0447.2011.01762.x

Wagner, S., Müller, C., Helmreich, I., Huss, M., & Tadić, A. (2015). A meta-analysis of cognitive functions in children and adolescents with major depressive disorder. *European Child & Adolescent Psychiatry*, *24*(1), 5–19. https://doi.org/10.1007/s00787-014-0559-2

Weavers, B., Riglin, L., Martin, J., Anney, R., Collishaw, S., Heron, J., Thapar, A., Thapar, A., & Rice, F. (2023). Characterising depression trajectories in young people at high familial risk of depression. *Journal of Affective Disorders*, *337*, 66–74. https://doi.org/10.1016/j.jad.2023.05.063

Wilkinson, P. O., & Goodyer, I. M. (2006). Attention difficulties and mood-related ruminative response style in adolescents with unipolar depression. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*(12), 1284–1291. https://doi.org/10.1111/j.1469-7610.2006.01660.x

Yang, Y., Shields, G. S., Zhang, Y., Wu, H., Chen, H., & Romer, A. L. (2022). Child executive function and future externalizing and internalizing problems: A meta-analysis of prospective longitudinal studies. *Clinical Psychology Review*, *97*. https://doi.org/10.1016/j.cpr.2022.102194

Zakzanis, K. K., Leach, L., & Kaplan, E. (1998). On the nature and pattern of neurocognitive function in major depressive disorder. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *11*(3), 111–119.

CHAPTER II: THE IMPACT OF HAPPINESS ON THE THEORY OF MIND ABILITY OF 3–4 ½ YEAR-OLD CHILDREN

Abstract

Background: The ability to reason about internal mental states (theory of mind) is essential for children's effective social functioning. However, mistakes caused by incorrectly relying on one's own egocentric knowledge can occur, leading to potential misunderstandings and conflict. In adults, incidental emotions (e.g., happiness) have been shown to influence egocentric biases during mental state reasoning. However, research has not yet investigated this in children. This study investigated the effect of induced happiness on theory of mind in 3-4 ½-year-old children.

Method: In total, 90 children between the ages of 3-4 ½ completed a classic 'change of location' false belief task. Prior to completing the false belief task, participants were pseudorandomised to either a happy or neutral mood condition, where the related mood was induced using relevant video clips. Mood manipulation was checked using a 4-point emotion rating scale.

Results: The mood manipulation was successful; children were significantly happier in the happy condition relative to the neutral condition. In the false belief task, a chi-square analysis showed that participants in the happy mood condition were significantly more likely to fail the false belief task than those in the neutral condition.

Conclusion: The results of this study suggest that children who are happier in mood may be less accurate at employing theory of mind processes. This may reflect happiness increasing an individual's reliance on their more readily accessible egocentric knowledge. The directions for future research and the possible clinical implications of these findings are discussed.

Introduction

Theory of Mind (ToM) is an element of social cognitive functioning that involves reasoning about one's own and other people's internal mental states, such as beliefs, desires, intentions, perspectives, knowledge etc (Epley & Waytz, 2010). Mood has been proposed to interact with a variety of cognitive processes, including ToM (Converse et al., 2008; Davis, 2016). There is a dearth of literature investigating these links in children. Exploring this interaction in children is particularly important as mood disorders are common in childhood and can lead to continued difficulties throughout adulthood (Kessler et al., 2001). This paper focuses on exploring the impact of mood (specifically happiness) on young children's ToM ability.

Theory of Mind in Children

To use ToM effectively, we must understand that others may possess different mental states to us, which can influence subsequent behaviour (Smogorzewska et al., 2019). ToM ability is essential for successful communication and social functioning in children and has been found to facilitate interpersonal relationships and prosocial behaviour (Brezack et al., 2021; Ghrear et al., 2021; Kravčenko & Šeibokaitė, 2018). However, ToM mistakes can frequently occur across the lifespan, resulting in potential miscommunication and conflict (Epley et al., 2004).

The ability to understand that other people can hold and act on false beliefs is a fundamental aspect of ToM, with experimental false belief tasks being extensively used as a

measure of ToM (Grosse Wiesmann et al., 2017; Liu et al., 2008). During a classic 'change of location' version of this paradigm, children are asked to specify where a protagonist will look for an item after it has been moved from one location to a different one, unbeknownst to the protagonist (Bernstein et al., 2011; Wimmer & Perner, 1983). This enables measurement of an individual's ability to disregard their own beliefs and infer another's mental state and subsequent actions (Begeer et al., 2012). A substantial body of literature has investigated ToM ability and its development in children, using these tasks. Indeed, a widely cited but less recent meta-analysis conducted by Wellman et al. (2001) reviewed 178 studies, finding that generally typically developing children are able to pass an explicit standard false belief task around the age of 4 years old. Conversely, children under the age of 4 years often perform at chance level or below, in which they particularly struggle to separate their own knowledge and beliefs from another's knowledge and beliefs, often termed egocentric bias (Epley et al., 2004; Kaysili & Acarlar, 2011).

There remains extensive debate within the literature regarding why this change observed around the age of four years old occurs (Barone et al., 2022; Samson & Apperly, 2010). Some researchers have proposed that even very young children have a conceptual understanding of ToM, however, do not possess the higher-level cognitive skills (e.g., executive function) needed to pass explicit false belief tasks (Leslie, 2005). Infancy research has been considered to support this theory, in which children as young as 15 months old have demonstrated a sensitivity to others' false beliefs when using more simple tasks and implicit measures, such as measuring eye gaze and spontaneous behaviour (Onishi & Baillargeon, 2005). Conversely, other researchers have argued that the data reported in infant studies may reflect agent-object-location associations or adoption of previously observed behavioural

patterns, as opposed to ToM ability (Perner & Ruffman, 2005). They instead suggest that there is a fundamental conceptual change in ToM ability around the age of 4 years, in which a more sophisticated understanding of ToM develops (Perner & Ruffman, 2005).

The two systems account of ToM proposed by Apperly and Butterfill (2009) could be considered an attempt to reconcile these two opposing lines of thought. Firstly, they propose an early-developing system that can seemingly track belief-like states efficiently and relatively automatically. However, this system is also suggested to be inflexible and limited in its ability to explicitly reason about complex propositional attitudes (e.g., beliefs and intentions) (Apperly & Butterfill, 2009). A distinct second cognitive process that is more deliberate, flexible, and effortful is then suggested to develop later, which enables more explicit mental state reasoning (Apperly & Butterfill, 2009). It is proposed that this effortful processing requires more general cognitive resources, such as executive function and language and thus develops with age and environmental experiences, such as interacting with siblings (Apperly & Butterfill, 2009). Explicit false belief tasks can be considered relatively cognitively demanding (Apperly et al., 2011) and can be prone to egocentric interference from a self-perspective (Birch & Bloom, 2004). They are therefore thought to require the latter more effortful processing system, which can help inhibit and overcome this often-initial egocentric interference (Apperly & Butterfill, 2009). Currently, the debate regarding the putative developmental transition in ToM ability at age 4 remains very much unresolved (Barone et al., 2022; Haskaraca et al., 2023). Thus, further exploration of factors that may influence young children's performance on false belief tasks is important in helping enhance and develop our understanding in this field.

Emotions and ToM

Research is currently growing in relation to the different target-based and/or perceiver-based factors that may influence ToM ability (Todd & Tamir, 2024). One particular perceiver-based factor being explored, that may impact mental state reasoning, is an individual's incidental emotional state (Todd et al., 2015). Indeed, alongside perspective taking, emotion has a key role in daily social interaction, with individuals developing an increased ability throughout childhood in their means of integrating emotional understanding with mental state reasoning (Wu et al., 2018; Zhai et al., 2021).

Currently, research has mainly focused on samples of adult participants when exploring the effect of emotional state on ToM ability. Broad evidence has indicated that different emotional states can impact the way 'other vs self,' information is processed in adults (Todd & Tamir, 2024). For example, compared to shame, a mood induction of guilt has been found to facilitate perspective taking ability (Yang et al., 2010). Whereas, compared to participants experiencing anger and disgust, anxious adult participants were found to show an increase in reliance on their egocentric perspective when mental state reasoning across both conceptual and perpetual tasks (Todd et al., 2015). A particularly early and influential study conducted by Converse et al. (2008), explored how incidental mood states of happiness and sadness effect ToM, using a university student sample. The authors found that participants who had a happy mood state induced were less likely to employ ToM accurately than participants who had a sad mood induced and thus demonstrated more egocentric thinking when asked to infer a less informed characters belief regarding the location of an item (Converse et al., 2008).

Converse et al's. (2008) findings align with research discussing how positive mood states can lead to more global-level processing that employs the use of heuristics, suggesting that a happy mood state may lead to an increased reliance on often more easily accessible defaults (Clore & Huntsinger, 2007), such as egocentric knowledge (Converse et al., 2008). On the other hand, it appears that a negative mood state may be associated with a more analytical and meticulous processing style, alongside a narrower attentional focus (Out et al., 2020). This may subsequently support the deliberate and effortful processing often required to overcome egocentrism and thus increases the ability to integrate knowledge about others when reasoning about mental states (Converse et al., 2008). A small number of studies in children also support the proposition that mood state impacts information processing. Schnall et al. (2008) found that children (6-11 years) who were induced to feel happy were also more likely to show global information processing and consequently impaired attention to detail on an embedded figures task, compared to those induced to feel sad. Davis (2016) focused on the regulation of sadness in 6–13-year-old children, finding that in the older children in their sample (10 years and older), effective regulation of sadness resulted in more global information processing (as in positive affect), on a global-local task. However, beyond the above studies, the current literature regarding the impact of emotions on children's information processing remains extremely limited, particularly in exploring any links with ToM processes and/or social functioning.

Broadly, theoretical conceptualisations regarding the mechanisms of affect remain the subject of ongoing debate within the literature (Barratt & Russell, 2015; Dejonckheere et al., 2021). Some theories propose that positive and negative affect represent mutually exclusive ends of a single continuum, where positive affect is the inverse of negative affect (Russell &

Carroll, 1999). By this view, negative and positive affect function effectively as part of a single scale, with "neutral" mood representing a singular mid-point between the two. Conversely, other theories consider positive and negative affect more independently, such that both can be measured separately and can co-exist (Watson & Tellegen, 1985). From this perspective, separable "neutral" states could exist, representative of the absence of positive or negative affect (i.e. that one is not experiencing positive affect is not informative as to whether one is experiencing negative affect). The background literature discussed above regarding affect and cognitive processes (Clore & Huntsinger, 2007; Converse et al., 2008) rarely addresses this question explicity, but tends to frame the relationship between affect and cognition within a linear continuum model: negative and positive affect (i.e. happiness and sadness) are purported to have quantifiably inverse effects on cognition, with both hypothesised to differ from neutral mood (Clore & Huntsinger, 2007). This perspective suggests that an increase in cognitive processing from one type of affect (e.g., sadness) would correspond with a decrease from another type of affect (e.g., happiness).

Despite the increasing evidence base focused on adults, there remains an absence of literature regarding the effect of emotions on ToM ability in children. Studying the impact of emotion on ToM is vital as emotionally evocative situations are common in day-to-day life, alongside emotional difficulties being particularly prevalent in various clinical conditions (Timmermans & Schilbach, 2014). Furthermore, extending this research to a younger population may enable further consideration regarding the potential underlying mechanisms contributing to the occurrence of egocentrism in ToM, by exploring if the incidental emotional state impact observed in adults is also present in children. Due to the novelty of this research,

it will aim to build on the initial evidence from Converse et al. (2008), by focusing specifically on the influence of positive affect.

Rationale and the Present Study

Accurate ToM processes are vital for navigating the social world (Brezack et al., 2021). Literature in adults has demonstrated that, in comparison to sadness, happiness increases the likelihood of egocentric responses on tasks designed to assess ToM ability (Converse et al., 2008). However, the potential impact of mood on ToM processes has not yet been tested in children. The present study aimed to address this by using a sample of 3 – 4 ½-year-old typically developing children and a classic 'change of location' false belief task to measure ToM. A between-subject design was employed, in which participants underwent either an incidental happiness or neutral mood induction process. The inclusion of children under the age of 4 years old will enable the addition of important new information regarding this critical time in children's ToM development.

