Doctorate in Professional Educational, Child and Adolescent Psychology

Programme Director: Vivian Hill

School-based mindfulness programmes and executive functions: a systematic review and meta-analysis of complex intervention studies.

Christopher Quinton

UCL Institute of Education
Doctorate in Professional Educational Child and Adolescent Psychology

Declaration

I, Christopher Quinton, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Word count (excluding abstract, impact statement, appendices, and list of references): 39, 252

Acknowledgements

I would like to thank my thesis supervisor, Zoe Gallant. You warned me this thesis was too much, and you were right! Thanks also to Rebecca Gordon, who agreed to supervise me at the last hour and shared her invaluable expertise on executive function and meta-analysis.

To my trainee cohort, colleagues, course tutors, lecturers, and supervisors, all the psychologists I've met along the way, thank you. The friendships I have made and the work that binds us have kept me going. It's a privilege to be one among you all.

Thank you, Mama, of course, for giving me life and making it worthwhile. Thank you, Grandma and Grandpa, your limitless belief in my abilities is misplaced, but I love it all the same. Thank you, Mad Aunt Clare, who knows more about psychology than I ever will. I carry you all in my heart.

Thank you, Luna and Mr Jubs, faithful, furry desk companions and distractions. And Jason, thank you for giving me perspective: doctorates are meaningless without you.

Abstract

This systematic review and meta-analysis investigated the differential effects of school-based mindfulness programmes (SBMPs) on cognitive processes associated with executive function (EF) in children and young people. The literature search yielded 30 controlled intervention studies that met the inclusion criteria. Sixty-nine outcomes from 27 studies were included in the random-effects meta-analyses exploring the impact of SBMPs on measures of cognitive flexibility, inhibition, working memory, attentional control, and global EF. Moderator analyses explored participant age and intervention characteristics as potential predictors of SBMP effectiveness.

The overall pooled effect size of SBMPs on all EF-related outcomes was significant ($p<0.0001$) and small-to-moderate ($g=0.33$). Significant differential effects were found between working memory ($g=0.48$) and cognitive flexibility ($g=0.21$). Participant age was identified as a moderator of SBMP effectiveness that approached significance for cognitive flexibility and inhibition, with greater benefits seen in younger children. Overall methodological quality and relevance, measured through critical appraisal, were also found to influence effect sizes. Except for inhibition outcomes, total intervention dosage did not moderate SBMP effectiveness.

The narrative synthesis highlighted the heterogeneity in SBMP interventions, which differed considerably in their programme components, dosage, administrator role, and fidelity of implementation. Intervention characteristics were inconsistently reported, which limited the replicability of the overall evidence base. This was also true of the reporting of SBMP acceptability and adherence to the intervention protocol. This review identified further limitations in research designs and how EF is measured.

The findings suggest that SBMPs can improve EF in children and young people, with differential effects on specific cognitive processes and developmental stages. However, the robustness of this modest effect is limited by inconsistent methodological rigour and potential publication bias. To navigate the complexity of school-based interventions, future research must employ rigorous empirical methods to reflect the true effectiveness of SBMPs and provide actionable insights for education stakeholders.

Impact statement

This thesis furthers our knowledge of mindfulness practices and executive function, focusing on school-based mindfulness programmes (SBMPs) and their potential to improve cognitive processes in children and young people. The rigorous systematic review and meta-analysis of twenty-seven controlled intervention studies presented here provides a comprehensive overview of the effectiveness of SBMPs on specific cognitive processes associated with executive function, the moderating role of age, and the influence of intervention characteristics. The findings have broad implications for educational psychology practice, research methodology, theoretical cognitive psychology, and young people's school functioning and well-being.

For educational psychologists, this research builds the capacity of practitioners to design and implement effective, school-wide, and developmentally appropriate interventions to support students' cognitive development. Findings relating to the differential effectiveness of SBMPs on associated cognitive processes support educational psychologists in designing more targeted and effective mindfulness interventions. The developmental perspective on SBMPs, gained from exploring the moderating effect of age, further supports practitioners and schools in maximising the benefits students receive from SBMPs. A narrative synthesis of intervention characteristics, such as programme content and dosage, contextualises the more generalisable findings of the meta-analysis. This supports practitioners in adapting research findings to their local educational contexts. This comprehensive overview of the evidence base can guide educational psychologists seeking to integrate mindfulness practices into the educational and family systems in their practice, ultimately improving outcomes for young people. This research can also support policymakers seeking to integrate mindfulness practices into curricula or teacher training programmes in wider systems.

This research enriches educational psychology research by adopting methodological approaches that are underrepresented in the applied field. Applying a multi-level meta-analytic approach uses relatively sophisticated statistical techniques, such as Correlated and Hierarchical Effects (CHE) modelling, that support the synthesis of complex intervention research. A thorough critical appraisal of the methodological quality of these studies also highlights the need for rigorous study design and reporting to advance evidence-based practice. It is hoped that by setting a high standard for research in this area, future studies will be encouraged to adopt more robust and comprehensive approaches to investigate the impact of educational interventions.

This study is the first meta-analysis to explore the differential effectiveness of SBMPs in improving cognitive processes associated with executive function. By adopting a dynamic model of executive function and considering intervention evidence in relation to developmental trajectories, this research also advances our conceptual understanding of executive function. The models that serve as a theoretical framework for this review advance the delineation of executive function as a psychological construct. In the longer term, by building our conceptual understanding of executive function and how it can be enhanced through intervention, this research can support young people with cognitive and emotional needs and promote resilience and success in school and beyond. The dissemination of this work through journal publication, conference presentations, and engagement with educational stakeholders can further translate this research into real-world impact.

Table of Contents

Table of tables

Table of figures

List of abbreviations

ACT	Acceptance and Commitment Therapy
ADHD	Attention Deficit Hyperactivity Disorder
AIC	Akaike Information Criterion
ANOVA	Analysis of Variance
BIC	Bayesian Information Criterion
BRIEF	Behaviour Rating Inventory of Executive Function
BRS	Behaviour Rating Scale
CBQ	Children's Behaviour Questionnaire
CBT	Cognitive-Behavioural Therapy
CCC	Cognitive Complexity and Control
CHE	Correlated and Hierarchical Effects
CI	Confidence Interval
cRCT	cluster Randomized Controlled Trial
CTRS-R	Conners' Teacher Rating Scale-Revised
CYP	Children and Young People
DCCS	Dimensional Change Card Sort
DCS	Distributed Control System
DECP	Division of Educational and Child Psychology
DERS	Difficulties in Emotion Regulation Scale
ECP	Executive-function-related Cognitive Process
EF	Executive Function
EFN-back	Emotional Faces N-back Task
EMA	Ecological Momentary Assessment
EP	Educational Psychologist
EPS	Educational Psychology Service
FWER	Familywise Error Rate
HTKS	Head, Toes, Knees, and Shoulders
KS	Key Stage
L2B	Learning to Breathe

LRT	Likelihood Ratio Test
MBI	Mindfulness-Based Intervention
MBSR	Mindfulness-Based Stress Reduction
MEFS	Minnesota Executive Function Scale
MiSP	Mindfulness in Schools Project
PBT	Performance-Based Task
PI	Prediction Interval
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
RCT	Randomized Controlled Trial
REML	Restricted Maximum Likelihood
ROBINS-I	Risk of Bias in Non-randomized Studies of Interventions
RVE	Robust Variance Estimation
SBMP	School-Based Mindfulness Programme
SDQ	Strengths and Difficulties Questionnaire
SE	Standard Error
SEL	Social and Emotional Learning
SEN	Special Educational Needs
SMD	Standardized Mean Difference
TAU	Treatment As Usual
TRS	Teacher Rating Scale
WISC-IV	Wechsler Intelligence Scale for Children, Fourth Edition
WM	Working Memory
WoE	Weight of Evidence
WoS	Web of Science

1. Introduction

This section begins by outlining the key concepts relevant to this systematic review and meta-analysis on the effectiveness of school-based mindfulness programmes (SBMPs) for improving executive function (EF) and EF-related cognitive processes (ECPs) in children and young people (CYP). This includes a working definition of EF and its relevance to school functioning, academic achievement, and life outcomes. It is followed by exploring mindfulness interventions, explicitly focusing on SBMPs and how they might support the development of EF skills. The role of Educational Psychologists (EPs) as a bridge between evidence-based research and intervention delivery in real-world educational settings is discussed, focusing on how EPs can use their consultancy skills to optimise SBMP design and implementation for specific contexts. The section concludes by presenting the intended focus and scope of this review.

1.1 Executive function

Executive function (EF) is central to educational and psychological discourse. Broadly speaking, EF is an umbrella term for a range of interrelated, higher-order processes necessary to perform goal-directed behaviours (Nigg, 2016). This goal orientation is key to understanding executive function, distinguishing it from processes associated with prepotent responses to the environment. In the dynamic environments in which we operate, it is frequently necessary to adapt to novel situations. These adaptive responses often conflict with more automatic responses and thus require effortful control to initiate and sustain. Executive function can be conceptualised as the skills necessary to exert control in the service of a goal. Definitions diverge depending on whether this control is supervisory and top-down or an emergent property of a complex network of distributed processes. This distributed control system (DCS; Zink et al., 2021) model is privileged in this research, justifying the exploration of distinct yet interrelated cognitive processes typically associated with EF. These include working memory updating, cognitive flexibility, and inhibition (Miyake et al., 2000) and are hereafter referred to collectively as executive cognitive processes (ECPs).

Despite challenges in reaching a consensus definition, there is general agreement that EF is implicated in a range of outcomes that are highly relevant to children and young people (Doebel & Muller, 2023). Executive function (EF) is considered foundational to many skills necessary for school readiness, such as reading (Follmer, 2018) and maths proficiency (Cragg & Gilmore, 2014). Furthermore, EF skills are associated with increased resilience (Masten et al., 2012), emotional regulation (Zelazo & Carlson, 2020), and social functioning (Clark et al., 2002). Evidence suggests that these childhood benefits are maintained across the

lifespan, with quality-of-life outcomes such as socioeconomic status also associated with EF (Rosen et al., 2020). Given the far-reaching impact of EF on CYP functioning and well-being, finding approaches that support the development of EF skills within school systems has become a focus for educators, psychologists, and policymakers (Zelazo et al., 2016).

This interest in interventions targeting EF has increased significantly over the past decade, driven by research that suggests it is more predictive of a young person's academic attainment than comparable fixed constructs such as intelligence quotients (Spiegel et al., 2021). EF might also be more malleable to environmental pressures (Zelazo & Carlson, 2012). This malleability makes EF a prime target for interventions to improve CYP's educational and life outcomes. However, despite this potential, results from approaches designed to directly target EF, such as computerised cognitive training (CCT), indicate that any improvements tend to be domain-specific (Cao et al., 2020) and do not generalise to meaningful contexts. Programmes that target EF implicitly, such as through martial arts or mindfulness, may show greater promise (Takacs & Kassai, 2019). In educational contexts, mindfulness interventions are currently enjoying significant research attention.

1.2 Mindfulness in schools

Mindfulness is defined by Bishop et al. (2004) as comprising two core components: 1. the self-regulation of thoughts, feelings, and sensations by regulating attention to the present moment; and 2. having an open, non-judgemental orientation to the experiences that arise during this process. Mindfulness-based interventions (MBIs) are therapeutic practices and programmes designed to develop mindfulness. These share a common ancestor in Jon Kabat-Zinn's (1997) Mindfulness-based Stress Reduction Programme (MBSR) and typically aim to improve psychological wellbeing through a combination of psychoeducation and meditation practices. Many of the practical elements of MBIs have existed for millennia (Analayo, 2019), but the term "intervention" hints at a decidedly more clinical flavour and "Western-centric" orientation.

The 21st century has seen a growing interest in how mindfulness practices can benefit children and young people (CYP) and how these practices can be incorporated into education systems, as seen in the 32-fold increase in mentions of "mindfulness in education" across all published English-language books between 2005 and 2019 (Roeser et al., 2023). This proliferation is not limited to academia, with global organisations such as Mindful Schools reporting a network of over 50,000 educators applying mindfulness practices in education settings. Locally, several organisations promoting mindfulness in British school systems, such as

the Mindfulness in Schools Project (MiSP), have emerged. At the level of individual schools, there is evidence to suggest that teachers and young people engage in mindfulness practices informally (Luiselli et al., 2017), although the prevalence and nature of this practice are not as well documented as school-based mindfulness programmes (SBMPs; Phan et al., 2022).

School-based mindfulness programmes (SBMPs) are interventions designed to develop the mindfulness skills of children and young people (CYP) within their educational settings. The exact practices that are used by SBMPs are heterogeneous, although they typically draw on the practices used in adult mindfulness programmes, such as the foundational mindfulness-based stress reduction (MBSR; Kabat-Zinn, 1990), adapted these to be more developmentally appropriate for their target CYP age range. These practices almost always include breathing exercises and body scans but might also include additional elements such as yoga (Thomas & Centeio, 2020), mindful eating (Pierson et al., 2019), and martial arts (Meixner et al., 2019). They often also include a psychoeducational element teaching the connections between thoughts, feelings, and the brain.

1.3 Rationale for the research focus

Although mindfulness in schools is becoming the norm rather than the exception (Roeser et al., 2023), the evidence base for the effectiveness of SBMPs has yet to be firmly established. A recent cluster randomised controlled trial (cRCT) involving over 8000 participants (Kuyken et al., 2022) found no effect on mental health outcomes, tempering some of the enthusiasm with which SBMPs have been adopted. The impact of SBMPs on cognitive functions is less extensive, but reviews of the evidence base suggest that these results may be more robust than those reported for mental health outcomes (Dunning et al., 2022; Roeser et al., 2023). Despite this burgeoning support, an understanding of the mechanisms by which SBMPs improve EF-related outcomes and the conditions that moderate this effect are yet to be addressed.

In previous meta-analyses of SBMPs focusing on cognitive outcomes, EF has been treated as a global supervisory system, with little research exploring how mindfulness programmes might differentially affect the subprocesses through which EF emerges, such as response inhibition, working memory updating, and cognitive flexibility (Miyake et al., 2002). Operationalising EF as a unitary construct in SBMP research may obscure the mechanisms through which SBMPs influence EF development. An improved understanding of how these interventions work may emerge from investigating the potential differential effects of SBMPs on specific ECPs. This can inform the design of targeted, process-specific interventions.

Exploring the moderating effects of participant age and intervention characteristics is also underrepresented in the literature. It is hypothesised that EF emerges over developmental time, with ECPs exhibiting distinct developmental trajectories (Zelazo & Carlson, 2020). A more granular investigation of how SBMPs interact with age may reveal different effects for specific ECPs at specific developmental stages. Additionally, SBMPs are complex interventions that depend on the interactions of multiple factors across a range of domains, including the context of the setting, population characteristics, intervention and implementation, and conceptual/theoretical underpinnings. Exploring intervention characteristics (e.g., total intervention dosage and the administrator's role) may help explain the differences between effect sizes reported by SBMP studies.

Additionally, a lack of methodological rigour is hypothesised to have led to overestimating the benefits of mindfulness interventions such as SBMPs (Van Dam et al., 2018). Reviews of SBMPs have found concerns that include a lack of appropriate control conditions, small samples, and the use of dubious outcome measures. Taken together, these may lead to inflated effect sizes and false conclusions regarding the effectiveness of SBMPs. Addressing these methodological concerns is necessary to ensure that the evidence base is sufficiently robust to warrant the adoption of school-wide approaches such as SBMPs.

1.4 Research aims and questions

The present systematic review and meta-analysis aim to address the issues raised here, by:

1. Mapping outcome measures to specific cognitive processes associated with EF to enable an examination of whether they are differentially affected by SBMPs. Included studies use either randomised control trials (RCTs) or controlled quasi-experimental designs to ensure a baseline level of methodological rigour.
2. Investigating the potential moderating role of participant age and intervention characteristics on SBMP effectiveness, using meta-regression and correlated and hierarchical effects (CHE) modelling.
3. Evaluating the methodological quality and relevance to practice of the SBMP studies exploring EF-related outcomes through critical appraisal and narrative synthesis of the evidence base.

These aims are organised around the following research questions:

1. Do school-based mindfulness programmes have a differential effect on cognitive processes associated with executive function?

2. Does age moderate the extent to which school-based mindfulness programmes improve measures of executive function and associated cognitive processes?
3. Which intervention characteristics influence the effectiveness of school-based mindfulness programmes to improve executive function and associated cognitive processes?
4. How can future research strengthen the evidence base for the effectiveness of school-based mindfulness programmes in improving executive function and associated cognitive processes?

1.5 Relevance to educational psychologists

The BPS Division of Educational and Child Psychology (DECP) lists strategic work with "schools and local authorities to improve all children's emotional wellbeing and experiences of learning" as a key function of educational psychology practice and EPs are often involved in the design, implementation, and monitoring of school-based interventions (Farrell et al., 2006). These interventions are complex, and success is typically dependent on a range of factors and the interactions between them (Petticrew et al., 2013). This study's results will help EPs navigate this complexity, adding to the profession's collective knowledge of the conditions leading to SBMPs effectiveness in improving EF-related outcomes.

This research enables EPs to give a more nuanced perspective when advising on how to adjust the SBMPs to maximise the benefits according to the developmental stage of the cohort. More granular knowledge of how mindfulness practices interact with specific ECPs may also improve the effectiveness of school-based interventions. Practice guidance from the British Psychological Society (BPS, 2017) indicates that applied psychologists working with children and young people will draw on "specialist training and experience" to understand their developmental needs. This developmental perspective is positioned as a core competency that aligns with exploring the potential moderating effect of age on SBMP effectiveness and comparing these results to the research literature on the developmental trajectories of ECPs. The research can develop knowledge of executive functioning as an emergent property of developmental trajectories within the profession.

The Health and Care Professions Council (HCPC, 2023) highlights the need for EPs to maintain current knowledge of theoretical frameworks, and this research's conceptualisation of executive function as an emergent state derived from a distributed control network aligns with this aim. Supporting executive function needs is a cornerstone of applied educational psychology (Meltzer, 2018). Developing the

profession's shared knowledge of EF is critical, especially since educational psychologists' theoretical understanding of EF is inconsistent (Price, 2023).

By triangulating the statistical results with a synthesis of contextual and study characteristics, the research aims to promote an understanding of school-based interventions as complex. This aligns with a key foundation of Educational Psychology Services (EPSs) outlined in BPS (2023) guidelines, which is to share an "interactionist understanding of diversity in development and learning." While this may typically be applied to special educational needs (SEN), it is this research's position that it should also inform systemic school development. The focus on universally targeted interventions supports calls within the EP community to work towards a "universal psychology" (MacKay, 2015) through systemic practice. As universal approaches require considerable resources, it is prudent to be as informed as possible on the potential benefits and limitations of SBMPs in schools, especially given the current "hype" surrounding mindfulness (Van Dam et al., 2018).

2. Literature Review

The literature review begins by evaluating competing psychological models to justify conceptualising EF as emerging from interrelated yet distinct processes and orienting the research towards cognitive domains. The differential developmental trajectories of inhibition, cognitive flexibility, and working memory are evaluated, with a developmental model of EF used as a theoretical justification for exploring the potential moderating effect of age on the effectiveness of SBMPs. A mechanism of action to explain how mindfulness improves executive function is provided, suggesting that it may affect individual ECPs differently. The chapter concludes with a review of the research on school-based mindfulness programmes (SBMPs). Differences found in SBMP effectiveness are explored, including the high heterogeneity between study contexts, intervention/implementation characteristics, and methodological quality.

2.1 Psychological models of executive function

Considering various competing models of EF is essential given this research's focus on whether SBMPs have a differential effect on ECPs. This focus assumes that EF can be operationalised as a cognitive construct and that it can be deconstructed into distinct subprocesses. However, both assumptions are contested in the research literature.

2.1.1 Unity and diversity in cognitive models

Variance in the conceptualisation of EF in the literature centres around whether EF is a unitary and supervisory system or an emergent property of interactions between several integrated yet distinguishable cognitive processes (Doebel & Muller, 2023). These differing conceptualisations have essential implications for SBMPs, which, depending on the underlying theoretical position, might take a holistic approach to intervention design or otherwise target specific cognitive processes through mindfulness practices.

The multicomponent conceptualisation of EF is heavily influenced by Miyake et al.'s (2000) seminal paper outlining the extraction of three latent variables through factor analysis of performance on a range of tasks commonly associated with EF. The three-factor model of EF that was developed from this analysis continues to be the most widely cited in the literature (Baggetta & Alexander, 2016). These factors are 1. inhibition, the cognitive process involved in stopping the execution of an impulsive response (Diamond, 2020); 2. updating, which is heavily associated with working memory (Best & Miller, 2010), is the process that allows for the management and monitoring of the limited working memory store (Ecker et al., 2010); and 3.

shifting, or cognitive flexibility, which enables shifting attention between tasks and adapting behaviour in response to novel information (Buttelmann & Karbach, 2017). This factor approach holds that although these processes are separable, they are highly interrelated and are typically activated concurrently when completing complex goal-directed behaviours. As such, EF is described as having both "unity and diversity" (Friedman & Miyake, 2017).

Although this diversity raises interesting research avenues, empirical research often fails to replicate distinct EF subcomponent factors. This may be particularly relevant to SBMPs as several confirmatory factor analyses comparing preschool children to adults suggest that EF may be a unitary construct in early years and only exhibit diversity at later stages of development (e.g. Lerner & Lonigan, 2014). However, this finding does not preclude the possibility that ECPs might have distinct developmental trajectories and longitudinal studies have shown evidence for differentiation of ECPs across childhood and adolescence (Brydges et al., 2014). This indicates that exploring separable ECPs may be relevant to CYP research. However, which cognitive processes are most implicated in EF and how they interact still need to be fully understood.

Subsequent factor analyses have questioned Miyake et al.'s (2000) model, with some indicating that the three factors can be reduced to two. For example, van der Ven et al. (2012) forward a two-factor model that keeps working memory as a separable EF process but subsumes cognitive flexibility and inhibition into a single factor. In other factor analyses, updating and inhibition have been identified as the two latent variables involved in EF (St Clair-Thompson & Gathercole, 2006). These differential results question the robustness of the three-factor model, which currently maintains the most robust empirical support and has been integrated into other models of cognition. For example, Himi et al. (2021) propose a five-factor model that integrates relational integration and divided attention with the three-factor model of EF to develop a broader model of general cognitive ability.

Other models adopt a hierarchical structure, suggesting that one cognitive process has a dominant influence. For example, through work with populations that exhibit differences in executive function (e.g. ADHD in Nigg et al., 2000), inhibition has been implicated as the primary process involved in executive function, acting as the informational gatekeeper for working memory. Conversely, working memory has also been hypothesised to be the dominant process. A review of research employing methods including functional magnetic resonance imaging (fMRI) and lesion studies (Munakata et al., 2011) argues that the prefrontal cortex (PFC) is most strongly associated with maintaining goal-related representations. In this model, the PFC projects to subcortical and archicortical areas, which activate inhibitory neurons that reduce

activity in those regions. From this perspective, working memory is the dominant process associated with executive function, which exerts top-down control over downstream processes such as inhibition. This contrasts with the view presented by Nigg et al. (2000), who suggest that observed difficulties with impulse control in conditions such as ADHD may be due to differences in how goal-related information is maintained rather than a direct inhibitory control deficit. These are but a few of the myriad competing models of EF which highlight how a more delineated exploration of how SBMPs might interact differently with separate ECPs could provide further insights into the structure of EF.

2.1.2 Integrated models

With a justification given for investigating separable ECPs, the second inherent assumption in the orientation of this research must now be considered, namely that there is value in exploring cognitive aspects of EF in relative isolation. This has been the case historically, where EF has typically been studied as a cognitive capacity (Welsh & Peterson, 2014). Miyake et al.'s (2000) seminal work exemplifies this approach, with their three latent variables derived from performance-based tasks (PBTs) conducted in controlled laboratory conditions. These tasks lack the emotional valence, task salience, and contextual relevance of EF as it emerges in daily life. Recent directions in EF research have moved towards integrating other non-cognitive systems into models of EF to better account for the differences seen between measures of EF through PBTs and observations of executive functioning in real-world scenarios.

Zelazo and Muller (2002) are often credited for introducing non-cognitive constructs to models of EF through their work distinguishing between "hot" and "cool" EF. Hot EF involves processes elicited under emotionally significant conditions that are personally meaningful to the individual. These include constructs such as delayed gratification, where, for example, resisting the urge to eat a marshmallow requires both inhibitory processes and the management of emotions. They contrast this to "cool" EF, the non-affective, cognitive EF associated with most PBTs. Some neuroimaging studies have supported this distinction, indicating that different brain networks are activated depending on whether EF is activated under affectively salient or neutral conditions (Zelazo & Carlson, 2012).

This hot/cool distinction presents a quandary for the focus of this research because EF is definitionally goal-oriented and top-down, so suggestions that EF may operate differently when completing abstract performance-based tasks (PBTs) compared with how it might emerge in home or school environments for CYP is potentially problematic. However, positive correlations between performance on hot and cool EF tasks (Willoughby et al., 2011) suggest that this distinction may not warrant discounting out of hand the

relevance of cool EF measures to real-world contexts. This is further supported by recent functional neuroimaging studies (Moriguchi & Phillips, 2023) exploring the hot/cool distinction in children, which have found that cortical pathways associated with emotional processing are activated when completing tasks traditionally considered purely cognitive (e.g. Stroop Task). Overall, while tasks that are primarily designed to measure cognitive constructs such as working memory and inhibition are likely to lack the same degree of emotional and motivational salience as EF that is activated in daily life, they are still deemed to be an acceptable way to operationalise a complex construct for an intervention study. It is also worth noting that this research includes data from behaviour-rating scales (BRSs), which are argued to be more reflective of "real-world" executive function (Veloso, 2022). Furthermore, this research's focus on SBMPs means that these studies are more meaningful contexts than might be seen in more clinical randomised controlled trials (RCTs) conducted in laboratory settings.

Doebel (2020) further expands on EF's emotional and motivational factors to include additional non-cognitive factors such as cultural beliefs and social norms. She suggests that EF may be better conceptualised as an emergent capacity from "the development of skills using control in the service of specific goals", which is differentially expressed based on contextual demands. This capacity depends on forms of knowledge, cultural beliefs, social mores, idiosyncratic preferences, and domain-general cognitive skills. Doebel's (2020) central thesis is that EF does not operate in a vacuum but is always dependent on the local context. From this perspective, a young person stopping themselves from taking a marshmallow is not simply activating inhibitory and emotional processes. Instead, they draw on stored knowledge, including memories of being admonished for taking food, alternative strategies such as asking permission, internalised values not to steal, etc. The model addresses the frequently replicated finding that cognitive measures of EF do not consistently correlate with behavioural measures, which are more context-dependent (Doebel, 2020). This systemic re-conceptualisation of EF is more inclusive of "hot" executive functions than the three-factor model, and further integrates other non-cognitive factors which have been found to influence EF, such as whether development has taken place within a more individualist or collectivist culture (Doebel & Muller, 2023).

From a complex systems perspective (Hilpert & Marchand, 2018), EF can be understood as an emergent state assembled from interactions between multiple cognitive and non-cognitive components. This state cannot be fully understood by reducing it to constituent components (e.g. inhibition) because EF emerges through dynamic interactions between processes, not the processes themselves. EF is not a purely cognitive process, nor the sum of a broader range of processes (e.g. emotional states, culture, etc.), but

rather a higher-order state emerging from myriad interactions. As such, Doebel (2020) highlights this issue by suggesting that "future work should avoid reifying executive function as components." However, the position taken here is that an appreciation of the overall holistic nature of EF is not incompatible with the investigation of specific ECPs. The position of this research is that the investigation of specific cognitive systems associated with EF does not deny that EF emerges from multiple non-cognitive systems, nor does it deny that a goal-oriented system is necessarily context-specific. Rather, this more granular approach enriches our understanding of EF's cognitive aspects.

2.2 Developmental models of executive function

Doebel's (2020) dynamic conceptualisation of EF as the "development of skills" indicates the importance of considering EF as evolving over lifespan. One of the benefits of exploring EF in relation to SBMPs is that comparison between studies targeting different age groups can develop our understanding of how EF evolves over time. This developmental approach may advance our knowledge of EF, with fresh insights informing the current conceptual debates in the research discourse. This perspective involves a shift from conceptualising EF as a static model to a dynamic system in a constant state of emergence. It evolves in response to the environment, organising itself into increasingly hierarchical structures that allow us to process more abstract and complex representations (Taylor et al., 2015).

2.2.1 Dynamic modelling

In one respect, the exact pathway to achieving a goal is never the same as, to draw on an oft-quoted aphorism, "You never step in the same river twice" (Heraclitus, c. 500 BC). Executive function is necessarily employed in novel situations because goal-directed behaviours do not exist in isolation but in the constant flux of various bioecological systems. While holding this to be accurate, it is also the case that some goals are more similar than others. Keeping track of a list of digits and then repeating them backwards (e.g. WISC-V Digit Span: Backwards) has more in common with keeping track of a conversation than resisting the urge to bite someone. These similarities are hypothesised to drive the dynamic development of specialised processes associated with a more generalised goal-oriented system (Ibbotson, 2023). From this dynamic perspective, EF is not modular per se but emerges from a plurality of potential processes that can serve towards attaining a goal. Comparable cognitive processes emerge due to similarities in universally shared environmental pressures constraining goal variance. These cognitive processes have a nebulous dual nature, partially separable from one another, but also highly interconnected (Friedman & Miyake, 2016).

An example of such a developmental account is cognitive complexity and control (CCC) theory, where EF emerges through the development of increasingly complex "if-then" conditional statements over time (Zelazo et al., 2003). For example, a young child might be able to recognise an emotion and engage in an appropriate response strategy by developing the conditional "If I'm feeling upset, then I should speak to a teacher." Developing a simple mindfulness technique might add a layer of complexity to this conditional: "If I'm feeling upset, I will take some deep breaths before deciding whether I should speak to a teacher." Further complexity provides greater control and indicates greater EF capacity, e.g. "If I'm upset, I will take some deep breaths and pay non-judgemental attention to my present experience to help me understand what I am feeling. I will then decide if I can cope without support. If I'm still overwhelmed, I will seek help."

These conditional statements are the product of predictive environments (Ibbotson et al., 2023), with our ability to predict which behaviours might elicit desired outcomes becoming more fine-tuned as we experiment with a broad range of strategies and then focus more narrowly on those that lead to goal attainment. This trial-and-error approach, also known as simulated annealing in the literature, shares assumptions with dynamic systems models of motor development (e.g. Kamm et al., 1990), both of which are grounded in the neuroconstructivism theoretical framework (Mareschal et al., 2007; Karmiloff-Smith et al., 2018).

Building on CCC theory, the Iterative Reprocessing (IR) model (Cunningham & Zelazo, 2007; Zelazo, 2015) suggests that the development of cognitive complexity is dependent on the reflective reprocessing of information in relation to specific problem-solving contexts. These reflective processes allow for the continued evaluation and re-evaluation of "if-then" conditional statements, leading to more refined cognitive rule structures. Crucially, the IR model posits that goal-directed behaviour is not solely the product of top-down cognitive processes but interactions between both top-down (e.g. goals, rules) and bottom-up (e.g. physiological states, sensory stimuli) components in context (Perone et al., 2021).

In the developmental model described here, the exact cognitive processes (e.g. cognitive flexibility) are not explicitly hardwired but driven by exploratory simulated annealing. Early in development, inhibition is a beneficial strategy to attain generalisable goals across populations (e.g. receiving comfort and food from a caregiver). As our goals become more complex, they become more entwined with context-specific features of our environments. These environments are uniquely intra-personal (e.g., genetics, cortical organisation, etc.) and extra-personal (e.g., caregiver attributes, cultural norms, resource availability). As such, this model predicts that older children should use a broader range of EF strategies than younger children. Results from

a recent meta-analysis (Messer et al., 2022) support this prediction, with age as an apparent moderator of EF differentiation.

The developmental model also offers a bottom-up approach to the development of EF, which addresses the "homunculus" problem, i.e. if EF is a supervisory system, what is it supervised by, and what would supervise that super-supervisory system *ad nauseam*. Replacing this hierarchical system with a distributed control system (DCS; Zink et al., 2021), which suggests that EF emerges from a complex network of distributed brain networks, is supported by findings from neuropsychological research showing that EF functioning can be preserved following lesion to the prefrontal cortex (PFC), which is commonly associated with EF (Alvarez & Emory, 2006). This model also provides a useful framework for understanding how ECPs exhibit both "unity and diversity". They are diverse processes developed in response to increasingly diverse goals. Still, they must also be employed in conjunction as the complexity of goals reaches a stage where they can only be achieved through the emergent efficacy of multiple interacting processes.

2.2.2 Differential developmental trajectories

This research's aim to investigate the moderating role of age on SBMP effectiveness in improving EF-related cognitive processes (ECPs) necessitates a review of how these processes are thought to develop through childhood and adolescence. Critically, evidence suggests that EF does not follow a unitary developmental trajectory but that associated cognitive processes develop at different rates and stabilise at various developmental stages. Any potential moderating effects found in the meta-analysis can be compared to the research literature on the differential developmental trajectories of ECPs, such as attentional control, inhibition, working memory, and cognitive flexibility. This allows for a richer understanding of how EF emerges across CYP development and can support educational psychologists in guiding the design and implementation of SBMPs to optimise their effectiveness for specific age groups.

2.2.2.1 Attentional control

Differentiation of executive function processes may emerge from the development of attentional control in early infancy. Colombo (2002) records that infants struggle to refocus attention away from a novel stimulus, such as an unusual shape or colour. As visual acuity and discrimination develop, this "sticky fixation" (Hunnius, 2004) is gradually replaced by enhanced attentional control, whereby infants are increasingly oriented towards stimuli they intrinsically prefer rather than being driven by extrinsic salience variables (e.g. brightness of colour; prominence of shape). Core attentional control in relation to visual

processing is argued to be relatively stable from around age seven (Oh-Uchi et al., 2010; Turoman et al., 2021), at least for neutral stimuli. However, for affective visual stimuli (e.g. emotional faces), a more graduated developmental trajectory is proposed that continues across middle childhood (Bigelow et al., 2021) and adolescence (Cohen et al., 2014), suggesting a possible "hot/cool" distinction (Zelazo & Carlson, 2012).

2.2.2.2 Inhibition

The research literature indicates that inhibition is a higher-order construct that manifests through interactions between skills, including sustaining attention, disengaging from stimuli, and blocking automatic responses. Recent research suggests that inhibition emerges as early as six months and develops rapidly in the first year of life (Holmboe et al., 2018). A further stage of rapid development is posited to occur between 3 and 5.25 years (Wiebe et al., 2012). Inhibitory control skills are thought to stabilise before adolescence, although empirical support for this is inconclusive (Best & Miller, 2023). If it is supposed that inhibitory skills do continue to develop into adolescence, this is hypothesised to emerge through increased connectivity with other cognitive functions, with supporting evidence provided via results from fMRI scans taken over years, during which participants completed a Go/No-Go task (Cope et al., 2020). A more diffuse pattern of increased cortical activity was observed as the participants aged, suggesting increased integration of multiple functional pathways.

2.2.2.4 Cognitive flexibility

Cognitive flexibility includes deactivating the cognitive processes involved in one task and initiating the cognitive processes required for a new task. It is proposed to be definitionally distinct from response inhibition because the shifting happens in response to environmental changes (Dajani & Uddin, 2015). It can be conceived as a more dynamic process that ensures sufficient fluidity to adapt behaviour when faced with the unexpected. The ability to switch between different cognitive processes to achieve a goal is also hypothesised to emerge later than inhibition, follow a slower developmental trajectory, and continue to develop into adolescence (Buttelmann & Karbach, 2017). Children are typically competent at three years in sorting objects into categories such as colour or shape. However, they need help to begin sorting by colour when they had previously been sorting by shape (Kirkham et al., 2003). This difficulty in shifting tasks is also seen in perspective-taking, where young children struggle to switch between different interpretations of an ambiguous figure drawing (Gopnik & Rosati, 2001). Although improvements in this

capacity are seen by around 4-5 years, it is not typically until around 7-9 that children evidence a consistent ability to adapt their responses from one trial to the next (Gupta & Karr, 2009).

2.2.2.3 Working memory

The development of working memory, or the ability to hold and manipulate visual and auditory information in the service of cognitive tasks, is suggested to have a slower trajectory (Crone et al., 2006). As with all the cognitive processes explored in this review, it is highly interconnected to other capacities, with its development hypothesised to be dependent on improvements in many other functions, including a faster processing speed and a broader crystallised knowledge base (Camos & Barrouillet, 2018). Improvements in working memory are apparent from around three years and continue to develop significantly until adolescence (Ahmed et al., 2022) and into young adulthood, peaking around 30 years (Ferguson et al., 2021). Compared to inhibition and cognitive flexibility, working memory correlates highly with adolescent IQ (Friedman et al., 2006).

These differences in developmental trajectories hint at the complex nature of the interconnected, yet possibly partially distinct, relationship between ECPs and suggest that any potential moderating effect of age on SBMP effectiveness may vary depending on which specific cognitive process is targeted or measured. The results of this research may inform the recommendations of educational psychologists consulting schools that are implementing programmes and interventions targeting EF, with a more nuanced understanding rooted in developmental trajectories serving to maximise effectiveness. Educational psychologists can also recommend appropriate outcome measures to best capture progress. For example, inhibition-related measures may be more suitable for SBMPs in Early Years as inhibition is hypothesised to experience rapid development during this developmental stage (Holmboe et al., 2018). On the other hand, adolescents may benefit from mindfulness practices targeting higher-order ECPs with longer developmental trajectories, such as working memory (Ahmed et al., 2022). This developmental perspective may inform the design of more precisely targeted EF interventions for young people with special educational needs (SEN). This may be particularly promising given that EF may be more malleable than IQ-based measures of cognitive functioning (Diamond, 2013) and more predictive of life outcomes and well-being (Moffitt et al., 2011).

2.3 Mindfulness and executive function

Enhancing attentional and executive control through mindfulness programmes is hypothesised to confer multiple benefits for children and young people, such as strengthening academic performance by supporting learning and knowledge retention (Blair & Raver, 2015). However, although cognitive processes are implicated in many of the reported benefits of mindfulness, research on the mechanisms by which mindfulness might result in improved cognitive outcomes for CYP needs better representation in the research base (Mak et al., 2018).

2.3.1 Mindfulness

Mindfulness has a documented history that can trace its roots as far back as 1st century BC through scriptures collectively known as the Pāli Canon. It is from these writings, which are considered foundational to the Theravada school of Buddhism (Anālayo, 2019a), that the concept of sati has entered the English lexicon as mindfulness, also translated as "attention, awareness, retention, and discernment" (Davidson & Kazniak, 2015). Whilst these translated terms track with commonly cited definitions of mindfulness in Western research literature, they are argued by some (e.g. Purser & Milillo, 2015) to be oversimplifications that obscure the true nature of sati, which in the Buddhist tradition might be more accurately conceived as non-reactive awareness that enhances recollections of the past and strengthen memory (Anālayo, 2019b). Whilst Buddhist definitions of sati are diverse, there is consensus amongst practitioners that mindfulness is a transformative spiritual process that is driven by the soteriological goal to achieve enlightenment (nirvana) and thus be liberated from the cycle of suffering (samsara). This occurs through an awakening (bodhi) in which a state of perfect understanding reveals the illusion of the self.

In the Western psychological literature, mindfulness is typically considered a higher-order process that emerges from interactions between cognitive constructs such as attention, awareness, and acceptance. This focus on cognitive processes has been considered overly reductionist, as sati has a holistic complexity that includes additional affective, behavioural, social, and ethical dimensions (Grossman & Van Dam, 2011). Whilst undoubtedly more faithful to the Buddhist roots of mindfulness, these additional dimensions introduce degrees of complexity that are ill-suited to the psychological approaches and tools currently available to experimental researchers. Even with a narrower focus on cognitive processes, there is highly varied use of the term mindfulness in the research literature, where it might be conceptualised as a temporary state, an enduring characteristic, or an intervention (Davidson, 2015). Navigating the full

complexity of mindfulness as a construct is beyond the scope of this review, which is focused on experimental or quasi-experimental research. As such, an appropriate operational heuristic is required.