Therefore, the identified research question for this study is: what is the impact of positive affect on ToM abilities in 3-4 ½-year-old children? Based on the previous research, it is hypothesised that relative to children in the neutral mood condition, children in the happy mood condition will make more errors on the false belief trial of a classic false belief task, due to an increased reliance on more readily accessible egocentric knowledge.

Method

The study was preregistered on the Open Science Framework (OSF; https://osf.io/tbzx3). Ethical approval was obtained from the University of Birmingham Science, Technology, Engineering and Mathematics ethics committee (ERN_0372) prior to the study commencing (see Appendix 5).

Piloting

Prior to formal testing, a pilot study was carried out to confirm the video clip choices for the mood induction process, alongside ascertaining if the manipulation check scale was understandable to participants. A class of 3–4-year-old children at a local nursery school (different to the one in which the final sample was recruited) were shown six different short video clips. Previous research has highlighted how showing brief video/film clips is an easy, effective, and reliable way of inducing required incidental mood states in children (Brenner, 2000; Fernández-Aguilar et al., 2019). The choice of video clips selected was predominantly guided by the current literature available validating emotionally evocative film clips suitable for young children (Gabel et al., 2019; von Leupoldt et al., 2007), alongside related studies (Davis et al., 2017; Siedlecka & Denson, 2019; Tan & Holub, 2018). Three video clips considered to induce a happy mood state were shown: the 'bare necessities' scene from *The Jungle Book*, the 'I just can't wait to be king' scene from *The Lion King* and the Pixar short film *For the Birds*. Three video clips considered to induce a neutral mood state were also shown: the opening scene of *The Last Unicorn*, the introduction to the forest episode from Netflix's *Our Planet* documentary and a YouTube video of pictures of everyday items e.g., a

chair, clock etc. After watching each video clip participants were asked to rate how the video made them feel using a 4-point cartoon face scale outlined further below. Generally, *The Jungle Book* clip was rated by participants as making them feel the happiest, whereas the *Our Planet* clip received the most ratings towards the neutral end of the scale. Therefore, these clips were selected to induce happy and neutral moods respectively in the formal testing. Overall, the manipulation check measure was acceptable to participants and the observed differences in scores for the different video clips validated children's ability to use the scale.

Participants

An a-priori G*Power calculation regarding the study's required effect size was conducted. The effect size was adapted from (Converse et al., 2008) which investigated the impact of mood on belief reasoning in adults (d = .93). Whilst this effect size likely reflects the best estimate of the true effect of mood on belief rating "scores" of this kind in adults, this study differs on two key parameters, the participant group (children) and analysis strategy (chi-square). Therefore, conservative choices were made. To have 95% power to identify a large effect (Cohen's w = .5), with a chi-square test (α < .05), 52 participants are recommended. To acknowledge the novelty of the question and potential variability of young children, we increased the sample size to reflect standard sample sizes in group-level designs with young children using an explicit false belief task (Baron-Cohen et al., 1985; Kaysili & Acarlar, 2011; Surian & Leslie, 1999, n = *72*) A final sample size, after attrition, was therefore selected *at* n = 80.

Ninety-two participants between the ages of 3-4 ½ years (46 boys and 46 girls) were recruited and included in the study. Over-recruitment with respect to the intended sample size

reflected testing children on a class-by-class basis, and conservativeness with respect to possible failures on the memory control trial of the false belief task. Two participants were excluded from the main analyses due to failing the memory control trial of the false belief task (one participant in each mood manipulation condition). Thus, ninety participants were retained in the final sample (44 boys and 46 girls). Table 2-1 provides a breakdown of demographics in each experimental group. All participants were recruited from a large local Birmingham (United Kingdom) nursery and infant school. Prior to testing, the school's Headteacher was provided an information sheet regarding the school's potential involvement in the research study (see Appendix 6), after which they provided informed consent for testing (in the position of loco parentis) (see Appendix 7). All parents/carers of eligible children were then sent an information sheet via email regarding the nature of the study (see Appendix 8), with a two-week window provided for them to contact the researcher or the school should they wish to withdraw their child from participating. Two parents made contact with the researcher prior to testing commencing, requesting that their child not participate.

Eligibility for participation in the study included children who were: a) between the ages of 3-4 ½ years old; b) English speaking and c) not experiencing any neurodevelopmental conditions/difficulties. Previous research has demonstrated that children with neurodevelopmental conditions can display increased difficulties in their ToM ability (Baron-Cohen et al., 1985; Smogorzewska et al., 2019). Therefore, as this was a novel study, to reduce any potential confounding factors at this stage, participants were excluded prior to testing if they were reported as having an intellectual disability and/or any neurodevelopmental conditions (e.g., autism spectrum disorder) by their class teacher and/or parent/carer.

Table 2-1

Participant Demographics for Each Group

		Happy condition	Neutral condition
Participants	*N* total	45	45
Sex	Boys:	21	23
	Girls:	24	22
Mean age and range		$\bar{x}$: 3.98 years SD: 0.54 years Range: 3 years – 4 years 6 months	$\bar{x}$: 3.99 years SD: 0.44 years Range: 3 years 1 month – 4 years 6 months

There were no significant group-level differences in participants age t (84) = -0.14, p = .90, or sex χ^2 (1) = 0.18, p = .67.

Mood induction and Manipulation Check

In line with the procedure outlined below, before completing the false belief task, participants underwent one of two mood induction procedures. Participants were pseudorandomised to either the happy mood or neutral mood experimental condition (n = 45 in each condition), in which they were alternately allocated to the different groups.

To induce positive affect, participants allocated to the happy mood condition watched a two-minute clip from the Disney *Jungle Book* cartoon film. The clip is a musical scene, in which the human character Mowgli and the bear character Baloo sing "the bare necessities" song. Participants allocated to the neutral mood condition instead watched a two-minute clip from the Netflix *Our Planet* nature documentary. This clip is the introduction to the episode

focused on forests and includes a narrator outlining the importance and characteristics of different forests, accompanied by various relevant video segments. Both video clips were presented on a 13-inch MacBook Pro laptop.

To determine the effectiveness of the mood induction procedure, a manipulation check, adapted from (Davis et al., 2017) was carried out. After watching the video clip, participants were presented with an A4 black and white printout of a simple 4-point emotion rating scale with gender-neutral facial expressions ranging from 'neutral' (score of 1) to 'very happy' (score of 4). See Appendix 9 for a copy of the scale. Each participant was asked to point to the face that indicated how the video had made them feel and their mood rating score was recorded by the primary researcher. The decision to use a 4-point rating scale measuring from 'neutral' to 'very happy', rather than a scale also including sadness, was primarily driven by practical decisions and the design of the study. This choice allowed for a more straightforward and child-friendly scale focusing specifically on positive and neutral affect. While the scale does not capture negative affect such as sadness, it aligns with this study's aim to specifically investigate the effects of positive affect compared to neutral affect.

False Belief Task

All participants then completed an age-appropriate change of location false belief task (Wimmer & Perner, 1983). Specifically, the 'Sally-Anne' false-belief task was adapted using procedures described by Baron-Cohen et al. (1985). False belief tasks are considered central in testing ToM (Wellman et al., 2001). The Sally-Anne task particularly, is well established for use with younger children and has been widely replicated within developmental psychology literature (Wu et al., 2018). The task included both a memory control trial and a false belief

trial and was visually demonstrated by the primary researcher using dolls and relevant props [3](e.g., miniature basket). See Appendix 10 for the test materials used.

During the task, participants were introduced to two characters, 'Sally' who had a basket and 'Anne' who had a box. The participants then witnessed a story in which 'Sally' put her toy marble in her basket whilst she went outside to play. Whilst 'Sally' was outside, 'Anne' took the marble from 'Sally's' basket and moved it to her box. At this stage, participants were asked to point to where the marble was at the start (memory control trial). 'Sally' then returned from playing outside and the participants were informed that she wished to play with her marble. The participants were asked to point to where 'Sally' would go to look for her marble (false belief trial).

Participants' responses to both the memory trial and false belief trial were recorded, in which responses were categorised by the researcher as either pass or fail (see Appendix 11 for the record form). To pass the memory trial, participants had to indicate that the marble was in 'Sally's' basket at the start of the task. Stating that it was in 'Anne's' box at the start resulted in participants failing the memory trial. In line with previous related literature, to be included in the main analysis, participants had to pass the memory control trial. This trial provides a means of ensuring participants have the cognitive abilities required (e.g., memory and comprehension) to assess their subsequent understanding of false belief scenarios (Ghrear et al., 2021). To pass the false belief trial, participants had to infer 'Sally's' false belief and indicate that she would look in the original location of her basket for the marble. Participants failed the false belief trial if they stated that 'Sally' would look in 'Anne's' box for the marble.

[3] This was adjusted from the OSF preregistration, in which the initial plan was to use animated clips to demonstrate the 'Sally-Anne' false belief task.

Procedure

Testing took place between November 2023 and December 2023 at the nursery/infant school. All participants were tested by the primary researcher on an individual basis in a quiet room free from distractions. The researcher sat next to the participants for the duration of the testing. Initially, the primary researcher spent a few minutes familiarising themselves with participants, alongside recording participant demographics (age and sex). At this stage, participants were also familiarised with the emotion rating scale, and their knowledge of what emotion the different faces may suggest was checked to ensure participants correctly understood the scale (Cummings & Rennels, 2014). Verbal assent was obtained from the participants prior to testing starting to ensure their participation in the study was voluntary.

Participants were informed that they would first be watching a brief video clip. Depending on their group allocation participants were asked to watch one of the two video clips outlined above, before rating their mood on the scale. The participants were then told that they would be playing a game, during which they completed the 'change of location' false belief task with the primary researcher. Finally, to avoid any potential carry-over effects into the classroom from the mood induction procedure, participants completed a brief non-emotionally stimulating puzzle (https://www.bbc.co.uk/cbeebies/puzzles/bluey-sticker-puzzle) on the same laptop the video clips were presented on, before being returned to their class. Overall, testing sessions for each participant took 10-15 minutes.

Results

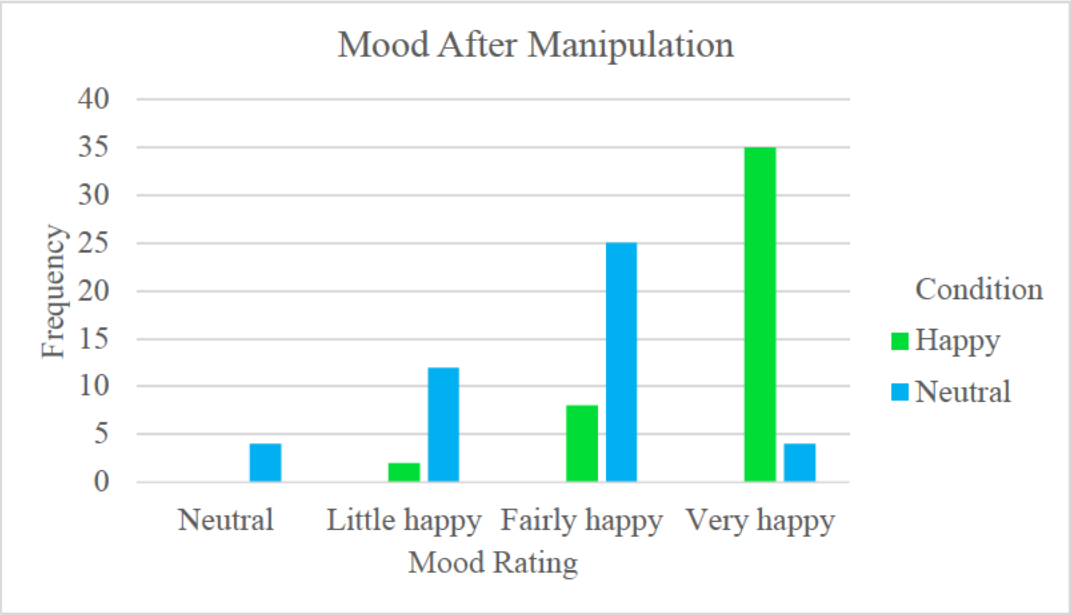
SPSS 29.0 was used to conduct all statistical analyses. Statistical significance was set at $p < 0.05$.