Several operational definitions have emerged to help navigate the conceptual heterogeneity found in mindfulness research. Kabat-Zinn (2003, p. 145), who is typically credited with pioneering the use of mindfulness practices in clinical psychology contexts, defines mindfulness as "the awareness that emerges through paying attention, on purpose, in the present moment, and non-judgmentally to the unfolding of experience moment by moment." This definition is further delineated in the two-component model of mindfulness (Bishop et al., 2004), which has emerged as one of the most frequently cited in the literature. This model is a helpful heuristic as it avoids wading too heavily into the conceptual complexities of mindfulness as a construct whilst presenting a definition in terms that are mutually intelligible with those found in the EF literature.

The first component of this model is defined as a process of sustained self-regulation to maintain attention on the immediate environment, allowing for enhanced awareness of experiential processes (e.g. bodily sensations, cognitive processes, emotions) in the present moment. The second component builds on this attention to the present moment by adding a particular orientation to the experience. This orientation is characterised by an openness and acceptance of the thoughts, feelings, and sensations that arise in our consciousness. As such, the primary goal of mindfulness is to be open to our experiences in the moment by accepting them as they are without trying to change them. The two-factor model also constructs mindfulness as a metacognitive skill that develops through practice, with early mindfulness practice associated with focusing attention, and the development of an equanimous orientation to the present moment, or "open monitoring style", associated with greater proficiency (Isbel & Summers, 2017). This second component delineates mindfulness from related practices such as relaxation techniques. For example, mindfulness is not only engaging in a breathing exercise to maintain focus on the present moment but also non-judgemental monitoring of thoughts and feelings as they arise during the process.

2.3.2 Cognitive models of mindfulness and executive function

There are two main pathways through which mindfulness is posited to improve executive functioning, which map directly onto Bishop et al.'s (2004) definition of mindfulness. Firstly, maintaining attention to the present moment is hypothesised to improve attentional control over time (Shapiro et al., 2006). Repeating exercises such as breathing exercises and body scans involve the gentle redirecting of attention back to a particular source, which over time are hypothesised to strengthen the underlying neural networks

associated with attentional control (Holzel et al., 2011). CYP may be particularly sensitive to this strengthening process as they are still developing their attentional abilities (Felver et al., 2014).

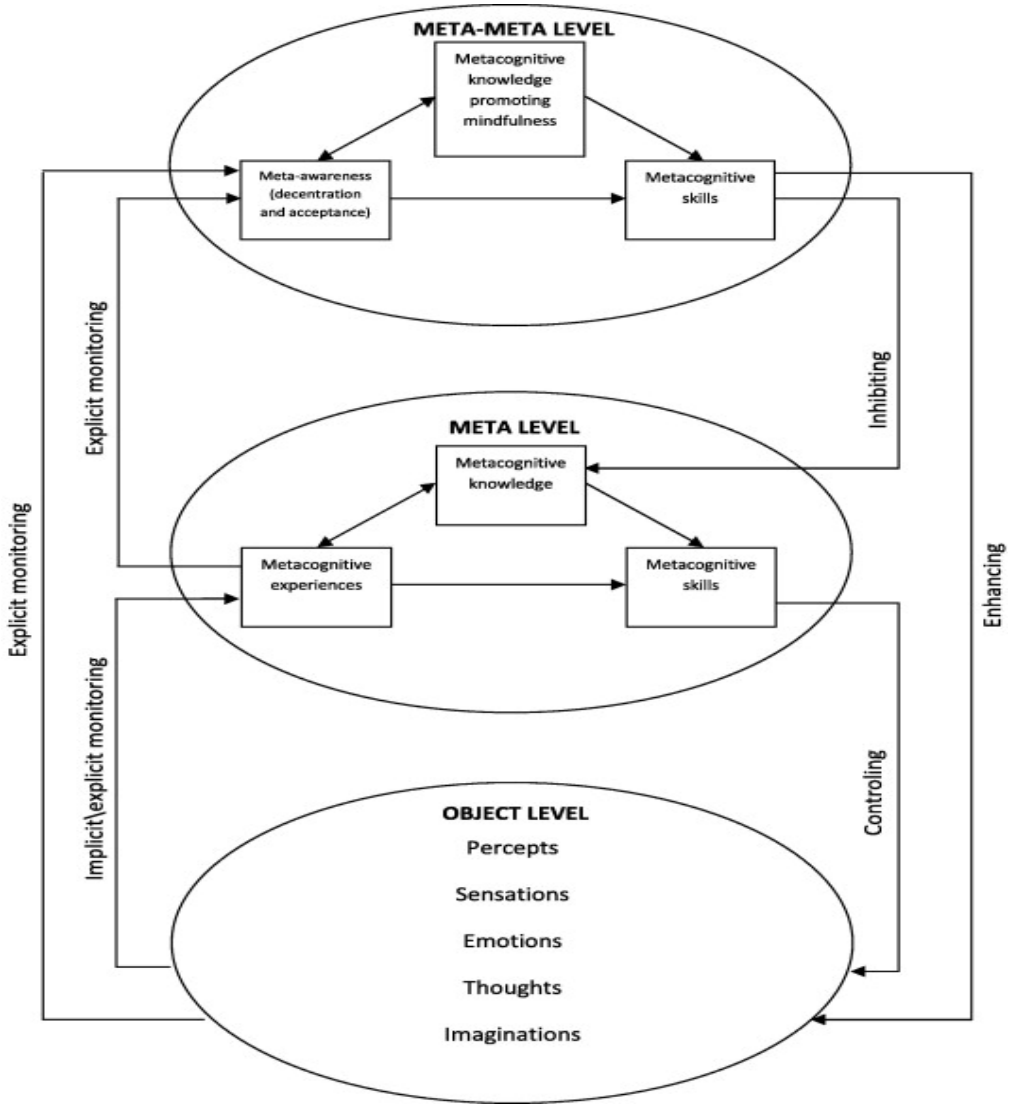
Pioneering modelling of attention using neuroimaging techniques has given rise to factor models of attention, such as the three-factor model (Posner & Petersen, 1990; Petersen & Posner, 2012) that includes an alerting network, which is linked to arousal and maintaining an alert state; an orienting network, which is concerned with the prioritisation of sensory input; and an executive control network, involved in conflict resolution and choosing between competing resources. By moving away from a monolithic conceptualisation of attention, a more nuanced understanding of how it is affected by mindfulness has emerged. For example, meta-analyses of RCTS on adult populations suggest that mindfulness practices can significantly improve executive control (Casedas et al., 2020) but do not influence the orienting network (Sumantry & Stewart, 2021). However, there is good evidence to suggest that mindfulness can, at least in adults, improve attentional and executive control, as well as executive functioning (Lodha & Gupta, 2022).

Secondly, the non-judgemental orientation to experience (Bishop et al., 2004) is hypothesised to support "dissociations between meta and object levels of cognition" (Holas & Jankowski, 2013). In other words, the top-down processes which enable disengagement from automatic thoughts and impulses are strengthened through mindfulness. Such activities can support the development of a meta-cognitive stance (Holas & Jankowski, 2014) that allows for monitoring internal states. These skills are more complex, emergent expressions of EF that may not be accessible to younger CYP, and SBMPs that are designed to cultivate a non-judgemental orientation to experience are therefore hypothesised to be more effective in later adolescence (Buttelmann & Karbach, 2017). This raises a potential concern for the many manualised SBMPs incorporating cognitive behavioural therapy (CBT) approaches. These can include activities that focus on reducing symptoms associated with negative emotional states, such as discussion of how to avoid reacting "badly" to situations (Thomas & Atkinson, 2016). Such practices may run contrary to a non-judgemental orientation if not carefully considered.

Combining the above cognitive connections between mindfulness and EF, Holas and Jankowski (2013; 2014) propose a model of mindfulness that centres on metacognitive processes operating on multiple levels of information processing (see Figure 1).

Figure 1

A cognitive model of mindfulness state together with its determinants.



Note. From "Jankowski, T., & Holas, P. (2014). Metacognitive model of mindfulness. Consciousness and cognition, 28, 64-80." The model describes mindfulness as emerging from dynamic interactions between metacognitive knowledge, skills, and experience. Monitoring the present moment (the object level) can induce a meta-cognitive experience (which might include insight, compassion, freshness, and curiosity). These experiences are maintained through meta-cognitive skills (such as sustaining attention, inhibition, and task-switching). Meta-knowledge enables a state of mindfulness, which refers to understanding how our beliefs, goals, and strategies interact with metacognitive experiences. This dynamic interplay achieves

a mindful state at the meta-meta level, involving monitoring our own experiences and our cognitive response to that experience.

This model proposes that "executive functions and attentional processes are fundamental in initiating and maintaining a mindfulness state". To focus attention on the present moment, cognitive processes must be engaged that support sustaining attention, as well as the inhibition of distractions and appropriate monitoring of when the mind is wandering from the present moment (Lutz et al., 2015). Holas & Jankowski (2013) argue that a further cognitive process is required to switch attention between multiple tasks, such as when switching attention from acknowledging a sensation to the present moment. These cognitive processes are the meta-cognitive skills component of this model, which are collectively referred to as working memory by the authors, map directly with the three-factor model proposed by Miyake et al. (2000) of EF as a higher-order process emerging from subprocesses engaged in response inhibition, shifting, and updating.

Although all cognitive processes explored in this research have been implicated in mindfulness, whether mindfulness practices have a differential effect on ECPs is less well-documented, especially concerning young people. Gallant (2016) mapped the impact of mindfulness practices against the three ECPs (updating, shifting, and inhibiting) by reviewing intervention studies in adult populations. Here, the most substantial effect was seen on inhibition-related outcomes, corroborated by recent research suggesting that inhibition is particularly associated with adult mindfulness training (Lodha & Gupta, 2022). Although the effect of SBMPs on inhibition is not yet fully known, improved inhibition is suggested to be a mechanism through which mindfulness can reduce classroom disruption (Zoogman et al., 2015).

Gallant (2016) found mixed results for the impact of mindfulness on working memory, with some studies showing improvements but also suggesting that this may be related to improvements in other domains, such as attentional control. There was even less support for cognitive flexibility, with the evidence suggesting that mindfulness only improved flexibility measures in older adults (Moynihan et al., 2013). This review indicates that mindfulness practices may have a differential effect on ECPs. However, because of the developmental trajectories of EF outlined previously, this differential effect cannot be generalised to CYP populations. The suggestion that age may be a moderator for the effectiveness of adult mindfulness interventions justifies exploring whether it also acts as a moderator for SBMPs.

2.4 SBMPs: the current state of the evidence

2.4.1 Effectiveness of SBMPs

Although SBMPs are experiencing a surge in use, there may be a need to temper this enthusiasm as this growth may be outpacing the evidence base (Weare, 2023). This criticism has gained traction through recent null findings published from large-scale, well-powered RCTs (Kuyken et al., 2022; Dunning et al., 2022), which call into question the effectiveness of SBMPs for improving mental health and well-being outcomes for CYP.

Although studies on the direct mental health effects of mindfulness training dominate the research landscape, SBMPs are posited to benefit other spheres, including various physical health outcomes, such as improved sleep (Bögels et al., 2014); and behavioural outcomes, such as reduced classroom disruption (Felver et al., 2013). Well-designed experimental studies have also found evidence for significantly improved social skills in CYP who have participated in an SBMP (Schonert-Reichl et al., 2015). Of specific interest to this review, SBMPs have been associated with improvements across several cognitive outcomes, such as attention and memory (Zoogman et al., 2015). A recent meta-analysis of high-quality SBMP studies conducted by Roeser et al. (2022) found that SBMPs also improve executive functioning.

This is suggested by Roeser et al. (2023) to run contrary to the findings published in the MYRIAD project (Dunning et al., 2022). However, in this case, it appears that the researchers have conflated executive functions with "affective executive control", which is the focus of this MYRIAD project study (Dunning et al., 2022). Affective executive control can be understood as emerging from the interactions between executive control and affective processes, more akin to the hot EF (Zelazo & Muller, 2014) previously discussed. Thus, affective control refers to a subset of EF functions implicated in the cognitive regulation of emotional responses (Schweiser, 2020). The additional affective dimension included in this MYRIAD study (Dunning et al., 2022) means that the results are difficult to generalise to the non-affective activation of EF. For example, inhibiting a non-affective response to a distraction (e.g. attending to a benign humming noise) is likely to be less cognitively demanding than inhibiting an impulsive emotional response (e.g. reacting aggressively when insulted). There appears to be a need to further delineate the relationship between SBMPs and the cognitive processes associated with executive function, as conflation between constructs such as 'response inhibition' and 'affective response inhibition' may lead to erroneous conclusions.

SBMP studies focussing on specific cognitive outcomes have generally found a positive effect on inhibition-related behavioural tasks (Juliano et al., 2020) and physiological correlates (Andreu et al., 2023). Positive outcomes have also been found relating to cognitive flexibility (Lassander et al., 2020; Lertladaluck et al., 2021) and working memory (Quach et al., 2016), although these benefits might not be maintained beyond the intervention period (Dunning et al., 2022). Whilst the above examples provide data on specific ECPs, this is atypical of the research field, and research has yet to be published that directly focuses on the potential differential effect of SBMPs on ECPs.

2.4.2 Age-related effects

Several reviews of the mindfulness literature have suggested that developmental perspectives are underrepresented (Van Dam et al., 2018; Roeser et al., 2023). Psychological research on SBMPs is predominately linked to educational (38.2%), clinical (35.6%), or other (12.5%) psychological fields (Roeser et al., 2023), with little input from developmental psychology. The distinct developmental trajectories posited for ECPs have implications for the optimal ways to conceptualise and operationalise SBMPs for different age groups. Ibbotson (2023) argues that a developmental model of EF must account for the variable trajectories through which a more domain-general EF emerges from associated cognitive functions. In SBMP research, Roeser et al. (2023) have identified that how mindfulness is conceptualised at varying stages of cognitive development needs to be better understood. Even less is known about how this age-dependent understanding of mindfulness interacts with EF, which is differentially expressed depending on developmental stage.

Whether participant age has a moderating effect on the impact of SBMPs on EF-related outcomes is not well documented in the literature. Regarding wellbeing-related outcomes, moderator analyses by Carsley et al. (2017) found that improvements were significantly greater in late adolescence compared with middle childhood, suggesting a positive correlation between age and SBMP effectiveness. Similarly, for emotional regulation, Pickerell et al. (2023) found a significant effect for participants aged ten years and over, with no significant SBMP effect found for younger children. However, the small number of studies included in this meta-analysis (k=8) indicates that any subgroup analysis would be well below the >10 threshold (Borenstein et al., 2011) typically recommended, and these results should be viewed with caution.

Furthermore, even if a tentative conclusion were drawn that there was a positive correlation between age and SBMP efficacy for well-being outcomes, there would be little justification to assume that the same correlation would exist for EF-related outcomes. Dunning et al. (2019) conducted a systematic review and

meta-analysis of mindfulness-based interventions for children and adolescents. They found significant improvements in cognitive flexibility across studies but did not find a significant moderating effect of age. The authors suggest that this may be due to the limited range of ages between the included studies and indicate that more research is needed to compare the effects of mindfulness interventions across different developmental stages.

2.4.3 SBMP heterogeneity

Despite equivocality over the differential impact of SBMPs in improving specific cognitive functions, there is consensus that they can positively impact general EF (Roeser et al., 2023). However, the magnitude of this effect is highly variable between studies, suggesting that contextual factors, participant characteristics, and intervention characteristics are likely to play a significant moderating role. In fact, given the high heterogeneity between studies concerning theoretical foundation, intervention, implementation, context, outcome measures, etc., it is reasonable to consider SBMPs as complex interventions (Petticrew, 2013).

Complex interventions emerge from multiple interactions to the extent that it is not necessarily possible to isolate single components as having a causal effect on outcomes. To illustrate this complexity, we can compare data from two RCTs on SBMPs collected through Likert scales to measure CYP acceptability of the intervention experience. Kuyken et al. (2022) found that CYP gave an average rating of 4.7/10 for their experience of the mindfulness programme, suggesting an overall lukewarm response at best. On the other hand, participants in Andreu et al. (2023) gave an average rating of 8.5/10. This considerable heterogeneity in intervention acceptability could be due to various factors. For one, these studies used different mindfulness programmes, with either psychologists (Andreu et al., 2023), or teachers (Kuyken et al., 2022), as the intervention administrators. Adding to this, Likert scale responses are often moderated by culture (Lee et al., 2002), and it may be the case that young people in Chile (Andreu et al., 2023) were more inclined to give a higher rating to the same valence than young people in England. The difference in the mean age of participants may also have influenced the results, with the older sample in Kuyken et al. (2022) perhaps less susceptible to social desirability bias.

Of course, an actual difference in subjective acceptability may exist. Still, the complexity of SBMPs means that this conclusion cannot be assumed without considering a wide range of factors. For this reason, Petticrew et al. (2013) suggest that systematic reviews of complex interventions, such as SBMPs, require a synthesis of the evidence base that explores a range of population, intervention, and methodological characteristics.

2.4.3.1 Programme structure

Whilst these school-based mindfulness programmes are sometimes treated as a homogenous category in systematic reviews (e.g. Fisher et al., 2016), the operationalisation of mindfulness practice is highly heterogeneous. Within the empirical literature, 20% of SBMPs operationalised mindfulness through a non-manualised programme (Roeser et al., 2022). Of the remaining 80% of intervention studies, 30% used a formal programme adapted from adult-targeted interventions such as MBSR, whilst the remaining 50% used novel programmes. In one review, only 32% of studies used an intervention that had been empirically evaluated (Felver et al., 2016).

The content of these programmes is highly variable and representative of the variation seen in broader CYP mindfulness literature, variously including elements from fields such as cognitive-behavioural therapy (CBT), social and emotional learning (SEL), and positive psychology. These differences in structure may impact the extent to which SBMP improvements are maintained over time, with one meta-analysis suggesting that although both manualised and novel SBMPs show significant effects post-intervention, these are only maintained at follow-up for interventions using a novel approach (Carsley et al., 2018).

2.4.3.2 Intervention components

Given the high heterogeneity seen in the operationalisation of SBMPs, a detailed exploration of the intervention and implementation protocols is warranted (Van Dam et al., 2018). Conflating interventions under the umbrella of the SBMP term may lead to erroneous conclusions when interventions with fundamentally different mechanisms of action and desired outcomes are grouped. For example, many SBMP reviews include mindful programmes that include elements of physical exercise, such as yoga, tai chi, and qigong (e.g. Zhang et al., 2023). Research suggests that acute physical exercise significantly improves executive functioning in CYP (Verburgh et al., 2014), but conclusions drawn from such reviews have limited generalisability to SBMPs that forego physical components. Furthermore, this conflation creates challenges when attributing a causal role to mindfulness alone. Therefore, it seems prudent to include coding of SBMP protocols based on the presence or absence of a physical component. Thus, through closer examination of intervention characteristics, the links between the attentional mechanisms involved in mindfulness practices and improved EF can be better isolated.

2.4.3.3 Administrator role

Differences in how SBMPs are administered add further complexity and vary between studies, with 40% of studies administered by external researchers (Phan et al., 2022), 28% by teaching staff, and 19% by mindfulness instructors. The role of the facilitator has been suggested to moderate the effectiveness of an SBMP, with Carsley et al. (2018) finding that teacher facilitators significantly improved mental health outcomes but not mindfulness. Interestingly, they found that the reverse was true for outside facilitators, who significantly improved the mindfulness, but not mental health, of the CYP involved. The role of the administrator as a predictive variable for SBMP efficacy has not been confirmed in subsequent meta-analyses, with Mettler et al. (2023) finding limited evidence that outcomes are moderated by whether the teacher is experienced in mindfulness practice, or even if the SBMP is administered through audio recordings. As Phan et al. (2022) suggest, exploring the potential moderating effect of the administrator role on SBMP effectiveness could help explain some of the variance between study effect sizes.

2.4.3.4 Intervention dosage

Total intervention dosage is highly variable between SBMPs, with Felver et al. (2016) reporting total intervention exposure times between 75 and 2160 minutes (M=396.7). Roeser et al. (2022) also report high variation between the length of individual sessions and time intervals between sessions. Despite this heterogeneity, whether intervention dosage and intensity moderate SBMP effectiveness in improving EF has yet to be explored. Findings from meta-analyses of adult mindfulness interventions suggest that increasing the number and duration of mindfulness sessions can improve performance on PBTs designed to measure inhibition (Verhaeghen, 2021). Whether this translates to CYP is uncertain, with Zoogman et al. (2015) suggesting that the intervention duration did not significantly affect intervention outcomes. However, this meta-analysis included studies conducted in clinical settings, and the findings may not be generalisable to SBMPs. Still, the finding that effectiveness was not mediated by dosage is interesting, given that the benefits of SBMPs are thought to derive from increased proficiency. This might assume that increased exposure would increase the likelihood of developing the mindfulness skills hypothesised to improve ECPs. Exploring the moderating effects of total dosage has been suggested as an avenue for future research in adult mindfulness interventions (Gallant et al., 2016). It would also contribute to our understanding of SBMPs.

2.4.4 Methodological concerns

A shared reflection emerging from reviews of SBMPs is that it is impossible to draw causal conclusions on the effectiveness of mindfulness interventions due to the low methodological rigour of the evidence base (e.g. Evans et al. 2018).

2.4.4.1 Control conditions

Felver et al. (2016) found that only half of the studies analysed in their review had a control condition, while Evans et al. (2018) found that only 19% of SBMP studies included in their review had a control condition. Of those that did include a control group, only 21% used a well-matched active comparison condition. As such, the low internal validity of studies lacking an appropriate control condition raises the likelihood that observed effects are artefacts of observed and unobserved confounding variables. The importance of a control condition is likely to be especially critical when considering mindfulness interventions, as many of the outcome measures are dependent on subjective self-report (Boot et al., 2013) or subjective reports from parents and teachers. This subjectivity makes these measures more susceptible to expectancy and placebo effects. The high prevalence of mindfulness in the media and the possible overestimation of its positive effects (Van Dam et al., 2018) may further increase expectancy effects. Furthermore, placebo effects have been associated with engaging in any structured routine (Moerman & Jonas, 2002), which raises doubts about whether daily mindfulness practice offers a better alternative to current provision, such as social and emotional learning (SEL).

2.4.4.2 Outcome measures

The outcome measures used to assess mindfulness interventions have been criticised as myopic (Bergomi, 2012), focusing primarily on self-report or parent/teacher questionnaires. Felver et al. (2016) found that the outcome measures of 96% of SBMP studies included the use of questionnaires; 82% relied solely on questionnaire data, and no studies included a cognitive task as an outcome measure. Although self-report measures of mindfulness evidence some validity in their ability to predict beneficial clinical outcomes (Khoury et al., 2015), alternatives to self-report are worth exploring (Davidson & Kaszniak, 2015). Regarding mindfulness programmes and cognitive functions, there is an increased interest in exploring the biological markers of mindfulness outcomes through MRI or EEG measures (e.g. Young et al., 2018; Doborjeh et al., 2019). However, as previously discussed, using these as proxies for cognitive processes is problematic (Kruger & Kruger, 2017). Alternative approaches use behavioural tasks such as the Stroop task (Davidson & McEwen, 2012) and Flanker task (Schonert-Reichl et al., 2015); or more comprehensive neuropsychological

batteries such as the NEPSY-II (Thomas & Atkinson, 2016). Such outcome measures are useful for triangulating data collected from questionnaires, which may provide a more ecologically valid measure of executive functioning but also carry a higher risk of bias.

2.4.4.3 Statistical power

Another methodological need emerging from the SBMP literature is more statistical power. Studies with low statistical power may lack the sensitivity to detect small to moderate changes following a mindfulness intervention, resulting in a type II (false-negative) error. In studies evaluating educational interventions with promising preliminary trial results, low power can lead to effect size inflation of 52% or more (Sims et al., 2022), resulting in a type M (magnitude) error. There is an increased probability that the benefits of school-based mindfulness interventions are currently overstated as inflated effect sizes are aggregated in the research literature through publication bias (Francis, 2012). Meta-analyses can effectively mitigate these concerns, as combining data from multiple studies increases the overall sample size. Publication bias adjustments can also be conducted through meta-analyses, which are typically neglected in educational research (Ropovik et al., 2021). As such, meta-analyses emerge as a useful tool to help ascertain whether SBMPs have a true positive effect on EF outcomes and thus increase the robustness of the evidence base.

2.4.4.4 Implications

A subsequent review (Roeser et al., 2022) has addressed these methodological concerns by limiting included studies to those that have employed a randomised (e.g. randomised controlled trials) or matched-groups design. To mitigate the exaggeration of effect sizes that might arise from studies with low statistical power, the same review also limited studies to those with a total sample size of at least 30. When correcting for these methodological concerns, the review found limited evidence suggesting that SBMPs effectively improved psychological well-being outcomes. This was consistent with published findings from the MYRIAD project (Kuyken et al., 2022), suggesting that meta-analyses limiting data to those collected through well-designed empirical studies may produce more reliable conclusions. Consolidation of data from high-quality studies in a meta-analysis focussing on executive function is needed to determine if the reported positive EF outcomes following SBMPs might have been overestimated due to a lack of methodological rigour.

2.5 Summary

Executive function (EF) is either a unitary supervisory system or a component system of cognitive processes such as inhibition, updating/working memory, and cognitive flexibility. This research conceptualises EF as a complex system emerging from the dynamic interplay of multiple processes (Zink et al., 2023), including non-cognitive factors such as culture (Doebel & Muller, 2023). Although EF is fundamentally holistic, investigating separable cognitive processes may enrich our understanding.

Developmental perspectives on EF and mindfulness are underrepresented in the literature. However, both are hypothesised to follow developmental trajectories that may significantly impact the effectiveness of SBMPs focused on EF-related outcomes. Although some meta-analyses have attempted to explore the moderating effect of age on SBMP effectiveness (Dunning et al., 2022), they have lacked the statistical power needed to draw robust conclusions from the data.

Metacognitive models of mindfulness hypothesise a differential effect on cognitive processes associated with EF (Holas & Jankowski, 2013; 2014), but limited research explores the relationship between SBMPs and separable ECPs (e.g., working memory, cognitive flexibility, and inhibition). Although some studies have explored separable processes in isolation, no published research has directly compared the potential differential effects of SBMPs on ECPs.

Although SBMPs are complex interventions, detailed syntheses of study characteristics are not typically found in systematic reviews and meta-analyses (Phan et al., 2022). Interventions differ with respect to content, administrator, total dosage, and outcome measures. Triangulation of these data with statistical analysis can provide a richer understanding of SBMPs and help explain the considerable heterogeneity of effect sizes between SBMP studies.

A lack of methodological rigour in the evidence base may lead to an overinflated estimate of SBMPs' effectiveness. Concerns include low statistical power, a need for appropriate control conditions, and over-reliance on BRSs or PBTs as outcome measures. These must be addressed to ensure that findings are sufficiently robust to warrant SBMP uptake.

3. Method

This section begins with a brief outline of systematic reviews, meta-analyses, and their relevance to Educational Psychology practice. This study's epistemological and ontological positions are provided, focusing on critical realism, pragmatism, and complexity theory to justify synthesising quantitative and qualitative data. The review questions developed through the initial literature scoping are presented next. The study selection process is then outlined, starting with an overview of the literature search and subsequent article screening process. An outline of how data has been extracted from studies is given, followed by details of the critical appraisal process using Gough's (2007) Weight of Evidence (WoE) framework. This is followed by an outline of the process by which complex intervention data will be collated into a narrative synthesis. The final section details the statistical analysis procedures. This describes how quantitative data has been synthesised through meta-analysis and how potential moderator variables have been explored through meta-regression. An outline of the processes for conducting heterogeneity, sensitivity, and power analyses is provided, alongside the process for examining publication bias.

3.1 Introduction

3.1.1 Systematic reviews and meta-analyses

A systematic review is a structured and comprehensive synthesis of the evidence base relevant to a particular research question. Systematic reviews allow for the evaluation of current evidence and can explore heterogeneity between studies, highlighting gaps in knowledge and providing a direction for future research (Liberati et al., 2009). They also facilitate the collation of large data sets, which can support concerns about the need for more statistical power in many mindfulness-based programme studies. Furthermore, systematic reviews allow for "emergent perspectives that transcend retrieved studies" (Heyvaert et al., 2013). This meta-perspective has the potential to enable faster implementation of research discoveries into practice (Gough et al., 2012; Moher et al., 2015).

A meta-analysis employs statistical methods to combine and analyse results from multiple studies to derive a pooled effect estimate. Conclusions drawn through meta-analyses of accumulated evidence can have increased reliability and accuracy due to the systematic methods used (Borenstein et al., 2022). These methods include corrections for potential bias through a critical appraisal process outlined above. This

approach can enhance the statistical power of smaller studies (Borenstein et al., 2022) and explore the potential factors contributing to variability between studies.

3.1.2 Methodological relevance to educational psychology

A systematic review and meta-analysis of controlled intervention studies is a notable departure from the qualitative approaches or less controlled quasi-experimental designs more typical of Educational Psychology research (Frederickson et al., 2008) on interventions. These latter approaches are essential to understanding the subjective experiences and contextual nuances that influence interventions. Still, they often have limited generalisability beyond the relatively narrow parameters of their research contexts. An over-reliance on idiographic approaches can limit the development of a shared, cumulative, and actionable evidence base for the EP profession.

Educational Psychology Services are well positioned to bridge the gap between research evidence and real-world practice through consultation (Davis, 2012), enhancing the reliability and impact of EP recommendations and ultimately improving outcomes for CYP. However, although psychologists working with CYP report evidence-based practice to some extent (Burnham, 2013), this is limited by practical constraints, including a lack of time and access to quality research (Cottrell & Barrett, 2015). By synthesising the data from studies that meet a minimum standard of methodological quality, this research can support EPs in grounding their work in the evidence base whilst managing the competing demands of an applied psychology role.

Overall, the relevance of systematic reviews of interventions to Educational Psychology practice lies in their ability to "globalise the evidence [and] localise the decision" (Eisenberg, 2002). The meta-analyses and moderator analyses help to globalise the evidence by pooling data from multiple studies. The narrative synthesis of study characteristics provides a contextual understanding of how this globalised evidence can be applied to "local" educational settings by exploring how studies compare in relation to setting, population, intervention, and implementation factors. Therefore, this combination of narrative synthesis and statistical analysis provides a macro-perspective that would otherwise be beyond the scope of doctoral research.

3.1.3 Ontological and epistemological positions

Education system research is highly complex (Hilpert & Marchand, 2018). As such, there is a rich plurality of paradigms through which such systems can be explored. Given the diversity in approaches, it is necessary

to outline the philosophical assumptions that guide research to avoid misinterpretation by those employing different paradigms (Bracken, 2010). As such, a justification for the ontological and epistemological positions taken in this review is needed, alongside the philosophical perspective of the researcher (Moon & Blackman, 2014).

Regarding ontology, this research uses a critical realism paradigm. This describes a shared external reality independent of our perceptions but posits that it is impossible to have a perfect understanding of this reality, as it can only be interpreted through our subjective experience (Bhaskar, 2020). Our biological systems, individual schemata, languages, social networks, cultural influences, etc. constrain this reality experience. RCTs might ostensibly be classed as nomothetic approaches that are designed to determine a shared, unbiased reality, but the position taken in this review is that although an objective reality may exist regarding the efficacy of SBMPs on EF, a truly objective evaluation is not possible. In fact, from a systems perspective, asking whether an SBMP causes improvements in a child's EF-related cognitive processes (ECPs) is grounded in a linear logic that fails to appreciate the interconnectedness of all influences acting on their bioecological systems.

Our understanding of the results presented in the studies is mediated by methodological and implementation choices based on the subjective perceptions of those involved and uncountable contextual factors, including staff buy-in, demographics of the student population, and school culture. A critical realist ontology conceptualises reality as stratified (Bhaskar, 2020) between the empirical, actual, and real domains. A deeper understanding is driven through the theoretical interpretation of these domains, as can be broadly seen in the EF literature (e.g. Diamond, 2016). Through this ontology, understanding the deeper structures and mechanisms through which observable phenomena emerge requires methodological pluralism (Kempster & Parry, 2011).

From an epistemological perspective, this research takes the position of pragmatism, drawing on both post-positivist and constructivist stances. A pragmatist approach is question-driven and assumes that taking multiple positions to best answer the research questions should be preferred over philosophical continuity (Creswell & Creswell, 2017). In line with the critical realist ontology, post-positivism assumes that absolute truth cannot be established, although an objective reality exists. From this perspective, the goal of knowledge-seeking is to attempt to best approximate this shared reality, which this study assumes to be through the synthesis of multiple approaches, including methodologically rigorous empirical studies.

In the context of this review, establishing whether SBMPs influence ECPs is not intended to imply causality per se. Rather, it aims to be a "proof-of-concept", with additional moderator analyses guiding when and where implementation of SBMPs might be advantageous. Ultimately, the review aims to support practitioners in deciding if, all things being equal, implementing a mindfulness intervention might improve executive functioning skills.

3.2 The review process

To answer the research questions, this systematic review and meta-analysis broadly follow the review process outlined by Pettigrew & Roberts (2006). The following sections explore this review's full process, including the literature search, article screening, and data extraction processes. The process is detailed to address reproducibility and transparency concerns levied against meta-analyses that do not report these procedures (Polanin et al., 2020).

Table 1

Overview of the procedural stages of this systematic review and meta-analysis

Stage	Brief Outline
Question formulation	Scoping questions defined during the initial review of the research base were refined using the SPICE framework.
Literature search & article screening	A systematic search of relevant literature was conducted through four databases. Comprehensive article screening followed Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines and used EPPI-Reviewer 6 software.
Data extraction	Data from the 30 included studies is extracted relating to setting, sample, intervention, implementation, and outcome measures. Descriptive statistical data is extracted for subsequent meta-analyses.
Critical Appraisal	All included studies have been appraised using a modified version of the Gersten et al. (2005) protocol to assess methodological quality. Review-specific appraisals of methodological and contextual relevance are based on criteria developed by the researcher.
Narrative synthesis of studies.	Textual data from studies has been qualitatively synthesised by the primary researcher to produce a narrative summary of study characteristics studies following three-stage process (Petticrew & Roberts, 2006).

Preliminary statistical analysis.	The initial analysis explores a composite outline of the evidence, which includes exploration of pooled effect size and between-study heterogeneity. Sensitivity analyses identify outliers and explore the fit of different models.
Meta-analyses of differential cognitive functions.	Five random effects meta-analyses explore the effect of SBMPs on measures of global EF, attention, response inhibition, working memory, and cognitive flexibility.
Moderator analyses	Random effects moderator analyses explore the potential influence of selected study variables on all five cognitive function meta-analyses. Both meta-regressions and Correlated and Hierarchical Effects (CHE) modelling are used.

3.2.1 Question formulation

The review questions have been formulated using the SPICE (Setting, Perspective, Intervention, Control, Evaluation) conceptual framework (Booth, 2006). Whilst some intervention factors can be quantitatively coded (e.g. total intervention exposure time; number of intervention components), others require thematic coding of qualitative data, such as descriptions of the specific mindfulness techniques employed in each programme. The meta-analysis will focus on elements that can be quantified. In contrast, the narrative synthesis will allow for a more exploratory investigation of the heterogeneity found between EF-focused SBMP intervention studies. As such, these questions emphasise the complexity of school-based interventions, where efficacy, as measured by empirical studies, can be considered as providing proof-of-concept evidence. However, the emergent effectiveness of SBMPs in real-life school settings depends on the interactions of myriad contextual factors. As identified in the scoping literature review, the factors that might be particularly relevant include population characteristics, such as the age and gender of participants; implementation characteristics, such as the programme administrator and intervention duration; and the overall quality and relevance of the included studies.

The review questions are:

1. Do school-based mindfulness programmes have a differential effect on cognitive processes associated with executive function?
2. Does age moderate the extent to which school-based mindfulness programmes improve measures of executive function and associated cognitive processes?
3. Which intervention characteristics moderate the extent to which school-based mindfulness programmes improve measures of executive function and associated cognitive processes?
4. How can future research strengthen the evidence base for the effectiveness of school-based mindfulness programmes in improving executive function and associated cognitive processes?

The post-positivist approach recognises the complexity of SBMPs, and the meta-regressions exploring potential moderating effects of various study characteristics aim to explore this complexity in further detail. A positivist stance might seek to establish a nomothetic causal role of SBMPs in improving EF. However, this is assumed to be practically impossible given the heterogeneity of reported outcomes outlined in the literature review and our current lack of computational power to model the complexity of SBMPs. For example, performance on a behavioural task measuring executive function emerges from countless interactions between forces operating at various levels of complexity within the individual (Karmiloff-Smith et al., 2017), and that operate outside of the individual from the microsystemic to the macrosystemic (Bronfenbrenner & Morris, 2007). A post-positivist perspective acknowledges that our understanding of SBMPs and EFs is in constant flux, and constructs such as mindfulness and working memory continue to evolve through an iterative process that also relies on qualitative data. Addressing the questions above in this study involves a degree of triangulation between the quantitative results from the statistical analyses and the qualitative data derived from the narrative synthesis. As such, all questions will draw on elements of implementation science (Bauer & Kirchner, 2020) by contextualising how, why, when, where, and with whom SBMPs improve EF.

3.2.2 Literature search

The literature was systematically searched between 22nd and 24th August 2023 using four online databases: EBSCO, PsycINFO, PubMed, and Web of Science (WoS).

Several search approaches were considered to balance the need for sufficient sensitivity whilst also ensuring enough specificity to make the screening process tenable within the constraints of this review. These included the PICO (Population, Intervention, Comparison and Outcomes) tool, highlighted in the Cochrane Handbook for Systematic Reviews of Interventions (Lefebvre et al., 2013), as well as an adapted version that increases specificity for qualitative or mixed methods reviews through an additional *Study Design* criterion (PICOS). A comparison study between the various approaches for qualitative reviews shows that PICO has greater sensitivity than PICOS (Methley et al., 2014). However, this is at the expense of specificity, which in the same study was significantly greater when using the PICOS tool.

Although the PICOS framework is preferred by PRISMA-CI guidelines (Guise et al., 2017), the SPICE (Setting, Perspective, Intervention, Control, Evaluation) framework (Booth, 2006) was also considered as it places greater emphasis on contextual factors, which may be particularly relevant to complex interventions. This framework separates the population into setting and perspective elements, which fits this review's focus

on school-based mindfulness programmes. These contextual factors are critical from a complex systems perspective (Booth et al., 2019), as is an evaluation of emergent aspects of how a system has changed following an intervention that linear outcomes cannot necessarily measure. For these reasons, and because this framework has already been used during the question formulation stage, this approach was deemed to have the most epistemological synergy with the main thrust of the review. It was selected to guide the search strategy (see Table 2).

Table 2

Search terms used in systematic literature search

SPICE	Search Term	Rationale
Setting	“school” or “education*”	This review aims to explore MBPs that are administered in an educational setting (e.g. school, alternative provision).
Perspective	“child*” OR “adolescen*” OR “youth” OR “student”	This review focuses on the impact of SBMPs on children and young people.
Intervention	“mindful*” OR “meditat*”	This review focuses on mindfulness-based programmes.
Comparison	“control*”	This review evaluates studies that include a control condition.
Evaluation	“executive" or "attention*" or "regulat*" or "working memory"	Review of studies included in previous SBMP meta-analyses identified these terms as sufficient to ensure adequate search sensitivity for EF-related SBMP studies.

Note. Boolean search modifiers used in this search include the truncation of root words with an asterisk (*) and the use of parentheses (“”) to preserve exact phrases.

3.2.3 Article screening

Searching the four databases yielded 1148 results (EBSCO=136; PsycINFO=179; PubMed=398; Web of Science=471). These articles were uploaded and cross-referenced using the systematic review software EPPI-Reviewer 6 [Evidence for Policy and Practice Information (EPPI) Centre, 2023]. Automatic tools were used to eliminate duplicates (n=347) and screen for 1. Type of Publication (n=49) 2. Date of publication (n=17), and 3. Language (n=8). A complete list of the exclusion criteria is presented in Table 3.

Table 3

Inclusion and exclusion criteria used to screen articles.

Criterion	Inclusion	Exclusion	Rationale
1. Type of publication	A peer-reviewed journal article.	Study is not published in a peer-reviewed journal article (e.g. dissertation).	The peer review process sets a minimum standard for research quality.
2. Date of publication	The publication date is between 01.01.2007 and 24.08.2023.	The publication date is before 01.012007.	SBMPs emerged c. 2007 (e.g. MiSP).
3. Language	The article is presented in English or Spanish.	The article is not presented in English or Spanish.	Advanced language skills are required to analyse the data.
4. Setting	The intervention is delivered in an educational setting (e.g. school, college, alternative provision).	The intervention is not delivered in an educational setting (e.g. clinic, home, web-based).	This review aims to evaluate interventions that have been administered in an educational setting.
5. Perspective	Participants are of school or preschool age (e.g. between 3-18 years) and have not been medically screened (e.g. ADHD).	Data is collected from participants that are not of school age (e.g. university students).	This review aims to evaluate interventions that have been conducted in mainstream classroom settings.
6. Intervention	a. The intervention practices self-regulation of attention and an open monitoring style. b. the intervention is universally targeted (i.e. administered to whole school population) c. Mindfulness is the central component of the intervention.	a. The intervention does not involve self-regulation of attention and an open monitoring style. b. the intervention is not universally targeted. c. Mindfulness is not the central component of the intervention (e.g. primarily involves yoga).	This review aims to evaluate studies that have operationalised mindfulness programmes according to the shared features of the most popular contemporary frameworks (e.g. Bishop et al., 2004; Isbel & Summers, 2017).
7. Control	The study includes a single intervention group and a control condition.	The study does not include a control condition.	Control conditions are necessary to account for confounding variables.

8. Evaluation	The main outcome measure relates to ECPs.	The main outcome measure does not relate to ECPs.	This review focusses on the cognitive outcomes of SBMPs.

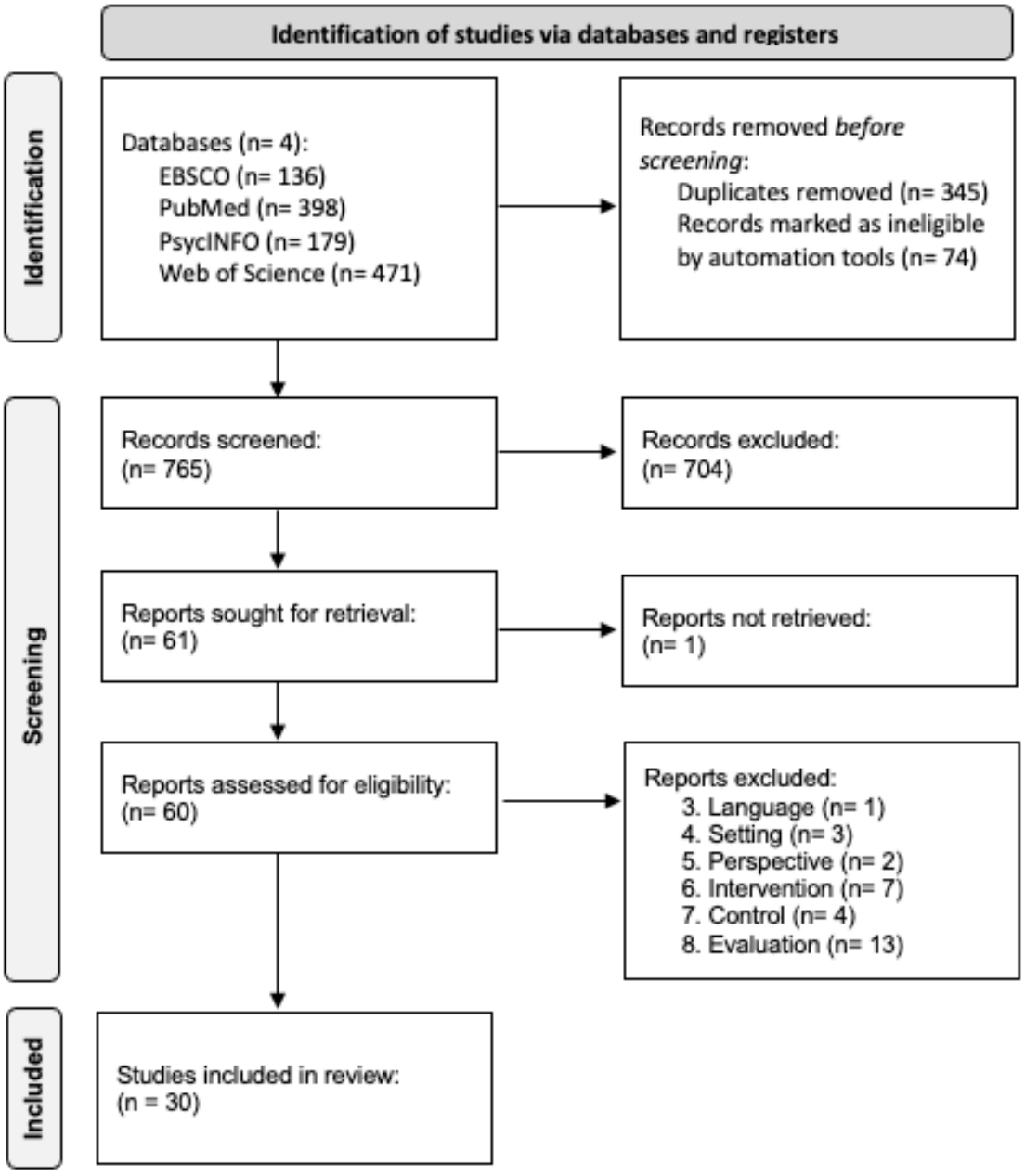
Note. MiSP=Mindfulness in Schools Project. ECP=EF-related Cognitive Process.

Ancestral and citation searches based on recent SBMP systematic reviews (Dunning et al., 2022; Phan et al., 2022; Rea et al., 2021; Roeser et al., 2022) yielded no additional articles, suggesting that the initial search parameters had sufficient sensitivity. Seven hundred sixty-five articles were selected for title and abstract screening. All searches were performed by the primary researcher, with the screening process facilitated by EPPI-Reviewer 6 software 6 (EPPI Centre, 2023).

The titles and abstracts of the remaining 765 studies were further screened against the inclusion criteria, excluding 705 articles. The remaining articles (n=60) were selected for full-text screening. One report was not retrieved, and 30 were excluded. The rationale for these exclusions is presented in Appendix A. Overall, 30 studies met the inclusion criteria, and their references are given in Appendix B. A PRISMA flow diagram outlines the full screening process (Moher et al., 2015) (see Figure 2).

Figure 2

PRISMA flow diagram showing the screening process of reviewed articles.



Note. Records screened (n=765) are those selected for title and abstract screening following removing duplicates and records marked as ineligible by automation tools. Reports assessed for eligibility (n=60) are those selected for full-text screening. PRISMA flow diagram produced with the PRISMA flow diagram tool (Haddaway, 2022).

3.2.4 Data extraction

Qualitative and quantitative data have been extracted from each study and presented in the systematic map of the evidence base. This data has been coded and collated using EPPI-Reviewer 6 (EPPI Centre, 2023), as recommended by Thomas and Harden (2008).

3.2.4.1 Coding study characteristics

This extracted data has been collated using a standardised form (see Appendix C). First, operationalisation data on SBMPs and the tools used to measure executive function have been collated and coded according to the best-fit theoretical (e.g. two-component) model. Demographic information includes the age range and mean age of participants, the ratio of males to females, and SEN characteristics (e.g. a diagnosis of autism). As the development of EF is a complex and dynamic process, the average age of participants is further categorised into developmental stages.

Data on the design and implementation of mindfulness programmes used in the included studies has been extracted to account for intervention complexity. The PRISMA-CI checklist (Guise et al., 2017) has informed the decision to extract intervention data based on the number and type of components, total intervention dosage and session length and frequency, replicability of intervention, theoretical foundation, and intervention incentives. Following practices outlined in a previous review (Phan et al., 2022), data on the role of the intervention administrator (e.g. teacher or external professional) was extracted, including information relating to training received by administrators and whether they were provided with ongoing supervision.

To further explore whether the reported efficacy of SBMPs on executive functioning is due to mindfulness-specific mechanisms or non-specific effects, data will be collected on the nature of the activities completed by control conditions in each study. This may be essential as mindfulness practices have entered the public discourse. Their popularity on social media may lead to heightened expectancy effects when engaging in interventions that appear 'mindful' (Van Dam et al., 2018). Control conditions have been categorised as 'active' if participants in the control group receive a form of intervention designed to control for non-specific factors (e.g. relaxation techniques) associated with SBMPs. Control conditions have been categorised as 'passive' if the participants are not exposed to an intervention, are in a treatment-as-usual (TAU) group or are in a waitlist group. A summary of the studies included in this study is presented in Table 4. A more detailed map of extracted data can be found in Appendix C.

Table 4

A brief overview of the 30 included studies

Authors	n.e	n.c	SBMP	Control condition	Mean age (years)	Admin.	Total dosage (minutes)
Andreu et al. (2023)	24	22	Crecer Respirando	SEL	9.8	External	450
Baena-Extremera et al. (2021)	156	164	Breathe through this	TAU	13.1	Teacher	310
Berti & Cigala (2023)	10	11	mindful play/ meditation	relaxation	5.5	External	360
Brann et al. (2023)	24	19	mindful-eating and yoga	TAU	4.4	External	360
Crescentini et al (2016)	16	15	mindful meditation	emotional awareness	7.4	External	468
Crooks et al. (2020)	261	323	MindUP	SEL	4.37	Teacher	187.5
Flook et al. (2010)	32	32	mindful awareness	silent reading	8.3	Teacher	480
Folch et al. (2021)	69	31	MBSR-based	TAU	10.4	Teacher	487.5
Frank et al. (2021)	131	124	Learning to Breathe	SEL	16	Teacher	720
Janz et al. (2019)	51	36	CalmSpace	SEL	6.5	Teacher	NA
Koncz et al. (2021)	31	30	story-based mindfulness	TAU	7	External	270
Lam & Seiden (2020)	45	51	Learning to Breathe	SEL	12.4	External	420
Lassander et al. (2020)	62	69	Dot.be	relaxation	13.5	External	405
Lertladaluck et al. (2021)	15	15	mindfulness training	TAU	4.39	External	960
Magalhaes et al. (2022)	28	29	audio-guided mindfulness	SEL	7.8	Mixed	600

Makmee et al. (2022)	30	30	mindful emotions	TAU	NA	NA	360
Milare et al. (2021)	111	96	mindful schools	SEL	8.95	External	480
Muller et al. (2021)	42	37	brief mindful exercises	classroom breaks	11.4	External	120
Quach et al. (2016)	52	51	MBSR-based	waitlist	13.1	External	360
Ricarte et al. (2015)	45	45	mindful emotion training	waitlist	8.9	Teacher	450
Schonert-Reichl et al. (2015)	51	48	MindUP	SEL	10.2	External	540
Shlomov et al. (2023)	15	13	mindfulness and kindness	dialogic reading	4.4	External	720
Suarez-Garcia et al. (2020)	33	40	Mindkeys	waitlist	8.1	Mixed	800
Thierry et al. (2016)	23	24	MindUP	SEL	4.6	Teacher	375
Thomas & Atkinson (2016)	16	14	Paws.be	waitlist	8.8	Teacher	360
Vickery & Dorjee (2016)	20	16	Paws.be	TAU	7.9	Teacher	360
Wimmer et al. (2016)	10	16	MBSR-based	TAU	10.8	Teacher	1200
Wood et al. (2017)	11	11	Mini-mind	waitlist	3.8	External	300
Zelazo et al. (2018)	72	58	mindfulness and reflection	TAU	4.8	Teacher	720

Note. n.e=sample size of experimental group; n.c.=sample size of control group; admin.=role of intervention administrator.

3.2.4.2 Categorising outcome measures

To explore whether SBMPs have a differential effect on individual ECPs, the outcome measures used in the included studies have been categorised as Attentional Control (e.g. Continuous Performance Task); Inhibition (e.g. Go/No-Go, Stop Signal Task); Cognitive Flexibility (e.g. Trail Making Test); Working Memory (e.g. N-back task); and Global EF (e.g. Minnesota Executive Function Scale). These have been determined where possible using categorisations provided by Gallant (2016) and Toplak et al. (2013). The research literature has been consulted on a case-by-case basis for other established outcome measures not covered by the above. Novel outcome measures (e.g. Likert Scales used in Wood et al., 2017) have been categorised according to the descriptions provided by the study authors.

Outcome measures are further categorised according to whether they are performance-based tasks (PBTs, e.g., a cognitive task completed by CYP) or behaviour rating scales (BRSs, e.g., self-report questionnaires and parent and teacher rating scales). Due to difficulties in interpreting neurophysiological evidence, measures using approaches such as electroencephalography (EEG) and functional Magnetic Resonance Imaging (fMRI) are not included in the review (Gonsalves & Cohen, 2010).

Descriptive statistics were collected from all EF-related outcome measures reported by included studies. Data inconsistencies were found in one study (Folch et al., 2021), resulting in the removal of an outcome measure. Three further studies (Makmee et al., 2022; Salmoraigo-Blotcher et al., 2019; Thomas & Atkinson, 2016) did not provide sufficient descriptive data to calculate effect sizes and were therefore excluded from the meta-analysis.

3.2.5 Critical appraisal

Assessing the quality of studies included in systematic reviews, or critical appraisal (Petticrew & Roberts, 2005), has been highlighted as especially pertinent given the low quality of many SBMP studies identified in the scoping literature review. Although some quality measures have been used during the screening process (e.g. peer-review, control condition research design), more is needed for the subsequent weighting of included studies in the meta-analysis (Newman & Gough, 2020). Previous systematic reviews exploring mindfulness-based programmes (e.g. Dunning et al., 2022) have opted for a critical appraisal framework that focuses on methodological quality only, such as the use of the Cochrane Collaboration's Risk-of-Bias Tool, V.2 (RoB-2; Sterne et al., 2019). As argued in Furlong and Oancea (2007), applied and practice-based research often have additional complexities that warrant appraisals that extend beyond abstract methodological criteria. These include review-specific judgments on the capacity of the research design to

answer the systematic review questions, the extent to which the setting/population of the study can be generalised to answer the review questions, and ethical considerations. This suggests the need for a more holistic approach, such as the Weight of Evidence (WoE) Framework (Gough, 2007; Gough, 2021). The WoE framework evaluates each study across three dimensions, the first of which (WoE A) can be considered analogous to a risk-of-bias tool. The other two dimensions are review-specific and explicitly focus on the appropriateness of the design (WoE B) and relevance of the evidence (WoE C).

3.2.5.1 Generic methodological quality

Weight of Evidence A (WoE A) is used as a generic measure of the study's methodological quality. It refers to how well the execution of the study aligns with the established standards of the method used.

Several protocols were considered to support calculating WoE A. This included the revised Cochrane Collaboration's Risk-of-Bias Tool, V.2 (RoB-2; Sterne et al., 2019), which remains the preferred method for appraising the methodological rigour of studies included in reviews of RCTs. Although this tool does provide supplementary material to support the appraisal of studies using cluster randomisation, the school-based nature of SBMPs has meant that many studies in this review have adopted a quasi-experimental design. As such, the Risk of Bias in Non-randomised Studies of Interventions (ROBINS-I) tool (Sterne et al., 2016) was considered as an alternative. While more sensitive to non-randomised designs, this tool was felt to similarly privilege studies that can exert a high degree of control over variables, such as clinical trials. This is not possible in the noisy, but more ecologically valid, school environment. Furthermore, these Cochrane tools require significant resources to implement effectively (Crocker et al., 2023) and are often misused by less experienced researchers (Puljak et al., 2020; Igelstrom et al., 2021). Due to these concerns, a tool was sought that was both relatively straightforward for an individual researcher to use and designed to accommodate the inherent complexity of social/psychological interventions.

The WoE A tool selected adapts the protocol developed by Gersten et al. (2005) to appraise the general methodological quality of the 30 included studies. This protocol provides an appraisal framework designed explicitly for experimental education research. Alongside more general methodological quality indicators (e.g. whether an active control condition was used over a treatment-as-usual or waitlist control group) seen in Risk-of-Bias protocols, the Gersten et al. (2005) protocol also includes indicators that cater to the complexity of interventions that are delivered on social and psychological levels, rather than the more biological levels that might be the case in randomised-controlled drug trials. These include the fidelity of

implementation and administrator characteristics (e.g. whether class teachers were trained in mindfulness techniques and provided with ongoing supervision).

This review uses a modified version of the Gersten et al. (2005) protocol, which includes amendments to reflect the universally targeted nature of SBMPs. In contrast, the original protocol was designed with a specific special education focus. The protocol appraises studies through essential and desirable quality indicators broadly organised around participants, intervention and implementation, outcome measures, and data analysis. A summary of WoE A scores for the 27 included studies is presented in Appendix D. The full protocol, alongside details of the review-specific amendments, is presented in Appendix E. An example of a completed coding protocol is given in Appendix F.

3.2.5.2 Appropriateness of design

Weight of Evidence B (WoE B) acts as a dimension that appraises the "appropriateness of the research methods of the included studies for answering the review question[s]" (Gough, 2021). The researcher has developed five criteria categories to assess the extent to which the methodological design used by each study is aligned with this review's research questions. The selected categories are research design, range of outcome measures, fidelity of implementation, reporting of moderator variables, and statistical analysis. Included studies were scored 1-3 for each category using descriptive criteria. These criteria, alongside a justification for the choice of criteria categories, are presented in Appendix G. A summary of WoE B scores is given in Appendix H.

3.2.5.3 Relevance of focus

Weight of Evidence C (WoE C) is used to appraise the suitability of the focus and context of the included studies to answer the research questions in this review. Where WoE B might focus on the review-specific appropriateness of the study design, WoE C can be considered a measure of the review-specific relevance of the evidence. Four criteria categories have been developed by the researcher to assess WoE C, focusing on the theoretical framework of the intervention, specificity of outcome measures, cultural relevance, and replicability of intervention in schools. As with WoE B, included studies have been rated 1-3 for each category. The rating criteria, alongside a justification for this review's choice of WoE C categories, are presented in Appendix I. A summary of WoE C scores is presented in Appendix J.

3.2.5.4 Overall appraisal of quality and relevance

Weight of Evidence D (WoE D) represents an overall evaluation of the potential of each study to answer the review questions. This is calculated by averaging scores across the three dimensions to produce a single numerical score (WoE D). A summary of the results from this process is presented in Table 5.

Table 5

Overall Weight of Evidence (WoE D) Score

Study	WoE A	WoE B	WoE C	WoE D
Andreu et al. (2023)	2	2	2.25	2.08
Baena-Extremera et al. (2021)	3	2	1.75	2.25
Berti & Cigala (2023)	3	1.6	1.75	2.12
Brann et al. (2023)	2	2	2	2
Crescentini et al (2016)	1	2	1.75	1.58
Crooks et al. (2020)	2	1.8	2.25	2.02
Flook et al. (2010)	1	2.2	2.75	1.98
Folch et al. (2021)	0	1.6	2	1.2
Frank et al. (2021)	3	2	2.5	2.5
Janz et al. (2019)	2	1.8	2.5	2.1
Koncz et al. (2021)	1	1.8	2.25	1.68
Lam & Seiden (2020)	3	2.4	1.75	2.38
Lassander et al. (2020)	3	2.4	2.25	2.55
Lertladaluck et al. (2021)	1	1.4	1.75	1.38
Magalhaes et al. (2022)	3	2	1.75	2.25
Makmee (2022)	0	1.4	1.25	0.88
Milare et al. (2021)	1	1.8	2	1.6
Muller et al. (2021)	1	1.6	1.75	1.45
Quach et al. (2016)	2	2.4	2	2.13
Ricarte et al. (2015)	2	1.4	1.75	1.72
Salmoirago-Blotcher et al. (2019)	3	2.2	2	2.4
Schonert-Reichl et al. (2015)	2	2.2	2.25	2.15
Shlomov et al. (2023)	2	2	2	2

Suarez-Garcia et al. (2020)	1	1.8	2	1.6
Thierry et al. (2016)	3	2	2.5	2.5
Thomas & Atkinson (2016)	2	1.8	2.25	2.02
Vickery & Dorjee (2016)	1	2	2.5	1.83
Wimmer et al. (2016)	2	2.2	2	2.07
Wood et al. (2017)	1	1.6	2	1.53
Zelazo et al. (2018)	3	2.4	2.25	2.55

Note. WoE B, C, and D descriptors use score ranges to convert raw scores into categorical data. Studies scoring less than 1.00 are categorised as very low; between 1.00 and 1.66 are low; between 1.66 and 2.33 are medium; and above 2.33 are high.

3.3 Synthesis and analysis

The analysis of the included studies takes a pragmatic approach. Firstly, qualitative data are synthesised to provide an overview of empirical SBMP intervention studies that primarily explore EF-related outcomes. Here, the heterogeneity between studies is treated as a strength, enabling rich comparison that may elucidate critical implementation factors and clarify "how the interventions work, why, and for whom" (Popay et al., 2006). This heterogeneity can later be triangulated against statistical data from meta-analyses exploring the differential effect of SBMPs on ECPs. This aims to answer the first review question:

1. Do school-based mindfulness programmes have a differential effect on cognitive processes associated with executive function?

Finally, meta-regressions and correlational and hierarchical effects (CHE) modelling are used to explore whether any implementation factors identified in the literature review and narrative synthesis are statistically significant predictors of SBMP efficacy. This aims to answer the second and third review questions:

2. Does age moderate the extent to which school-based mindfulness programmes affect measures of executive function and associated cognitive processes?
3. Which intervention characteristics influence the effectiveness of school-based mindfulness programmes to improve executive function and associated cognitive processes?

These findings are cross-referenced against the critical appraisal of methodological quality and relevance of included studies to the aims of this research. This aims to answer the final review question:

4. How can future research strengthen the evidence base for the effectiveness of school-based mindfulness programmes in improving executive function and associated cognitive processes?

3.3.1 Narrative synthesis

The initial scoping literature review highlighted significant heterogeneity across intervention studies exploring the impact of SBMPs on executive function. The purpose of the narrative synthesis is twofold. It aims to complement the meta-analysis by exploring the richness of the research field and providing an overview of how the results of the statistical analyses can be localised to specific educational settings.

The primary researcher has qualitatively synthesised textual data from studies to produce a narrative summary and analysis of included studies that follow a two-stage process adapted from Petticrew and Roberts (2006). The first stage involved organising studies into logical categories based on relevant study characteristics. These data are assumed to moderate the efficacy of SBMPs and, given the complexity inherent in school-based interventions, a comprehensive categorisation of study characteristics was considered appropriate. These data were collected in a standardised format, with a summary provided in Appendix C. Data categories included:

- The study's region and educational setting, including relevant details (e.g., socioeconomic development or whether it was rural or urban).
- The population sample, including the sample size, mean age, gender ratio, ethnicity, and other participant characteristics (e.g. above-average percentage of highly educated parents).
- The research design, number and timing of data collection points, number and type of conditions, and allocation procedure were also recorded.
- Intervention characteristics, following TIDIER guidelines (Hoffman et al., 2014) for complex interventions, including the title of the SBMP, whether it was custom-designed or based on an established programme, the main intervention components, and the theoretical foundation of the intervention.
- Implementation characteristics, such as the role of the administrator, the intervention intensity and total dosage, the replicability of the intervention, adherence, and acceptability.

The second stage synthesises these data to make between-study comparisons. This involved qualitatively exploring the similarities and differences between all 30 included studies. The extracted data from all studies was combined for each category according to the categories outlined above. A summary of key

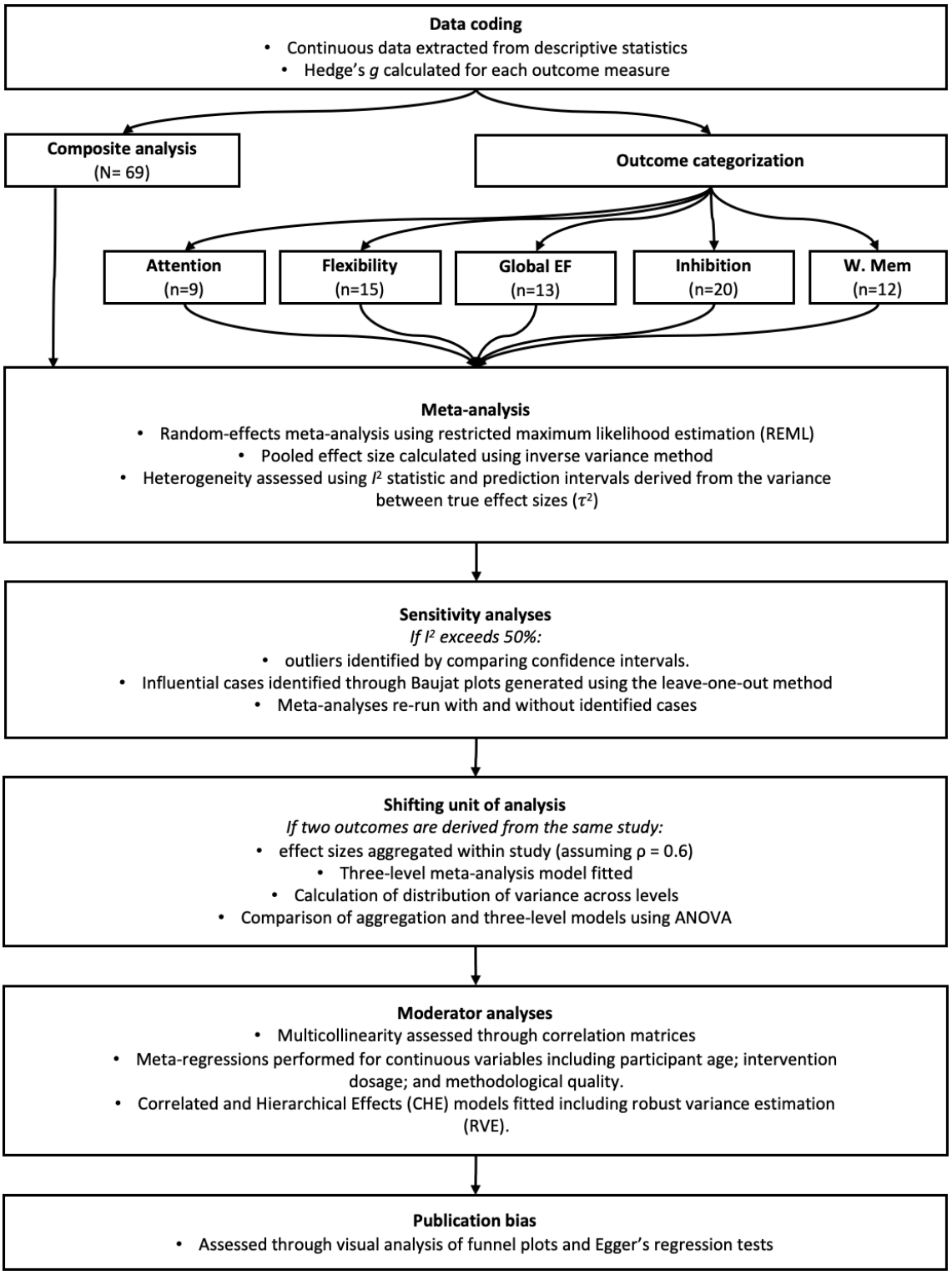
commonalities and differences for each category was then developed and presented in the results. A comparative map of the local context of each study was used to triangulate the results of the statistical analyses. For example, when comparing the meta-regression results exploring the moderating effect of age, these are considered in relation to whether SBMPs targeting the same age groups shared similar intervention components.

3.3.2 Statistical analysis

This section outlines the rationale for selecting the statistical approaches used in the meta-analyses and moderator analyses. Save for the conversion of specific descriptive data (e.g. combining the results that had been split by gender), all analyses were performed using the R statistical package (R Core Team, 2021). A flow diagram summarises the statistical analysis process (see Figure 3).

Figure 3

Flow diagram summary of the statistical process and methods used for the meta-analysis.



3.3.2.1 Effect size calculation

The effect sizes of all 69 EF-related outcomes from the 27 studies included in the meta-analysis were calculated from descriptive statistics, using Hedge's g (Hedge, 1981) to better account for potential biases in studies with small sample sizes (Pigott et al., 2021).

All effect sizes used in this meta-analysis have been derived from descriptive statistics of all EF-related outcome measures reported by included studies to ensure consistency across studies. Inconsistencies were found in data reporting in one study (Folch et al., 2021), resulting in the removal of one outcome measure. A further three studies (Makmee et al., 2022; Salmoraigo-Blotcher et al., 2019; Thomas & Atkinson, 2016) needed more descriptive data to calculate effect sizes and were therefore excluded from the meta-analysis.

These data included the sample size, mean, and standard deviation for the pre-test and post-test scores for both the experimental and control groups. The differences in pre-test and post-test outcome results for each group, or standardised mean gain, were calculated from these. To account for differences in scoring systems, outcomes where a reduced score indicated improvement (e.g. BRIEF measures) were inverted. For example, decreases in BRIEF-2 scores were given a positive value, and increases were given a negative value.

Standardised between-group mean differences ($SMD_{between}$) have been calculated to allow comparison between the many different value scales used by outcome measures included in this meta-analysis. This is achieved by dividing the between-group mean difference by the pooled standard deviation. The pooled standard deviation of the experimental and control conditions is calculated using the formula:

$$s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}}$$

This standardised mean difference, known as Cohen's d, expresses the mean between-group difference in terms of standard deviation, where $SMD_{between}$ =1 indicates that the change in the experimental group is one standard deviation greater than the control group. The hedge's g correction is then applied to his statistic using the following formula:

$$g = d(1 - \frac{3}{4n - 9})$$

Where *d* is Cohen's *d*, and N is the total sample size of both groups. This adjusts effect sizes to account for the inflation in studies using smaller samples. To facilitate this process and minimise the potential for

human error, all effect sizes have been calculated using the R statistical package (R Core Team, 2021) and cross-checked using *psychometrica* software (Lenhard & Lenhard, 2022).

3.3.2.2 The unit-of-analysis problem

When calculating the overall effect of SBMPs on EF, consideration of how to correct for issues relating to dependent effect sizes is needed. Dependent effect sizes can occur when studies contribute multiple effect sizes from the same sample (Harrer, 2021), such as those calculated from the three EF-related outcome measures used by Lassander et al. (2020). If all three of these measures were included in a meta-analysis without any attempt to account for their being derived from the same sample, the assumption of independence (Higgins et al., 2019) would be violated. This can lead to artificially reduced heterogeneity and increase the likelihood of a Type I error (Cheung, 2014).

Shifting unit of analysis

To account for this, alternative approaches that involve "shifting the unit of analysis" (Scammacca et al., 2014) have been employed. To return to the example of Lassander et al. (2020), the three outcomes (WISC-IV Backward Digit Span, NEPSY-II Inhibition A, and the DKEFS Trail Test) are designed to measure working memory, response inhibition, and cognitive flexibility respectively. As identified in the literature review, the specificity of EF-related measures is often overlooked in meta-analyses (e.g. Dunning et al., 2022; Roeser et al., 2023), likely due to the variance in how primary researchers have deployed them. In many cases, a false equivalence is established between more targeted measures of specific ECPs (e.g. working memory) and broader measures of a more global EF construct.

Shifting the unit of analysis changes the level at which data is collated and analysed. This review explores the differential effect of SBMPs on specific cognitive functions. Shifting the unit of analysis to that of specific cognitive functions has the additional benefit of partially correcting for the unit-of-analysis problem. As five separate meta-analyses will be conducted, the three outcomes from Lassander et al. (2020) contribute to their respective meta-analyses as independent effect sizes.

Aggregation of effect sizes

The issue of dependent effect sizes can also be corrected using aggregation (Borenstein et al., 2022), where combining multiple within-study effect sizes produces a single statistic. This well-established approach is recommended by Cochrane guidelines (Higgins et al., 2019). Aggregation might be particularly appropriate in cases where some equivalence between outcome measures can be assumed. Specifically, the aggregation of same-study outcome measures of a specific cognitive function can be considered more

parsimonious than that of less similar measures. Using the three outcome measures provided by Lertladaluck et al. (2021) as an example, it is assumed that there will be a higher degree of correlation between the "Bear and Lion Task" and the "Peg Tapping Task" (both coded as measures of inhibition) than between either and the "Missing Scan Task" (coded as a measure of working memory). Therefore, aggregating the two inhibition tasks in an inhibition-specific meta-analysis will likely introduce less methodological bias than if all measures are aggregated for an overall composite meta-analysis. This accounts for the dependency between effect sizes that remains despite shifting the analysis unit to the cognitive function level.

The within-study correlation needs to be estimated to aggregate multiple outcome measures nested within a single study. This assumption holds for all studies reporting various outcome measures. This meta-analysis assumed a moderate to large correlation (ρ=0.6) due to the hypothesised increase in specificity from the categorisation process by cognitive function leading to greater within-study correlation.

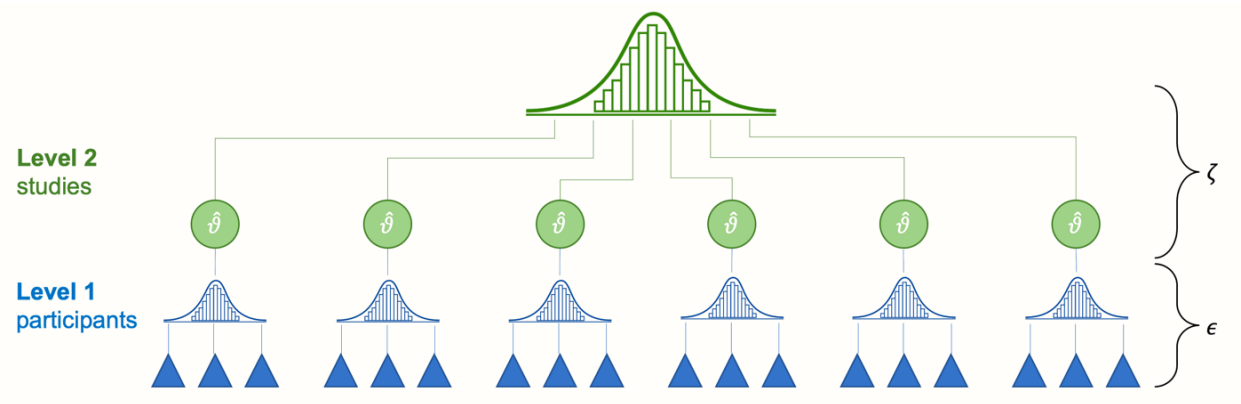
3.3.2.3 Hierarchical modelling

A limitation of the aggregation approach is that the within-study correlation must be estimated and assumed to be the same for all studies in the meta-analysis. This can lead to imprecision, given the expected variance in within-study correlations, and this heterogeneity is not accounted for in an aggregation model. This aggregation process can also lead to an oversimplification of results as it ignores any within-study heterogeneity that is not due to sampling error. The variance between two measures of the same ECP could provide useful information on how different aspects of said process are affected by SBMPs. For example, a lower score on a behaviour rating scale (BRS) of inhibition than a performance-based task (PBT) might suggest that the intervention is less effective at providing far-transfer effects to real-world situations that are more emotionally charged than a cognitive task.

Three-level meta-analysis. A three-level meta-analysis can correct this oversight by modelling the nested nature of dependent outcome measures within individual studies. Although often described as a multi-level meta-analysis, this is a slight misnomer as all meta-analyses have at least two levels: the lowest level accounts for sampling error within studies, and the second level accounts for between-study heterogeneity. The variance between individual participants on Level 1 is, therefore, nested within studies on Level 2 (see Figure 4).

Figure 4

Multilevel Structure of conventional random effects model



Note. From "Harrer et al. (2021). Doing meta-analysis with R: A hands-on guide."

A third level can be introduced by shifting Level 2 to the individual effect sizes of separate outcome measures. These, in turn, can be nested within clusters, denoted by κ, which represent individual studies. This three-level model is fitted using the *metafor* package (Viechtbauer, 2010). Random effects are assigned to individual outcome measures on Level 2 and the larger cluster of individual studies within which they are nested on Level 3. Analysis of Variance (ANOVA) can compare the three-level model with a reduced model that constrains the Level 3 (between-study) heterogeneity to zero. A multilevel allocation of I^2 can be calculated to determine the distribution of variance using the "var.comp" function of the *d.metar* package (Harrer et al., 2019).

Meta-analyses by cognitive function. This review is primarily interested in exploring whether SBMPs have a differential impact on cognitive processes associated with executive function. Therefore, the results derived from individual outcome measures must be pooled according to the specific function they are designed to measure. However, the process by which effect sizes are pooled can introduce bias leading to false negatives, false positives, and magnitude errors. One criticism often levied against meta-analyses, known as the "apples and oranges" problem (Purgato & Adams, 2011), questions the validity of pooled statistics based on studies that are not readily comparable. This is a pertinent issue in psychological research where constructs such as mindfulness and EF lack consensus regarding their conceptual or operational definitions.

Some attempt to correct for potential error emerging from this "apples and oranges" problem is provided by shifting the analysis unit to specific cognitive functions, as discussed earlier in this chapter. However,

given the plethora of outcome measures used to define EF-related outcomes, this step by no means assumes that anything approaching equivalency between measures could be established. Adding to this between-study variance, some heterogeneity in effect sizes is assumed to be at least partially attributable to the inherent complexity of SBMPs and the contexts in which they are deployed.

In other words, differences in true effect sizes are unlikely to be due to sampling error alone, necessitating the adoption of a random effects model to account for variance attributable to another source of error, denoted as ζ_k. To account for ζ_k, the variance between true effect sizes τ^2 must be estimated. Several approaches are available to estimate τ^2, with this meta-analysis opting to use the Restricted Maximum Likelihood (REML) procedure recommended for continuous outcomes by Veroniki et al. (2016). The random effects meta-analyses are calculated using the "*metagen*" function of the *meta* statistical package (v.7.0-0, Schwarzer, 2024). This uses the generic inverse variance method (Borenstein et al., 2010) to calculate pooled effect sizes.

3.3.2.4 Heterogeneity analysis

The expected variance between studies warrants careful consideration of how heterogeneity within the meta-analyses will be analysed.

Assessing between-study variance. One popular approach, Cochran's Q statistic (Higgins & Thompson, 2002), is highly influenced by the number of studies (K) included in a meta-analysis and is therefore not considered a reliable tool for investigating between-study heterogeneity (Harrer et al., 2022). As this meta-analysis consists of 27 studies, Cochran's Q statistic may artificially inflate assessments of heterogeneity. Alternatively, the I^2 statistic (Higgins & Thompson, 2002) is less influenced by K and is, therefore, a preferred measure of between-study heterogeneity for meta-analyses (von Hippel, 2015). The I^2 statistic measures the total variance of the between-study heterogeneity that is not due to sampling error or chance. Whilst more fit-for-purpose than Cochran's Q, the I^2 statistic is still influenced by individual studies' sample sizes (precision). Studies with larger sample sizes will tend to increase I^2 values even when between-study differences remain constant.