Manipulation Check

The distribution of mood rating scores for each experimental condition on the emotion rating scale is presented in Figure 2-1, in which higher scores indicated increased levels of happiness. Non-normality in the distribution of the data was confirmed by a significant Shapiro-Wilk test, w (90) = 0.81, p = < .001. Therefore, a non-parametric, Mann-Whitney U test was performed on these scores to confirm if the mood induction procedure was successful. The analysis indicated that participants in the happy condition rated themselves significantly happier (M = 3.73, SD = 0.54) than those in the neutral condition (M = 2.64, SD = 0.77), U = 272, $p < .001$.

Figure 2-1

Distribution of Mood Ratings in Each Condition



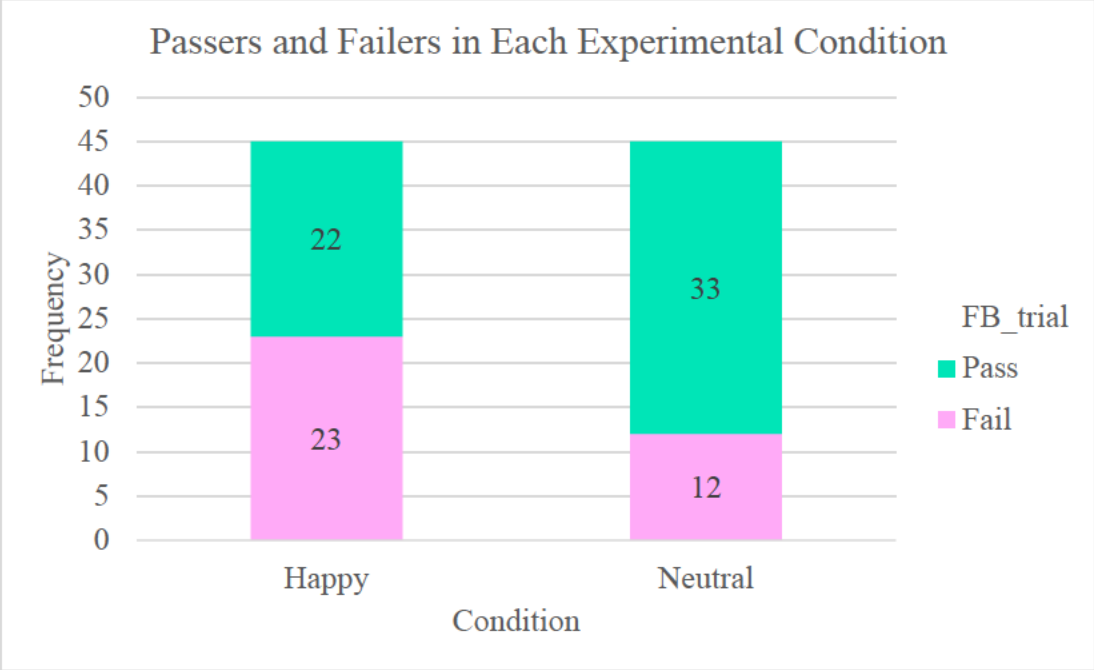
False Belief Task

For the false belief task, Figure 2-2 highlights the frequency of pass and fail responses in the false belief trial for the two experimental conditions. A 2 (condition; happy, neutral) x 2 (false belief trial; pass, fail) chi-square test of independence was performed to evaluate differences in performance on the false belief trial between conditions. Participants in the neutral mood condition were significantly less likely to fail (27%) compared to participants in the happy mood (51%) condition, $\chi^2(1) = 5.66$, $p = .017$. As a point of comparison, a chi-square test of independence was also performed to evaluate differences in performance on the memory control trial between the two conditions. Participants' performance on the memory control trial did not significantly differ across the two experimental mood manipulation

conditions, χ^2 (1) = .00, p = 1.00. Due to 2 cells (50%) having expected counts less than 5, there was potential for the chi-square test assumptions to be violated. To address this, Fisher's Exact Test was additionally undertaken, confirming the lack of significant association, p=1.000 (two-sided).

Figure 2-2

Frequency of Participants Passing or Failing the False Belief Trial in Each Experimental Condition



Additional Analyses

Further analysis was undertaken to explore a) the mood ratings of those who passed the false belief trial compared to those who failed the false belief trial; b) the effect of age on

performance on the false belief task and c) the effect of sex on performance on the false belief task.

To explore the mood ratings in those who passed the false belief trial compared to those who failed the false belief trial, a Mann-Whitney U test was performed. The analysis indicated that there was no significant difference in the average mood rating between those who passed (M = 3.15) and those who failed (M = 3.26) the false belief trial, $U = 889.5$, $p = .516$.

To explore the differences in participant age across performance on the false belief task a further Mann-Whitney U test was performed. The analysis indicated that there was no significant difference in the average age between participants who passed the false belief trial (M = 4.1 years) and participants who failed (M = 3.9 years) the false belief trial, $U = 743$, $p = .068$.

Finally, a 2 (sex; boy, girl) x 2 (false belief trial; pass, fail) chi square test of independence was performed to evaluate differences in performance on the false belief task between the different participant sex groups. The analysis indicated that there was no significant difference in participants' success on the false belief trial based on being a boy (57%) or a girl (65%), $\chi^2 (1) = 0.67$, $p = .414$.

Discussion

This study investigated the impact of induced happiness on young children's ToM ability, using a classic 'change of location' false belief task (Baron-Cohen et al., 1985).

Analysis of the mood manipulation check suggested that the induction of happiness was successful, such that participants were significantly happier in the happy condition relative to the neutral condition. In line with the hypothesis, this study found evidence that positive affect can impact children's ToM. Participants were less accurate when asked to infer a protagonist's false belief when they were induced to feel happy, compared to when they were induced to feel neutral in mood.

Understanding the Impact of Happiness on Children's False Belief Reasoning

One explanation for failure on a false belief task is the difficulty associated with disregarding one's own privileged belief, reflective of egocentrism (Begeer et al., 2012). As hypothesised, from this perspective, happiness seems to have contributed to increased egocentric biases in the children in this study. These findings in young children extend the work conducted in adults by Converse et al. (2008), who also found that participants induced to feel happy, were more likely to demonstrate egocentric thinking during a ToM task. These results could also be understood as consistent with evidence from domain-general tasks, regarding the depth of information processing in different mood states (Clore & Huntsinger, 2007). Therefore, similar to adults, children who are happier in mood may also engage in more global, top-down processing that tends to rely on more easily accessible defaults (Schnall et al., 2008), such as egocentric knowledge. Consequently, children who are happier may be less likely to engage in more deliberate and effortful processing, which can help overcome egocentric biases and enhance the integration of knowledge about others. Thus, leading to the ToM ability required to explicitly reason about another's false belief being significantly impaired.

Of note, there is growing literature suggesting that mood impacts a broad range of cognitive abilities in children, including areas such as attention, working memory and inhibition (Chiew & Braver, 2011; Diamond, 2013). This study specifically investigated the impact of mood on young children's ToM, with the findings remaining relatively neutral regarding whether mood has a unique impact on ToM or also affects other cognitive processes that contribute to it. Although our findings indicated that happiness impaired children's performance on the false belief trial of the experimental task, it did not affect their performance on the memory control trial of the task relative to the neutral condition. This suggests that the happy mood condition did not induce general cognitive impairment severe enough to affect simpler memory recall tasks. Notably, though, participants' performance on the memory control trial were at ceiling, rendering it only a very crude measure of working memory. The absence of a mood effect on the memory control trial may provide some insight into the specificity of mood impacts, suggesting that ToM might be more sensitive to mood variations than basic working memory. However, these findings do not rule out the possibility that mood subtly influences other cognitive functions integral to theory of mind, such as working memory and attention. As discussed in the limitations and future research section below, including more sensitive measures of these broader cognitive abilities in any further studies may help provide additional clarity into the underlying mechanisms of how mood impacts cognitive performance in children. This would help in determining how far mood effects generalise across different cognitive domains and enhance our understanding of if/what specific cognitive processes are influenced by mood in children.

These findings also suggest that ToM in children under the age of 4 years is impacted by mood state. This is highly consistent with the early emergence hypothesis (Leslie, 2005)

regarding how ToM develops. In considering the conceptual change hypothesis (Perner & Ruffman, 2005), these findings may contribute to a better understanding of the difficulties children have when in the initial stages of developing a ToM concept. From the perspective of the two systems account of theory of mind development (Apperly & Butterfill, 2009), this could be seen as evidence that in the early stages of effortful and deliberative ToM abilities developing, the application of this system is less consistent in younger children. They may be more liable to making errors as they acquire knowledge of how and when to use this latter system effectively (Samson & Apperly, 2010).

Additional analyses indicated that there was no evidence of a significant difference in the distribution of mood ratings of participants who passed vs. failed the false belief trial. Therefore, although the mood induction procedure appeared to be sufficient enough to predict performance on the false belief task and the emotion rating scale sensitive enough to pick up group-level changes related to receiving the induction, individual differences in mood were not directly associated with performance on the false belief task. This raises questions regarding the underlying mechanism of the impact of the mood induction procedure.

In line with the predicted effects and theoretical background, the primary results have been interpreted with reference to a broad change in processing patterns in the happy condition (meaning that a subset of children who would under neutral conditions have passed the false belief task, failed), but alternative explanations cannot be entirely ruled out. It is possible that something other than the intended emotional impact of the video clips impacted participants' performance on the false belief task. Alternative hypotheses could include the nature documentary clip priming attention and/or concentration and thus improving children's ToM performance as opposed to the happiness condition impairing it. Although this

alternative hypothesis cannot be completely excluded, based on the available literature, there is no clear evidence of similar manipulations producing such results. Broadly, studies have shown that that neutral affect typically serves as a control condition because it does not significantly influence cognitive processing (Schwarz, 1990). Notably, performance on the memory trial did also not appear to be impacted, which would also be predicted by the above hypothesis. Alternatively, performance in the happiness condition could appear to be consistent with chance responding, rather than a systematic effect of mood. Again, though, this seems a less likely alternative. Performance was broadly in line with the meta-analytic date reviewed by Wellman et al. (2001), which is typically considered to be reflective of a systematic shift in performance (as younger children perform significantly below chance). The existing literature supports the notion that positive affect can lead to a less deliberate and more heuristic information processing style generally (Clore & Huntsinger, 2007; Schnall et al., 2008) and impair adult's ToM performance specifically (Converse et al., 2008); happiness leading to chance responding in children is therefore a less parsimonious explanation. Thus, the consistent pattern observed in our study aligns well with established and well-grounded theories. Again, the memory control trial was also not significantly impacted, which would likely be expected if participants were guessing in their responses.

Taken as a whole, the most likely explanation is that the primary results were driven by the mood manipulation as predicted. However, the findings for the distribution of mood ratings of participants who passed vs. failed the false belief trial still require further exploration. Most likely is that the emotion rating scale was not sufficiently sensitive to account for the individual variability that can exist, thus making it difficult to discern any

within-group differences. Future research could focus on improving scale measurements and testing a broader range of videos.

Additional analyses also indicated that neither participant sex nor participant age, differed in the those who passed the false belief trial and those who failed the false belief trial. Based on the findings outlined by Wellman et al. (2001), it could have been expected that older children would be more likely to pass the false belief task than younger children. Research regarding sex differences is more mixed, with some studies finding no significant differences (Hughes & Dunn, 1998) and others finding that female children are more competent on ToM tasks (Walker, 2005). The absence of any significant differences for these two factors may again suggest that the presence of variability among individuals can impact the ability to detect within-group differences.

Limitations and Future Research

It is important to consider this study's various limitations, which may also serve to inform potential directions for future research. Firstly, it is generally assumed that the major reason children fail explicit false belief tasks is due to relying on their own privileged knowledge and thus making an incorrect egocentric response (Begeer et al., 2012). This is well supported by the fact that the errors children make are predominantly consistent with their own perspective as opposed to random errors (Flavell et al., 1981; Piaget & Inhelder, 1956) However, explicit false belief tasks are complex, and research has indicated how other processes may additionally be required, such as linguistic ability, attention, working memory and/or inhibition (Baillargeon et al., 2010; Drayton et al., 2011; Rubio-Fernández & Geurts, 2013). These processes have also been shown to be impacted by mood in some research

(Tornare et al., 2017; Wagner et al., 2015), and as a result may have influenced the participant's subsequent performance on the false belief task. Although, the memory control trial attempted to account for certain extraneous factors that may impact performance, further structurally controlled studies including tasks specifically examining these other potentially relevant factors may be beneficial. It is acknowledged that this may be complex to undertake, however, it may also enable further strengthening of the conclusions drawn in this study.