A further statistic, the variance between true effect sizes (τ^2), can be used to account for this. This is not systematically influenced by the number of studies or sample size but can be challenging to interpret (Harrer et al., 2021). Prediction Intervals (PIs) provide the interval within which the effect sizes of future studies are predicted to lie. PIs around effect sizes can be calculated using between-study heterogeneity

variance (τ^2) and provide a statistic on between-study heterogeneity that is easier to interpret than τ^2 alone. PIs are calculated using the formula:

$$\mu \pm t_{K-1,0.975}\sqrt{SE_{\mu}^2+\tau^2}$$

$$\mu \pm t_{K-1,0.975}SD_{PI}$$

This review assesses heterogeneity between studies using the I^2 statistic and 95% prediction intervals. Wherever possible, these are presented alongside effect sizes for ease of comparison.

Outliers. As a general "rule of thumb," it is common practice to check for outliers when I^2 exceeds 50% (Harrer et al., 2021). Any attempt to remove outliers requires a strict a priori approach to mitigate the potential for researcher bias. Therefore, outliers were only investigated for meta-analyses where the proportion of the total variation between studies not due to sampling error (I^2) is greater than 50%.

Initial sensitivity analyses were deliberately conducted under the false assumption that all effect sizes are independent. This was to determine whether removing a specific outcome measure might be more appropriate than an entire study. Outliers that might warrant removal from the dataset were first identified by comparing the 95% confidence intervals of individual effect sizes with the 95% confidence interval of the pooled effect size. If a particular effect size's upper bound is lower than the pooled effect size's lower bound, it is identified as an outlying case (Viechtbauer & Cheung, 2010).

Influential cases. Given the range in sample sizes between studies, specific outcome measures may exert a greater influence on the overall pooled effect size. Such influential cases contribute more significantly to overall heterogeneity when their effect sizes deviate considerably from the pooled effect size.

Influenced cases were identified using the "*InfluenceAnalysis*" function of the *d.metar* (v.0.1.0, Harrer et al., 2019) package. This function generates plots using the leave-one-out method, where the results of the original meta-analysis are recalculated K times, each time leaving out one case. These data can then be used to produce influence diagnostics to determine which outcome measures exert the greatest influence on the overall effect size and whether said influence distorts this pooled effect (Viechtbauer & Cheung, 2010). For this meta-analysis, the Baujat plot (Baujat et al., 2002) generated through the "*InfluenceAnalysis*" function is used to assess influential cases. These provide a visual diagnostic that plots each study's overall contribution to heterogeneity (measured using Cochran's Q) against its influence on the pooled result.

3.3.2.5 Moderator analyses

From a systems perspective, current Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA; Moher et al., 2009) guidelines are argued to place too little emphasis on contextual factors (Booth et al., 2019). This is partly addressed in supplementary guidance for complex intervention systematic reviews (PRISMA-CI; Guise et al., 2017) which recommend as essential the reporting of the following intervention items: "the number and description of all components; necessary vs. optional discretionary components; intervention intensity; intervention frequency; replicability of intervention; theoretical foundation for the intervention; intervention incentives; and a priori vs. final components." This review assumes SBMPs are complex interventions as they often have multiple components; they are hypothesised to result in multiple beneficial outcomes, which emerge from multiple causal pathways. They are also conducted in multiple settings and implemented in multiple ways to target multiple populations. Understanding the emergent properties of complex educational interventions cannot be fully understood "by simply asking whether they work" (Petticrew, 2015).

Therefore, a meta-analysis of effect-size data from RCT trials may be necessary, but not sufficient, for a complete evaluation of the effectiveness of SBMPs. These studies address whether, ceteris paribus, SBMPs may improve executive function. However, given the inherent complexity of school systems, this information may not benefit individual schools in deciding between a mindfulness approach and alternatives. The high variance seen between studies regarding SBMP efficacy is hypothesised to be mediated by population and intervention characteristics. Meta-regression analyses will be used to explore the potential influence of specific variables on the main treatment effect.

The moderators investigated include the categorical variables of outcome measure modality (i.e. PBT or BRS), the nature of the control condition, the culture group, and the role of the intervention administrator. The following continuous variables will also be explored: mean age of the sample, total duration of intervention exposure, and the quality and relevance of the study, as measured through the WoE D rating. To address the potential for predictors to exhibit multi-collinearity, the "*chart.correlation*" function of the *PerformanceAnalytics* package (Peterson & Carl, 2020) has been used to visualise the correlation matrix between continuous variables.

Meta-regression. Moderator analysis methods have been selected based on the model that is deemed most appropriate to account for the unit-of-analysis problem. Meta-regression models will be used for function-specific effect sizes (i.e. attentional control, cognitive flexibility, global EF, inhibition, and working memory)

that have used aggregation to correct for lack of case independence. Here, the moderators, or covariates, act as fixed effects which moderate the effect sizes derived from the random effects meta-analysis model. Their value as a predictor of effect size can be measured by the percentage of variation between studies that can be explained by the model (R^2). For continuous predictors that are either significant (i.e. $p<0.05$) or approach significance, bubble plots are presented to visualise the meta-regression results. These show the estimated regression slope by plotting the SMD of each study against the covariate (e.g. mean age).

CHE modelling. Correlated and Hierarchical Effects (CHE) models (Pustejovsky & Tipton, 2021) have been used for analyses where multilevel modelling has been used. One issue with the 3-level model used above is that it assumes that the sampling errors of effect sizes within a study are independent. This assumption should be revised since the effect sizes are derived from the same sample. Using a CHE model accounts for the likelihood that within-study effect sizes based on the same sample will have some degree of sampling error correlation. A large correlation coefficient ($\rho=0.6$) is assumed for the CHE models used in this meta-analysis. This enables the calculation of a variance-covariance matrix for each study using the "*impute_covariance_matrix*" function of the *clubSandwich* statistical package (Pustejovsky, 2022), which can be used to generate variance-covariance matrices. These can then be fitted to generate CHE models using the covariates of interest.

Robust Variance Estimation (RVE) was added to the CHE models to further increase the results' robustness. RVE approaches provide more stringent estimates of variance and standard error in clustered data (i.e. separate outcomes nested within a single study) (Pustejovsky, 2022). This decreases the risk that moderator analyses overestimate the influence of covariates. These robust estimates were calculated using the "*coef_test*" function, which was also from the *clubSandwich* package.

3.3.2.6 Publication bias

Although the meta-analysis would ideally collate all relevant empirical data on SBMPs, only published data is included in this study. The process by which studies meet the criteria for publication introduces the potential for bias because studies that report large effect sizes or unexpected positive results are more likely to be published (Fanelli et al., 2012). This publication bias has been found in school psychology journals (McClain et al., 2021), and the resulting inflated effect sizes can lead to over-enthusiastic uptake of educational interventions.

This study will use visual analysis of funnel plot graphs and statistical Egger's tests. Effect sizes are plotted on the x-axis in a funnel plot and measured against precision (standard error) on the y-axis. Assuming an

ideal dataset, larger studies (i.e. lower standard error) would cluster around the mean effect size near the top of the funnel. Smaller studies will exhibit a greater range of effect sizes due to higher standard error. Therefore, results should fan out symmetrically from the mean estimate as standard error increases. This creates a pyramidal or inverted funnel. Any asymmetry when actual results are plotted against the ideal funnel model would indicate potential publication bias.

As a second measure, Egger's tests are also calculated to quantify potential publication biases. This statistical method uses regressions of standardised effect size estimates against their precisions. Although this approach is potentially less sensitive to publication bias than funnel plot analysis, using a statistical approach compliments the more subjective visual analysis involved in funnel plot analysis. Using multiple measures of publication bias increases the robustness of findings, particularly in the fields relevant to this review, where multiple measures have only been observed in 27% of school psychology meta-analyses (McClain et al., 2021).

3.3.2.7 Power analysis

Many earlier studies exploring the impact of school-based mindfulness programmes have been criticised for lacking the statistical power to detect small-moderate-sized effects (Sims et al., 2022). Borenstein et al. (2011) estimate that the median number of studies typically included in a Cochrane systematic review is six. This review, which includes 69 effect sizes from 27 studies, greatly exceeds this median estimate. As the composite analysis uses the entire dataset, the meta-analysis will likely have sufficient statistical power.

However, as most analyses in this review focus on a specific cognitive function associated with EF, they depend on subsets of the overall data. In some of these subsets, there may be an insufficient number of independent cases (i.e. studies) to detect small-moderate effect sizes. Given the diffuse nature of SBMPs and other variables that impact executive functioning, small effect sizes are anticipated. Therefore, the function-specific analyses that include fewer independent cases may lead to Type II error (β), where the null hypothesis is confirmed despite a true effect.

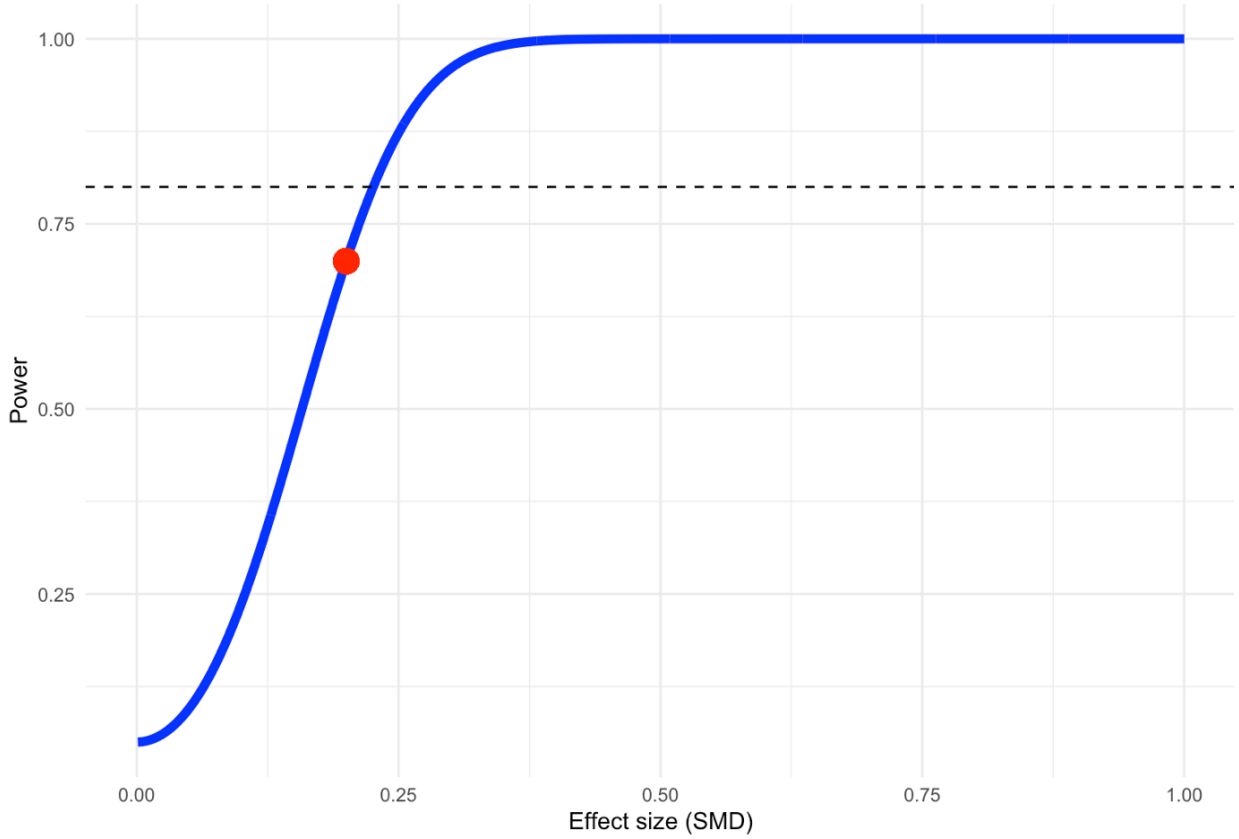
Conducting a power analysis is therefore justified to explore the minimum number of studies needed to detect small-moderate effect sizes. A random-effects model has been employed since moderate heterogeneity between effect sizes is expected. Using guidelines from Hedges and Piggott (2001), moderate heterogeneity has been calculated as:

$$\sigma_{\theta}^{*} = \sqrt{1.67 \times \frac{\sigma_{\theta}^{*}}{K}}$$

The overall mean sample size for experimental conditions (n=52.2) and control conditions (n=51.3) have been used to conduct power analyses through the *power.analysis* function of the *d.metar* package (v.0.1.0, Harrer et al., 2019), assuming a meta-analysis that includes ten independent cases (see Figure 5).

Figure 5

A plot showing the statistical power of a meta-analysis with data from 10 studies.



Note. The red dot indicates the power to detect small effect sizes (g=0.2) when k=10. This is below the conventional threshold of 80% (where β=0.2, and power is defined as 1-β), suggesting susceptibility to Type II error for small effects. The graph indicates robust power for effect sizes equal to or greater than 0.23.

4. Results

4.1 Narrative synthesis

The narrative synthesis provides a descriptive overview of the similarities and differences found between the included studies. It also presents a summary overview of the field, using examples to illustrate key commonalities and differences. Appendix C provides a more comprehensive summary of population and intervention characteristics, and outcome measures and results are in Appendix K.

4.1.1 Participants

3,308 CYP participated in the studies included in this review (M=110.3), with data from 2826 CYP contributing to the meta-analysis. Individual sample sizes ranged from 21 (Berti & Cigala, 2020) to 584 (Crooks et al., 2020). Ten studies had a sample size less than 50; 12 studies had a sample size between 50 and 100; and eight had sample sizes that exceeded 100.

4.1.1.1 Age

A wide range of ages is represented across the included studies. The youngest mean age reported by an included study was 3.75 years (Wood et al., 2017), with 16.0 years as the eldest (Frank et al., 2021). 25.9% of studies included in the meta-analysis focused on children in the Early Years Foundational Stage (EYFS) of the National Curriculum. 11.1% of studies involved Key Stage 1 (KS1) children, and 40.7% involved Key Stage 2 (KS2) children. Primary-school-aged children represent 51.8% of the total meta-analysis data population. 18.5% of studies investigated secondary-aged young people. This is split between four that investigated predominately Key Stage 3 students (KS3) and one study focused on Key Stage 4 (KS4) students.

4.1.1.2 Gender

Overall, studies reported relatively balanced gender ratios, with the total percentage of females in each study ranging from 34.8% (Lam & Seiden, 2020) to 83.3% (Quach et al., 2016). Almost all studies reported no significant differences in gender ratios between experimental and control conditions. Janz et al. (2019) reported a significantly greater ratio of girls in the control group (67% girls) compared to the intervention group (38% girls). Conversely, Quach et al. (2016) reported significantly fewer girls in the control group (52.8%) compared to the intervention group (83.3%). The comparability across conditions was reduced in

the two studies reporting significant gender differences between experimental and control groups, resulting in a lower WoE A grading.

4.1.1.3 Socio-economic status (SES)

Several studies focused on schools serving lower SES communities (e.g. Baena-Extremera et al., 2021; Crooks et al., 2020; Janz et al., 2019; Thierry et al., 2016; Zelazo et al., 2018). In one study (Andreu et al., 2023), a more targeted approach specifically recruited children "in a high-risk context". At the other end of the spectrum, some studies reported many participants as having parents with advanced degrees (e.g. Berti & Cigala, 2020; Shlomov et al., 2023; Wood et al., 2017). Studies predominately inferred socioeconomic status through demographic data on household income and the highest level of parental educational attainment.

4.1.1.4 Ethnicity

Overall, there was a high degree of variance in the ethnicity data reporting. Studies conducted in regions classified as English-speaking (e.g. United States, Canada, Australia) were significantly more likely to report on ethnicity, perhaps reflecting greater population heterogeneity.

4.1.2 Settings

Fifteen countries are represented in the 30 studies included in this review. The range of educational settings used to deliver SBMPs broadly follows the range of ages of the participants.

4.1.2.1 Regions

The regions represented in this review are located across five of the six habitable continents. In order of frequency, countries represented by more than one study in the meta-analysis (K=27) are the United States (US; k=7); Spain (k=4); Canada (k=2); Germany (k=2); and Italy (k=2). The remaining countries, represented by one study each, are Australia, Brazil, Chile, Finland, Hong Kong, Hungary, Israel, Portugal, Thailand, and the United Kingdom (UK).

Culturally, the studies were organised into the following four broad categories: English-speaking, or Anglophone, countries (Australia, Canada, UK, and US); Mediterranean countries (Israel, Italy, Portugal, and Spain); Latin American countries (Brazil and Chile); Northern and Central European countries (Finland, Germany, and Hungary); and Southeast and East Asian countries (Hong Kong and Thailand).

4.1.2.2 Educational settings

The interventions were delivered in a preschool or kindergarten setting in eight studies. Fifteen studies were based in primary schools, and seven were in secondary schools. Many studies reported on the relative population density of the schools in which the studies took place, with a range of urban (e.g. Lassander et al., 2020), rural (e.g. Ricarte et al., 2015), and mixed (e.g. Baena-Extremera et al., 2021) settings described. Most educational settings were defined as mainstream and state-funded, although academically selective settings (e.g. Wimmer et al., 2016) and university-affiliated schools (e.g. Wood et al., 2017) are also represented.

Almost all studies conducted the interventions in regular classrooms. A notable exception is the protocol Berti and Cigala (2020) used, in which a wooden hut was designed to create a calming environment for mindfulness practice. In two further studies (Andreu et al., 2023; Zelazo et al., 2018), classrooms were split into small groups for intervention delivery.

4.1.3 Research Designs

The exclusion of studies that did not employ a control condition during the screening process has led to broadly similar research designs used by the 30 included studies. Most studies (n=20) employed a quasi-experimental design with non-random assignment of participants to intervention and control groups. In these studies, allocation to groups was based on various methods, including matching, volunteer interest, and convenience sampling. Most of the remaining ten studies used a cluster-randomization method, randomly assigning whole classes to the intervention or control group. Full randomisation was possible for some studies where the interventions were conducted in small groups (e.g. Zelazo et al., 2018). Although inconsistently reported across the studies, randomisation methods included coin flips, random number generators, and hat draws.

4.1.3.1 Conditions

All studies included at least two conditions, including an SBMP condition and a control condition. Eighteen studies employed a passive control condition. In 14 of these studies, young people either engaged in regular school activities (treatment-as-usual, TAU) or received no additional intervention. In four studies, a waitlist control condition was used. These studies received a reduced rating on the "research design" criteria category of the WoE B rating. The remaining 12 studies used active control conditions that included social-emotional learning (SEL) programmes (e.g. Crooks et al., 2020; Magalhaes et al., 2022; Schonert-Reichl et

al., 2015), relaxation or yoga interventions (Baena-Extremera et al., 2021; Quach et al., 2016), or other educational and physical activities (e.g. Shlomov et al., 2023; Wimmer et al., 2016).

Five studies employed more than two conditions. Folch et al. (2021) included two experimental conditions, one administered by a trained mindfulness teacher and the other by an untrained teacher. Four studies included both active and passive control conditions. For the meta-analysis, the passive waitlist was selected as the comparison condition in two cases due to potentially confounding similarities between SBMPs and Hatha Yoga (Quach et al., 2016) and concentration training (Wimmer et al., 2016). In one case (Andreu et al., 2023), the active condition was selected as the skills-for-life programme was deemed to have little conceptual overlap with mindfulness practices. Baena-Extremera included three experimental conditions for each of the three different school settings used in the study. These three experimental groups were aggregated from descriptive statistics into a single comparison condition for the meta-analysis.

4.1.3.2 Data collection points

All studies in this review included at least one pre-test and one post-test data collection point. The timing of the post-test data collection point varied, with most studies collecting data immediately after the intervention (e.g. Flook et al., 2010; Lam & Seiden et al., 2020) or soon after. In some studies (e.g. Magalhaes et al., 2022; 2022; Muller et al., 2021), the latency period between the end of the intervention and the post-test assessment was not reported. Eight studies included a follow-up assessment to assess whether changes in executive function had been sustained over time. Follow-up assessments were administered between 4-6 weeks (Zelazo et al., 2018) and six months (Lassander et al., 2020; Salmoirago-Blotcher et al., 2019) after the post-intervention assessment.

4.1.4 Interventions

As expected, considerable heterogeneity existed between the mindfulness interventions employed by the included studies. In just under half of the interventions (n=14), the researchers developed a custom mindfulness programme for the study. The remaining 40% of studies (n=12) used an established mindfulness intervention or some adaptation thereof. These adaptations were often justified because content needed to be more developmentally appropriate. (e.g. Andreu et al., 2023). The most common adaptations were to MBSR programmes (e.g. Quach et al., 2016; Wimmer et al., 2016). The most popular established programmes included the manualised mindfulness-based SEL programme “MindUP” (Crooks et al., 2020; Schonert-Reichl et al., 2015; Thierry et al., 2016); the “.b” and “paws.b” programmes

(Lassander et al., 2020; Thomas & Atkinson, 2016; Vickery & Dorjee, 2016); and the "Learning 2 Breathe" programme (Frank et al., 2021; Lam & Seiden, 2020).

4.1.4.1 Intervention components

Intervention components across studies exhibited some clear similarities, as well as a high degree of diversity. Nearly all interventions involved teaching and practising methods designed to cultivate awareness of the present moment. This involved the near-universal inclusion of breathing exercises, where awareness is focused on the breath. Activities that draw awareness to the physical body, such as "body scanning", were also included in almost all intervention protocols. Frequently, activities designed to develop awareness of other senses were employed. These included mindful eating exercises (e.g. Baena-Extremera et al., 2021) and mindful listening (e.g. Frank et al., 2021).

There was a broad attempt to adapt intervention components to the developmental needs of the target population. Studies on earlier childhood typically incorporated play-based elements and were considerably more interactive than later years' components. Developing knowledge of abstract concepts such as bodily awareness was frequently supported using concrete objects and visuals. Story-based approaches using fictional characters were also used to help younger children develop their mindfulness knowledge base. There was also an explicit acknowledgement of the need to support younger children's self-regulation during mindfulness interventions by incorporating physical movement activities, including child-friendly yoga practices (Brann et al., 2022).

Interventions for CYP in middle childhood typically followed a mixed approach, combining play-based activities and introducing more formal mindfulness practices, such as extended meditation. Although less common than in early years interventions, creative use of physical objects was also employed to support children's understanding of abstract concepts. These included the "hula-hoops", "gratitude balls", and "generosity buttons" used by Milare et al. (2021). Group discussions on thoughts, feelings, and the brain were more frequent. In some instances, researchers felt that some mindfulness activities used in these interventions were not fully developmentally appropriate. For example, Wimmer et al. (2016) noted that the young people in their study (M=10.8 years) struggled to self-regulate during a total body scan activity.

Interventions for older children and adolescents were typically more formal. These included practical elements such as seated meditation practices and body scans. Psychoeducation was typically a significant

component of the intervention for this age group and included the discussion of more complex concepts, such as the relationship between thought patterns and stress.

Despite commonalities, the interventions included in this review were notable as much for their heterogeneity. In some cases, the components included diverged considerably from what might be considered core mindfulness practices. Interventions that incorporated elements of cognitive-behavioural therapy (CBT) approaches were common. Other studies incorporated practices from different fields such as positive psychology (e.g. Schonert-Reichl et al., 2015; Crooks et al., 2020). The MindUP curriculum explicitly includes gratitude and kindness to others, as seen in many custom-designed SBMPs. Homework components, which encouraged participants to practice mindfulness outside the classroom, were expressly mentioned by just under half of the included studies.

4.1.4.2 Theoretical Foundation

The included studies draw on various theories to operationalise mindfulness, although some broad trends emerge. Firstly, most studies drew on a secularised approach to mindfulness practice grounded in cognitive psychology, pioneered by MBSR (Kabat-Zinn, 1990). Only two of the included studies (Baena-Extremera, 2021; Lertladaluck et al., 2021) explicitly referenced the Buddhist roots of mindfulness practices. This is unsurprising given that a focus on executive function orients this review towards a more cognitive conceptualisation of mindfulness.

Again, as might be expected, many of the included studies have taken a developmental perspective, often using this approach to inform the choice of intervention and how it has been adapted to meet the developmental needs of the target population. This was especially evident in earlier childhood studies, such as the theoretical grounding in the Iterative Reprocessing Model (Zelazo, 2015) used in two studies to explain how mindfulness can support self-regulation by developing top-down control processes critical to EF.

The most frequently cited conceptualisation of mindfulness was Bishop et al.'s (2004) two-factor model. Many studies explicitly referenced how their interventions promoted the regulation of attention to the present moment and cultivated an open, non-judgemental orientation to experience. Alternative theoretical frameworks used by studies in this review generally aligned with the essential characteristics of the two-factor model, whilst also including additional elements. For example, Lertladaluck et al. (2021) used

a three-factor model (Shapiro et al., 2006), which builds on the two-factor model by adding an "intention" component to highlight the importance of individual motivation to engage in mindfulness practice.

4.1.4.3 Administrators

86.7% of the included studies reported the role of the administrator. Four studies did not explicitly state the role of the administrator (Flook et al. 2010; Makmee et al. 2022; Ricarte et al. 2015; Thomas & Atkinson, 2016), although in three cases (Flook et al., 2010; Ricarte et al., 2015; Thomas & Atkinson, 2016), the administrator could be assumed as the study adopted an established SBMP that is teacher-administered.

Thirteen studies used teaching staff as the sole administrators of the intervention. Teacher training was highly variable, ranging from a 9-week mindfulness course involving four hours of mindfulness-based learning and practice a week (Baena-Extremera et al., 2021) to half-day training (Janz et al., 2019). Training conducted over weeks was coded as "extensive" (k=4) compared with "brief" training ranging from ½ a day to two days (k=6). In some cases, the training regimen for school staff was not reported and could not be inferred from established programme practices (e.g. Wimmer et al., 2017). As the role of the administrator was selected as a moderator variable for this review, these studies scored lower on the WoE B rating.

An external instructor administered fourteen studies. Of these, eight used professional mindfulness instructors to deliver the training. Research assistants were used in three studies, and in one study, an experienced psychologist and a mindfulness instructor were used (Andreu et al., 2023). In two studies, the intervention administrator was also an author of the study (Lam & Seiden et al., 2020; Lertladaluck et al., 2021).

4.1.4.4 Dosage

Overall, the mean total intervention dosage was 496.3 minutes (SD=239.6). Total dosage was highly variable across studies, ranging from 120 minutes (Muller et al., 2020) to 1200 minutes (Wimmer et al., 2016). The overall duration of the trial phase was similarly highly variable, ranging from two weeks (Muller et al., 2020) to an entire school year (Crooks et al., 2020), as was the average length of individual sessions. However, they typically lasted between 15 and 60 minutes. An overview is presented in Appendix L.

4.1.5 Fidelity of Implementation

60% of studies provided sufficient detail to support the replicability of the intervention by either using manualised programmes or supplementary materials. Measures of adherence to the intervention protocol

were reported in 40% of the studies included in the review. Only five studies (16.7%) formally assessed the acceptability of the SBMP from the perspective of CYP, teachers, and parents. Two other studies reported the informal assessment of intervention acceptability.

4.1.5.1 Replicability

Across all studies, intervention replicability was variable. Studies using external, manualised programmes well-established as SBMPs were typically considered more replicable than studies based on novel intervention approaches. These established programmes included Learning to Breathe (L2B); dot be (.b) and its younger years variant, paws.b, and the MindUP curriculum. Although not all studies using these programmes reported the interventions according to TIDIER guidelines (Hoffman et al., 2014), sufficient information on the intervention protocol can easily be found through the associated programmes' websites. In studies using SBMPs adapted from other mindfulness programmes, these adaptations were typically made to make them appropriate for the target age group. However, the exact details of these adaptations were not reported in several studies (e.g. Andreu et al., 2023).

Studies that developed entirely novel programmes required detailed reporting of intervention characteristics to facilitate replicability, and some studies in this category provided sufficient detail, including supplementary material outlining the protocol in detail and examples of the materials used (e.g. Flook et al. 2010). This was done according to TIDIER guidelines (Hoffman et al., 2014) in some cases, such as Baena-Extremera et al. (2021), which provided a detailed description of the custom SBMP based on "Breathe through this: Mindfulness for parents and teenagers" (Snel et al., 2015) and included pre-recording audio tracks which further facilitated replication. However, in other cases, the level of detail in which novel interventions were described was insufficient. For example, Makmee et al. (2022) did not report the role of the administrator despite designing a novel SBMP for their study. Two further studies reported insufficient data, such as total intervention dosage (e.g. Janz et al., 2019).

4.1.5.2 Adherence

Adherence to intervention protocols was inconsistently reported across all studies (40%). Studies providing relevant data generally reported high levels of adherence, such as teacher reports that 98% of critical intervention components were delivered as planned (Wood et al., 2020). Brann et al. (2022) also used teacher reports, finding some issues with engagement, although this was not further elaborated.

Independent coders were used in some studies. For example, Frank et al. (2021) found that fidelity was maintained across 78.6% of sessions, with 92% inter-rater reliability. In other studies, fidelity measures were described, but the results were not reported. Salmoirago-Blotcher et al. (2019), who audio-recorded sessions for review, and Magalhaes et al. (2022), who used independent observers to monitor implementation fidelity, describe the protocol but do not report the results.

Where data on adherence to homework elements was shared, degrees of practice were varied. Vickery and Dorjee (2016) report that 21.2% of participants in their study never practised mindfulness outside of school, and 39.4% rarely practised.

4.1.5.3 Acceptability

Most studies (76.7%, k=23) did not report how much participants enjoyed the SBMP and felt it was helpful. The studies that did report acceptability generally reported positive feedback from participants. Andreu et al. (2023) report a mean rating of 8.45/10 for participant experience of the SBMP. High acceptability was also found in Vickery and Dorjee (M=7.9), finding that 76% of participants liked or extremely liked the intervention. In Lam and Seiden (2020), the feedback was more equivocal, with 52.8% of the participants viewing the programme as useful. In other studies (e.g. Janz et al., 2019; Ricarte et al., 2015), teacher and parent feedback were used to measure acceptability.

4.1.6 Outcome Measures

In total, 69 outcome measures of executive function from 27 studies were included in the meta-analysis. Seven studies included only a single measure of EF-related processes. Six studies included a mix of PBT and BRS measures, receiving higher WoE B ratings. A summary of the 69 outcome measures and their effect sizes can be found in Appendix K.

4.1.6.1 Behaviour rating scales

42.0% of outcomes (n=29, k=14) included behaviour rating scales (BRSs) as an indirect measurement of EF. Of these, behaviour ratings scored by teachers (n=15) represented 51.7% of all BRS data. Parent reports were used for five outcome measures (k=3). The use of this approach appeared to coincide with the age of participants in the study, with two studies using children in the EYFS key stage (Brann et al., 2023; Zelazo et al., 2018) and one study involving children in the KS2 group (Flook et al., 2021). For the remaining BRS outcomes, self-reports from the CYP participating in the studies were used (n=7, k=3).

The most common BRSs were from the Behaviour Rating Inventory of Executive Function (BRIEF) family. This included both the first edition (Gioia et al., 2000), the second edition (BRIEF-2; Gioia et al., 2015), and the preschool version (BRIEF-P; Gioia et al., 1996). In total, 12 effect sizes were derived from BRIEF tools, representing 60% of all BRS outcome measures. All studies used well-established behaviour rating scales standardised against CYP populations, save for the researcher-developed Likert scales used in Wood et al. (2017), for which this study received lower WoE A and WoE B ratings.

In some cases, BRSs were selected based on their standardisation against a specific age group. The use of the Children's Behaviour Questionnaire (CBQ, Rothbart et al., 2001), a parent-scored report designed to provide a profile of children's temperament aged 3-7, was used by one study whose participants had a mean age of 4.77 years (Zelazo et al., 2018). Similarly, the Early Adolescent Temperament Questionnaire-Revised (EATQ-R; Ellis & Rothbart, 2001), designed for children aged 9-15, was used in a study where the mean age is 9.8 years (Andreu et al., 2023). The use of the "Attention" subscale of the Youth Self-Report (YSR) by Lam and Seiden (2020) also specifically targets the study's sample population. A description of the BRS measures used by the included studies is presented in Appendix M.

Some BRS outcome measures had a more diagnostic function (YSR, CTRS-R, SDQ, SENA, V-TRS) and broadly focused on inattention and hyperactivity. In these cases, scales focused on inattentive behaviours were categorised as "Attention", and those focused on hyperactivity were categorised under "Inhibition".

4.1.6.2 Problem-based Tasks

The remaining 58.0% of outcome measures (n=40, k=19) involved direct measurement of ECPs through performance-based tasks (PBTs). An overview is presented in Appendix N.

PBT measures of inhibition were the most frequently cited by studies included in the meta-analysis, followed by cognitive flexibility, working memory, attention, and global EF, respectively. The use of comprehensive PBT measures of global executive function was almost non-existent, with only the MEFS measure used by Zelazo et al. (2018) providing a holistic measure of global EF through multiple tasks designed to assess composite functioning. This contrasts with BRS measures like the BRIEF family of assessments, which are structured to provide a global executive composite score.

The literature base has been consulted where there is considerable ambiguity regarding the categorisation of PBTs. In the case of the HTKS Task, McClelland et al. (2022) make a convincing case that the task is a measure of general executive functioning through experimental comparison with PBTs that have an

established “function specificity”, such as the Dimensional Change Card Sort (DCCS) task as a measure of cognitive flexibility. In all cases, the same cognitive task has been categorised under the same cognitive function group. Therefore, the HTKS task has been classified as a “Global EF” measure in all the cases in which it is used, regardless of whether this aligns closely with the theoretical perspective of the study (as in Zelazo et al., 2018) or not (as in Brann et al., 2023).

4.1.6.3 Excluded outcome measures

Some studies reported on outcome measures unrelated to executive function, which were automatically excluded from the meta-analysis. In other cases, the lack of EF-specificity of some outcome measures warranted their exclusion from the meta-analysis. For example, Folch et al. (2021) used the Rey-Osterreith Task to measure global executive functioning, although this multidimensional neuropsychological measure is not explicitly designed to measure EF. It is often used with adult populations as a dynamic assessment of planning and goal-directed behaviour, but the multiple confounding variables involved in the complexity of this task make it problematic as a measure of EF, particularly for CYP.

Several studies presented multiple effect sizes from the same outcome measure. For example, Frank et al. (2021) reported both the proportion of hits on the EFNBACK task and the proportion of false alarms. While both data can measure working memory, “hits” data is likely to be a more direct measure of working memory. “False alarm” rates refer to the number of incorrect responses given, which can be due to a lack of inhibitory control. In these instances, the score best aligned with the target ECP was selected for the meta-analysis.

4.2 Composite meta-analysis

The composite meta-analysis explored the full dataset, including the effect sizes of all 69 outcomes, calculated from descriptive data provided by the 27 studies. These are presented in Appendix K. This meta-analysis examined the heterogeneity in the data and conducted various sensitivity analyses to identify potential outliers and compare different models accounting for the unit-of-analysis problem. The section also presents the results of CHE models used to explore moderator variables.

4.2.1 Heterogeneity analyses

An initial random-effects model was conducted under the false assumption of independence between all 69 outcome measures (see Table 6).

Table 6

A summary of pooled effect size and heterogeneity for all 69 outcomes.

Model	g	95% CI	p	95% PI	I^2	95% CI
Independent outcomes	0.33	0.25, 0.41	<0.0001	-0.13, 0.79	53.2%	38.5%, 64.4%

Note. I^2 is the proportion of total variation in observed effect sizes not due to sampling error.

Between-study variance estimated as τ^2=0.05, 95% CI [0.024, 0.11]. τ^2 was significantly greater than zero, indicating that the data contains moderate-large between-study heterogeneity. This supported the use of a random effects model. The lower bound of the 95% prediction interval was -0.13, indicating that negative intervention effects cannot be ruled out for future studies.

The I^2 value (53.2%, 95% CI [38.5%, 64.4%]) suggests the presence of moderate-large heterogeneity in the data, which is supported by the value of Cochran's Q=145.36, which is significant (p<0.0001). As I^2 is greater than 50%, further investigation of potential outliers and influential cases was warranted.

4.2.1.1 Outliers

The find.outliers function detected ten outliers. When excluding these studies, I^2 is reduced from 55.4% to 7.3%, and the confidence interval around τ^2 (0.01) now includes 0 (95% CI [0.00, 0.03]). Cochran's Q statistic (Q=61.36, p=0.36) is no longer significant, which has resulted in the narrowing of the prediction interval (95% CI [0.09, 0.52] of the pooled effect size (see Table 7).

Table 7

Comparison of heterogeneity between all outcomes and with outliers removed.

Analysis	g	95% CI	p	95% PI	I^2	95% CI
All 69 outcomes	0.33	0.25-0.41	<0.0001	-0.13-0.79	53.2%	38.5%- 64.4%
Outliers Removed[1]	0.31	0.25-0.37	<0.0001	0.09-0.52	5.5%	0.00%-30.8%

[1] Removed: Baena-Extremera d2 Test (total effectiveness); Berti & Cigala Go/NoGo Task; Frank N-back Task (hits); Janz SDQ (hyper/attention); Lam & Seiden BRIEF-2 Shift; Lassander NEPSY-II Inhibition A; Schonert-Reichl Heart & Flowers Task; Shlomov Bear & Lion Task; Suarez-Garcia SENA-Hyperactivity; Wimmer Stroop Inference Test (RT).

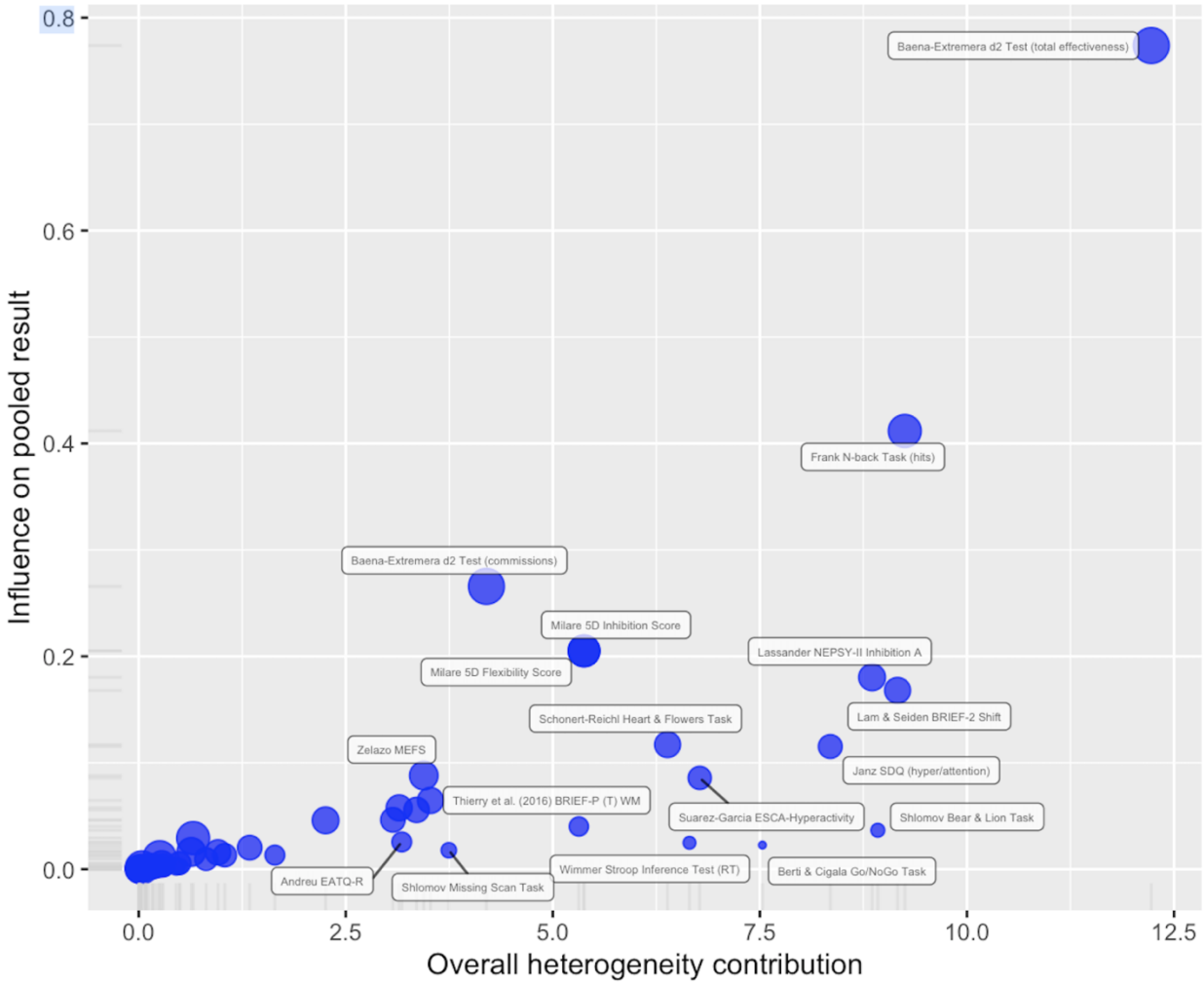
Removing outliers significantly reduces I^2 to 5.5%, 95% CI [00.0%, 30.8%], whilst minimally impacting the pooled effect size. The loss of 14.5% (n=10) of outcome measures was considered an over-correction, especially as a reasonably high degree of heterogeneity between outcome measures is expected due to the complexity of SBMPs.

4.2.1.2 Influential Cases

A Baujat plot derived from an influence analysis on all 69 outcome measures identified several influential cases considered for removal (see Figure 6).

Figure 6

Baujat plot of the influence of all 69 outcome measures on the pooled effect size.



Note. Overall heterogeneity contribution is measured using Cochran's Q. Influence on pooled results is

calculated using the leave-one-out method. The plot identifies six potential cases: Baena-Extremera d2 Test (total effectiveness); Frank N-back Task (hits); Lassander NEPSY-II Inhibition A; Lam & Seiden BRIEF-2 Shift; Jans SDQ (hyperactivity); and Shlomov Bear & Lion Task.

Baena-Extremera d2 Total Effectiveness contributed the most heterogeneity (Q≈12.25) and exerted the greatest influence on the pooled result. Due to the study's large sample size (N=320) and high methodological quality rating using the Gersten et al. (2005) protocol, this result was not removed. For similar reasons, two other high heterogeneity contributors, Lam & Seiden BRIEF-2 Shift and Lassander NEPSY-II Inhibition A, were not removed.

The second most influential case, Frank EFN-back Task (hits), was considered for removal due to the potentially confounding differences between the Emotional Faces N-back Task (EFN-back) and the traditional N-back Task. This task is designed to measure both emotional regulation and working memory by using images of emotional faces to stimulate an emotional reaction to the task.

Janz et al.'s (2019) SDQ outcome was also considered for removal. Eight outcome measures used such diagnostic tools, although only two were high contributors to heterogeneity. To maintain consistency and conceptual transparency, all eight "diagnostic" outcome measures were excluded in the "diagnostic measures removed" sensitivity analysis.

A comparison of sensitivity analyses suggests that the pooled effect size is minimally influenced by potential outliers and influential cases (see Table 8).

Table 8

Comparison of pooled effect and heterogeneity

Analysis	n	g	95% CI	p	95% PI	I^2	95% CI
Independent outcomes	69	0.33	0.25, 0.41	<0.0001	-0.13, 0.79	53.2%	38.5%, 64.4%
Outliers removed*	59	0.31	0.25, 0.37	<0.0001	0.09, 0.52	5.5%	0.00%, 30.8%
Diagnostic measures removed**	61	0.31	0.22, 0.40	<0.0001	-0.15, 0.77	53.2%	37.3%, 65.0%
Influential case removed***	68	0.34	0.26, 0.42	<0.0001	-0.11, 0.78	50.5%	34.4%, 62.6%

Note. *Baena-Extremera d2 Test (total effectiveness); Frank N-back Task (hits); Janz SDQ (hyper/attention); Lam & Seiden BRIEF-2 Shift; Lassander NEPSY-II Inhibition A; Schonert-Reichl Heart & Flowers Task; Suarez-Garcia ESCA-Hyperactivity; Wimmer Stroop Inference Test (RT); Berti & Cigala Go/NoGo Task; Shlomov Bear & Lion Task. ** Crescentini et al. (2016) CTRS-R Cog/Inattention; Janz SDQ (hyper/attention); Lam & Seiden YSR Attention; Magalhaes Vanderbilt TRS (attention); Suarez-Garcia SENA-Attention; Suarez-Garcia SENA-Hyperactivity; Frank DERS Goal; Frank DERS Impulse; Lam & Seiden DERS Goal. ***Frank EFN-back Task (hits).

Removing Frank EFN-back (hits) has a negligible effect on the overall pooled effect size (g=-.34). Whilst I^2 is reduced, there is still considerable variance (i.e. >50%) between effect sizes that is not attributable to sampling error.

4.2.2 Unit of analysis

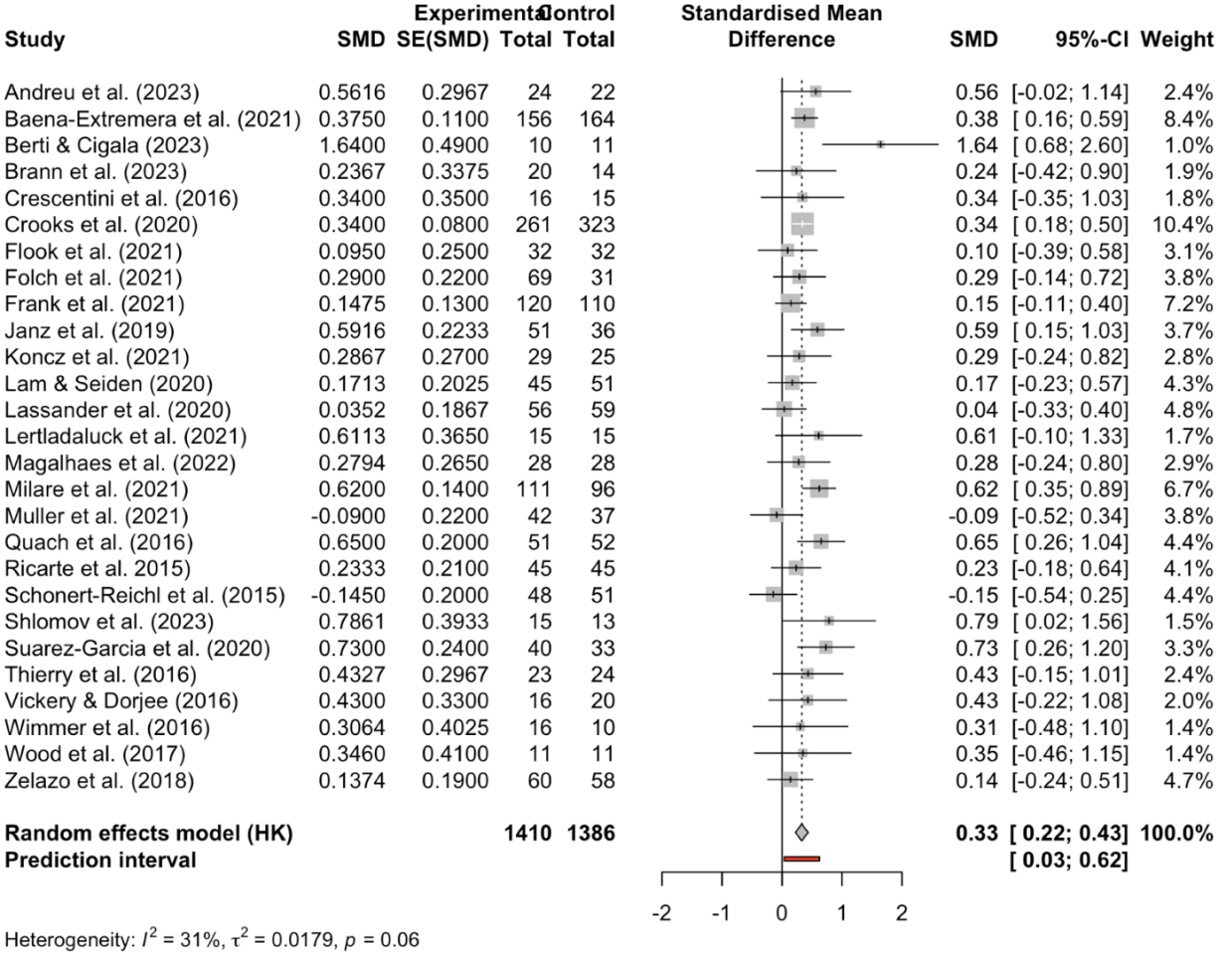
Both aggregation and multi-level models were trialled to account for the lack of independence between all effect sizes, given that many studies reported multiple outcome measures.

4.2.2.1 Aggregation

The 69 outcome measures included in the meta-analysis were aggregated into 27 studies included in the meta-analysis using the aggregate function of the metafor package (v.4.4-0., Viechtbauer, 2010), with the within-study correlation estimated as $\rho = 0.6$ (see Figure 7).

Figure 7

Forest plot showing aggregated effect sizes for every study included in the meta-analysis.



Note. Aggregated effect sizes have been calculated assuming a moderate-large correlation (ρ=0.6) between effect sizes nested within the same study.

Using this aggregation method, the composite effect size (g=0.33, p<0.0001) did not differ from the overall effect size when all outcome measures were treated as independent. Sensitivity analyses adjusting the assumed correlation to $\rho = 0.4$ has a negligible effect on the overall pooled effect (See Table 9).

Table 9

Comparison of pooled effect and heterogeneity

Analysis	n	g	95% CI	p	95% PI	I^2	95% CI
Independent outcomes	69	0.33	0.25, 0.41	<0.0001	-0.1, -0.79	53.2%	38.5%, 64.4%

Aggregated ($\rho = 0.6$)	27	0.33	0.22, 0.43	<0.0001	0.03, 0.62	31.1%	0.0%, 57.0%
Aggregated ($\rho = 0.4$)	27	0.33	0.22, 0.44	<0.0001	0.03, 0.63	32.5%	0.0%, 57.8%

Note. I^2 is the proportion of total variation in observed effect sizes not due to sampling error.

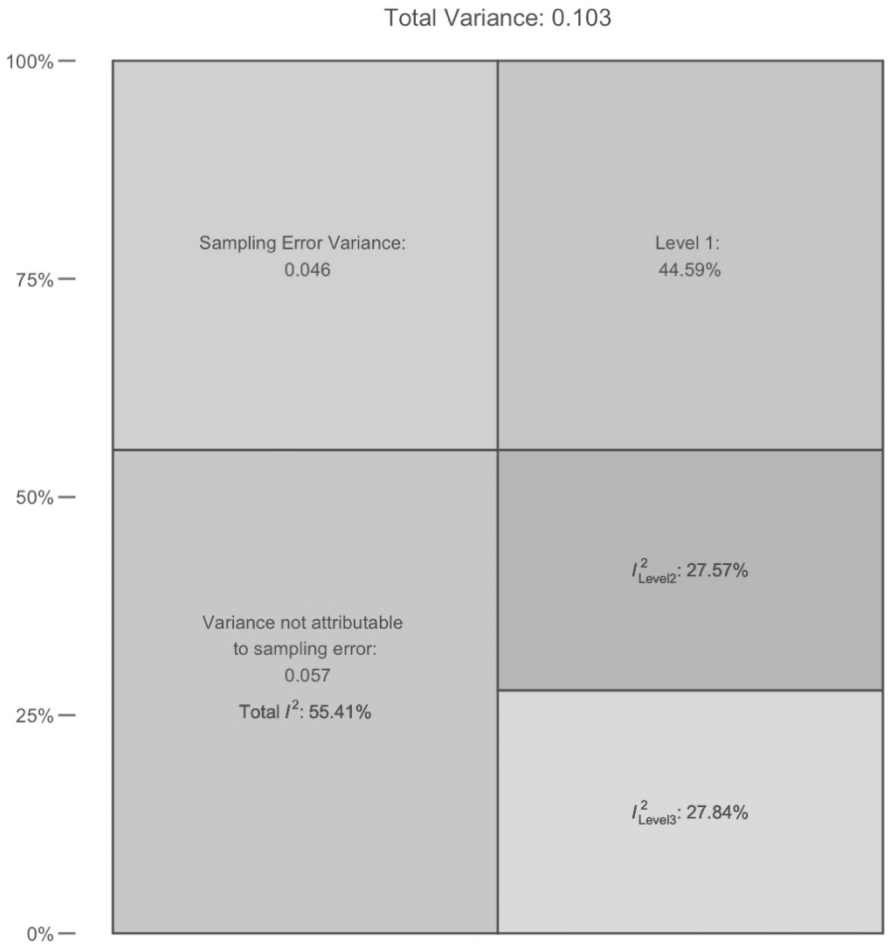
The confidence interval around τ^2 (0.02; 95%CI [0.00, 0.10]) now includes zero, increasing the robustness of the 95% prediction interval [0.03, 0.62]. The total variance not due to sampling error (I^2=31.1%) indicates no need to investigate potential outliers based on a priori conditions.

4.2.2.2 Three-level modelling

The pooled effect correlation using the three-level meta-analytic model was r=0.33 (95%CI: 0.24, 0.45; p<.0001). The proportion of total variance attributable to within-study differences between effect sizes (27.57%) was approximately the same as the total variance explained by between-study differences (27.84%) (see Figure 8).

Figure 8

Distribution of total variance across the three levels in a multivariate meta-analysis model.



ANOVA results indicated that the three-level model did not provide a significantly better fit than the two-level model (X2/1=3.57; p=0.059) (see Table 10).

Table 10

Results of ANOVA comparing three-level and two-level models.

Analysis	Correlation (r)	AIC	BIC	LRT	p
Three-level model	0.33	54.11	60.76		
Two-level model	0.32	55.68	60.11	3.57	0.059

Note. Correlations are converted from z estimates. AIC=Akaike Information Criterion; BIC=Bayesian Information Criterion; LRT=Likelihood Ratio Test.

4.2.3 Composite moderator analyses

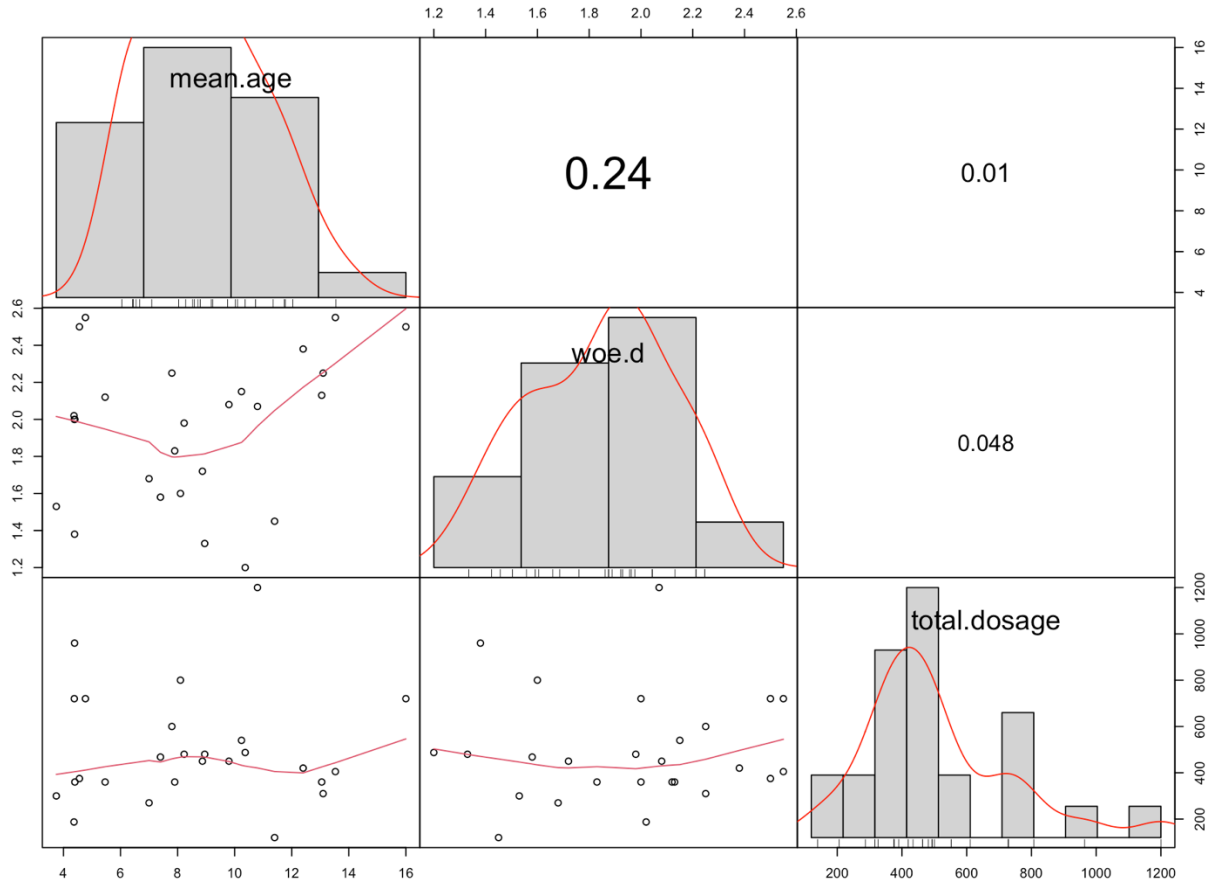
Moderator analyses were explored on the full dataset to explore the relationship between variables of interest and the effectiveness of SBMPs to improve EF and ECPs.

4.2.3.1 Multicollinearity

To assess whether multi-collinearity might exist between continuous variables, correlation matrices were explored between WoE D scores, mean age of participants, and total intervention dosage (see Figure 9).

Figure 9

Visualisations of the correlation matrices between potential moderator covariates



There was little overall correlation between all three continuous variables (r=0.01), indicating that the three continuous variables were not different faces of a larger construct. There was a small correlation (r=0.24) between participant age and the study's quality/relevance, which did not indicate a concerning degree of multicollinearity. This analysis suggests that all three continuous variables can be used in subsequent meta-regressions.

4.2.3.2 Correlated and hierarchical effects modelling

As the aggregation method involved combining effect sizes from potentially dissimilar outcome measures (e.g. a measure of working memory with a measure of inhibition), initial moderator analyses were conducted using three-level modelling (see Table 11).

Table 11

Correlated and Hierarchical Effects (CHE) models including all 69 outcomes.

Covariate	estimate	SE	z-value	95% CI	p-value
Participant age	-0.024	0.013	-1.80	-0.05, 0.002	0.066.
Total dosage	0.0001	0.0002	0.52	-0.00036, 0.00062	0.56
Total sessions	0.0033	0.0032		-0.0046, 0.011	0.34
Mean session duration	-0.0019	0.0013		-0.005, 0.0013	0.20
Intervention weeks	0.0012	0.003		-0.0097, 0.012	0.72
WoE D	-0.23	0.10	-2.26	-0.45, -0.0008	0.049*
Sample size	-0.0002	0.00028		-0.0012, 0.0008	0.59

Note. P-values use robust variance estimation (RVE).

To test whether SBMPs have a differential effect on ECPs, CHE models with robust variance estimation (RVE) were fitted. Attention-related outcomes were omitted due to a potential lack of equivalency.

Additionally, subgroup analyses were conducted on the aggregated dataset to preliminarily explore the between-study heterogeneity across categorical data. These are presented in Appendix O.

As many studies reported on multiple ECPs, exploring this covariate would not be viable using the aggregated dataset. Therefore, a three-level model was fitted for the subgroup analysis exploring this category. Attentional outcomes were removed from this analysis due to a potential lack of equivalency and theoretical overlap (see Table 12).

Table 12

Subgroup analysis exploring separate ECPs using a three-level model.

ECP	estimate	SE	p-value	$p_{subgroup}$
All 69 outcomes				0.13
Flexibility	0.20	0.09	0.034*	
Global EF	0.40	0.12	0.09.	
Inhibition	0.39	0.10	0.07.	
Working Memory	0.45	0.12	0.029*	
Outlier removed				0.018
Flexibility	0.21	0.08	0.037*	

Global EF	0.37	0.10	0.12
Inhibition	0.35	0.09	0.12
Working Memory	0.56	0.11	0.002*

Note. Outlier removed: Janz et al. (2015) EFN-back (hits).

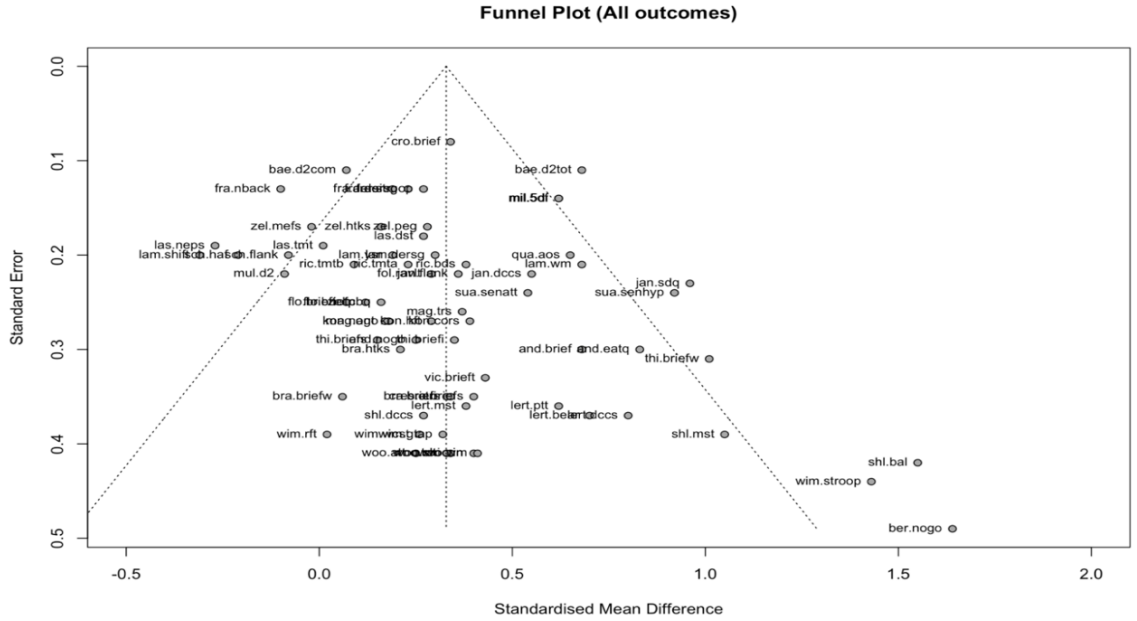
The regression matrix indicates a significant difference between working memory and flexibility outcomes (p=0.029), although ECP was a non-significant moderator overall (F=1.98; p=0.13). This same analysis performed with the Frank EFN-back outlier removed found that ECP was a significant moderator coefficient (F=3.63; p=0.018).

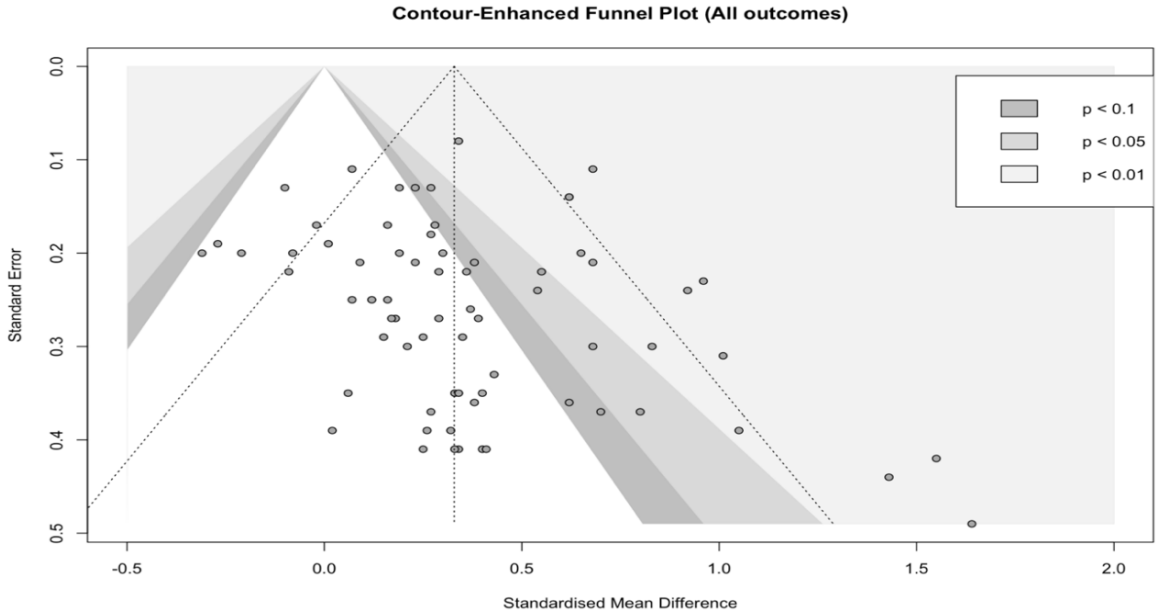
4.2.4 Publication bias

Overall, publication bias was assessed by exploring potential bias at the level of individual outcome measures and individual studies. Visual analysis of the funnel plot suggested a degree of publication bias (see Figure 10).

Figure 10

A funnel plot and contour-enhanced funnel plot of all outcome measures



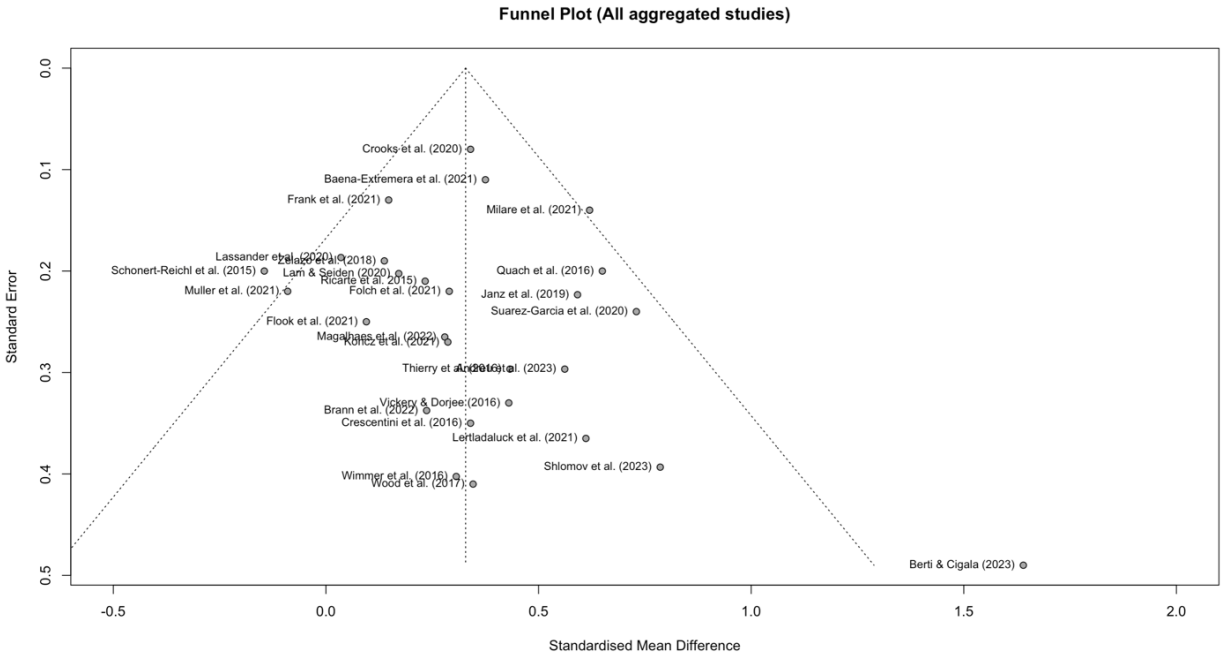


Note. The vertical dotted line represents the pooled SMD. The contour funnel plot is shaded to represent levels of significance. Cases in dark grey represent results approaching significance ($p<0.1$); medium grey represents significant results ($p<0.5$); and light grey represents highly significant results ($p<0.01$).

The funnel plot highlights a cluster of three outcome measures (Shlomov Bear and Lion Task; Wimmer Stroop Inference Test (RT); Berti & Cigala Go/NoGo Task) that exhibit both significant effect sizes ($p<0.01$) and large standard errors. For comparison, funnel plots for effect sizes aggregated within individual studies were also reviewed (see Figure 11).

Figure 11

Funnel plot showing the SMD and SE of aggregated outcomes from all 27 studies



Overall, the funnel plot did not indicate asymmetry, except for Berti and Cigala (2023), which contributes the largest aggregated effect size and standard error. Egger's tests support this lack of asymmetry (see Table 13).

Table 13

Egger's test to indicate the presence publication bias across all outcomes and studies.

	intercept	95% CI	t	p
Independent outcomes (n=69)	0.75	-0.08, -1.58	1.77	0.08
Aggregate analysis (k=27)	0.43	-0.61, -1.46	0.80	0.43

Note. Egger's tests on meta-analyses with less than ten studies may lack the statistical power to detect publication bias.

Results did not indicate the presence of publication bias, assuming an alpha level of 0.05. However, Egger's test results for independent outcomes (t=1.77, p=0.08) approached significance.

4.3 Differential ECP meta-analyses

As the primary focus of this review is on the potentially differential effect of SBMPs on EF-related cognitive processes (ECPs), meta-analysis and moderator analysis results have been organised according to cognitive function.

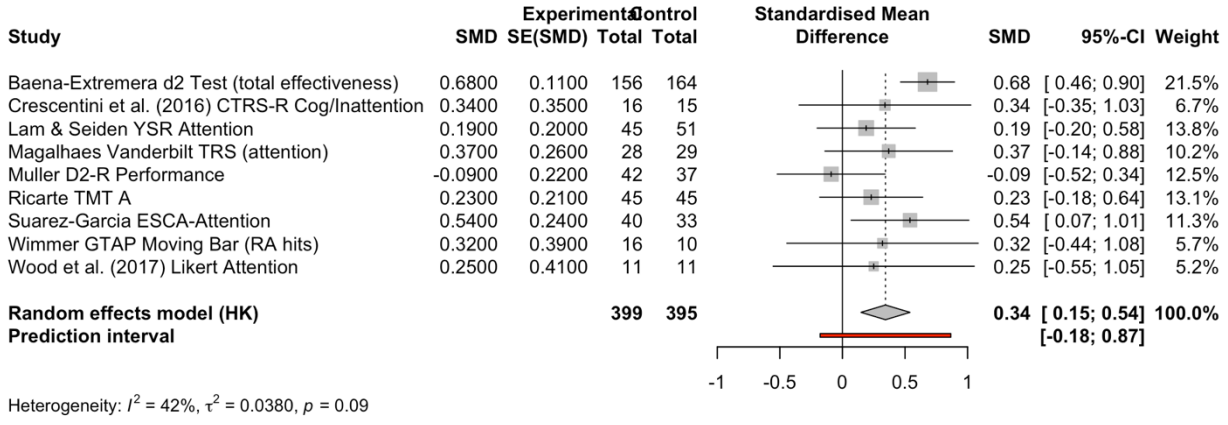
4.3.1 Attentional control

Nine outcome measures were coded as measures of attentional control. These were derived from nine separate studies and can thus be treated as independent effect sizes without further modelling to account for the unit-of-analysis problem. Outcome measures were split between BRSs (n=5) and PBTs (n=4). No participants from KS1 or KS4 are represented in the data. Interventions were delivered by external administrators in four studies, by teachers in three studies, and by a mix of both in two studies. Total intervention dosage ranged from 120 minutes to 1200 minutes, a tenfold difference.

The pooled effect size of all attentional control effect sizes was considered small-moderate (g=0.34, 95%CI [0.15, 0.54) and significant (p<0.004) (see Figure 12).

Figure 12

Forest plot of the effect sizes of attention-coded outcomes.



Heterogeneity is below the outlier and influential case threshold $(\mathrm{I}^2 = 42\%, \tau^2 = 0.038, \mathrm{p} = 0.09)$. The 95% prediction interval [-0.18, 0.87] passes through zero, indicating that negative effect sizes for future SBMP studies focusing on attention-related outcomes cannot be ruled out.

4.3.1.1 Moderator analyses

Due to the small number of studies included in this sub-analysis (k=9), no moderator analyses using categorical data were conducted on this dataset. As the conventional k≥10 rule is more flexible for meta-regressions using continuous data covariates (Borenstein et al., 2011), three were performed (see Table 14).

Table 14

Meta-regression analysis of continuous variables on aggregated flexibility effect sizes.

Covariate	estimate	SE	p-value	95% CI	R^2
Participant age*	0.01	0.03	0.71	-0.07, 0.09	0.00%
Total dosage	0.0002	0.0003	0.48	-0.0005, 0.001	0.00%
WoE D	0.32	0.23	0.21	-0.23, 0.86	21.1%

Note. R^2=unaccounted heterogeneity explained by covariate.

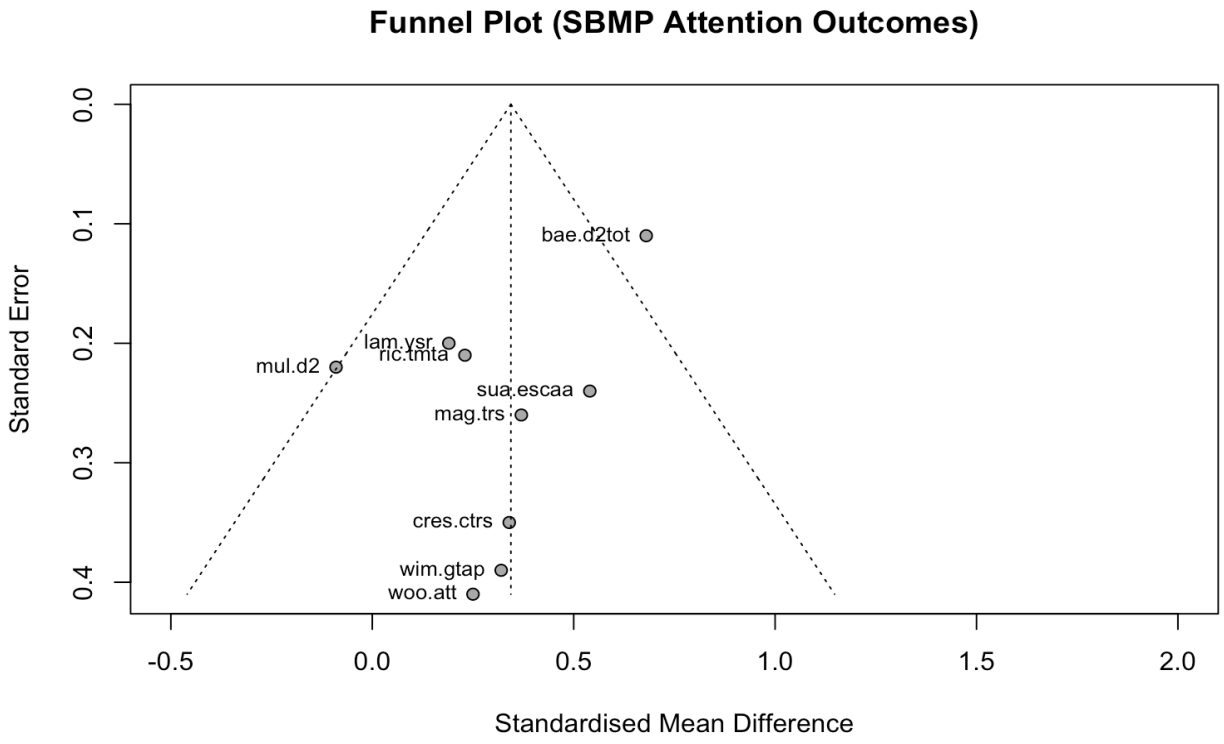
In all cases, the 95% CI ranges between zero, indicating that these are poor predictors of attentional outcomes. Two covariates, participant age and total dosage, are estimated to account for none of the between-study heterogeneity (R^2=0.00%). WoE D rating was not identified as a significant moderator (p=0.21), although it was estimated to account for approximately a fifth of the between-study heterogeneity.

4.3.1.2 Publication bias

Funnel plot analysis did not indicate publication bias (see Figure 13).

Figure 13

Funnel plot showing SMD of attentional control outcomes plotted against standard error.



4.3.2 Flexibility

Fifteen outcome measures were coded as measures of cognitive flexibility, derived from 14 separate studies. Outcome measures were predominately PBTs (n=11). Participants from KS4 were not represented in this dataset. Interventions were delivered by external administrators in 11 studies, by teachers in two studies, and a mix of both in the remaining study. Total intervention dosage ranged from 270 to 1200 minutes.

4.3.2.1 Model fitting

The two outcome measures derived from Wimmer et al. (2016), the Wisconsin Card Sorting Task (g=0.02) and the Reversible Figures Task (g=0.26) were aggregated. A comparison with an analysis treating all outcome measures as independent suggests this aggregation had no impact on the pooled effect size, indicating it may be an appropriate model for subsequent moderator analyses (see Table 15).

Table 15

Comparison of pooled effect and heterogeneity between models for flexibility outcomes

Analysis	g	95% CI	p	95% PI	I^2	95% CI	τ^2	95% CI
Independent outcomes	0.21	0.03, 0.37	0.02	-0.34, 0.75	48.0%	5.4%, 71.4%	0.06	0.001, 0.16
Aggregated Model	0.21	0.03, 0.40	0.03	-0.35, 0.78	51.3%	10.2%, 73.6%	0.06	0.003, 0.18

Note. τ^2=between-study variance in true effect sizes.

A three-level model was also fitted as a potential alternative to the aggregated model. An ANOVA was used to compare it with a model that treats all effect sizes as independent (see Table 16).

Table 16

Comparison of 2- and 3-level model fit for flexibility-related effect sizes.

Analysis	k/n	df	AIC	BIC	AICc	LRT	p
3-level Flexibility Model	14	3	14.69	16.61	17.09	n/a	n/a
2-level Flexibility Model	15	2	12.74	14.02	13.82	0.05	0.82

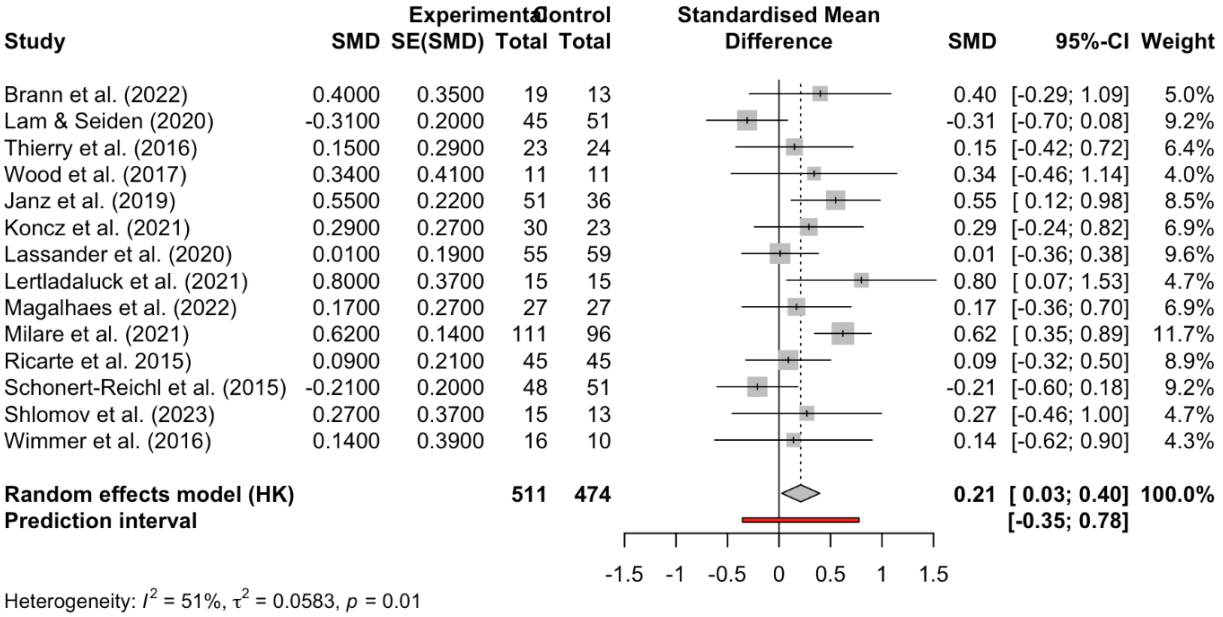
Note. AIC=Akaike Information Criterion; BIC=Bayesian Information Criterion; LRT=Likelihood Ratio Test.