Secondly, in considering the self-report mood rating scale used, this scale has been employed in previous research with children (Davis et al., 2017) and offers a practical, efficient way of ascertaining participants' mood. However, as previously outlined, the rating scale consisted of only 4 points to reduce its complexity, but which may have also reduced its overall sensitivity. Furthermore, self-report scales may be vulnerable to the impact of demand characteristics (Westermann et al., 1996), resulting in inaccurate ratings and participants potentially overstating their level of happiness. The current measure appeared to be sensitive enough to demonstrate the impact of the mood induction, however, may not necessarily be sensitive enough to determine the underlying mechanism. Therefore, future research may consider exploring a more sensitive mood rating measure for young children to use, alongside including additional and more objective means of assessing participant mood (e.g., behavioural observational measures).

Due to this study aiming to build on the work conducted by Converse et al. (2008), an induction of happiness and neutral mood was specifically focused on. According to the conceptual framework discussed (Clore & Huntsinger, 2007; Converse et al., 2008), positive and negative affect are often considered to be at opposing ends of a single continuum with differentially identifiable effects on cognitive processes (Russell & Carroll, 1999). However,

as this study only included neutral and happy mood conditions, we were not able to address the impact of negative affect on children's ToM directly. One hypothesis from the literature suggests that sadness may improve children's ToM ability by promoting more deliberate and detailed thinking (Clore & Huntsinger, 2007; Converse et al., 2008; Davies et al., 2016). However, other studies, such as Schnall et al. (2008) found that while sad mood altered information processing relative to happy mood, it did not significantly differ from neutral mood, potentially indicating that it is positive mood that impairs performance rather than negative mood improving it. This higlights the need for further research to explore the impact of negative affect on children's ToM ability, whilst comparing this to positive affect. This will enable a more direct comparison to the literature focused on adults, alongside helping elucidate whether the effects of mood are linear or more distinct for positive and negative affects. Future research could additionally extent into other incidental mood states that have also been shown to impact ToM ability in adults (e.g., anxiety and guilt). This may help to increase the ability to isolate the impact of specific emotions, better consider the effect of differing emotions with the same valance and thus further explore potential underlying mechanisms (Todd et al., 2015).

Finally, although a sufficient sample size was recruited for this study, as indicated by the power analysis, all participants were from one specific age group, alongside being recruited from only one nursery/infant school in one UK location. Therefore, replicating this study across a wider range of childhood ages and within a wider area/geographical location would enable more broader generalisation of the findings to the childhood period. To increase generalisability further, it may also be interesting for future research to attempt to replicate

this study's findings across different ToM tasks (e.g., appearance-reality tasks) and using different mood induction procedures (e.g., music-based mood induction).

Clinical Implications

The findings observed in this study have indicated that thinking about mood is important when considering children's ToM ability. This may have implications within the field of clinical psychology and mental health/wellbeing. Firstly, although positive affect was specifically investigated in this study, sadness has been proposed to reflect an opposing emotional state (Tay & Kuykendall, 2017). Therefore, there may be merit in tentatively relating this study's findings to the clinical condition of depression, due to depressive disorder being particularly characterised by significant low mood (Gruenberg et al., 2005). In line with Converse et al. (2008), individuals experiencing mild-moderate symptoms of depression (e.g., dysphoria) have been found to be more accurate on ToM and related social cognitive tasks in comparison to controls (Harkness et al., 2011). Those findings have been discussed in relation to the theory previously outlined, that elevated levels of negative affect may be characterised by a more analytical and deliberative processing style, with reduced reliance on heuristic defaults (e.g., egocentric knowledge) when mental state reasoning (Mangardich et al., 2022). However, despite the more deliberate ToM processing and better inhibition of egocentric interference observed in low mood, leading to reduced errors on specific social cognitive tasks, mild-moderate depression is still associated with particular challenges in everyday social and interpersonal functioning (Hirschfeld et al., 2000; Mangardich et al., 2022). Mangardich et al. (2022) discussed how not using accessible defaults when appropriate (e.g., in familiar relationships with shared understandings) and consistently recruiting particularly

effortful ToM processes may lead to more stilted and inefficient social interaction. This could result in potential conflict, tension and rejection within relationships (Mangardich et al., 2022). In line with the findings of this study, the absence of regular positive affect in day-to-day life may consequently increase the use of unnecessary cognitive resources, whilst reducing the use of more efficient, global processing, potentially leading to further interpersonal stressors that can result in more severe depressive experiences, such as major depressive disorder (Harkness et al., 2011). It will be important to continue exploring the complex impact of positive and negative affect on social understanding/functioning. This may enable further development of social functioning models in children with mood disorders.

Additionally, many mental health interventions for young children are often particularly game/play based in nature, which amongst various benefits can help children have more positive emotional experiences (Francis et al., 2022). However, mental health interventions (e.g., cognitive behavioural therapy), often involve complex thinking and can require the ability to effectively reason about mental states (Bateman & Fonagy, 2006). Therefore, based on the findings of this study, making children particularly happy in therapy may reduce the overall depth of processing they are able to engage in, alongside potentially impairing aspects of their ToM ability. Consequently, clinicians should hold an awareness of the impact of mood, alongside ensuring there is an appropriate balance in influencing children's emotional state during different aspects of therapeutic work. This will hopefully enable more meaningful engagement in any psychological intervention being undertaken.

Conclusions

In conclusion, this novel study found that young children's ToM ability is influenced by the incidental emotional state of happiness. Participants were more likely to make errors on the false belief trial of a classic false belief task when happy in mood, compared to when neutral in mood. This may reflect happiness increasing an individual's reliance on their often more easily accessible egocentric knowledge. However, further research is needed to better understand the underlying mechanisms and potential links to clinical conditions. Improvements for future research may also focus on using a broader range of measures, alongside exploring different emotions, and increasing the overall generalisability of the findings.

References

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*(4), 953–970. https://doi.org/10.1037/a0016923

Apperly, I. A., Warren, F., Andrews, B. J., Grant, J., & Todd, S. (2011). Developmental continuity in theory of mind: Speed and accuracy of belief–desire reasoning in children and adults. *Child Development*, *82*(5), 1691–1703. https://doi.org/10.1111/j.1467-8624.2011.01635.x

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*(3), 110–118. https://doi.org/10.1016/j.tics.2009.12.006

Barrett, L., F., & Russell, J., A. (2015). *The psychological construction of emotion.* (1st ed.). The Guilford Press.

Baron-Cohen, S., Leslie, A., & 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

Barone, P., Wenzel, L., Proft, M., & Rakoczy, H. (2022). Do young children track other's beliefs, or merely their perceptual access? An interactive, anticipatory measure of early theory of mind. *Royal Society Open Science*, *9*(10). https://doi.org/10.1098/rsos.211278

Bateman, A. W., & Fonagy, P. (2006). *Mentalization based treatment for borderline personality disorder: A practical guide* (1st ed.). Oxford University Press.

Begeer, S., Bernstein, D. M., van Wijhe, J., Scheeren, A. M., & Koot, H. M. (2012). A continuous false belief task reveals egocentric biases in children and adolescents with

autism spectrum disorders. *Autism: The International Journal of Research and Practice*, *16*(4), 357–366. https://doi.org/10.1177/1362361311434545

Bernstein, D. M., Thornton, W. L., & Sommerville, J. A. (2011). Theory of mind through the ages: Older and middle-aged adults exhibit more errors than do younger adults on a continuous false belief task. *Experimental Aging Research*, *37*(5), 481–502. https://doi.org/10.1080/0361073X.2011.619466

Birch, S. A. J., & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Sciences*, *8*(6), 255–260. https://doi.org/10.1016/j.tics.2004.04.011

Brenner, E. (2000). Mood induction in children: Methodological issues and clinical implications. *Review of General Psychology*, *4*(3), 264–283. https://doi.org/10.1037/1089-2680.4.3.264

Brezack, N., Meyer, M., & Woodward, A. L. (2021). Three-year-olds' perspective-taking in social interactions: Relations with socio-cognitive skills. *Journal of Cognition and Development*, *22*(4), 537–560. https://doi.org/10.1080/15248372.2021.1901713

Chiew, K. S., & Braver, T. S. (2011). Positive affect versus reward: Emotional and motivational influences on cognitive control. *Frontiers in Psychology*, *2*(279). https://doi.org/10.3389/fpsyg.2011.00279

Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences*, *11*(9), 393–399. https://doi.org/10.1016/j.tics.2007.08.005

Converse, B. A., Lin, S. H., Keysar, B., & Epley, N. (2008). In the mood to get over yourself: Mood affects theory-of-mind use. *Emotion*, *8*(5), 725–730. https://doi.org/10.1037/a0013283

Cummings, A. J., & Rennels, J. L. (2014). How mood and task complexity affect children's recognition of others' emotions. *Social Development*, *23*(1), 80–99. https://doi.org/10.1111/sode.12038

Davis, E. L. (2016). An age-related mechanism of emotion regulation: Regulating sadness promotes children's learning by broadening information processing. *Child Development*, *87*(5), 1617–1626. https://doi.org/10.1111/cdev.12552

Davis, E. L., Parsafar, P., Quiñones-Camacho, L. E., & Shih, E. W. (2017). Psychophysiological assessment of the effectiveness of emotion regulation strategies in childhood. *Journal of Visualized Experiments*, *120*. https://doi.org/10.3791/55200

Dejonckheere, E., Mestdagh, M., Verdonck, S., Lafit, G., Ceulemans, E., Bastian, B., & Kalokerinos, E. K. (2021). The relation between positive and negative affect becomes more negative in response to personally relevant events. *Emotion*, *21*(2), 326–336. https://doi.org/10.1037/emo0000697

Drayton, S., Turley-Ames, K. J., & Guajardo, N. R. (2011). Counterfactual thinking and false belief: The role of executive function. *Journal of Experimental Child Psychology*, *108*(3), 532–548. https://doi.org/10.1016/j.jecp.2010.09.007

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40*(6), 760–768. https://doi.org/10.1016/j.jesp.2004.02.002

Epley, N., & Waytz, A. (2010). Mind perception. In S. Fiske, D. Gilbert, G. Lindzey, & A. Jongsma (Eds.), *The Handbook of Social Psychology* (5th ed., pp. 498–541). John Wiley & Sons. https://doi.org/10.1002/9780470561119.socpsy001014

Fernández-Aguilar, L., Navarro-Bravo, B., Ricarte, J., Ros, L., & Latorre, J. (2019). How effective are films in inducing positive and negative emotional states? A meta-analysis. *PLOS ONE*, *14*(11). https://doi.org/10.1371/journal.pone.0225040

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the level 1–level 2 distinction. *Developmental Psychology*, *17*(1), 99–103. https://doi.org/10.1037/0012-1649.17.1.99

Francis, G., Deniz, E., Torgerson, C., & Toseeb, U. (2022). Play-based interventions for mental health: A systematic review and meta-analysis focused on children and adolescents with autism spectrum disorder and developmental language disorder. *Autism & Developmental Language Impairments*, *7*, 1–44. https://doi.org/10.1177/23969415211073118

Gabel, L., Daoust, A., Salisbury, M., Grahn, J., Durbin, E., & Hayden, E. (2019). Development and validation of a battery of emotionally evocative film clips for use with young children. *Psychological Assessment*, *31*(8), 1040–1051. https://doi.org/10.1037/pas0000726

Ghrear, S., Baimel, A., Haddock, T., & Birch, S. A. J. (2021). Are the classic false belief tasks cursed? Young children are just as likely as older children to pass a false belief task when they are not required to overcome the curse of knowledge. *PLoS ONE*, *16*(2). https://doi.org/10.1371/journal.pone.0244141

Grosse Wiesmann, C., Friederici, A. D., Singer, T., & Steinbeis, N. (2017). Implicit and explicit false belief development in preschool children. *Developmental Science*, *20*(5). https://doi.org/10.1111/desc.12445

Gruenberg, A., Goldstein, R., & Pincus, H. (2005). Classification of depression: Research and diagnostic criteria: DSM-IV and ICD-10. In J. Licinio & M. Wong (Eds.), *Biology of Depression: From Novel Insights to Therapeutic Strategies* (pp. 1–12). Wiley VCH.