The Akaike (AIC) and Bayesian (BIC) Information Criteria are higher for the 3-level model, suggesting that this is a worse fit for the data. Further indicating the relative paucity of the 3-level model, the likelihood ratio test (LRT) is non-significant (x2/1=0.05, p=0.82). The fitting of a three-level multivariate analysis model found that all the variance not attributable to sampling error (48.05%) was accounted for by between-study heterogeneity, and none was attributed to within-study heterogeneity.

Therefore, the aggregated model was deemed the most appropriate dataset to report the effects of SBMPs on flexibility-related outcomes and conduct subsequent meta-regressions. An overview of the results of aggregated effect sizes is provided through a forest plot (see Figure 14).

Figure 14

Forest plot of aggregated flexibility outcomes.



The pooled effect size of flexibility-related effect sizes was small (g=0.21), representing a smaller effect than seen in the composite analysis. The between-study heterogeneity not accounted for by sampling bias increased from 48.0% to 51.3%, crossing the threshold to consider potential outliers. No outliers were identified as deviating from 95% confidence parameters.

Influential case analyses indicated that one result, Milare 5D Flexibility Score, contributed the most overall heterogeneity (Q=8) and exerted the greatest influence on the pooled effect (around 2.15). This study received a low rating using the Gersten et al. (2005) protocol. However, the Five Digit Test (5D) flexibility index is an established measure of flexibility, with the Brazilian version used in this study reporting good-excellent internal consistency (a=0.88-0.90; Sedo et al., 2004). A comparison of the aggregated model before and after the removal of Milare 5D is presented in Table 17.

Table 17

Sensitivity analyses of different models for the effect of SBMPs on flexibility

Analysis	g	95% CI	p	95% PI	I^2	95% CI	τ^2	95% CI
Full aggregated model	0.21	0.03, 0.40	0.03*	-0.35, 0.78	51.3%	10.2%, 73.6%	0.06	0.003, 0.18

Influential case removed	0.15	-0.04, 0.33	0.11	-0.30, 0.59	27.0%	0.00%, 62.2%	0.03	0.00, 0.16

Note. τ^2=between-study variance in true effect sizes.

4.3.2.2 Moderator analyses

Although the number of studies exceeded ten (k=14), they were unevenly split between categories of interest (e.g., only two studies used teachers as an administrator; only four BRS outcomes). As such, only moderator analyses of continuous variables have been presented (see Table 18).

Table 18

Meta-regression analysis of continuous variables on aggregated flexibility effect sizes.

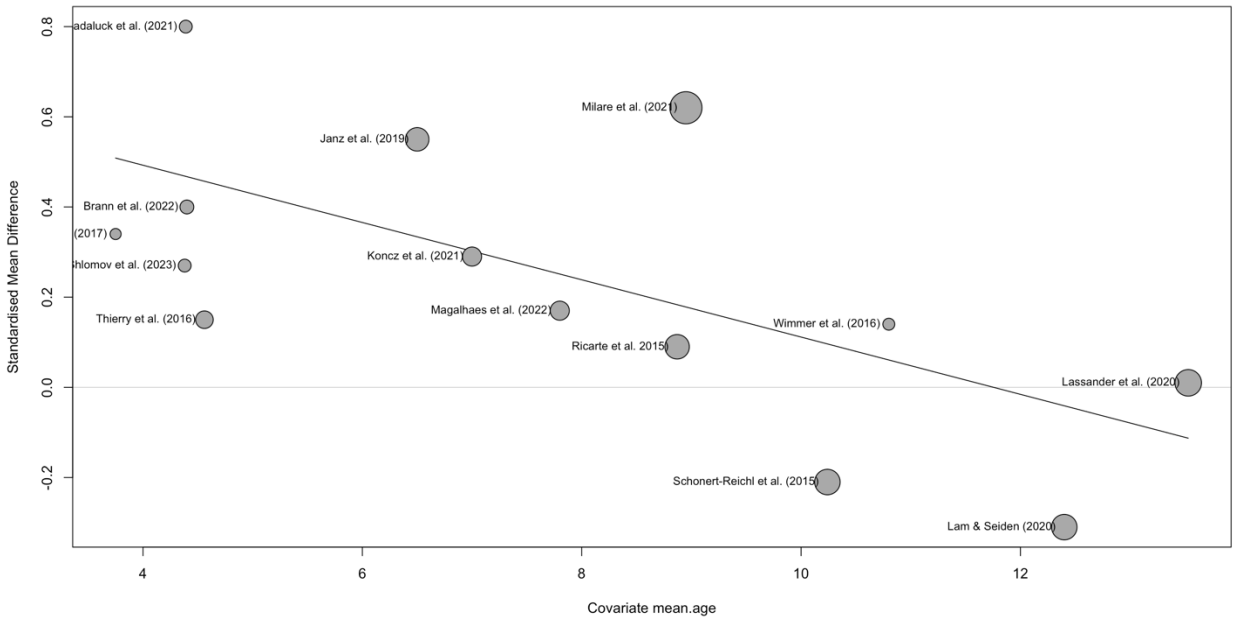
Covariate	estimate	SE	p-value	95% CI	R^2
Full aggregated model					
Mean age	-0.06	0.02	0.02*	-0.12, -0.01	38.2%
Total dosage	-0.00	0.0004	0.53	-0.0006, 0.0012	0.00%
WoE D	-0.55	0.16	0.005**	-0.88, -0.21	73.9%
Influential case removed					
Mean age	-0.07	0.018	0.004**	-0.11, -0.03	100%
Total dosage	0.0003	0.0004	0.41	-0.0005, 0.0012	0.00%
WoE D	-0.48	0.22	0.052·	-0.96, 0.004	37.87%

Note. Janz et al. (2019) did not report sufficient dosage data for this meta-regression.

Mean participant age (p=0.02) and WoE D (p=0.005) were significant covariates with effect size for flexibility outcomes. The regression slope for age as a moderator indicates an inverse relationship between participant age and SBMP effectiveness for flexibility-related outcomes (see Figure 15).

Figure 15

Bubble plot showing the correlation between flexibility effect sizes and mean age.

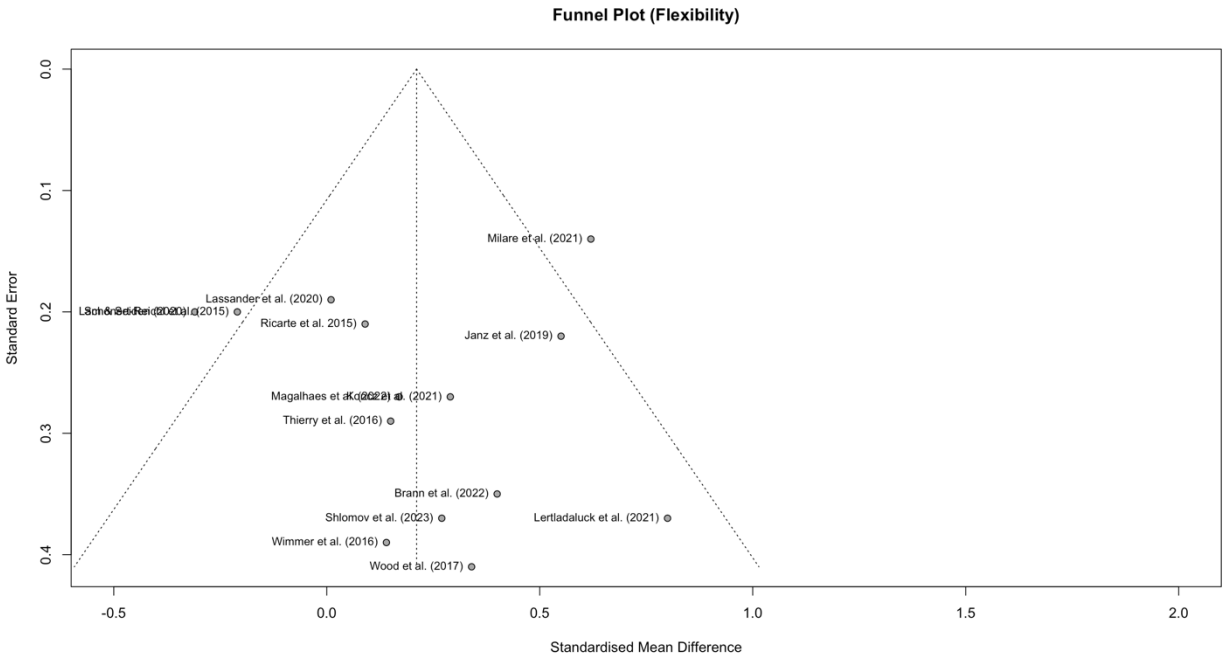


4.3.2.3 Publication bias

The contour-enhanced funnel plot identifies Milare et al. (2021) as the sole highly significant ($p<0.01$) result (see Figure 16).

Figure 16

Funnel plot showing of aggregated flexibility outcomes



Overall, the funnel plot does not appear to indicate asymmetry in the studies included in this review. This is corroborated by the non-significant (p=0.97) Egger's test (-0.05; 95% CI [-2.55, 2.46]).

4.3.3 Global EF

Thirteen effect sizes were recorded as measures of Global EF, derived from 10 studies (k=10, n=13). Ten effect sizes were based on BRS measures, with the remaining three derived from PBT measures. This global EF meta-analysis also represents all the Key Stages (EYSF-KS4) in the overall review. The total intervention dosage ranged from 187.5 to 480 minutes.

Initial sensitivity analyses assuming independence of all Global EF outcomes detected Janz et al. (2018) as a potential outlier. However, as the total variance not attributable to sampling error (I^2=37.9%) was below 50%, it was felt there was little justification for removal.

4.3.3.1 Aggregation

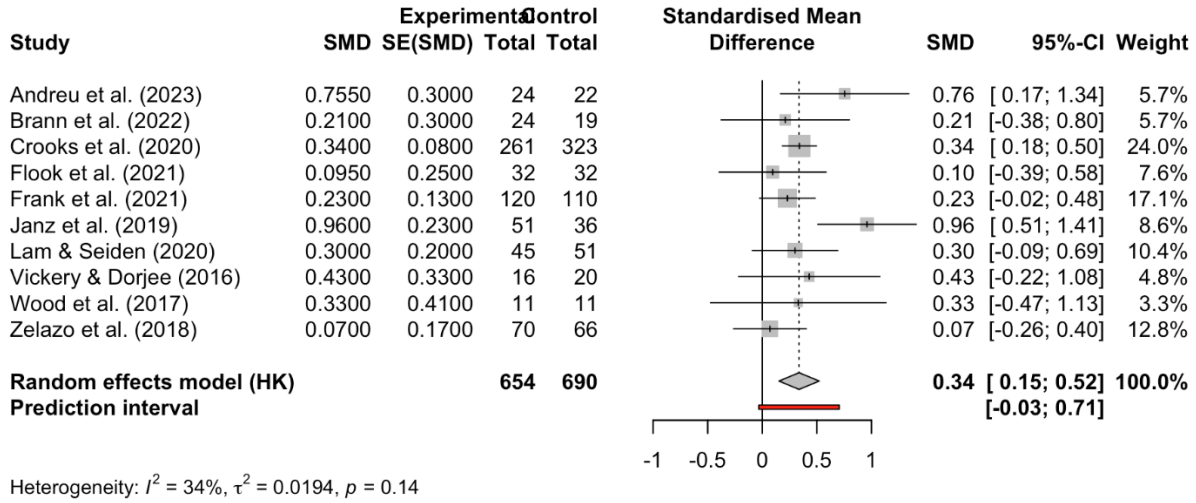
Three studies included multiple outcomes measuring Global EF. Flook et al. (2021) contributed data from both the teacher version of the BRIEF-2 (g=0.07) and the parent version (g=0.12). As both measures are derived from the same rating scale, and the effect sizes are similar there was justification for aggregating

these results. Andreu et al. (2023) contributed data from two behaviour rating scales, the EATQ-R (g=0.83) and the BRIEF-2 (T) (g=68). Both BRS measures capture Global EF and were considered sufficiently comparable for aggregation.

Aggregating the results of Zelazo et al. (2018) was more challenging. Although both are PBT measures, the MEFS (g=-0.02) is designed explicitly as an Early Years measure of global executive functioning. The HTKS task (g=0.16), as discussed in the narrative review, is considered both a measure of inhibition and a measure of Global EF. With these concerns in mind, an aggregated model indicated a small-moderate overall pooled effect (g=0.34) (see Figure 17).

Figure 17

Forest plot of aggregated outcomes measuring Global EF



4.3.3.2 Three-level modelling

A random-effects three-level meta-analysis model was fitted by defining the individual outcome measures as lower-order variables nested within the larger cluster of independent studies. The three-level model estimates the same pooled effect estimation (g=0.34) as the aggregated model (g=0.34) (see Table 19).

Table 19

Comparison of pooled effect and heterogeneity between models of Global EF.

Analysis	effect est.	95% CI	p	95% PI	I^2	95% CI	τ^2	95% CI

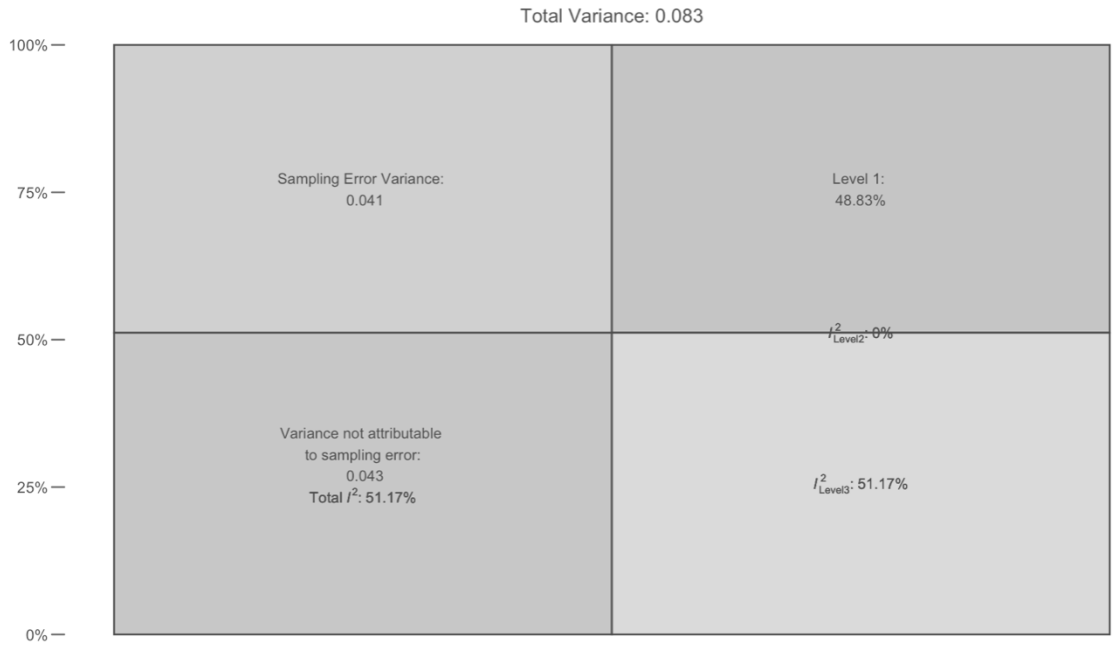
Three-level Global EF	0.32	0.15,	0.001	-0.09,	38.7%	0.0%,	0.029	0.00,
Model		0.48		0.73		67.8%		0.10
Aggregated Global EF	0.34	0.15,	0.003	-0.03,	33.9%	0.0%,	0.02	0.00,
Model		0.52		0.71		68.5%		0.21

Note. Effect est.=effect estimate. τ^2=between-study variance in true effect sizes.

A multilevel version of I^2 (Cheung, 2014) was calculated to determine the distribution of total variance across levels. The results of this analysis are summarised in Figure 18.

Figure 18

A diagram of total variance distribution across the three levels in a Global EF model.



Sampling error variance accounts for 48.8% of the total heterogeneity between Global EF effect sizes. The total heterogeneity variance explained by Level 2 is 0%. The greatest percentage of total variance (I^2=51.17) is attributed to Level 3, or between-study variance. An ANOVA was performed to compare the fit of two-level and three-level modelling to the Global EF dataset (see Table 20).

Table 20

Comparison of the fit of 2- and 3- level models for Global EF.

Analysis	df	AIC	BIC	AICc	LRT	p
3-level Global EF Model	3	8.96	10.41	11.96	n/a	n/a
2-level Global EF Model	2	9.04	10.01	10.37	2.08	0.15

Note. AIC=Akaike Information Criterion; BIC=Bayesian Information Criterion; LRT=Likelihood Ratio Test.

Small differences in the Akaike (AIC) and Bayesian (BIC) Information Criteria between the three-level and two-level models suggest that the two models are similar fits for the data. This is supported by the nonsignificant (p=0.15) likelihood ratio test (LRT).

As 0% of the total variance between effect sizes has been attributed to within-study differences, and as there is reasonable synergy between the multiple outcome measures used by each study (e.g. either both the same measure, both BRSs, or both PBTs), the aggregated model was considered a more parsimonious approach to account for the unit-of-analysis problem and was selected for subsequent moderator analyses.

4.3.3.3 Moderator analyses

The Global EF dataset may have the statistical power to detect moderator effects through meta-regression. A summary of key meta-regression results suggests that methodological quality and relevance, measured by WoE D scores, is not a significant predictor of SBMP effectiveness for Global EF outcomes (see Table 21).

Table 21

Meta-regression analysis of continuous variables on aggregated Global EF effect sizes.

Covariate	estimate	SE	p-value	95% CI	R^2
Participant age	-0.003	0.21	0.88	-0.05, 0.05	0.00%
Total dosage	-0.0003	0.0002	0.13	-0.0007, 0.0001	0.00%
WoE D	-0.32	0.30	0.31	-1.01, 0.37	37.24%

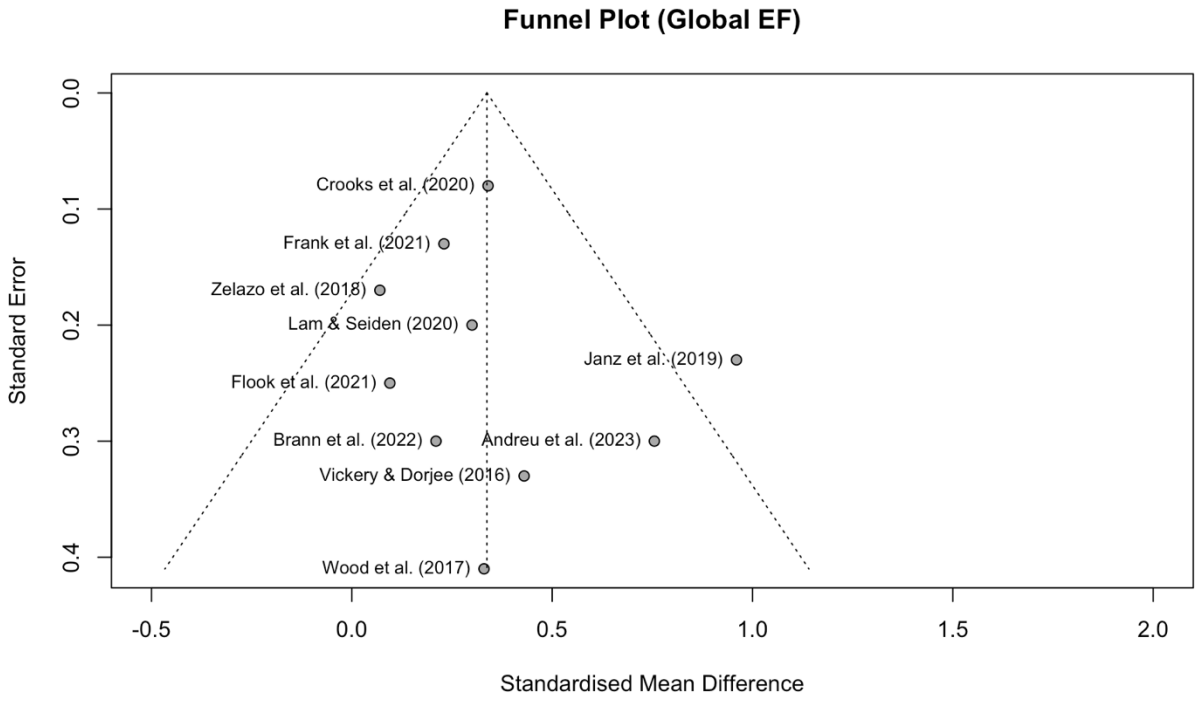
Note. Janz et al. (2019) did not sufficiently report on intervention dosage for these data to be included in the meta-analysis.

4.3.3.4 Publication bias

Overall, funnel plots did not indicate asymmetry, suggesting that publication bias is not significantly present in the data (see Figure 19).

Figure 19

Funnel plot showing SMD of global EF studies plotted against standard error.



Despite potential skewing from the Janz et al. (2018) effect size, Egger's test was non-significant (0.44; p=0.61).

4.3.4 Inhibition

Twenty effect sizes were categorised as loading onto "Inhibition", derived from 17 studies (k=17, n=20). BRS measures (n=6) represent 30% of this sub-analysis, with the remaining 70% of effect sizes based on PBT measures (n=14). This inhibition-specific meta-analysis also represents all the Key Stages (EYSF-KS4) in the review. In 10 studies, external administrators were used; six used class teachers, and one study used a mix of both. Total intervention dosage ranged from 270 to 1200 minutes.

4.3.4.1 Heterogeneity analyses

An exploratory meta-analysis treating all inhibition outcome measures as independent found a small-moderate significant effect (g=0.39, 95% CI [0.18, 0.60], p=0.001). The between-outcome heterogeneity not attributable to sampling error was moderate to large (I^2=65.6%). The initial "brute force" approach detected three potential outliers. Removing outliers reduced the overall effect size to 0.34, and I^2 to 46.5% (see Table 22).

Table 22

Comparison of pooled effect and heterogeneity between inhibition datasets.

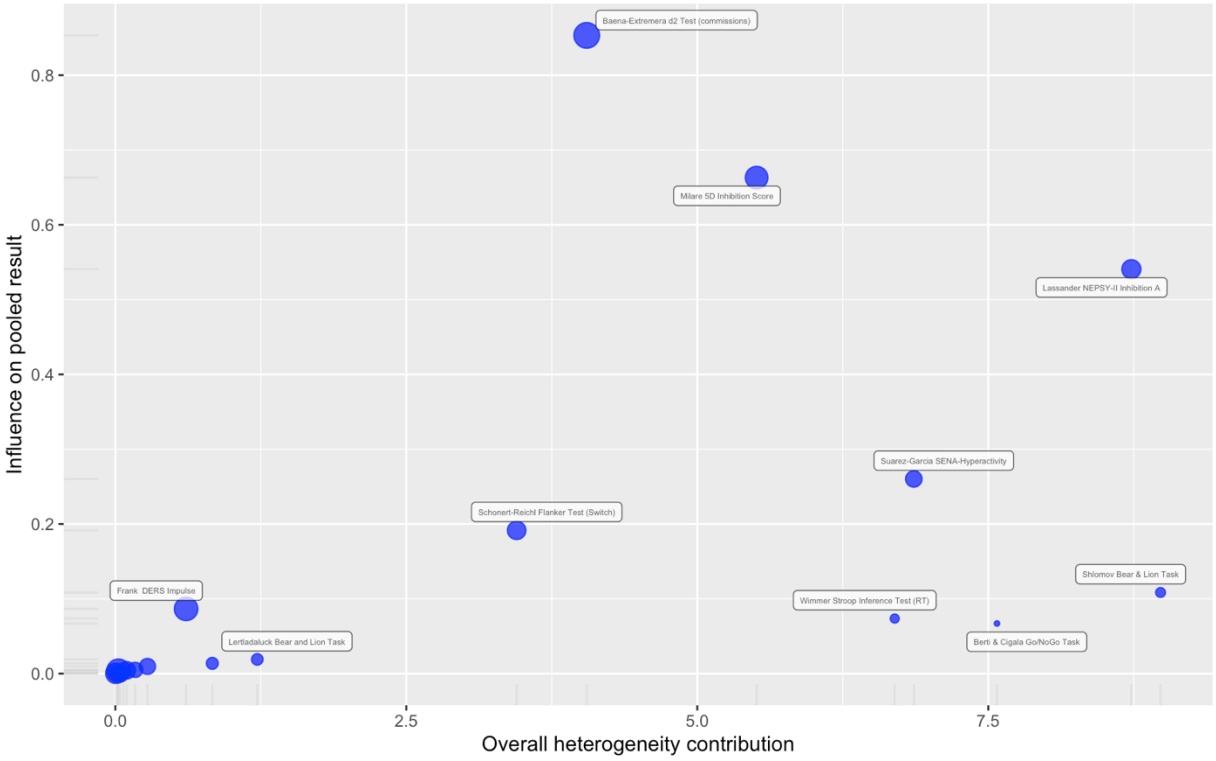
Analysis	n	g	95% CI	p	95% PI	I^2	95% CI	τ^2	95% CI
Independent Outcomes	20	0.39	0.18, 0.60	0.001	-0.30, 1.09	65.6%	44.8%, 78.6%	0.10	0.04, 0.44
Outliers removed	17	0.34	0.18, 0.49	0.0003	-0.08, 0.76	46.5%	5.9%, 69.6%	0.034	0.001; 0.19

Note. The three outliers removed are Berti & Cigala Go/NoGo Task, Lassander NEPSY-II Inhibition A, and Shlomov Bear & Lion Task.

A Baujat plot exploring the relative influence of inhibition-related outcome measures indicates that two of the previously identified outliers (Berti & Cigala Go/NoGo Task; Shlomov Bear & Lion Task) have little overall influence on the pooled result (see Figure 20).

Figure 20

Baujat plot of the influence of all inhibition-related outcomes on pooled effect size.



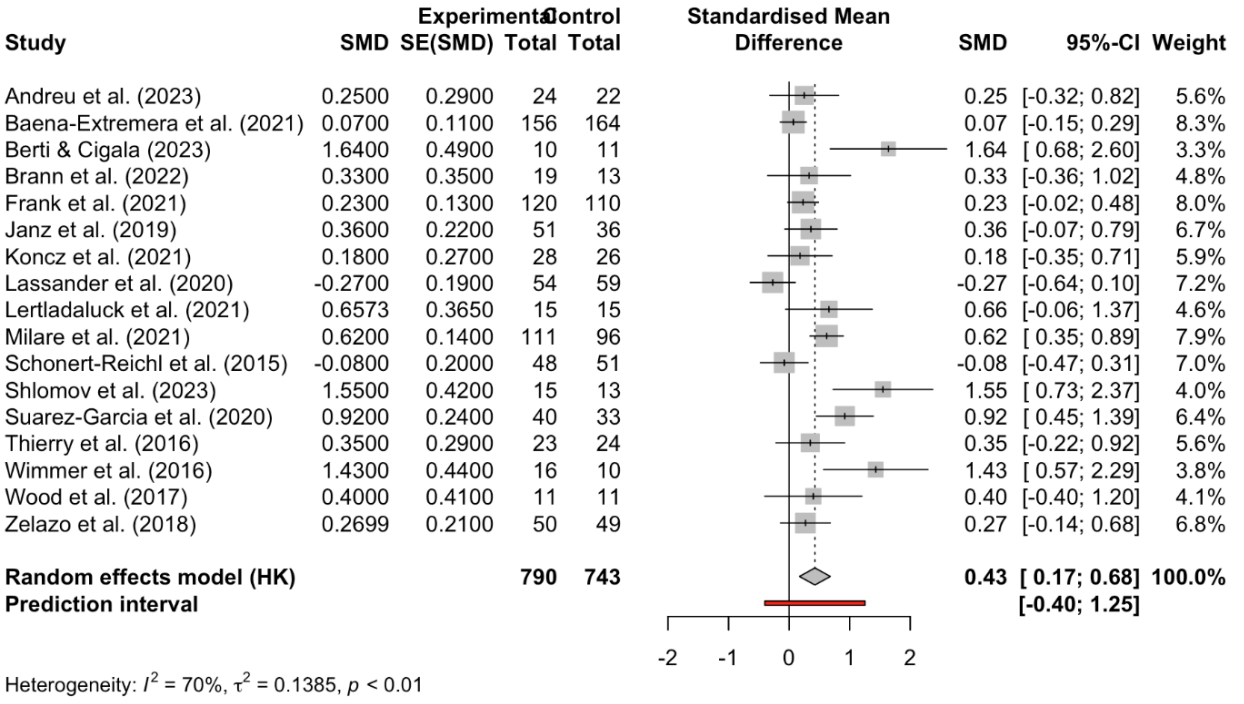
Lassander et al.'s (2020) effect size (g=-0.27) exerts more influence but was not removed due to the study's overall high methodological quality rating and the established validity of the NEPSY-II measure.

4.3.4.2 Unit of analysis

Three studies included multiple outcomes measuring inhibition. These outcomes were aggregated within their respective studies (see Figure 21).

Figure 21

Forest plot of aggregated outcomes measuring the effect of SBMPs on inhibition



Comparison with the meta-analysis assuming independence of all outcomes shows a small increase in pooled effect size. The percentage of total variance not attributable to sampling error increased from I^2=65.6% to 70.1% (see Table 23).

Table 23

Comparison of pooled effect and heterogeneity between Inhibition models.

Analysis	g	95% CI	p	95% PI	I^2	95% CI	τ^2	95% CI
Independent outcomes	0.39	0.18, 0.60	0.001	-0.30, 1.09	65.6%	44.8%, 78.6%	0.10	0.04, 0.44
Aggregated model	0.43	0.17, 0.68	0.003	-0.40, 1.25	70.1%	50.9%- 81.8%	0.37	0.23, 0.76

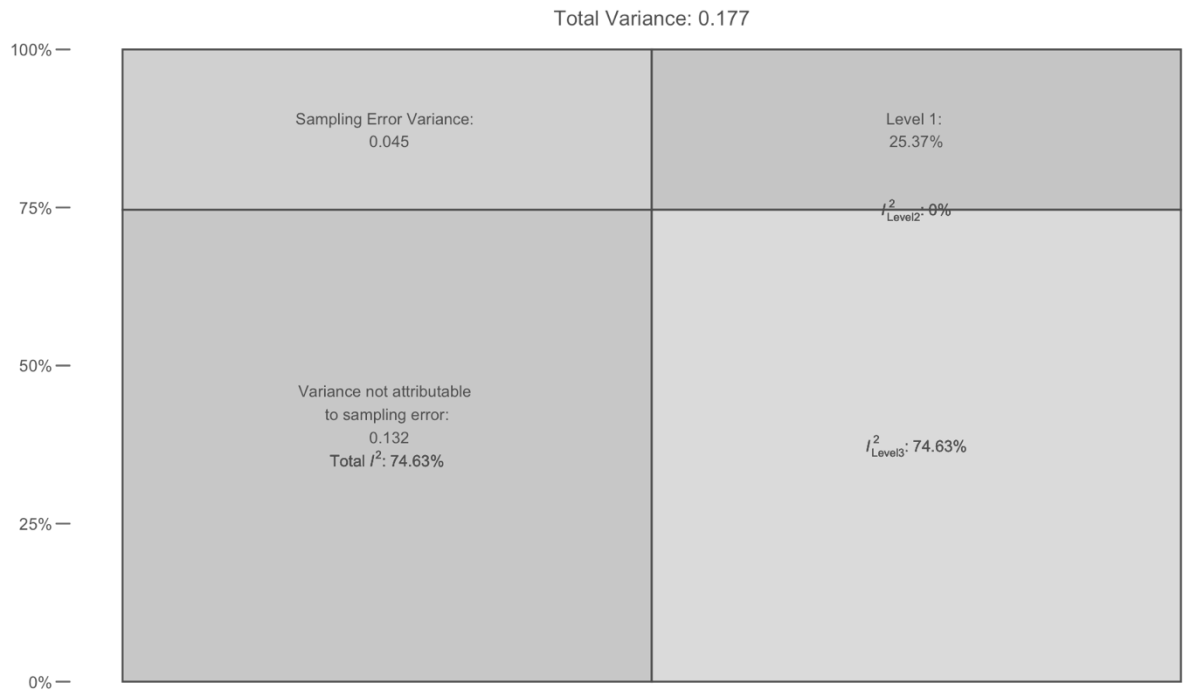
Note. τ^2=between-study variance in true effect sizes.

A random-effects three-level meta-analysis model was fitted to explore this heterogeneity in the dataset further. The three-level model estimated a slightly reduced pooled effect (z=0.34) to the independent outcomes model (g=0.39) and the aggregated model (g=0.43).

A multilevel allocation of I^2 (Cheung, 2014) was calculated to determine the distribution of total variance across levels (see Figure 22).

Figure 22

Diagram of total variance distribution across three levels in an inhibition model.



As with previous analyses, Level 2 explained 0% total heterogeneity variance. In other words, within-study variance accounts for 0% of the total heterogeneity, suggesting a significant degree of dependence between inhibition outcome measures derived from the same study. The greatest percentage of total variance (I^2=74.63) is attributed to Level 3, or between-study variance.

An ANOVA was performed to compare the fit of two-level and three-level modelling to the inhibition dataset (see Table 24).

Table 24

Comparison of the fit of 2- and 3- level models for inhibition outcomes.

Analysis	df	AIC	BIC	AICc	LRT	p
3-level Global EF Model	3	28.26	31.09	29.86	n/a	n/a
2-level Global EF Model	2	28.96	30.84	29.71	2.70	0.10

Note. AIC=Akaike Information Criterion; BIC=Bayesian Information Criterion; LRT=Likelihood Ratio Test.

The 3-level model has a slightly lower value for the Akaike Information Criterion (AIC) and a slightly higher value for the Bayesian Information Criterion (BIC). This suggests that both models provide a similar fit for the data. This is supported by the nonsignificant (p=0.10) likelihood ratio test (LRT), although the p-value could be argued to be approaching significance.

4.3.4.3 Moderator analyses

As the number of studies (k=17) is greater than 10, the inhibition dataset may have the statistical power to simultaneously detect multiple variables without risking over-fitting. Meta-regressions were performed using participant age, total dosage, and WoE D as covariates (see Table 25).

Table 25

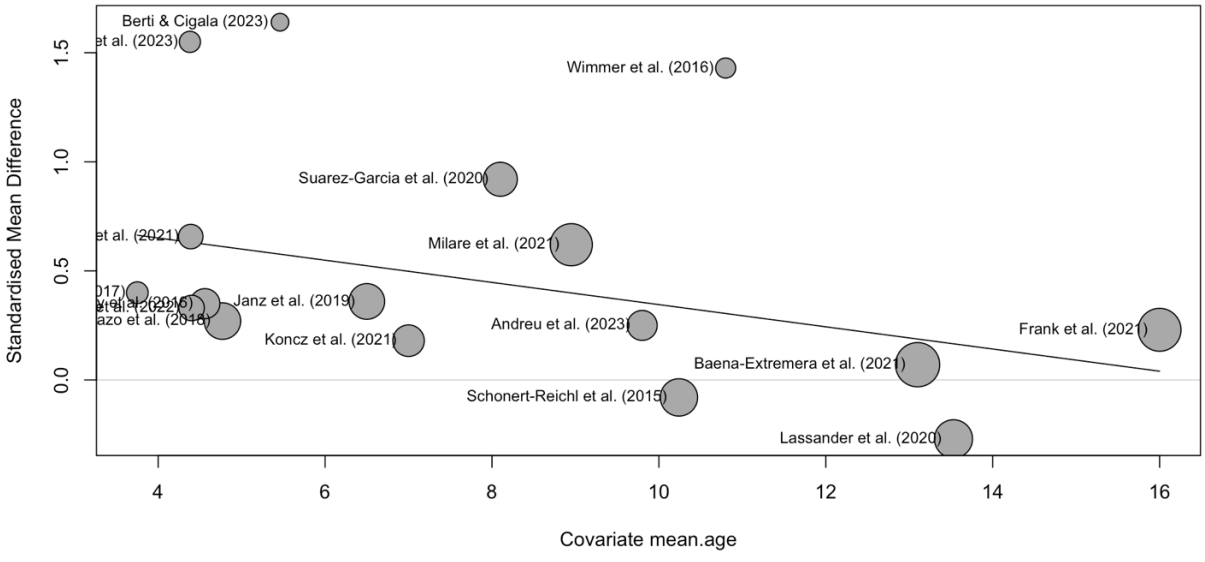
Results of meta-regressions of continuous variables on aggregated inhibition effect sizes.

Covariate	estimate	SE	p-value	95% CI	R^2
Participant age	-0.05	0.03	0.10	-0.11, 0.01	27.5%
Total dosage	0.001	0.0005	0.06	-0.00, 0.002	32.8%
WoE D	-0.47	0.27	0.10	-1.05, 0.1	39.2%

Note. Janz et al. (2019) did not report sufficient dosage information to be included in this meta-regression.

Figure 23

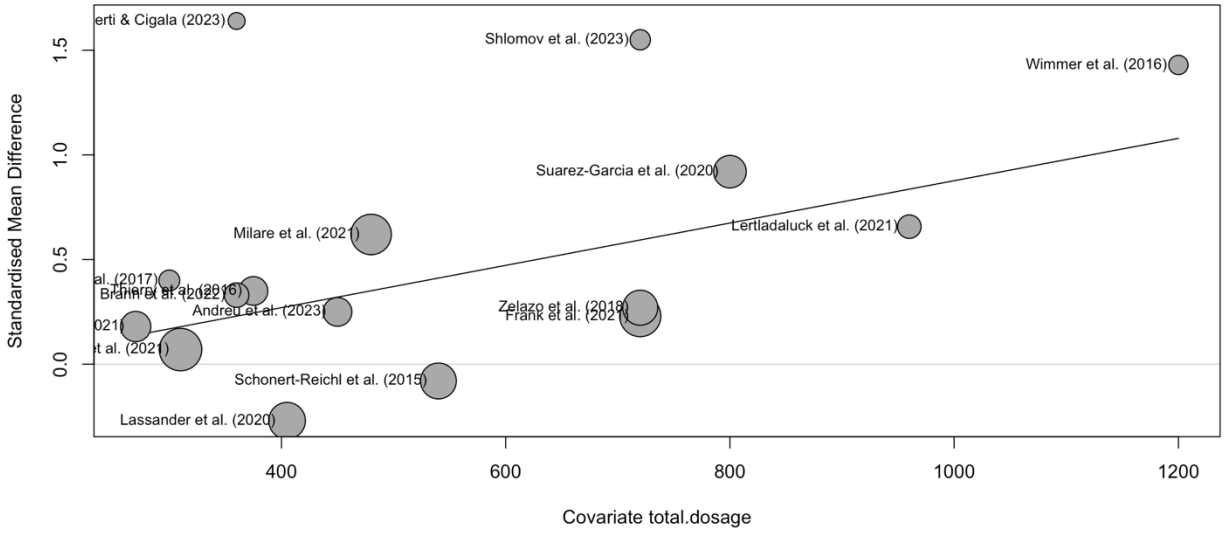
A bubble plot showing the correlation between inhibition-related outcomes and mean age.