Harkness, K., Washburn, D., Theriault, J., Lee, L., & Sabbagh, M. (2011). Maternal history of depression is associated with enhanced theory of mind in depressed and nondepressed adult women. *Psychiatry Research*, *189*, 91–96.

Haskaraca, F. N., Proft, M., Liszkowski, U., & Rakoczy, H. (2023). How robust are egocentric and altercentric interference effects in social cognition? A test with explicit and implicit versions of a continuous false belief task. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1142302

Hirschfeld, R. M., Montgomery, S. A., Keller, M. B., Kasper, S., Schatzberg, A. F., Möller, H. J., Healy, D., Baldwin, D., Humble, M., Versiani, M., Montenegro, R., & Bourgeois, M. (2000). Social functioning in depression: A review. *The Journal of Clinical Psychiatry*, *61*(4), 268–275. https://doi.org/10.4088/jcp.v61n0405

Hughes, C., & Dunn, J. (1998). Understanding mind and emotion: Longitudinal associations with mental-state talk between young friends. *Developmental Psychology*, *34*(5), 1026–1037. https://doi.org/10.1037//0012-1649.34.5.1026

Kaysili, K., & Acarlar, F. (2011). The development of theory of mind according to false belief performance of children ages 3 to 5. *Educational Sciences*, *11*(4).

Kessler, R. C., Avenevoli, S., & Ries Merikangas, K. (2001). Mood disorders in children and adolescents: An epidemiologic perspective. *Biological Psychiatry*, *49*(12), 1002–1014. https://doi.org/10.1016/S0006-3223(01)01129-5

Kravčenko, K., & Šeibokaitė, L. (2018). The effect of joy and anxiety on egocentric decisions among young adults. *International Journal of Psychology*, *22*, 55–68. https://doi.org/10.7220/2345-024X.22.3

Leslie, A. M. (2005). Developmental parallels in understanding minds and bodies. *Trends in Cognitive Sciences*, *9*(10), 459–462. https://doi.org/10.1016/j.tics.2005.08.002

Liu, D., Wellman, H. M., Tardif, T., & Sabbagh, M. A. (2008). Theory of mind development in Chinese children: A meta-analysis of false-belief understanding across cultures and languages. *Developmental Psychology*, *44*(2), 523–531. https://doi.org/10.1037/0012-1649.44.2.523

Mangardich, H., Tollefson, N., Harkness, K., & Sabbagh, M. (2022). Theory of mind in dysphoric and non--dysphoric adults: An ERP study of true-- and false--belief reasoning. *Social Neuroscience*, *17*(1), 73–85. https://doi.org/10.1080/17470919.2021.2005678

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*, 255–258. https://doi.org/10.1126/science.1107621

Out, C., Goudbeek, M., & Krahmer, E. (2020). Do speaker's emotions influence their language production? Studying the influence of disgust and amusement on alignment in interactive reference. *Language Sciences*, *78*. https://doi.org/10.1016/j.langsci.2019.101255

Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science*, *308*, 214–216. https://doi.org/10.1126/science.1111656

Piaget, J., & Inhelder, B. (1956). *The child's conception of space*. London: Routledge & Kegan Paul.

Rubio-Fernández, P., & Geurts, B. (2013). How to pass the false-belief task before your fourth Birthday. *Psychological Science*, *24*(1), 27–33. https://doi.org/10.1177/0956797612447819

Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin*, *125*(1), 3–30. https://doi.org/10.1037/0033-2909.125.1.3

Samson, D., & Apperly, I. A. (2010). There is more to mind reading than having theory of mind concepts: New directions in theory of mind research. *Infant and Child Development*, *19*(5), 443–454. https://doi.org/10.1002/icd.678

Schnall, S., Jaswal, V. K., & Rowe, C. (2008). A hidden cost of happiness in children. *Developmental Science*, *11*(5), F25–F30. https://doi.org/10.1111/j.1467-7687.2008.00709.x

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. Higgins T. & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 527–561). New York: Guilford Press.

Siedlecka, E., & Denson, T. (2019). Experimental methods for inducing basic emotions: A qualitative review. *Emotion Review*, *11*(1), 87–97.

Smogorzewska, J., Szumski, G., & Grygiel, P. (2019). Theory of mind development in school environment: A case of children with mild intellectual disability learning in inclusive

and special education classrooms. *Journal of Applied Research in Intellectual Disabilities*, *32*(5), 1241–1254. https://doi.org/10.1111/jar.12616

Surian, L., & Leslie, A. (1999). Competence and performance in false belief understanding: A comparison of autistic and normal 3-year-old children. *British Journal of Developmental Psychology*, *17*, 141–155. https://doi.org/10.1348/026151099165203

Tan, C. C., & Holub, S. C. (2018). The effects of happiness and sadness on Children's snack consumption. *Appetite*, *123*, 169–174. https://doi.org/10.1016/j.appet.2017.12.021

Tay, L., & Kuykendall, L. (2017). Why self-reports of happiness and sadness may not necessarily contradict bipolarity: A psychometric review and proposal. *Emotion Review*, *9*(2), 146–154. https://doi.org/10.1177/1754073916637656

Timmermans, B., & Schilbach, L. (2014). Investigating alterations of social interaction in psychiatric disorders with dual interactive eye tracking and virtual faces. *Frontiers in Human Neuroscience*, *8*, 758. https://doi.org/10.3389/fnhum.2014.00758

Todd, A. R., Forstmann, M., Burgmer, P., Brooks, A. W., & Galinsky, A. D. (2015). Anxious and egocentric: How specific emotions influence perspective taking. *Journal of Experimental Psychology*, *144*(2), 374–391. https://doi.org/10.1037/xge0000048

Todd, A. R., & Tamir, D. I. (2024). Factors that amplify and attenuate egocentric mentalizing. *Nature Reviews Psychology*, *3*(3), 164–180. https://doi.org/10.1038/s44159-024-00277-1

Tornare, E., Cuisinier, F., Czajkowski, N. O., & Pons, F. (2017). Impact of induced joy on literacy in children: Does the nature of the task make a difference? *Cognition and Emotion*, *31*(3), 500–510. https://doi.org/10.1080/02699931.2015.1132682

von Leupoldt, A., Rohde, J., Beregova, A., Thordsen-Sörensen, I., zur Nieden, J., & Dahme, B. (2007). Films for eliciting emotional states in children. *Behavior Research Methods*, *39*(3), 606–609. https://doi.org/10.3758/bf03193032

Wagner, S., Müller, C., Helmreich, I., Huss, M., & Tadić, A. (2015). A meta-analysis of cognitive functions in children and adolescents with major depressive disorder. *European Child & Adolescent Psychiatry*, *24*(1), 5–19. https://doi.org/10.1007/s00787-014-0559-2

Walker, S. (2005). Gender differences in the relationship between young children's peer-related social competence and individual differences in theory of mind. *The Journal of Genetic Psychology*, *166*(3), 297–312. https://doi.org/10.3200/GNTP.166.3.297-312

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, *98*(2), 219–235. https://doi.org/10.1037//0033-2909.98.2.219

Wellman, H., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655–684. https://doi.org/10.1111/1467-8624.00304

Westermann, R., Spies, K., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, *26*(4), 557–580. https://doi.org/10.1002/(SICI)1099-0992(199607)26:4<557::AID-EJSP769>3.0.CO;2-4

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

Wu, Y., Haque, J., & Schulz, L. (2018). Children can use others' emotional expressions to infer their knowledge and predict their behaviors in classic false belief tasks. *Proceedings of the 40th Annual Meeting of the Cognitive Science Society*.

Yang, M.-L., Yang, C.-C., & Chiou, W.-B. (2010). When guilt leads to other orientation and shame leads to egocentric self-focus: Effects of differential priming of negative affects on perspective taking. *Social Behavior and Personality*, *38*(5), 605–614. https://doi.org/10.2224/sbp.2010.38.5.605

Zhai, J., Xie, J., Chen, J., Huang, Y., Ma, Y., & Huang, Y. (2021). The presence of other-race people disrupts spontaneous level-2 visual perspective taking. *Scandinavian Journal of Psychology*, *62*. https://doi.org/10.1111/sjop.12751

CHAPTER III: PRESS RELEASE FOR THE LITERATURE REVIEW

Depression Linked to Executive Dysfunction in Children and Adolescents

In comparison to healthy children and adolescents, executive function is significantly worse in children and adolescents with depression, emphasising the need for adapted treatment and support in this population.

Depression is increasingly common in children and adolescents (Erskine et al., 2017). Executive function is a complex set of mental skills required to manage and carry out day-to-day tasks (Majeed et al., 2023). It involves skills like avoiding the most obvious answer ("inhibition"), switching between tasks ("set shifting") and, holding in mind and manipulating information ("working memory"; Miyake et al., 2000). It is well established in adults with depression that their executive function ability is significantly impaired (Snyder, 2013). However, it is less clear if this is the case for children and adolescents with depression, in which there is a large amount of inconsistency within the current scientific literature (Goodall et al., 2018; Wagner et al., 2015). Therefore, this review aimed to conclude if children and adolescents with depression do show impaired executive function abilities.

Researchers at the University of Birmingham undertook a comprehensive review of 28 studies published in the current scientific literature. It was found that compared to healthy children and adolescents, children and adolescents with depression performed significantly worse on executive function tasks assessing the three core areas described above, with the greatest difficulty found in the working memory skill.

It is currently not clear why working memory is particularly impacted in children and adolescents with depression. In thinking about the overall executive function difficulties found, one idea considers how common symptoms of depression (e.g., negative thinking) can lead to people finding it difficult to allocate their cognitive resources to tasks requiring substantial mental effort, such as executive function tasks (Nuño et al., 2021). Another possibility is that impairments in executive function may be a risk factor for depression, making it harder for people to regulate their negative emotions and thoughts, which can then lead to the development of symptoms of depression (Nelson et al., 2018). More research conducted with children with depression over longer periods of time will help us better understand the relationship between childhood/adolescent depression and executive function.

The review also explored whether the relationship between executive function and childhood/adolescent depression was affected by factors such as the age, sex and intellectual ability of participants, the presence of other co-occurring mental health difficulties in participants with depression and the use of mental health medication in participants with depression. Additionally, the impact of executive function tasks assessing accuracy compared to those assessing reaction time was explored. Interestingly, the only factors found to have an impact were the age and intellectual ability of participants with depression. This indicated that executive function difficulties in children and adolescents with depression was more pronounced in younger children and those with lower intellectual ability. However, it was highlighted that future research would be required to better understand the impact of these factors.

Variability in the different executive function tasks used in the reviewed studies, and how the studies were conducted may have influenced the findings reported. It is recommended that additional research be carried out to address these potential limitations.

For media enquiries please contact Bryony Fenton, School of Psychology, University of Birmingham, email:

References

Erskine, H. E., Baxter, A. J., Patton, G., Moffitt, T. E., Patel, V., Whiteford, H. A., & Scott, J. G. (2017). The global coverage of prevalence data for mental disorders in children and adolescents. *Epidemiology and Psychiatric Sciences*, *26*(4), 395–402. https://doi.org/10.1017/S2045796015001158

Goodall, J., Fisher, C., Hetrick, S., Phillips, L., Parrish, E. M., & Allott, K. (2018). Neurocognitive functioning in depressed young people: A systematic review and meta-analysis. *Neuropsychology Review*, *28*(2), 216–231. https://doi.org/10.1007/s11065-018-9373-9

Majeed, N., Chua, Y., Kothari, M., Kaur, M., Quek, F., Ng, M., Ng, W., & Hartanto, A. (2023). Anxiety disorders and executive functions: A three-level meta-analysis of reaction time and accuracy. *Psychiatry Research Communications*, *3*. https://doi.org/10.1016/j.psycom.2022.100100

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

Nelson, T. D., Kidwell, K. M., Nelson, J. M., Tomaso, C. C., Hankey, M., & Espy, K. A. (2018). Preschool executive control and internalizing symptoms in elementary school. *Journal of Abnormal Child Psychology*, *46*(7), 1509–1520. https://doi.org/10.1007/s10802-017-0395-1

Nuño, L., Gómez-Benito, J., Carmona, V. R., & Pino, O. (2021). A systematic review of executive function and information processing speed in major depression disorder. *Brain Sciences*, *11*(2), 147. https://doi.org/10.3390/brainsci11020147

Snyder, H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: A meta-analysis and review. *Psychological Bulletin*, *139*(1), 81–132. https://doi.org/10.1037/a0028727

Wagner, S., Müller, C., Helmreich, I., Huss, M., & Tadić, A. (2015). A meta-analysis of cognitive functions in children and adolescents with major depressive disorder. *European Child & Adolescent Psychiatry*, *24*(1), 5–19. https://doi.org/10.1007/s00787-014-0559-2

CHAPTER IV: PRESS RELEASE FOR THE EMPIRICAL RESEARCH PAPER

Happiness Can Impact Children's Ability to Understand Others

A research study revealed that being made to feel happy led to 3-4 ½ - year-old-children making more errors when they had to think about someone else's beliefs.

Theory of Mind (ToM) is a skill that helps us understand and reason about what someone else might be thinking or what they might believe (Epley & Waytz, 2010). ToM is needed for effective socialising in children (Brezack et al., 2021). However, mistakes in ToM are common. One particular mistake results from people struggling to separate their own knowledge and beliefs from another's knowledge and beliefs, often called egocentric bias (Epley et al., 2004). It is important to find out what factors may make this difficulty worse. In adults, mood states such as happiness have been found to influence the level of egocentric bias an individual makes (Converse et al., 2008). However, there is limited evidence on the impact of mood on ToM ability in children. This study aimed to address whether happiness in children impacted their ToM, specifically their likelihood of incorrectly relying on their own knowledge/beliefs.

Researchers at the University of Birmingham recruited ninety, 3-4 ½-year-old children from a local nursery and infant school. All the participants completed an experimental false belief task (Baron-Cohen et al., 1985). False belief tasks are often used as a measure of ToM (Grosse Wiesmann et al., 2017) and this specific task has been well used within the field. The task requires the child to conclude a character's false belief about the location of an item that has been moved without the character knowing, but that the child has witnessed being moved.

Before completing the task, the participants were made to feel either happy or neutral in mood by watching relevant brief video clips.

The results showed that children made to feel happy were more likely to be influenced by their own knowledge/belief of knowing where the object was and incorrectly stated that the character would look in the location the object had been moved to as opposed to the original location. Why might this be the case? One potential explanation is that happiness may result in a broader thinking style (Clore & Huntsinger, 2007). This thinking style can offer less attention to detail and increase the use of information that is more easily accessed, such as our own knowledge (Clore & Huntsinger, 2007). Using a broader thinking style may then lead to children being less likely to use their more focused thinking, which is needed to help overcome a reliance on our own knowledge (Apperly & Butterfill, 2009).

What might this mean? Well, for clinicians or mental health staff working therapeutically with children, they may need to consider the impact of influencing a child's mood, particularly as child psychological interventions are often game/play based. Secondly, these findings may help contribute to an increased understanding of social difficulties/functioning in children with mood disorders, such as depression.

This study offers initial evidence regarding how mood may impact children's ToM ability. However, more research is required to further understand why these findings occurred, broaden testing procedures, and extend the findings to other emotional states and childhood age groups.

For media enquiries please contact Bryony Fenton, School of Psychology, University of Birmingham, email: [redacted]

References

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*(4), 953–970. https://doi.org/10.1037/a0016923

Baron-Cohen, S., Leslie, A., & 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

Brezack, N., Meyer, M., & Woodward, A. L. (2021). Three-year-olds' perspective-taking in social interactions: Relations with socio-cognitive skills. *Journal of Cognition and Development*, *22*(4), 537–560. https://doi.org/10.1080/15248372.2021.1901713

Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences*, *11*(9), 393–399. https://doi.org/10.1016/j.tics.2007.08.005

Converse, B. A., Lin, S. H., Keysar, B., & Epley, N. (2008). In the mood to get over yourself: Mood affects theory-of-mind use. *Emotion*, *8*(5), 725–730. https://doi.org/10.1037/a0013283

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40*(6), 760–768. https://doi.org/10.1016/j.jesp.2004.02.002

Epley, N., & Waytz, A. (2010). Mind perception. In S. Fiske, D. Gilbert, G. Lindzey, & A. Jongsma (Eds.), *The Handbook of Social Psychology* (5th ed., pp. 498–541). John Wiley & Sons. https://doi.org/10.1002/9780470561119.socpsy001014

Grosse Wiesmann, C., Friederici, A. D., Singer, T., & Steinbeis, N. (2017). Implicit and explicit false belief development in preschool children. *Developmental Science*, *20*(5). https://doi.org/10.1111/desc.12445

CHAPTER V: APPENDICES

Appendix 1: PRISMA checklist

Appendix 2: Output from the Three-Level Meta-Analytic Model

Appendix 3: Sensitivity Analysis Output

Appendix 4: Weighted Standardised Mean Differences for Each Effect in the Primary Studies

Appendix 5: Ethical Approval Letter from the Ethics Committee

Appendix 6: Headteacher Information Sheet

Appendix 7: Headteacher Consent Form

Appendix 8: Parent/Carer Information Sheet

Appendix 9: Mood Manipulation Check Scale

Appendix 10: Perspective Taking Task Materials

Appendix 11: Perspective Taking Task Record Form

Appendix 1 – PRISMA Checklist

Section and topic	Item #	Checklist item	Location where item is reported
Title			
Title	1	Identify the report as a systematic review.	Title page
Abstract			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Abstract
Introduction			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Introduction (rationale)
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Introduction (rationale)
Methods			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Method (inclusion and exclusion criteria) and Table 1-2
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Method (search of electronic databases and paper selection)
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Method (search of electronic databases) and Table 1-1
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Method (paper selection) and Figure 1-1
Data collection processes	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any	Method (data extraction)

		processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Method (data extraction)
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Method (data extraction)
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Method (risk of bias assessment) and Table 1-3
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Method (data analysis strategy)
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Method (inclusion and exclusion criteria) and Table 1-2
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Method (data analysis strategy)
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Figure 1-3
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of	Method (data analysis strategy)

		statistical heterogeneity, and software package(s) used.	
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Method (data analysis strategy)
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Method (data analysis strategy)
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	N/A
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Method (data analysis strategy)
Results			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Results (results of the search) and Figure 1-1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Results (results of the search) and Figure 1-1
Study characteristics	17	Cite each included study and present its characteristics.	Table 1-5
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Results (risk of bias review) and Table 1-6
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Figure 1-3, Table1-9, Table 1-10 and Table 1-11.
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Table 1-5 and Table 1-6
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and	Results (meta-

		its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	analysis results)
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Results (meta-analysis results, further subgroup analysis and meta-regression)
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Results (meta-analysis results) Appendix 2 and 3
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	N/A
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Results (overall executive function
Discussion			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Discussion (comparison with related reviews)
	23b	Discuss any limitations of the evidence included in the review.	Discussion (limitations)
	23c	Discuss any limitations of the review processes used.	Discussion (limitations)
	23d	Discuss implications of the results for practice, policy, and future research.	Discussion (future research directions and clinical implications)
Other information			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Method (opening paragraph)

	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Method (opening paragraph)
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Method (opening paragraph)
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Not currently but will be made available on submission of final paper.
Competing interests	26	Declare any competing interests of review authors.	Not currently but will be made available on submission of final paper.
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Not currently but will be made available on submission of final paper.

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj n71

Appendix 2 – Output from the Three-Level Meta-Analytic Model

The below tables report the standardised mean differences derived from the three level meta-analytic model. For all below analyses, the estimated standardised mean differences were not meaningfully different to when calculated with the two-level random effects model.

Executive Functioning Subdomain

	K	SMD	95% CI	tau^2	Q	I^2
subgroup = Shifting	17	0.2787	[0.0771; 0.4803]	0.1331	36.84	56.6%
subgroup = Working memory	24	0.5000	[0.3966; 0.6034]	0.0003	17.93	0.0%
subgroup = Inhibition	19	0.2404	[0.0778; 0.4031]	0.0417	29.32	38.6%
subgroup = Combined	1	1.1543	[0.4675; 1.8411]	--	0.00	--

Executive Function Task Outcome Measure:

	K	SMD	95% CI	tau^2	Q	I^2
subgroup = Accuracy	42	0.3812	[0.2909; 0.4714]	0.0154	51.22	19.9%
subgroup = Reaction Time	18	0.3283	[0.1143; 0.5423]	0.2272	44.60	61.9%

Medication Use in Participants with Depression:

	K	SMD	95% CI	tau^2	Q	I^2
subgroup = Included	22	0.3416	[0.1499; 0.5333]	0.1614	43.34	51.5%
subgroup = Excluded	33	0.4021	[0.2597; 0.5444]	0.0705	50.35	36.4%

Presence of Comorbidities in Participants with Depression:

	K	SMD	95% CI	tau^2	Q	I^2
subgroup = Included	43	0.3381	[0.2170; 0.4591]	0.0890	68.88	39.0%
subgroup = Excluded	14	0.4760	[0.1923; 0.7597]	0.1808	26.40	50.7%

Meta-Regression Results

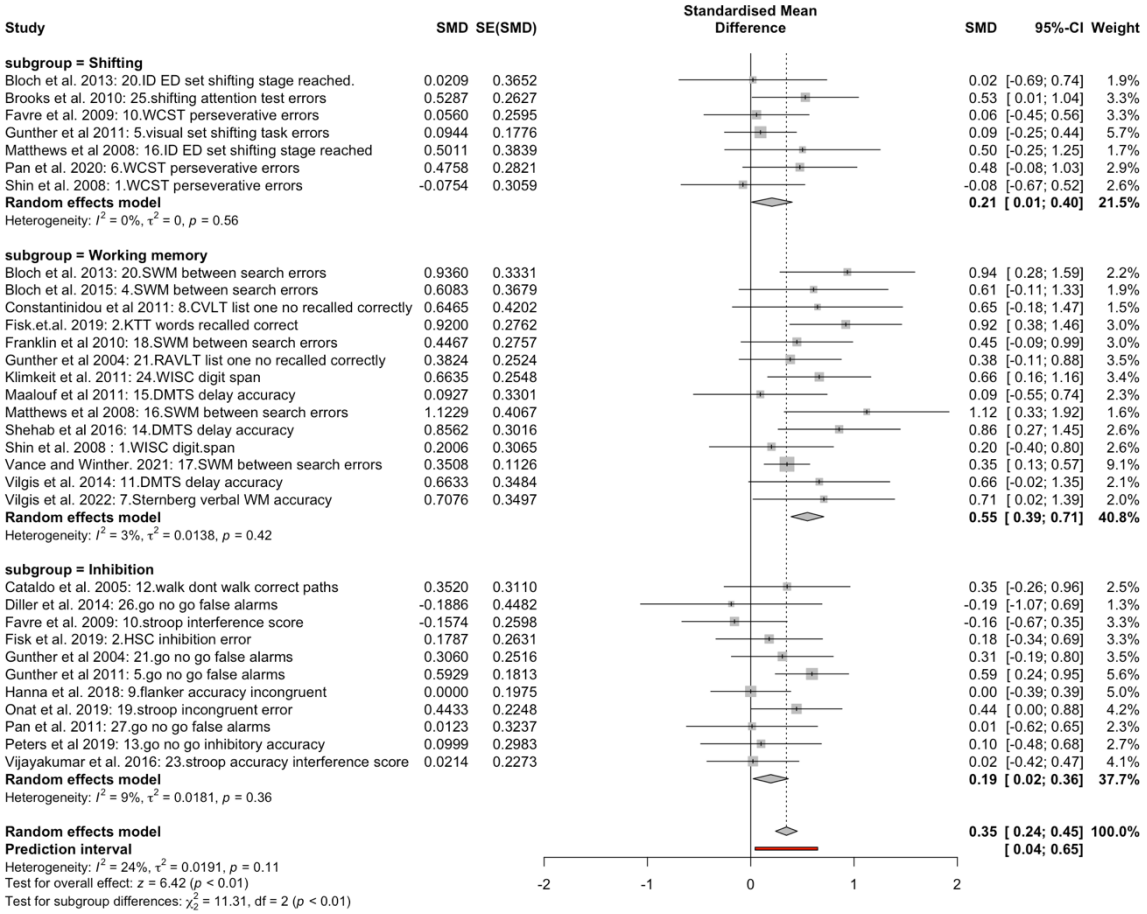
	Coefficient	SE	Z	p
Age	-0.0628	0.0270	-2.3288	0.0199
FSIQ	-0.0166	0.0091	-1.8234	0.0682
% male	0.0012	0.0022	0.5296	0.5964

Appendix 3 – Sensitivity Analysis Output

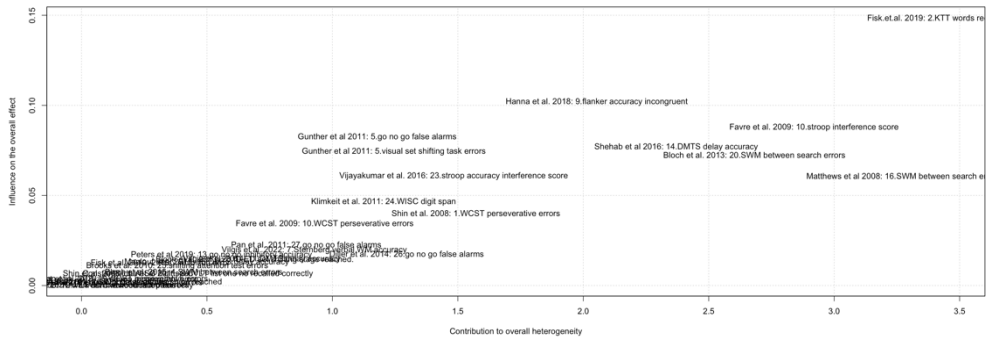
Output when only one effect, per executive functioning subdomain, per study was included. These analyses were separated into executive function tasks reporting accuracy data and tasks reporting reaction time data.