Unlike the other ECP analyses, total dosage significantly affects inhibition outcomes (see Figure 24).

Figure 24

Bubble plot showing the correlation between inhibition effect sizes and dosage.



Two further CHE models were fitted using "the total number of sessions" and the "mean duration of sessions" as covariates to further explore the potential moderating impact of intervention dosage on inhibition-related outcomes. This indicated that the length of individual sessions did not impact inhibition effect sizes. The total number of sessions accounted for 11.39% of the variance between studies, although this was non-significant (see Table 26).

Table 26

Results of meta-regressions using dosage variables on aggregated inhibition effect sizes.

Covariate	estimate	SE	p-value	95% CI	R^2
Number of sessions	0.09	0.26	0.15	-0.008, 0.05	11.39%
Mean duration	-0.007	0.009	0.44	-0.03, 0.01	0.00%

Note. R^2=unaccounted heterogeneity explained by covariate.

CHE models fitted with robust variance estimation found similar results to those derived from the aggregated dataset (see Table 27).

Table 27

Correlated and Hierarchical Effects (CHE) modelling for all inhibition outcomes

Covariate	estimate	SE	95% CI	p-value
Participant age	-0.050	0.024	-0.11, 0.009	0.08.
Total dosage	0.001	0.0004	0.00012, 0.0019	0.03*
Total sessions	0.019	0.010	-0.0076, 0.045	0.13
Mean session duration	-0.0068	0.0085	-0.028, 0.14	0.45
WoE D	-0.46	0.15	-0.84, -0.090	0.02*

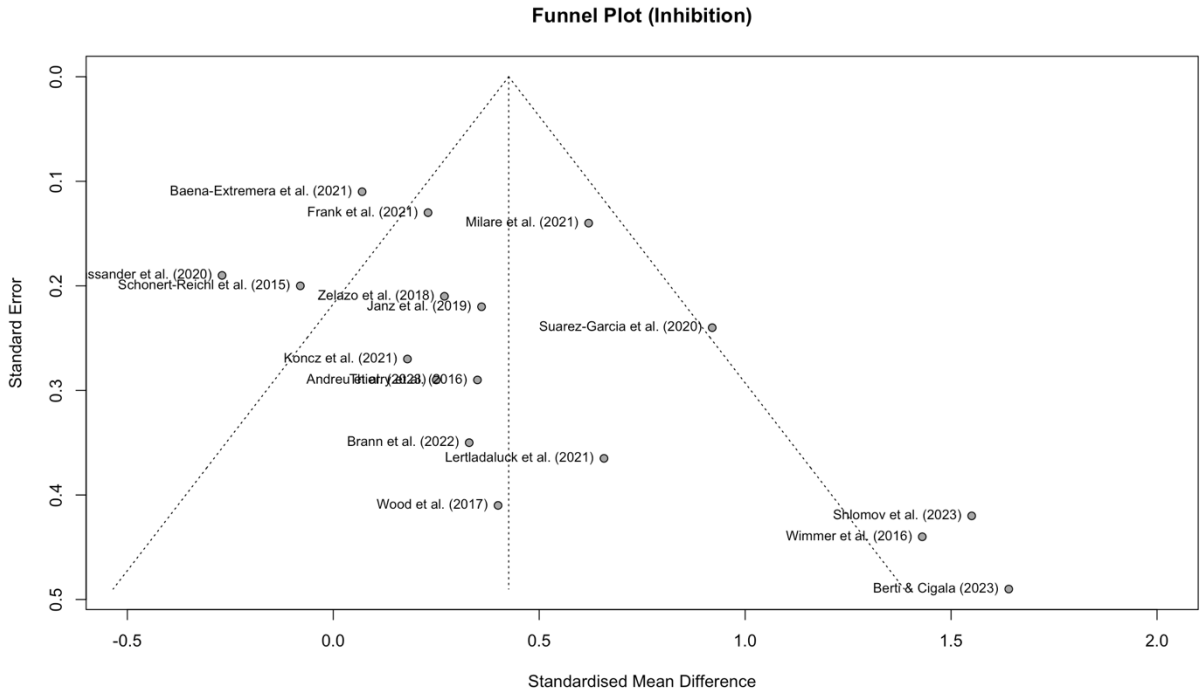
Note. P-values use robust variance estimation. Janz et al. (2019) did not report sufficient data on three covariates: total dosage, total sessions, and mean session duration.

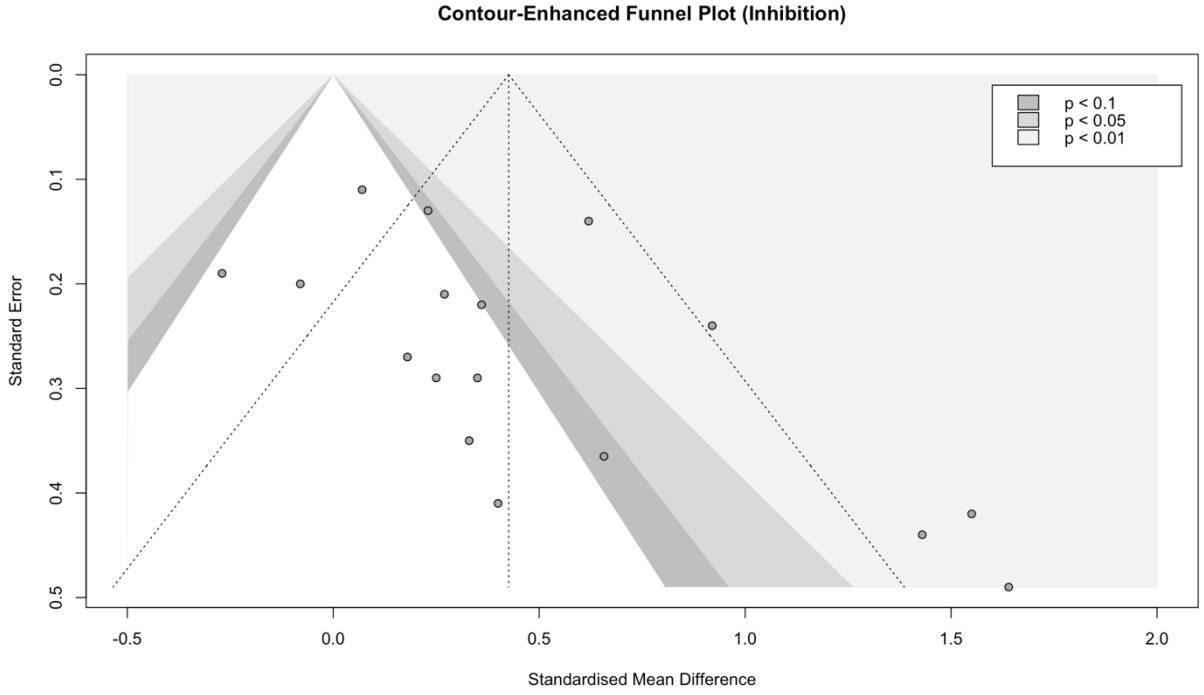
4.3.4.4 Publication bias

Funnel plot analysis indicated asymmetry in the dataset. Three studies (Berti & Cigala, 2023; Shlomov et al., 2023; Wimmer et al., 2023) had the greatest effect sizes and the largest standard error. There is no equivalent for these smaller studies on the other side of the funnel plot, indicating publication bias (see Figure 25).

Figure 25

Funnel plot and contour-enhanced funnel plot of inhibition studies.





Egger's tests confirmed the likelihood of publication bias, both within the aggregated dataset (2.24; 95%CI [0.42, 4.05]; p=0.03), and when assuming independence between all outcomes (2.09; 95%CI [0.53, 3.65]; p=0.02).

Overall, there are strong indications that the inhibition-related outcomes dataset contains publication bias that inflates overall g. This is seen when comparing the pooled effect size of the full aggregated model (g=0.43, 95%CI [0.17, 0.68]) with an "outliers removed" (g=0.28, 95%CI [0.10, 0.46]) model excluding the three studies identified as outliers in the funnel plot (see Table 28).

Table 28

Comparison of pooled effect and heterogeneity between inhibition models.

Analysis	k	g	95% CI	p	95% PI	I^2	95% CI	τ^2	95% CI
Aggregated Model	17	0.43	0.17, 0.68	0.003	-0.40, 1.25	70.1%	50.9%, 81.8%	0.37	0.23, 0.76
Outliers Removed	14	0.28	0.10, 0.46	0.005	-0.27, 0.82	56.0%	20.0%, 75.8%	0.06	0.006, 0.19

Note. τ^2=between-study variance in true effect sizes.

4.3.5 Working Memory

Twelve effect sizes from 12 separate studies were derived outcomes coded as measures of "working memory" (k=12, n=12). As all outcome measures were derived from independent studies, there was no need to account for the unit-of-analysis problem in this sub-analysis. BRS measures (n=4) represent 33.3% of this sub-analysis, with the remaining 66.6.% of effect sizes based on PBT measures (n=8). Working memory outcomes represent all the Key Stages (EYSF-KS4) in the review.

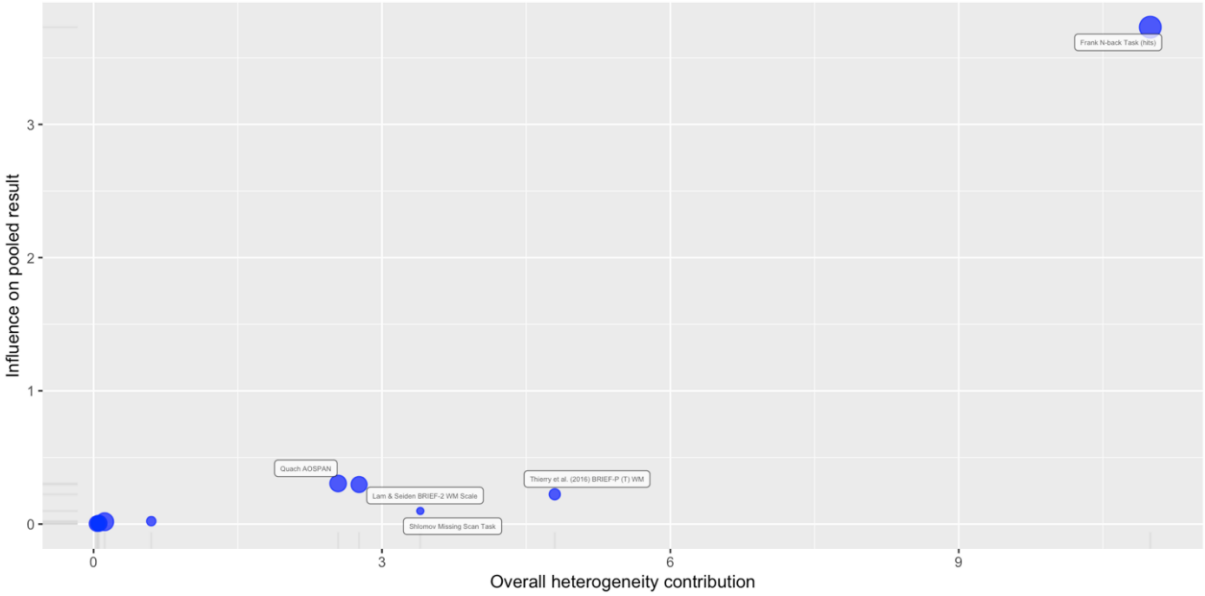
4.3.5.1 Sensitivity analysis

Sensitivity analyses detected Frank N-back Task (hits) as a potential outlier. The total variance not attributable to sampling error (I^2=56.7%) was above 50%, indicating that it may be appropriate to exclude cases.

A Baujat plot exploring the relative influence of working memory outcomes indicates that Frank N-back Task (hits) exerts the most influence on the pooled result and contributes the most heterogeneity (see Figure 26).

Figure 26

Baujat plot showing the influence of all working memory outcomes on pooled effect size.



As discussed in Section 4.2.1.2, removing Frank et al.'s (2021) EFN-back outcome might be justified based on a priori exclusion criteria. Removing the outlier reduced the percentage of variance not explained by sampling error from I^2=56.7% to I^2=6.4%. The overall pooled effect size increased from g=0.41 to g=0.48. (see Table 29).

Table 29

Comparison of pooled effect and heterogeneity between working memory models

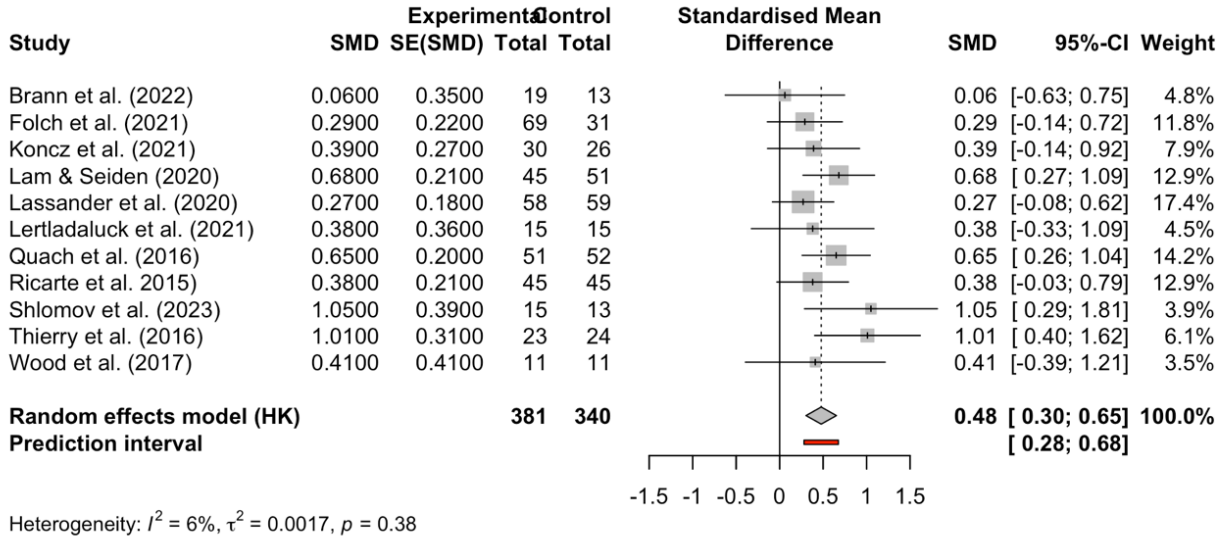
Analysis	k	g	95% CI	p	95% PI	I^2	95% CI	τ^2	95% CI
All cases	12	0.41	0.20, 0.62	0.001	-0.18, 1.01	56.7%	17.5%, 77.3%	0.06	0.005, 0.26
Outlier removed*	11	0.48	0.30, 0.65	0.0001	0.28, 0.68	6.4%	0.0%, 62.8%	0.002	0.00, 0.19

Note. *Outlier removed: Frank et al. (2021) EFN-back (hits).

Whereas the "all cases" analysis indicates that negative effects could not be ruled out (95%PI [-0.18, 1.01]), this is not the case when the outlier is removed as the prediction interval no longer passes through zero (95%PI [0.28, 0.68]). A forest plot of the working memory analysis with outlier removed is presented in Figure 27.

Figure 27

Forest plot of working memory effect sizes, minus EFN-back outlier



The overall pooled effect size and between-group heterogeneity are highly influenced by one case, Frank et al. (2021). Alone, this would not necessarily justify removal. Still, as it can be argued that this emotional version of the typical N-back test does not meet a priori inclusion criteria, it was determined that there was sufficient justification for removing the results from Frank et al. (2021) from subsequent moderator analyses.

4.3.5.2 Moderator analyses

As the number of studies (k=11 or 12) is greater than 10, conducting separate meta-regressions was considered appropriate. These included key meta-regressions using continuous variables (see Table 30).

Table 30

Continuous covariate meta-regressions on working memory outcomes.

Covariate	estimate	SE	p-value	95% CI	R^2
All cases					
Participant age	-0.036	0.022	0.13	-0.08, 0.013	33.4%
Total dosage	-0.0005	0.0005	0.39	-0.0017, 0.0007	17.18%
WoE D	0.18	0.22	0.94	-0.49, 0.52	0.00%
EFN-back outlier removed					

Participant age	-0.009	0.024	0.71	-0.06, 0.044	0.00%
Total dosage	-0.0001	0.0006	0.85	-0.0012, 0.0014	0.00%
WoE D	0.19	0.17	0.31	-0.21, 0.58	0.00%

Note. R^2=unaccounted heterogeneity explained by covariate.

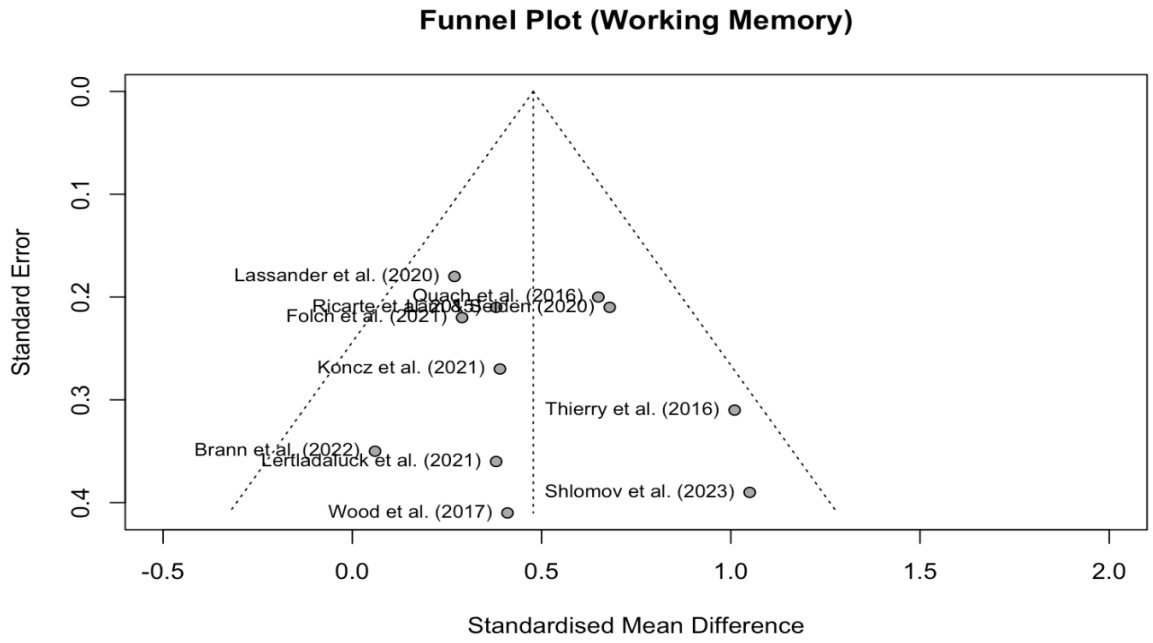
For working memory-related outcomes, no covariates were significantly correlated with effect size. Although participant age was estimated to account for 33.4% of between-study variance in effect sizes, this was not significant (p=0.13).

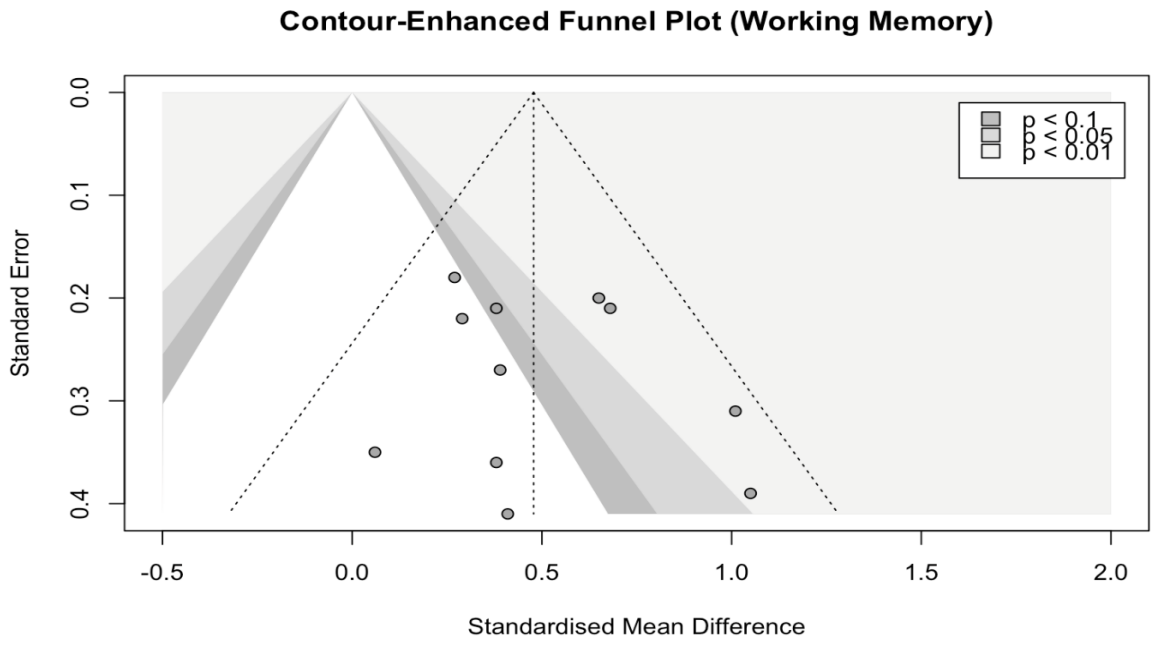
4.3.5.3 Publication bias

The funnel plots of the working memory meta-analysis did not indicate asymmetry. The contour-enhanced plot suggests that all four studies reporting effect sizes with alpha values less than 0.01 lie within the funnel (see Figure 28).

Figure 28

Funnel plot and contour-enhanced funnel plot of inhibition studies





Supporting visual analysis of the funnel plot for WM effect sizes (outlier removed), Egger's test does not indicate the presence of publication bias (0.76; 95%CI [-1.59, 3.1]; p=0.54). Egger's test performed on a working memory meta-analysis that includes Frank et al. (2021) does indicate publication bias (2.57; 95%CI [0.4, 4.74]; p=0.04).

5. Discussion

The present meta-analysis aimed to investigate the effectiveness of school-based mindfulness programmes (SBMPs) on executive function (EF) and associated cognitive processes in CYP. The overall composite effect size of SBMPs on EF outcomes was significant ($p<0.0001$) and estimated to have a small-moderate effect using both multilevel modelling ($r=0.33$) and aggregate modelling ($g=0.33$). This finding suggests that SBMPs positively impact executive functioning in CYP, although the magnitude of this impact is relatively modest.

The small-moderate effect size is consistent with other meta-analyses of SBMPs exploring cognitive outcomes (e.g. Dunning et al., 2022) and compares favourably against recent null findings for mental health and well-being outcomes (Kuyken et al., 2022). However, an intervention effect of one-third of a standard deviation is hardly a resounding endorsement for the universal implementation of SBMPs. This small-moderate effect is surprising given that the theoretical synergy between mindfulness and executive functioning suggests that SBMPs may hold more potential to improve EF and related cognitive processes than is currently being realised. However, this relatively modest composite effect size may obscure essential differences in how SBMPs interact with specific EF-related cognitive processes (ECPs) and certain age groups.

This discussion chapter explores these potential differential effects and considers how mindfulness might be better operationalised in SBMPs to optimise EF-related outcomes for young people. It is organised around the four research questions, synthesising the narrative synthesis results with the meta-analyses. The results relevant to each research question are compared to the existing literature to explore implications for future research and EP practice.

5.1 Differential effects

This section triangulates results from meta-analyses and with the narrative synthesis of included studies to answer the first research question: Do school-based mindfulness programmes have a differential effect on cognitive processes associated with executive function?

The overall effect of ECP as a moderating variable influencing the effectiveness of SBMPs is non-significant ($p=0.13$) using CHE models with robust variance estimation (RVE). Similarities between ECP outcomes are seen in the comparable effect sizes found for inhibition ($g=0.41$) and working memory ($g=0.45$) compared

to more global EF measures (g=0.32). However, SBMPs resulted in a positive intervention effect for all ECPs, and cognitive flexibility was found to have a significantly smaller effect size (g=0.20) than working memory. This suggests that SBMPs may have a differential impact on some, albeit not all, ECPs.

There is further cause to suggest that SBMPs have a differential effect on ECPs when examining the distribution of total variance in the three-level model that includes all outcome measures. In this case, level 2 (i.e. within-study variance) is estimated to account for 27.57% of the total variance in effect size, with level 3 (i.e. between-study variance) accounting for a similar 27.84% (Figure 8). This result from the composite three-level model contrasts with the within-study variance found in three-level models fitted for separate ECPs, with both the inhibition model and the working memory model indicating that within-study variance accounted for 0% of the total variance in their respective datasets. This suggests that the within-study variance seen in the composite model may be due to the inclusion of multiple outcome measures that target different ECPs. The following sections explore the separate ECP results to elucidate the potential differential effects of SBMPs further.

5.1.1 Attentional control

The overall pooled effect size of all attentional control measures was considered small-moderate (g= 0.34, 95%CI [0.15, 0.54] and significant (p<0.004). This represents a negligible difference from the pooled effect size derived from the dataset, which includes all EF-related outcome measures aggregated within studies. Although the significant effect finding suggests that SBMPs positively impact attentional control, the 95% prediction interval [-0.14, 0.54] crosses zero, suggesting that iatrogenic effects cannot be conclusively ruled out under certain conditions. However, the overall effect size suggests that the risk of harm for EF-related outcomes is less than has been implied by large-scale studies examining the impact of SBMPs on well-being outcomes (e.g. Kuyken et al., 2022).

The attentional control category was included to account for EF-related outcome measures not associated with one of the tripartite components of EF (Miyake et al., 2000). As a result, the outcome measures were more variable than other meta-analyses, and 44.4% (n=4) could be considered diagnostic tools. These included the Conners' Teacher Rating Scale-Revised (CTRS-R) and the Vanderbilt ADHD Teacher Rating Scale (Vanderbilt TRS). These tools are primarily designed to screen for and assess symptoms of attention deficit hyperactivity disorder (ADHD) and other behavioural problems rather than to measure attentional control as a component of EF precisely. These scales often include elements that focus on emotional regulation and might, therefore, have dubious construct validity as measures of the "cool" executive functioning,

which is the focus of this meta-analysis. The heterogeneity in the outcome measures may explain the moderate heterogeneity (I2=41.5%) found between studies, especially as none of the covariates explored in the moderator analyses (age, total dosage, methodological quality) were significant predictors.

This potential lack of comparability between outcome measures suggests caution when comparing the results from the attentional control analysis to the other ECPs investigated in the meta-analysis. However, sensitivity analyses on a dataset with diagnostic outcome measures (n=8) removed show a negligible change in overall pooled effect (g= 0.31, 95%CI [0.22,0.40] compared with the full dataset (g= 0.33, 95%CI [0.25, 041], indicating that these measures did not significantly influence the overall effect. It has been tentatively concluded that these diagnostic measures may serve as an acceptable proxy for more specific measures of EF, such as the BRIEF-2.

5.1.2 Flexibility

The results of the meta-analysis indicate a small but significant effect size. (g= 0.21, 95%CI [0.03, 0.40], p=0.03). As with attentional control, the 95%PI [-0.35, 0.78] crosses zero, indicating that negative SBMP effects on cognitive flexibility cannot be ruled out. For this analysis, an aggregate model was selected to address the dependency of two variables nested within the same study (Wimmer et al.,2016). Although this ignores potential within-study variance, aggregation was deemed the most parsimonious approach as both measures were PBTs specifically designed to measure cognitive flexibility and had similar effect sizes.

CHE modelling suggests that improvements in flexibility-related outcomes following SBMPs are significantly smaller (p=0.029) than working memory outcomes. This difference approached significance when comparing flexibility with inhibition (p=0.07) and Global EF (p=0.09) outcomes. This difference could be due to how mindfulness has been operationalised in many SBMPs. Activities such as focusing on breathing have a clear link to sustained attention. Such activities may engender many opportunities to inhibit responses to sensory stimuli but fewer opportunities to switch between tasks. Whilst mindfulness practices using a two-factor definition (Bishop et al., 2004) should include an open monitoring style that allows for noting sensations and thoughts as they arise, such elements were not nearly as common as breathing exercises and body scans. Holas and Jankowski (2013) argue that the nonjudgemental awareness element gives rise to cognitive flexibility processes, such as switching attention from an acknowledgement of a sensation and back to the present moment. This is a nuanced difference connected to the meta-cognitive model of mindfulness (Holas & Jankowski, 2013) discussed in the literature review (Section 2.3.2). Mindfulness

emerges at the "meta-meta level" through a meta-awareness of the cognitive processes involved in accepting our thoughts, feelings, and sensations. Simply inhibiting our thoughts and feelings to return to focus on breathing does not activate this higher-order state.

Overall, although there is a small positive effect on flexibility outcomes, consistent with previous research (Heeren et al., 2019), the evidence presented does not provide a resounding indication that mindfulness can improve cognitive flexibility better than alternative approaches like task-switching training (Karbach & Kray, 2009). This may be because cognitive flexibility is a complex process, and the operationalisation of mindfulness in SBMPs may be too simplistic to activate higher-order processes.

5.1.3 Inhibition

The overall pooled effect size for inhibition outcomes (three-level: g=0.39; aggregated: g=0.43, p=0.003) suggests a significant small-moderate effect of SBMPs on improving inhibition-related outcomes. This significant positive effect aligns with the metacognitive model of mindfulness (Holas & Jankowski, 2014), where inhibitory skills are considered necessary for maintaining a mindful state. Inhibition outcomes were the most widely reported of all ECPs, which may reflect its pre-eminence as a foundational skill for developing more complex ECPs like working memory and cognitive flexibility (Diamond, 2013).

However, the inhibition results may be less robust as this was the only ECP dataset indicating publication bias. Removing the three cases identified from the publication bias analysis reduced the pooled effect size from g=0.43 to g=0.28, suggesting that the true effect of SBMPs on inhibition-related outcomes may be smaller than indicated in the published research base. Furthermore, the 95% prediction interval [-0.40, 1.24] also suggests that the chance of SBMPs resulting in iatrogenic effects under specific conditions cannot be conclusively ruled out.

The inhibition dataset had the highest between-study heterogeneity not attributable to sampling error (I^2 =65.6%, increasing to 70.1% following aggregation). This suggests that the effect of SBMPs on inhibition may be especially sensitive to the influence of study characteristics. This is tentatively confirmed through moderator analyses, where inhibition outcomes were the only ECPs found to be moderated by intervention dosage. This greater heterogeneity between studies may also be an artefact of poor correlations between PBTs and BRSs as measures of inhibition (Friedman & Miyake, 2004).

One limitation of the aggregation approach used to account for the dependency of effect sizes nested within the same study is the lack of equivalency between outcome measures. This was less concerning in

some cases, such as the Peg Tapping Task (g=0.62) and the Bear and Lion Task (g=0.70) used by Lertladaluck et al. (2021) as measures of inhibition, where both tasks are PBTs designed to measure inhibition, and the effect sizes are congruent (both moderate-large). However, aggregating data from the other two multiple-outcome studies was more problematic. Frank et al. (2021) used both a self-report measure (BRS-S) through the DERS Impulse Scale (g=0.19) and a performance-based (PBT) measure through the Stroop Task (g=0.27). These differing measurement approaches might indicate aggregation is a suboptimal approach, although both measures report similar effect sizes.

Consideration of the aggregation of the Zelazo et al. (2018) outcomes added further complexity. As with Frank et al. (2021), both measures recorded similar small effect sizes but included one BRS tool, the CBQ Inhibitory Control Scale (g=0.16), and one PBT tool, the Peg Tapping Task (g=0.28). In addition, the samples of each measure represented different populations, with parents (n=61) contributing to the CBQ data and children (n=136) contributing to the Peg Tapping data. This lack of equivalency between outcome measures can introduce bias and highlight the subjective decision-making involved in the meta-analysis process. This serves as a reminder that the statistical approach taken in the present study does not necessarily imply that the results are freer of investigator bias than qualitative approaches.

5.1.4 Working memory

The results of the working memory meta-analysis are arguably the most robust. When removing Frank et al.'s (2021) EFN-back (hits) outcome, the heterogeneity between studies not attributable to sampling error is 6.4%. Furthermore, all outcome measures were derived from independent studies, so there was no need to account for the unit-of-analysis problem. The results can, therefore, be argued to be derived from a more parsimonious model. With the outlier removed, the overall pooled effect was moderate (g=0.48; p=0.0001), representing the highest effect size when comparing ECPs. The 95% prediction interval [0.28, 0.68] does not cross zero, suggesting that adverse effects on working memory outcomes following SBMP intervention are unlikely. Finally, unlike other ECPs, methodological quality and relevance to school practice were not a significant moderator of SBMP effectiveness (p=0.31). They explained 0.00% of unaccounted heterogeneity, regardless of whether the EFN-back outcome was included. This suggests that the working memory results are more robust in the face of methodological criticisms levied against mindfulness research in general (Van Dam et al., 2018).

These working memory results represent the most substantial evidence for a positive SBMP effect of all ECPs. This is a promising result for educational psychology as working memory is considered a critical

cognitive capacity underlying a range of academic abilities essential to school functioning, including reading comprehension and mathematical problem-solving (Gathercole et al., 2006). Findings from cognitive neuroscience (e.g. Munakata et al., 2011) suggest that working memory may be the preeminent process involved in EF, suggesting that improvements in working memory may have the broadest application in real-world contexts compared with the other ECPs.

An interesting finding emerging from the working memory meta-analysis is the effect of removing the emotional n-back (EFN-back) outlier, which raises a future research avenue to compare the effects of SBMPs on "hot" and "cool" EF processing. The Emotional Faces N-back Task (EFN-back) includes images of emotional faces to stimulate an emotional response in the task performer, designed to increase the emotional valence of the task and thus measure emotion regulation and working memory. As such, it can be considered more reflective of "hot" processing than non-emotional tasks. The fact that the EFN-back task produced a negative intervention effect and reduced the pooled effect size from g=0.48 to g=0.41 could be used to indicate that SBMPs may target "cool" processing over "hot" processing.

This hot/cool distinction remains speculative as it is derived from a single outlier. However, the potential difference between the effectiveness of SBMPs in improving working memory under emotional conditions warrants further investigation within educational psychology, as executive functioning in school systems is often associated with concurrent emotional regulation, especially for CYPs with special educational needs (Berlin et al., 2004). It may be the case that "cool" EF processes exhibit greater plasticity and are thus more malleable to the influence of SBMPs, as "hot" processing is hypothesised to be more embedded in life experiences and individual temperament (Zelazo & Carlson, 2012).

However, as the inclusion of the outlier still produces a significant and small-moderate pooled effect (g=0.41), this potential concern does not significantly diminish the effectiveness of SBMPs on working memory outcomes. Furthermore, as discussed in Section 2.1.2 (paragraph 3), there are indications that emotional processing still occurs in traditional PBTs, such as the digit span backwards task (see Appendix N), and that there are positive correlations between hot and cool EF tasks (Willoughby et al., 2011). This meta-analysis also included behaviour rating scales (BRSs) as measures of working memory, which may better reflect working memory functioning in emotionally charged contexts, such as school environments. Furthermore, the pooled effect of attentional control outcomes (g=0.34), which typically included more diagnostic measures that included elements of emotional regulation, was essentially the same as the

pooled effect size of all outcome measures using a three-level model (r=0.33). As such, the improvements in experimental outcomes following the SBMP may be translatable to real-world outcomes.

5.2 Moderating effect of age

Meta-regressions and CHE modelling with RVE were used to answer the second research question: Does age moderate the extent to which school-based mindfulness programmes improve measures of executive function and associated cognitive processes?

5.2.1 Overall moderating effect of age

The CHE model indicates that participant age emerged as a predictor covariate that approaches significance ($p= 0.066$). The regression weight of participant age was 0.024, suggesting that the effect of SBMPs on composite EF reduces by 0.024 for every one-year increase in the average age of participants. This result does not reach the conventional significance threshold of $p < 0.05$, suggesting that the moderating effect of age on SBMP effectiveness could be due to chance. However, CHE modelling, which accounts for the dependency of effect sizes nested within the same study, reduces the risk of biased estimates and Type I error. Further increasing the trustworthiness of this finding, a robust variance estimation (RVE) was employed for this moderator analysis. Using RVE reduces the degree to which the analysis results depend on normal distribution assumptions, increasing resilience to outliers and heteroscedasticity. Again, this reduces the risk of false positives often associated with meta-regressions (Hedges & Olkin, 2014).

It is also possible that this correlation between effectiveness and participant age is due to another associated variable that correlates with age. In studies targeting younger children, greater adjustments to established protocols were typically required to ensure the suitability of the content, such as adapting outcome measures to engage young children. For example, in the narrative element of the “Bear and Lion Task” (used by (Lertladaluck et al., 2021; Shlomov et al., 2023), participants were told that the “bear was good so they should always do exactly what the bear told them to” (McAlister & Peterson. 2006). This may introduce an element of emotion regulation absent in the more established EF tasks typically employed to measure EF-related outcomes in older participants.

Therefore, the possibility remains that the moderating effect of age is a result of multi-collinearity with methodological quality and relevance. In other words, the age effect may reflect a hypothetical possibility that studies with older participants can better control extraneous variables. However, correlation matrices

between participant age and WoE D rating found only a small correlation (r=0.24), suggesting that the moderating effect of age is not solely an artefact of co-linearity between participant age and study quality. Whilst this tentatively supports the conclusion that age moderates the effectiveness of SBMPs on EF-related outcomes, multi-collinearity may exist between age and another unknown variable not investigated in this review.

5.2.1 Differential moderating effect of age

The results of further moderator analyses suggest that age was a significant predictor of the size of the effect of SBMPs for some, but not all, cognitive processes associated with executive function. This suggests a complex interaction between age and SBMP intervention, which the differential developmental trajectories of ECPs may mediate.

5.2.1.1 Cognitive flexibility and inhibition

Meta-regressions indicate that age is a significant moderator for cognitive flexibility outcomes. The regression weight of participant age is estimated to be -0.06, suggesting that as age increases by one year, the SMD is predicted to decrease by 0.06. The model predicts that participant age accounts for 38.2% of the difference between true effect sizes. This moderating effect was also found for inhibition-related outcome measures, with a –0.05 estimate for the regression weight of participant age. In other words, a ten-year age difference would represent a reduction in effect size of 0.5 or half a standard deviation.

However, the observed age-related differences in SBMP effectiveness for cognitive flexibility and inhibition are consistent with the developmental trajectories of these EF components. *Inhibitory control* is a foundational EF-related process that emerges comparatively early, with rapid development occurring between 3 and 5.25 years (Wiebe et al., 2012). The meta-regression results suggest that this developmental trajectory may represent a critical period for the development of inhibitory control, during which there is a heightened sensitivity to environmental influences, such as an SBMP.

The developmental trajectory of cognitive flexibility peaks later than inhibition and continues to develop into adolescence (Dajani & Uddin, 2015), with improvements in task-switching paradigms observed from 4 to 13 years (Davidson et al., 2006). Within this period, accelerated improvement is posited to occur between 6 and 10 years, and the upper bound of this range may represent a ceiling for the effectiveness of SBMPs in improving cognitive flexibility. This upper bound of 10 years old has an interesting parallel with

the findings of the meta-regression. Visual analysis of the bubble plot (Figure 15) indicates a trend towards negative effect sizes for participants aged ten and over.

This negative correlation between age and SBMP effectiveness runs contrary to predictions made by Dunning et al. (2019), who hypothesised that the benefits of mindfulness practices would be best realised in more developmentally mature individuals with the cognitive capacity to achieve higher-order mindfulness states. This discrepancy may indicate that the observed age effects in this meta-analysis may not be reflective of the potential of mindfulness practices for this age group but rather indicate a limitation of the operationalisation of SBMPs for older CYP, which may not generate sufficient challenge to engender a meta-metacognitive state (Holas & Jankowski, 2014). Simple breathing exercises may support the re-organisation of inhibitory systems at earlier developmental stages but lack the complexity needed to build cognitive flexibility skills in adolescence.