Accuracy Data

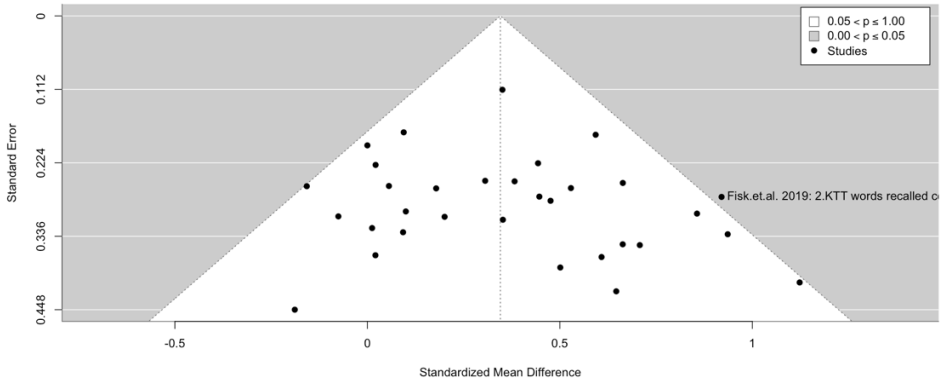
Forest Plot



Baujat Plot



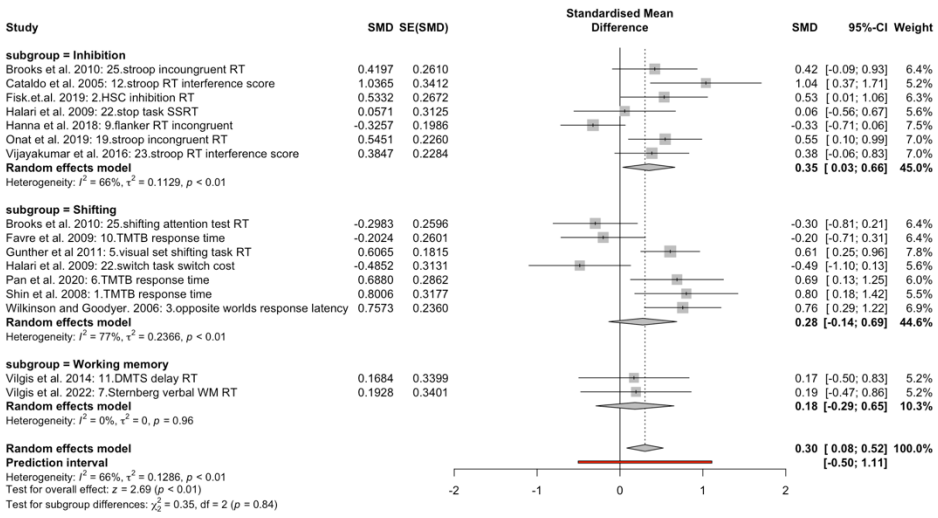
Funnel Plot



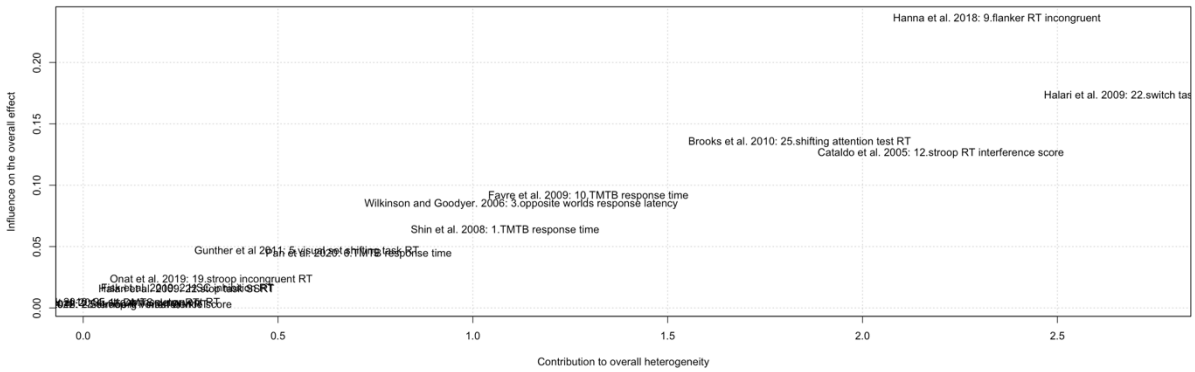
Egger's test of asymmetry was non-significant ($t = 0.83$, $p = 0.41$).

Reaction Time Data

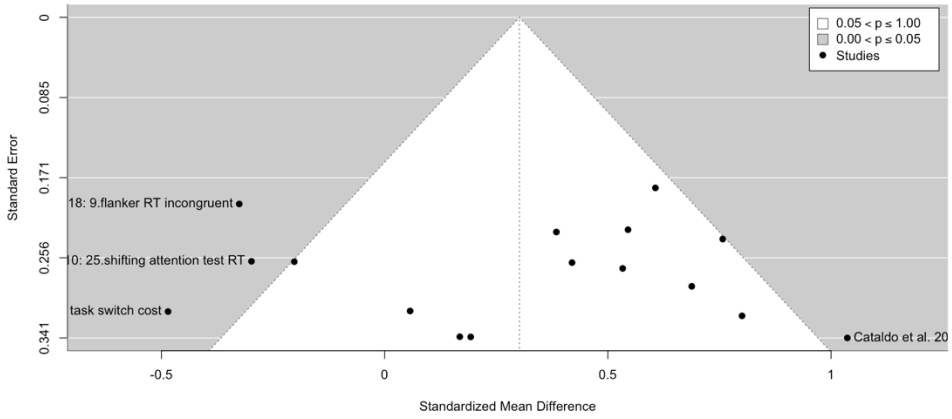
Forest Plot



Baujat Plot



Funnel Plot



Egger's test of asymmetry was non-significant ($t = 0.05$, $p = 0.96$).

Appendix 4 – Weighted Standardised Mean Differences for Each Effect in the Primary Studies

Study	Effect	Std.Er	CI lower	CI upper	Random effects weighting
Bloch et al. 2013: 20.ID ED set shifting error no	-0.004202	0.3651	-0.719881	0.71148	5.425
Bloch et al. 2013: 20.ID ED set shifting stage reached.	0.020935	0.3652	-0.694761	0.73663	5.425
Bloch et al. 2013: 20.SWM between search errors	0.935999	0.3331	0.283148	1.58885	6.175
Bloch et al. 2013: 20.WM SSP span score	0.495952	0.3211	-0.133299	1.1252	6.491
Bloch et al. 2015: 4.SWM between search errors	0.608277	0.3679	-0.112755	1.32931	5.367
Bloch et al. 2015: 4.WM SSP span score	0.371801	0.3629	-0.339519	1.08312	5.473
Brooks et al. 2010: 25.shifting attention test errors	0.528733	0.2627	0.013906	1.04356	8.335
Brooks et al. 2010: 25.shifting attention test RT	-0.298343	0.2596	-0.807211	0.21053	8.446
Brooks et al. 2010: 25.stroop incoungruent RT	0.419718	0.261	-0.091884	0.93132	8.395
Cataldo et al. 2005: 12.stroop RT interference score	1.036492	0.3412	0.367789	1.7052	5.974
Cataldo et al. 2005: 12.walk dont walk correct paths	0.352023	0.311	-0.257502	0.96155	6.77
Constantinidou et al 2011: 8.CVLT list one no recalled correctly	0.646456	0.4202	-0.177052	1.46996	4.395
Diller et al. 2014: 26.go no go false alarms	-0.18857	0.4482	-1.067039	0.6899	3.97
Favre et al. 2009: 10.stroop interference score	-0.15738	0.2598	-0.66661	0.35185	8.439
Favre et al. 2009: 10.TMTB error no	0.24827	0.2604	-0.262062	0.7586	8.419
Favre et al. 2009: 10.TMTB response time	-0.202353	0.2601	-0.712067	0.30736	8.43
Favre et al. 2009: 10.WCST perseverative errors	0.055966	0.2595	-0.452615	0.56455	8.452
Fisk et al. 2019: 2.HSC inhibition error	0.178687	0.2631	-0.337051	0.69442	8.318
Fisk.et.al. 2019: 2.HSC inhibition RT	0.533216	0.2672	0.009438	1.05699	8.17
Fisk.et.al. 2019: 2.KTT words recalled correct	0.92001	0.2762	0.378754	1.46127	7.859
Franklin et al 2010: 18.SWM between search errors	0.446654	0.2757	-0.093753	0.98706	7.873
Franklin et al 2010: 18.WM SSP span score	0.852551	0.2844	0.29506	1.41004	7.582
Gunther et al 2004: 21.go no go false alarms	0.306009	0.2516	-0.187079	0.7991	8.75
Gunther et al 2004: 21.RAVLT list one no recalled correctly	0.382373	0.2524	-0.112312	0.87706	8.719
Gunther et al 2011: 5.go no go false alarms	0.592927	0.1828	0.234598	0.95126	11.847
Gunther et al 2011: 5.visual set shifting task errors	0.09438	0.179	-0.256526	0.44528	12.042
Gunther et al 2011: 5.visual set shifting task RT	0.606528	0.183	0.24785	0.96521	11.838
Halari et al. 2009: 22.simon task simon effect	0.055025	0.3087	-0.549947	0.66	6.837
Halari et al. 2009: 22.stop task SSRT	0.057075	0.3125	-0.555422	0.66957	6.727
Halari et al. 2009: 22.switch task switch cost	-0.485246	0.3131	-1.098941	0.12845	6.71
Hanna et al. 2018: 9.flanker accuracy incongruent	0	0.1975	-0.38713	0.38713	11.111
Hanna et al. 2018: 9.flanker RT incongruent	-0.32572	0.1986	-0.71495	0.06351	11.059
Klimkeit et al. 2011: 24.WISC digit span	0.66355	0.2548	0.164086	1.16301	8.626
Maalouf et al 2011: 15.DMTS delay accuracy	0.092686	0.3301	-0.55422	0.73959	6.253

Study	Effect	Std.Er	CI lower	CI upper	Random effects weighting
Matthews et al 2008: 16.DMTS delay accuracy	0.783743	0.3922	0.015033	1.55245	4.883
Matthews et al 2008: 16.ID ED set shifting stage reached	0.501088	0.3839	-0.251244	1.25342	5.042
Matthews et al 2008: 16.SWM between search errors	1.12288	0.4067	0.325841	1.91992	4.622
Matthews et al 2008: 16.WM SSP span score	0.286273	0.3799	-0.458309	1.03085	5.12
Onat et al. 2019: 19.stroop incongruent error	0.443313	0.2248	0.002733	0.88389	9.851
Onat et al. 2019: 19.stroop incongruent RT	0.545131	0.226	0.102173	0.98809	9.798
Pan et al. 2011: 27.go no go false alarms	0.012268	0.3237	-0.622118	0.64665	6.421
Pan et al. 2020: 6.TMTB response time	0.687981	0.2862	0.126965	1.249	7.523
Pan et al. 2020: 6.WCST perseverative errors	0.475791	0.2821	-0.077036	1.02862	7.66
Pandina: 28.EF composite score	1.154265	0.3504	0.467456	1.84107	5.754
Peters et al 2019: 13.go no go inhibitory accuracy	0.099875	0.2983	-0.484815	0.68457	7.144
Shehab et al 2016: 14.DMTS delay accuracy	0.856204	0.3016	0.265056	1.44735	7.044
Shin et al. 2008 : 1.WISC digit.span	0.200628	0.3065	-0.400114	0.80137	6.9
Shin et al. 2008: 1.TMTB error.no	0.963355	0.3229	0.330467	1.59624	6.441
Shin et al. 2008: 1.TMTB response time	0.800621	0.3177	0.177946	1.4233	6.582
Shin et al. 2008: 1.WCST perseverative errors	-0.075359	0.3059	-0.674814	0.5241	6.919
Vance and Winther. 2021: 17.SWM between search errors	0.350802	0.1126	0.130037	0.57157	15.705
Vance and Winther. 2021: 17.WM SSP span score	0.483631	0.1132	0.26172	0.70554	15.673
Vijayakumar et al. 2016: 23.stroop accuracy interference score	0.02141	0.2273	-0.424138	0.46696	9.741
Vijayakumar et al. 2016: 23.stroop RT interference score	0.384719	0.2284	-0.063017	0.83246	9.693
Vilgis et al. 2014: 11.DMTS delay accuracy	0.663254	0.3484	-0.019692	1.3462	5.8
Vilgis et al. 2014: 11.DMTS delay RT	0.168356	0.3399	-0.497849	0.83456	6.005
Vilgis et al. 2022: 7.Sternberg spatial WM accuracy	0.576982	0.3462	-0.101651	1.25562	5.852
Vilgis et al. 2022: 7.Sternberg spatial WM RT	0.370789	0.3422	-0.299896	1.04147	5.949
Vilgis et al. 2022: 7.Sternberg verbal WM accuracy	0.707643	0.3497	0.022257	1.39303	5.771
Vilgis et al. 2022: 7.Sternberg verbal WM RT	0.192776	0.3401	-0.473792	0.85934	6.001
Wilkinson and Goodyer. 2006: 3.opposite worlds response latency	0.757268	0.236	0.294781	1.21976	9.375

Appendix 5 – Ethical Approval Letter from the Ethics Committee

Dear Andrew Surtees,

RE: Children's mood and Theory of Mind
Application for Ethical Review: ERN_0372-Apr2023

Thank you for your application for ethical review for the above project, which was reviewed by the Science, Technology, Engineering and Mathematics Committee.

On behalf of the Committee, I confirm that this study now has ethical approval.

Any adverse events occurring during the study should be promptly brought to the Committee's attention by the Principal Investigator and may necessitate further ethical review.

Please ensure that the relevant requirements within the University's Code of Practice for Research and the information and guidance provided on the University's ethics webpages (available at https://intranet.birmingham.ac.uk/finance/accounting/Research-Support-Group/Research-Ethics/Links-and-Resources.aspx) are adhered to.

Please be aware that whilst Health and Safety (H&S) issues may be considered during the ethical review process, you are still required to follow the University's guidance on H&S and to ensure that H&S risk assessments have been carried out as appropriate. For further information about this, please contact your School H&S representative or the University's H&S Unit at healthandsafety@contacts.bham.ac.uk.

Kind regards,

The Co-Chairs of the Science, Technology, Engineering and Mathematics Committee

E-mail: ethics-queries@contacts.bham.ac.uk

Appendix 6 – Headteacher Information Sheet

UNIVERSITY OF BIRMINGHAM

The effect of mood on children's perspective taking

Dear Headteacher,

Our names are Bryony Fenton and Pardis Hashmezadeh and we are postgraduate psychology students from The University of Birmingham's School of Psychology. Please read this information sheet carefully before deciding whether you are happy for the children in your school/nursery to take part in the study - 'The effect of mood on children's perspective taking'.

What does the research involve?
The research project aims to explore how emotions impact children's ability to take the perspectives of others. Perspective taking is a key component of successful social interaction. Most studies of perspective taking involve simple tasks in neutral scenarios. However, in the real world, we often have to engage with others in emotional situations. Our study looks at what difference that makes to children.

Prior to testing, class teachers will be asked to provide the initials, class name/number and year and month of birth for each child participating. They will also be asked to indicate if the child has an intellectual disability and/or any neurodevelopmental disorders (to their knowledge). We will provide you with letters to inform parents about the study and give them the option to request that their child does not take part.

The research will involve children watching a brief age-appropriate film clip, with the aim to induce either a happy or neutral mood. Children will then complete a simple age-appropriate task with the experimenter. During the task they will be asked to try and take another's perspective. It is anticipated that testing will take between 10-15 minutes for each child, to minimise the impact on their classroom activities. The task is designed to be fun and engaging and children will be offered a sticker for taking part. Children also have the choice to stop the task at any point.

Who can take part?
Children between the ages of three and ten years old are invited to take part. Because children tend to enjoy testing, we invite all children of the above ages to participate. However, children who have an intellectual disability and/or any neurodevelopmental disorders would not be able to have their scores included in the data analysis.

Is the data anonymous?
All records will be kept confidential. Any personal details (e.g., child initials) will be kept separately from any other data in an encrypted electronic folder. Participants will be identified through the study by an ID number. Any hard copies will be stored in a locked file cabinet at The University of Birmingham. Electronic copies of data will be kept on secure University computer systems. Only the researchers and supervisor will have access to the data. At the end of the study, any personal details will be destroyed. Ten years after the end of the study, we will destroy all anonymous data collected during the study.

Once the study is completed, an anonymised version of the data, in which no child could be individually identified, will be made publicly available in line with good practice in open research.

Personal identifying 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. 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).

Can children withdraw from the study?
Children's participation is voluntary, and they are free to stop taking part in the study at any point. Children can be withdrawn from the study up to 14 days following participation, without giving a reason and we will destroy all their data. After this point it may not be possible to withdraw a child as any records we hold of their personal details may have been destroyed. This means that we would no longer be able to trace a child's results back to them and withdraw them from the study. If you wish to withdraw a child, you may contact Pardis or Bryony (details below).

What are some of the potential risks of taking part?
Participating in this research will not expose children to any greater risks than in their everyday environment.

What do we have to do as a school/nursery?

If you are happy for the children in your school/nursery to take part in the study, please read and complete the attached consent form.

We will then contact you to arrange a set of dates where we are able to visit and carry out testing. Prior to testing we would ask you to please send out provided parental information sheets to the parents/carers of all children who are the eligible age to participate

Further information
Please do not hesitate to contact a member of the research team if you have any questions or require any more information. If you would prefer a verbal explanation of the research, please

contact Bryony or Pardis who will be happy to help with this. A summary of the study results will be provided to you on its completion.

Yours sincerely,

Bryony Fenton and Pardis Hashmezadeh

Contact details

Bryony Fenton, Trainee Clinical Psychologist)

Pardis Hashmezadeh, Trainee Clinical Psychologist

Dr Andrew Surtees, PhD, ClinPsyD

This study has been approved by the University of Birmingham Science Technology Engineering and Mathematics Ethical Review Committee. If you have any concerns about the study, then please contact the Head of School of Psychology, Professor Ed Wilding on

Thank you for taking the time to read this information sheet.

Appendix 7 – Headteacher Consent Form

UNIVERSITY OF BIRMINGHAM

The effect of mood on children's perspective taking

I confirm that I have read and understood the attached information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

I understand that participation of all children is voluntary and that I am free to withdraw consent without giving any reason. ☐

I understand that I can contact the researchers up to 14 days after participation in the study to withdraw any child's data. If I do this the child's data will be destroyed. ☐

I understand that all information collected during the study will be confidential. Only members of the research team will know who has participated in the study. All information collected during the study will be stored in locked or password protected storage that only members of the research team will have access to. No names will be published in any reports. Anonymous datasets (with all personal information removed) will be made publicly available. Information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 2018. ☐

I understand that my contact details will only be used by the research team for the purpose of this study alone. ☐

I agree to distribute parental information sheets to the parents/carers of all children who are eligible to participate in the study. ☐

I consent for the eligible children in my school to take part in the study 'The effect of mood on children's perspective taking'. ☐

Print Name: Signature:_____

Date:
Name of school/nursery:

Address:

Email:

Telephone number:

Relationship to participants:

Appendix 8 – Parent/Carer Information Sheet

UNIVERSITY OF BIRMINGHAM

The effect of mood on children's perspective taking

Dear Parent/guardian

We are writing to inform you about an upcoming research study that will be taking place at your child's school. Your child's school has kindly accepted for your child to take part in this research. In this letter, you will be given information about the study and a choice to withdraw your child's participation.

What does the research involve?
The research project aims to explore how emotions impact children's ability to take the perspectives of others. Perspective taking is a key component of successful social interaction. Therefore, this research aims to contribute to the developing knowledgebase of how emotions may shape the processes involved in perspective taking and at what point in the lifespan.

Children will be invited to watch an age-appropriate short film clip, before being asked to complete a task with the experimenter. The task will involve listening to a brief story and answering questions about perspective taking. It is anticipated that testing will take 10-15 minutes and should not disrupt your child's learning. The task is designed to be fun and engaging. Children also have the choice to stop the task at any point.

Who can take part?
Children between the ages of three and ten years old are invited to take part. As children tend to enjoy taking part in new activities, we invite all children to participate, however if your child has an intellectual disability and/or any neurodevelopmental conditions their data will not be included in the analysis. This is because previous studies have demonstrated that 'neurodivergent' children can display more difficulties in perspective taking (Baron-Cohen et al., 1985; Smogorzewska et al., 2019) which may impact our findings.

Is the data anonymous?
All records will be kept confidential. We will record your child's initials, class name/number and year and month of birth. This will help us identify your child if you decide to withdraw them from the study. At the end of the study, all personal details will be destroyed. Ten years after the end of the study, we will destroy all anonymous data collected during the study. Once the study is completed, an anonymised version of the data will be made publicly available.

Personal identifying 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. 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).

Can I withdraw from the study?
Children's participation is voluntary, and they are free to stop taking part in the study at any point. Children can be withdrawn from the study up to 14 days following participation, without giving a reason and we will destroy all their data. After this point it may not be possible to withdraw your child as any records we hold of their personal details may have been destroyed. This means that we would no longer be able to trace a child's results back to them and withdraw them from the study.

What are some of the potential risks of taking part?
Participating in this research will not expose children to any greater risks than in their everyday environment.

Further information
Please do not hesitate to contact a member of the research team if you have any questions or require any more information. If you would prefer a verbal explanation of the research, please contact Bryony or Pardis who will be happy to help with this. Individual children's results on the task may not be meaningful to share and therefore we will not be able to provide individual feedback. A summary of the study results will be provided to you on its completion.

If you would rather your child did not take part, please let us know via email using the below details or talk with your child's teacher.

Yours sincerely,

Bryony Fenton and Pardis Hashmezadeh

Contact details
Bryony Fenton, Trainee Clinical Psychologist

Pardis Hashmezadeh, Trainee Clinical Psychologist

Dr Andrew Surtees, PhD, ClinPsyD

This study has been approved by the University of Birmingham Science, Technology, Engineering and Mathematics Ethics Review Committee. If you have any concerns about the study, then please contact the Head of the School of Psychology, Professor Ed Wilding

Thank you for taking the time to read this information sheet.

Appendix 9 – Mood Manipulation Check Scale

How did the video make you feel?

Date –

Initials –

Age –

Sex -

Condition –

Rating score (1-4) –

Appendix 10 – Perspective Taking Task Materials

Appendix 11 – Perspective Taking Task Record Form

Date:
Initials:
Age:
Sex:
Condition:

Trial	Pass	Fail
Memory		
False belief		