5.2.1.2 Working Memory and Global EF

The moderating effect of age on working memory is difficult to determine. When no outliers were removed from either dataset, age predicted 33.4% of the variance between studies with a regression coefficient estimated at -0.0036, although this was not significant (p=0.13). Removing the Frank EFN-back outlier reduced overall I2 to 6.4%, and age no longer explained any of the variance in this analysis. However, removing this outlier also removed any study representing older adolescents in this sample.

The potential moderating effect of age depends on the inclusion/exclusion of the Frank et al. (2021) EFN-back outcome. It may be that "hot" executive functioning (Zelazo & Muller, 2002) is less moderated by age. This view has tentative support as attentional control outcomes were not moderated by age, and this dataset included multiple diagnostic measures that screened for emotional regulation. Global EF, also not moderated by age, was predominately measured through BRSs whose rating scales included emotionally salient elements, such as the BRIEF. This "hot" and "cool" distinction can be linked to developmental theory, where experimental evidence comparing hot and cool PBTs suggests that hot executive functioning follows a delayed developmental trajectory, with the greatest improvements seen in adolescents (Prencipe et al., 2011).

The results from the Global EF analysis were less susceptible to the influence of a single outlier, had sufficient statistical power for meta-regressions using single covariates, and included at least one study covering all key stages. Moderator analyses found that age predicted 0% of the variance between studies

measuring Global EF outcomes. This result indicates that age acts as a differential moderating variable on the effectiveness of SBMPs, which may predict a significant proportion of the between-study variance for separate ECPs but not for a global EF construct.

The poor predictive power of age as a moderator for the effectiveness of SBMPs on global EF outcomes can be understood by considering the distinct developmental trajectories of the main cognitive processes through which it emerges. As discussed earlier, the current evidence suggests that inhibition develops rapidly in early childhood and is relatively stable from 6-7 years onwards (Best & Miller, 2010). Cognitive flexibility exhibits a similar spike in development, although this is more graduated and peaks around 8-10 years, with shifting skills stabilising after 13-15 years (Davidson et al., 2006). Working memory, which may be a higher-order process that has a supervisory function over processes such as inhibition (Munakata et al., 2011), exhibits a flatter maturational curve that extends throughout adolescence and does not reach full maturity until adulthood (Huizinga et al., 2006). The results of the age moderator analyses on ECPs support the claim that they have distinct developmental trajectories.

5.2.2 Implications for theory and practice

The differential age-related effects observed in this meta-analysis could suggest that some ECPs undergo developmental trajectories that include distinct sensitive periods. These periods are characterised by accelerated development where the neural substrates underlying each ECP exhibit heightened plasticity and are thus more malleable to the influence of SBMPs. These findings support a dynamic model of executive function (Zink et al., 2023) over a static, three-factor model. Considering EF's unity and diversity as it develops across the lifespan is hypothesised to lead to improved effectiveness for interventions that target EF. Future research would benefit from longitudinal approaches to build our understanding of the developmental trajectories of ECPs and how EF emerges from them.

The moderating effect of age has direct implications for EPs supporting the design of SBMPs, where approaches that more directly target inhibition may be more effective for younger CYP. In contrast, SBMPs targeting older CYPs should focus on those ECPs that continue to mature through adolescence, such as cognitive flexibility and working memory. Future research would also benefit from consideration of how to ensure that an appropriate level of challenge is integrated into SBMPs that target more developmentally mature CYP. This could be achieved by including practices such as "thoughts as clouds", where participants are encouraged to mentally visualise their thoughts as clouds passing in the sky. This heightened level of

abstraction can engage cognitive flexibility skills whilst providing a suitable level of challenge linked to cognitive developmental stages (Steinberg, 2005).

There are other, more practical explanations for this observed moderating effect of age. Firstly, Early Years and Primary school settings may be inherently more "mindful" than secondary school settings. Mindfulness requires psychological safety that may be harder to engender in the more stressful secondary school environment. The physical space is often more adaptable in Early Years settings than in Secondary schools. This difference is exemplified in Berti & Cigala (2022), where a specially designed and decorated hut was used for the SBMP in this study. This may have some explanatory power as to why this study produced by far the largest aggregate effect size (g=1.64) of any in the meta-analysis.

A fundamental limitation in interpreting the age-related effects found in this study relates to how measures of EF are adapted to assess different developmental stages. In younger children, measures such as the HTKS and Hearts and Flowers tasks were frequently employed as child-friendly measures of ECPs. As the tasks used to measure ECPs in the included studies differed depending on developmental age, it is difficult to determine whether the age-related effects may be due to true developmental differences or specific task effects. This might also explain the differential moderating effect of age in this meta-analysis. Karr et al. (2018) found that some working memory measures, such as the Digit Span Backward, were used consistently across ages. However, measures of inhibition such as the Stroop task often varied (e.g. Boy/Girl Stroop; Day/Night Stroop; Colour/Word Stroop). Therefore, the moderating effect of age on inhibition outcomes may be due to task effects emerging from adapting PBTs for different age groups.

Another limitation of the present review is the lack of representation from KS4 and KS5 CYP. Only two studies included participants over 14 years old, and none included CYP in the KS5 age range (16-18 years). This limited representation of older adolescents restricts the extent to which the negative moderating effect of age found in this meta-analysis can be extrapolated beyond early adolescence. The promising impact of mindfulness programmes on university-aged young people (e.g. Galante et al., 2017) hints at a potential U-curve relationship between age and SBMP effectiveness. A speculative hypothesis based on this observation might predict that SBMPs have a greater benefit for two distinct age groups: younger children, who have increased EF-network, and older adolescents, who have reached a stage of cognitive maturation that enables them to engage with more complex iterations of mindfulness. Many studies exploring MBIs for university-aged young people were identified through the screening process for this review (see Appendix A). However, there appears to be a relative scarcity of SBMP research focusing on

the 14-18 age range. Such studies could help to further elucidate the relationship between the developmental trajectories of EF and mindfulness practices.

5.3 Intervention characteristics

This section aims to answer the third review question: Which intervention characteristics influence the effectiveness of school-based mindfulness programmes to improve executive function and associated cognitive processes?

5.3.1 Intervention components

As outlined in the narrative synthesis, age-related differences were found in the components of SBMPs, with programmes aimed at younger children typically including more physical and engaging activities, which were gradually replaced with a greater emphasis on psychoeducational components in adolescence. In the studies that sought feedback from CYP, acceptability was typically higher for younger children, which may be due to more enjoyable intervention components. Enjoyment of the SBMP components will have a significant impact on motivation to practice mindfulness, with this intention being one of the core components in the three-factor model (Shapiro et al., 2006) of the mindfulness construct itself, used as the theoretical basis for one of the studies included in this review (Lertladaluck et al., 2021). This element may be particularly relevant to SBMPs, where participants' intrinsic motivation must be cultivated, as students are initially motivated by compliance with school expectations and have not made a self-directed choice to engage in mindfulness practices. These differences in intervention components could explain the observed age effects described in the previous section.

While many theoretical frameworks used to justify the intervention components of SBMPs were broadly harmonious, such as the two-factor (BIshop et al., 2004) and three-factor (Shapiro et al., 2006) models, other models were conflicting. For example, some approaches included elements of psychoeducation grounded in CBT. Although mindfulness-based CBT emphasises a non-judgemental orientation, including practices such as identifying "negative" thinking patterns introduces an element of judgment. Elements of gratitude and kindness to others were included in other cases, such as the MindUP curriculum. These activities may support the development of self-awareness, but they also have the corollary effect of establishing a hierarchy of thoughts and feelings. This hierarchy, where some experiences are privileged as "better" than others, could contradict mindfulness's "non-judgmental orientation". This may reflect a broader conflict between openness to orientation and the goals of an SBMP focusing on EF, which are

typically driven by the goals of the school systems in which they are employed, such as improved academic performance.

Some SBMPs convincingly integrated adjunct approaches within a mindfulness theoretical framework. For example, the acceptance and commitment therapy (ACT) elements in the L2B curriculum complement the "open orientation to experience" mindfulness component. However, in all cases where interventions have diverged from core practices, introducing additional variables acts as a barrier to better understanding the unique contributions of mindfulness-specific practices. This intervention diversity indicates a potential need for a more granular understanding of how mindfulness can improve executive functioning.

Given the considerable heterogeneity in how SBMPs are operationalised, a key focus for research moving forward is the isolation of the specific mindfulness components that support the development of executive functioning. Furthermore, it remains to be established whether the mindfulness-specific components have a greater impact on EF-related outcomes than more parsimonious approaches such as relaxation techniques. For example, the negative intervention effect for an inhibition-related (g=-0.27) outcome reported by Lassander et al. (2020) may be linked to this study's more specific isolation of mindfulness components by developing a relaxation programme for their control condition. Most studies did not use a truly 'active' control condition, so it is difficult to ascertain whether the mindfulness components significantly improve outcomes over and above less resource-intensive interventions.

The results of this research raise several considerations for EPs consulting schools in the design and implementation of SBMPs and related mindfulness interventions. Chiefly, interventions should be tailored to developmental stage. The present review indicates that this is most successfully achieved for SBMPs targeting younger children, where intervention components often include a broad range of engaging sensory and physical activities to motivate CYPs to practice mindfulness. For adolescents, age-related adaptations frequently mean a greater focus on psychoeducation, with sessions that are more embedded within a regular curriculum structure. Secondary schools should ensure that SBMPs are sufficiently challenging to cultivate higher-order mindfulness while ensuring that programme content is engaging. This is a difficult balance that current SBMPs are not effectively addressing. The vast heterogeneity in SBMP intervention components means it is difficult to isolate the elements that are uniquely "mindful" and positively affect EF. Future research should determine whether SBMPs are more beneficial than straightforward relaxation techniques.

5.3.2 Administrator role

Exploratory subgroup analyses found little difference between the pooled effect size for studies using external administrators (g=0.34) and those using teacher administrators (g=0.30). Studies (n=2) that used a mix of teachers and administrators had a larger effect (g=0.52). However, the limited number of studies in this group suggests that this is likely due to chance, and the differences between all subgroups were non-significant (p=0.59).

This was surprising as the role of the administrator was hypothesised to have a moderating effect on the effectiveness of SBMPs (Phan et al., 2022). However, the results of this study do support those found in previous SBMP meta-analyses (Mettler et al., 2023). Interpreting this null finding is challenging, given the approach taken in this study. Firstly, the role of the administrator was operationalised in the meta-analysis through relatively crude categorisation depending on whether the administrator was a class teacher, external to the school, or a mix of both. Following the narrative synthesis, it became clear that these categories were highly heterogeneous. Within the teacher category, the training protocols used by SBMPs differed considerably. The teachers in Baena-Extremera et al. (2021) received four hours of mindfulness training a week for nine weeks, compared to half a day for the teachers in Janz et al. (2019).

The external administrator category was also heterogeneous and included research assistants, the study authors, and trained mindfulness professionals. Based on this meta-analysis's results, such diversity within these categories makes it difficult to assess the moderating effect of the administrator role. In two studies, the intervention administrator was also an author (Lam & Seiden et al., 2020; Lertladaluck et al., 2021). Whilst this was likely necessary for practical purposes, this dual role introduced an additional risk of investigator bias.

Results indicate that teachers can be effective administrators, supporting SBMP viability in school settings. The relational closeness between teachers may counteract the relative expertise of an external administrator. Upskilling teachers in mindfulness may increase the likelihood that practices will be embedded into the school culture. Although there was limited follow-up data collection from the studies in this meta-analysis, an argument could be made that effects will be better maintained over time when trained class teachers administer SBMPs.

5.3.3 Total intervention dosage

The overall moderating effect of intervention dosage on all outcomes was 0.0001 (95% CI [-0.00036, 0.00062]) and non-significant (p=0.56) using CHE modelling with RVE. The 95% confidence interval indicates this is a poor predictor of SBMP effectiveness. This is a surprising result because the amount of time spent engaging in mindfulness practice is hypothesised to be a critical determiner of mindfulness proficiency (Parsons et al., 2017). This meta-analysis showed a ten-fold difference in intervention dosage between the shortest SBMP (120 minutes) and the longest (1200 minutes). This reflects the high dosage variability in previous SBMP reviews (Roeser et al., 2022). However, none of the meta-analyses, save for inhibition outcomes, found total intervention dosage to be a significant moderator of SBMP effectiveness.

A non-linear relationship between SBMPs and dosage may explain the lack of a moderating effect, as found in previous SBMP meta-analyses (Zoogman et al., 2015). SBMPs may have a minimum threshold for effectiveness, with limited benefit from additional practice beyond this point. Most of the studies included in this meta-analysis had intervention dosages exceeding 300 minutes, which may be a sufficient threshold for improvements at post-intervention data collection points. Although purely speculative, it is interesting that the shortest SBMP (Muller et al., 2021) was one of only two studies that reported a negative aggregate effect size (g=-0.09). However, the idea of a minimum threshold does not seem congruent with the theoretical conceptualisation of mindfulness as a more complex skill developed through practice and the accrual of meta-cognitive knowledge and skills (Holas & Jankowski, 2014). Nor does it fit well with the dynamic model of EF advocated in this research. Here, it is considered more likely that the lack of dosage effects is reflective of the limited effectiveness of SBMPs in cultivating sufficiently complex mindfulness practices to improve EF past a certain point.

For inhibition outcomes, total intervention dosage was found to have a significant positive correlation with effect size (0.001; p=0.03) using CHE modelling with RVE. This finding is consistent with previous adult MBI studies (Verhaegen, 2021). A meta-regression using the aggregated inhibition model found that dosage predicted 32.8% of the heterogeneity in effect sizes between studies. Although this finding could suggest that inhibitory skill development may benefit from longer sessions, which provide more opportunities to inhibit emotions such as boredom, further meta-regressions indicated that the mean duration of each session explains 0.00% of the variance. This suggests that dosage had a greater effect on inhibition outcomes when it was “little and often” and was unrelated to the length of individual sessions. Alternatively, it could be the case that other ECPs, which may be more complex than inhibition, are more sensitive to the

quality or type of mindfulness instruction than SBMPs. This hypothesis would fit with the dynamic model of EF and suggests that higher-order ECPs might be less activated by SBMPs past a certain threshold.

These surprising results may be difficult to apply to EP practice, and total intervention dosage is an imperfect measure of the extent to which CYP engages with mindfulness practices both within and outside SBMP sessions. Some speculative evidence from the meta-analysis could suggest that SBMPs can have positive effects at relatively short dosages, with diminishing marginal returns after around 300 minutes. There are also indications that interventions may benefit from a more dispersed intervention schedule, with more sessions spread out over an extended period preferable to longer-duration sessions. Higher-order ECPs, such as cognitive flexibility and working memory, may be more sensitive to the quality and content of SBMPs than total exposure. However, total dosage is an imperfect measure of CYP engagement with SBMPs, and any true dosage effects may be obscured without proper control for differences in intervention quality and fidelity of implementation.

5.4 Strengthening the evidence base

This section explores some of the methodological concerns raised by previous reviews by comparing the present narrative synthesis and meta-analysis results to address the fourth research question: How can future research strengthen the evidence base for the effectiveness of school-based mindfulness programmes in improving executive function and associated cognitive processes?

Methodological quality and relevance to the aims of this review, as measured by the overall Weight of Evidence (WoE) D score, was found to be a significant moderator of SBMP effectiveness on EF-related outcomes ($p=0.049$), with an estimate of –0.23. In other words, for every point increase in WoE D score (scores ranged from 1.2 to 2.55 for studies included in the meta-analysis), the effect size is predicted to decrease by –0.23. This appears to support concerns raised in previous meta-analyses suggesting that low methodological quality may result in inflated effect sizes and lead to over-zealous adoption of SBMPs by educational settings (Roeser et al., 2022).

This moderating effect was not found for Global EF (-0.32, $p=0.31$) or working memory (0.18, $p=0.94$) outcomes, suggesting that these ECP-specific results may be less susceptible to Type I error emerging from a lack of methodological rigour. Nevertheless, the overall moderating effect of methodological quality and relevance concerns EPs and questions whether SBMPs can be recommended as an evidence-based intervention.

5.4.1 Research Design

The nested nature of participants within school classes means that proper randomisation would be unfeasible and limit the ecological validity of the results. As such, the critical appraisal protocol used for this research did not penalise non-randomisation to the same extent as a Cochrane Risk of Bias tool. Nevertheless, it was encouraging to note that a significant minority of studies used cluster randomisation, where whole classes were randomly assigned to the intervention or control group. Full randomisation was possible for a few further studies where the interventions were conducted in small groups (e.g. Zelazo et al., 2018). These studies typically received higher WoE A ratings and indicate that the SBMP empirical research field has shifted towards more rigorous research designs than seen in earlier mindfulness intervention studies (Van Dam et al., 2018).

Sample size. One main criticism against SBMP research is that many studies with small samples lack sufficient statistical power (Roeser et al., 2022). As an evidence base built on multiple small-scale studies is more susceptible to publication bias, there was assumed to be a negative correlation between sample size and effect size. Surprisingly, this was not the case, with CHE modelling finding a non-significant correlation ($p=0.59$) between total sample size and standardised mean difference. This indicates that the small samples found in many SBMP studies exploring EF-related outcomes may not have overestimated effect sizes as much as previously thought.

Control conditions. Although it was initially hoped that control conditions could be explored as a potential moderating factor, isolating how these might influence intervention effectiveness proved challenging as SBMPs are nested within school curricula. It was not always clear how to best determine whether a control condition was active or passive. For example, in some studies, the SBMPs were administered during lesson time that was otherwise dedicated to social-emotional learning (SEL). In these cases (e.g. Janz et al., 2019; Lam & Seiden, 2020), the control condition was listed as "passive". In other cases, the control condition was exposed to a SEL programme that was not part of the standard curriculum and was therefore listed as "active". TAU SEL programmes may include psychoeducational elements that are like those found in the SBMPs and, in many cases, are likely to be better controls for confounding variables than "active" controls given a dissimilar task, such as a dialogic reading programme (Shlomov et al., 2023). This lack of consistent comparison exemplifies the tension between the ecological validity of conducting studies on school-based interventions and the capacity to control confounding variables. It is considered critical for complex intervention studies to provide detailed descriptions of control conditions moving forward to support the replicability of findings.

Given the considerable heterogeneity in how SBMPs are operationalised, a key focus for research moving forward is the isolation of the specific mindfulness components that support the development of executive functioning. Furthermore, it remains to be established whether the mindfulness-specific components have a greater impact on EF-related outcomes than more parsimonious approaches such as relaxation techniques. For example, the negative intervention effect for an inhibition-related (g=-0.27) outcome reported by Lassander et al. (2020) may be linked to this study's more specific isolation of mindfulness components by developing a relaxation programme for their control condition. Most studies did not use a truly 'active' control condition. Therefore, it is difficult to ascertain whether the mindfulness components significantly improve outcomes over and above less resource-intensive interventions.

5.4.2 Statistical analysis

This meta-analysis used multilevel modelling to ensure that the assumption of independence is not violated. However, information on how dependency was accounted for was lacking in 60% of the included studies, which all received a reduced WoE B rating. SBMPs are nested within school systems, and most studies used pre-existing classes to allocate participants to experimental or control conditions. Students within the same class are likely to exhibit more similarities than students between classes due to their shared environment and experiences, introducing a level of dependence. Future research on SBMPs, indeed any school-based intervention study using whole classes as groups, would benefit from incorporating random effects multilevel modelling to more accurately estimate intervention effects.

Another statistical issue found across studies was the use of multiple outcome measures, which increases the risk of Type I errors. This was especially true in studies with hypotheses related to EF outcomes and other areas of functioning, such as well-being. In some instances, no attempt was made to control this increased risk of a false positive by correcting the familywise error rate (FWER). Although studies were penalised for this using the WoE criteria, future research should attempt to mitigate the risk of Type I error by statistically correcting for FWER.

5.4.3 Fidelity of Implementation

Despite the complexity of mindfulness interventions and classroom environments, most included studies did not report implementation fidelity in detail. This was a significant contributor to reduced WoE A scores. Especially for studies that use custom SBMPs, it is critical to provide a detailed intervention protocol to support the replicability of the intervention. This was not done consistently by the included

studies, and whilst word limits may limit the extent to which interventions can be described in the journal article, supplementary materials should be made available to support more comprehensive application of the SBMP to non-research contexts.

There was inconsistent reporting of adherence to intervention protocols across all studies, reducing the relevance of the evidence base to support EPs in the design and implementation of SBMPs within their professional contexts. Over half the studies included in this review did not report adherence measures, although those that did generally reported high levels of adherence. However, these were often derived from teacher reports (e.g. Brann et al., 2022), which were considered a suboptimal measure of adherence, especially in studies using teacher-administered SBMPs where the risk of self-reporting bias and social desirability bias may be higher. Inconsistent adherence to intervention protocols may explain why total intervention dosage was found to be a non-significant predictor of SBMP effectiveness, and future research should endeavour to include a measure of adherence, ideally using independent coders.

Most studies (76.7%) did not report how much participants enjoyed the SBMP and felt it was worthwhile. This limited reporting on acceptability raises concerns regarding the extent to which the implementation of SBMPs can be confidently recommended to schools. The lowest CYP acceptability rate (Lam and Seiden, 2020) is notably derived from one of the few studies that reported a negative effect size. This hints at a link between acceptability rates and SBMP effectiveness and is unsurprising, given that the benefits of mindfulness are hypothesised to be best realised when practices become embedded in daily routines (Amundsen et al., 2020). As such, the limited attempt to collect the views of CYP regarding their experiences of SBMPs is a significant gap in the literature base that needs to be addressed. Although most studies reporting acceptability were favourable, it is an ethical responsibility for EPs to ensure that recommendations are based on a holistic understanding of CYP: small-moderate improvements in executive functioning would not justify an intervention with low acceptability that may cause iatrogenic harm in other areas of functioning.

5.4.4 Outcome measures

Seven studies included a single measure of EF-related processes, receiving a reduced WoE B rating. Studies that included multiple measures designed to target specific cognitive processes associated with EF were considered to have the most suitable breadth and depth to answer the research questions in this review, especially when these included both PBTs and BRSs.

Although an overreliance on BRSs in the research literature has been criticised as "myopic" (Bergomi, 2012), this study found that they were useful as measures of the far-transfer effects of SBMPs in real-world situations. These BRSs were most relevant to the research aims when they used standardised approaches specifically designed to measure executive functioning, such as the BRIEF-2. However, many studies used BRSs more oriented towards clinical screening, which was considered suboptimal. Although inattention and hyperactivity scales have been used as a proxy for attention and inhibition, respectively, these scales often aim to capture broader constructs. For example, hyperactivity is associated with behavioural and cognitive issues that make the inhibitory elements of these results more difficult to disentangle. Behaviours such as fidgeting may be captured on a hyperactivity scale, although this does not necessarily imply cognitive inhibition. In many cases, fidgeting can be employed as a self-regulatory tool to enhance response inhibition (Farley et al., 2013).

In general, PBT the construct validity of PBTs included in the meta-analysis was well-reported and typically high. These measures had greater specificity than BRSs concerning their reported ability to target specific ECPs. The high percentage of PBT measures counteracted criticism of the research field on the inherent subjectivity in rating scales. However, this may have come at the cost of ecological validity. For example, the Stroop Task is frequently used to measure response inhibition (n=2) but inhibiting whether a word or a colour is verbalised has limited relevance or significance to everyday life. The process of inhibition in the Stroop Task context is likely different from inhibition, which occurs under more emotionally charged or stressful conditions. The lack of task salience in traditional PBT measures limits the generalisability of these results to real-world contexts, especially as the hot/cool distinction (Zelazo & Muller, 2014) may significantly mediate SBMP effectiveness.

5.4.5 Assessing executive function

The heterogeneity in measures of executive function and the potential lack of concept validity of said measures are significant concerns for the present study and limit the interpretation of the findings. This issue is not unique to the present review but reflects a broader challenge in EF research, where the conceptualisation and operationalisation of EF are highly variable (Baggetta & Alexander, 2016; Karr et al., 2018).

Interpretation of the differential SBMP effects on ECPs is limited by an outcome measure's capacity to isolate separable cognitive processes. For example, the "Head, Toes, Knees, and Shoulders" (HTKS) Task, used by Brann et al. (2023) and Zelazo et al. (2018), requires children to perform the opposite action to

verbal instructions and is often used as a measure of response inhibition. However, the HTKS task is also argued to involve working memory and cognitive flexibility skills, as new rules are sequentially introduced and must be stored and applied effectively. This multidimensional nature of the HTKS Task suggests that it may be more representative of a higher-order, general EF construct rather than a specific ECP. This concern is reflective of the broader debate in the literature regarding EF's dual unity and diversity. This study has proposed using a dynamic model (Zink et al., 2023) to account for this conceptual lack of clarity. Still, it remains to be seen how this model can be effectively operationalised.

The limitations in concept validity and task salience directly affect educational psychology practice in evaluating SBMPs and how EPs assess EF in their practice. This often involves using standardised measures such as the BRIEF-2 (Gioia et al., 2015). However, the findings from this review indicate that such measures may not be sufficient to capture EF as a complex, emergent state, especially in older CYP. Perhaps performance-based tasks that include emotional components, such as the EFN-back task, are better indicators of executive function as they emerge in meaningful school contexts. However, introducing "emotional" elements may be unethical if they carry a risk of causing psychological distress. This could be mitigated through naturalistic observation of CYP in various environments using an EF-related coding protocol. This might be an effective strategy for individual assessments but is unlikely to be feasible for intervention studies, which might benefit from advances in virtual reality to simulate real-world scenarios or use ecological momentary assessment (EMA) (Shiffman et al., 2008) for older CYP.

5.5 Conclusion

From this research, fresh perspectives on the complex relationship between SBMPs and executive function have emerged regarding the differential effects of SBMPs on EF-related cognitive processes, the potential moderating role of age, and the influence of intervention characteristics and research design.

Combining the narrative synthesis of 30 studies and a meta-analysis of 69 EF-related outcomes from 27 controlled intervention studies, the results suggest that SBMPs have a small-to-moderate positive effect on executive functioning. There is also some evidence to suggest a differential effect on separable ECPs, with SBMPs appearing to have a significantly greater impact on working memory and a significantly smaller impact on cognitive flexibility. These findings have extended the evidence base for SBMPs by providing a more granular understanding of how they influence EF-related cognitive processes and are moderated by

differential developmental trajectories. These results align with a metacognitive model of mindfulness (Holas & Jankowski, 2014) that posits that mindfulness practices may engage ECPs to varying degrees.

SBMP effects were also found to be differentially moderated by age, with younger children showing increased benefits for inhibition and cognitive flexibility outcomes. The results are also broadly supportive of a dynamic model of EF as an emergent property of lower-order processes (Zink et al., 2023) and that their distinct developmental trajectories may reveal sensitive periods where the neural networks underlying ECPs are more plastic and therefore malleable to influence from intervention. This more nuanced understanding highlights the importance of tailoring interventions to developmental stage and the value of integrating mindfulness into Early Years education. Educational Psychologists are key in advocating, developing, and implementing developmentally appropriate mindfulness interventions to support executive functioning. Longitudinal studies would be instrumental to further explore the complex interplay between mindfulness practices, neurocognitive development, and meaningful outcomes for CYP.

The critical appraisal and narrative synthesis of the studies included in this review have revealed limitations in the current evidence base. SBMP intervention studies varied considerably in their intervention characteristics, which made it difficult to readily compare effectiveness across studies and hindered attempts to isolate the specific mindfulness components that lead to improvements in EF-related outcomes. These outcomes were not consistently applied, and research exploring separable ECPs is limited by a lack of construct validity in both the PBT and BRS measures used to operationalise EF.

Moderator analyses suggest that methodological quality may lead to inflated estimates of SBMPs' effectiveness and underscore a call to future research to adopt more rigorous approaches, including the use of active controls, follow-up data collection, and full reporting of intervention protocols and fidelity measures. By addressing these concerns, a robust and generalisable evidence base for SBMPs can be developed, ultimately supporting schools in implementing effective interventions to support cognitive skills critical to school and life success.

Despite these challenges, these results offer evidence of SBMPs' potential to support the development of EF in CYP above and beyond the modest effects currently reported in the literature. By targeting specific ECPs and age groups, the benefits of mindfulness interventions can be enhanced. Educational psychologists have a critical role in contributing to our rapidly evolving understanding of executive function and mindfulness through applied research and dissemination of knowledge to schools and families.

Ultimately, an EP's capacity to affect change in school systems and improve outcomes for all stakeholders depends on their ability to translate quality research into complex environments. The present review therefore advocates a commitment to evidence-based practice built on rigorous empirical research.

References

Ahmed, S. F., Ellis, A., Ward, K. P., Chaku, N., & Davis-Kean, P. E. (2022). Working memory development from early childhood to adolescence using two nationally representative samples. Developmental Psychology, 58(10), 1962.

Allen, M., Dietz, M., Blair, K. S., van Beek, M., Rees, G., Vestergaard-Poulsen, P., ... & Roepstorff, A. (2012). Cognitive-affective neural plasticity following active-controlled mindfulness intervention. Journal of Neuroscience, 32(44), 15601-15610.

Ames, C. S., Richardson, J., Payne, S., Smith, P., & Leigh, E. (2014). Mindfulness-based cognitive therapy for depression in adolescents. Child and Adolescent Mental Health, 19(1), 74-78.

Amundsen, R., Riby, L. M., Hamilton, C., Hope, M., & McGann, D. (2020). Mindfulness in primary school children as a route to enhanced life satisfaction, positive outlook and effective emotion regulation. BMC psychology, 8, 1-15.

Anālayo, B. (2019a). Adding historical depth to definitions of mindfulness. Current opinion in psychology, 28, 11-14.

Anālayo, B. (2019b). Mindfulness-Based Interventions and the Four Satipaṭṭhānas. Mindfulness, 10(4), 611-615.

Andreu, C. I., García-Rubio, C., Melcón, M., Schonert-Reichl, K. A., & Albert, J. (2023). The effectiveness of a school mindfulness-based intervention on the neural correlates of inhibitory control in children at risk: A randomized control trial. Developmental Science, e13403.

Baggetta, P., & Alexander, P. A. (2016). Conceptualization and operationalization of executive function. Mind, Brain, and Education, 10(1), 10-33.

Baelen, R. N., Gould, L. F., Felver, J. C., Schussler, D. L., & Greenberg, M. T. (2023). Implementation reporting recommendations for school-based mindfulness programs. Mindfulness, 14(2), 255-278.

Baena-Extremera, A., del Mar Ortiz-Camacho, M., Marfil-Sánchez, A. M., & Granero-Gallegos, A. (2021). Improvement of attention and stress levels in students through a mindfulness intervention program. Revista de Psicodidáctica (English ed.), 26(2), 132-142.

Bauer, M. S., & Kirchner, J. (2020). Implementation science: What is it and why should I care?. Psychiatry research, 283, 112376.

Baujat, Bertrand, Cédric Mahé, Jean-Pierre Pignon, and Catherine Hill. 2002. “A Graphical Method for Exploring Heterogeneity in Meta-Analyses: Application to a Meta-Analysis of 65 Trials.” Statistics in Medicine 21 (18): 2641–52.

Bentley, N., Hartley, S., & Bucci, S. (2019). Systematic review of self-report measures of general mental health and wellbeing in adolescent mental health. Clinical Child and Family Psychology Review, 22(2), 225-252.

Bergomi, C., Tschacher, W., & Kupper, Z. (2013). The assessment of mindfulness with self-report measures: Existing scales and open issues. Mindfulness, 4, 191-202.

Berlin, L., Bohlin, G., Nyberg, L., & Janols, L. O. (2004). How well do measures of inhibition and other executive functions discriminate between children with ADHD and controls?. Child Neuropsychology, 10(1), 1-13.

Berti, S., & Cigala, A. (2022). Mindfulness for preschoolers: Effects on prosocial behavior, self-regulation and perspective taking. Early Education and Development, 33(1), 38-57.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. Child development, 81(6), 1641-1660.

Bigelow, F. J., Clark, G. M., Lum, J. A., & Enticott, P. G. (2021). The development of neural responses to emotional faces: A review of evidence from event-related potentials during early and middle childhood. Developmental cognitive neuroscience, 51, 100992.

Bishop, S. R., Lau, M., Shapiro, S., Carlson, L., Anderson, N. D., Carmody, J., ... & Devins, G. (2004). Mindfulness: A proposed operational definition. Clinical Psychology: Science and Practice, 11(3), 230-241.

Bhaskar, R. (2020). Critical realism and the ontology of persons. Journal of Critical Realism, 19(2), 113-120.

Blair, C., & Raver, C. C. (2015). School readiness and self-regulation: A developmental psychobiological approach. Annual Review of Psychology, 66, 711-731.

Bögels, S. M., Hellemans, J., van Deursen, S., Römer, M., & van der Meulen, R. (2014). Mindful parenting in mental health care: Effects on parental and child psychopathology, parental stress, parenting, coparenting, and marital functioning. Mindfulness, 5(5), 536-551.

Booth, A. (2006). Clear and present questions: Formulating questions for evidence-based practice. Library Hi Tech, 24(3), 355-368.Booth, A., Noyes, J., Flemming, K., Moore, G., Tunçalp, Ö., & Shakibazadeh, E. (2019).

Booth, A., Noyes, J., Flemming, K., Moore, G., Tunçalp, Ö., & Shakibazadeh, E. (2019). Formulating questions to explore complex interventions within qualitative evidence synthesis. BMJ global health, 4(Suppl 1), e001107.

Borenstein, M. (2022). Comprehensive meta-analysis software. Systematic reviews in health research: meta-analysis in context, 535-548.

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

Bracken, S. (2010). Discussing the importance of ontology and epistemology awareness in practitioner research. Worcester Journal of learning and teaching, (4).

Brann, L. S., Razza, R. A., & Smith, C. S. (2024). The Feasibility and Preliminary Effectiveness of a Mindfulness Intervention on Preschooler's Executive Function and Eating Behaviors. Early Education and Development, 35(2), 395-412.

British Psychological Society. (2017). Practice Guidelines. https://www.bps.org.uk/sites/bps.org.uk/files/Policy/Policy%20%20Files/BPS%20Practice%20Guidelines%20(Third%20Edition).pdf

British Psychological Society. (2023). Standards for the accreditation of doctoral programmes in educational psychology: England, Wales and Northern Ireland. Leicester, UK: BPS.

Bronfenbrenner, U., & Morris, P. A. (2007). The bioecological model of human development. In Handbook of child psychology: Theoretical models of human development (Vol. 1, pp. 793-828). John Wiley & Sons.

Brydges, C. R., Fox, A. M., Reid, C. L., & Anderson, M. (2014). The differentiation of executive functions in middle and late childhood: A longitudinal latent-variable analysis. Intelligence, 47, 34-43.

Buttelmann, F., & Karbach, J. (2017). Development and plasticity of cognitive flexibility in early and middle childhood. Frontiers in psychology, 8, 1040.

Calkins, S. D., & Marcovitch, S. (2010). Emotion regulation and executive functioning in early development: Integrated mechanisms of control supporting adaptive functioning.Camos, V., & Barrouillet, P. (2018). Working memory in development. Routledge.

Cao, Y., Huang, T., Huang, J., Xie, X., & Wang, Y. (2020). Effects and moderators of computer-based training on children's executive functions: a systematic review and meta-Analysis. Frontiers in Psychology, 11, 580329.Carsley, D., Khoury, B., & Heath, N. L. (2018). Effectiveness of mindfulness interventions for mental health in schools: A comprehensive meta-analysis. Mindfulness, 9, 693-707.

Cásedas, L., Pirruccio, V., Vadillo, M. A., & Lupiáñez, J. (2020). Does mindfulness meditation training enhance executive control? A systematic review and meta-analysis of randomized controlled trials in adults. Mindfulness, 11, 411-424.

Cherkin, D. C., Sherman, K. J., Balderson, B. H., Cook, A. J., Anderson, M. L., Hawkes, R. J., ... & Turner, J. A. (2016). Effect of mindfulness-based stress reduction vs cognitive behavioral therapy or usual care on back pain and functional limitations in adults with chronic low back pain: A randomized clinical trial. JAMA, 315(12), 1240-1249.

Cheung, M. W. L. (2019). A guide to conducting a meta-analysis with non-independent effect sizes. Neuropsychology review, 29(4), 387-396.

Chiesa, A., & Serretti, A. (2009). Mindfulness-based stress reduction for stress management in healthy people: a review and meta-analysis. The journal of alternative and complementary medicine, 15(5), 593-600.

Clark, C., Prior, M., & Kinsella, G. (2002). The relationship between executive function abilities, adaptive behaviour, and academic achievement in children with externalising behaviour problems. Journal of Child Psychology and Psychiatry, 43(6), 785-796.

Cohen Kadosh, K., Heathcote, L. C., & Lau, J. Y. (2014). Age-related changes in attentional control across adolescence: how does this impact emotion regulation capacities?. Frontiers in Psychology, 5, 111.

Colombo, J. (2002). Infant attention grows up: The emergence of a developmental cognitive neuroscience perspective. Current Directions in Psychological Science, 11(6), 196-200.

Cope, L. M., Hardee, J. E., Martz, M. E., Zucker, R. A., Nichols, T. E., & Heitzeg, M. M. (2020). Developmental maturation of inhibitory control circuitry in a high-risk sample: a longitudinal fMRI study. Developmental Cognitive Neuroscience, 43, 100781.

Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. Trends in neuroscience and education, 3(2), 63-68.

Crescentini, C., Capurso, V., Furlan, S., & Fabbro, F. (2016). Mindfulness-oriented meditation for primary school children: Effects on attention and psychological well-being. Frontiers in psychology, 7, 805.

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

Crocker, T. F., Lam, N., Jordão, M., Brundle, C., Prescott, M., Forster, A., ... & Clegg, A. (2023). Risk-of-bias assessment using RoB2 was useful but challenging and resource-intensive: observations from a systematic review. Journal of clinical epidemiology.

Crone, E. A., Wendelken, C., Donohue, S., van Leijenhorst, L., & Bunge, S. A. (2006). Neurocognitive development of the ability to manipulate information in working memory. Proceedings of the National Academy of Sciences, 103(24), 9315-9320.

Crooks, C. V., Bax, K., Delaney, A., Kim, H., & Shokoohi, M. (2020). Impact of MindUP among young children: Improvements in behavioral problems, adaptive skills, and executive functioning. Mindfulness, 11, 2433-2444.

Cunningham, W. A., Zelazo, P. D., Packer, D. J., & Van Bavel, J. J. (2007). The iterative reprocessing model: A multilevel framework for attitudes and evaluation. Social Cognition, 25(5), 736-760.

Dajani, D. R., & Uddin, L. Q. (2015). Demystifying cognitive flexibility: Implications for clinical and developmental neuroscience. Trends in neurosciences, 38(9), 571-578.

Davidson, R. J., & Kaszniak, A. W. (2015). Conceptual and methodological issues in research on mindfulness and meditation. American Psychologist, 70(7), 581-592.

Davidson, R. J., & McEwen, B. S. (2012). Social influences on neuroplasticity: Stress and interventions to promote well-being. Nature Neuroscience, 15(5), 689-695.

Davis, T. S. (2012). Mindfulness-based approaches and their potential for educational psychology practice. Educational psychology in practice, 28(1), 31-46.

Diamond, A. (2020). Executive functions. In Handbook of clinical neurology (Vol. 173, pp. 225-240). Elsevier.

Doebel, S. (2020). Rethinking executive function and its development. Perspectives on Psychological Science, 15(4), 942-956.

Doebel, S., & Müller, U. (2023). The Future of Research on Executive Function and Its Development: An Introduction to the Special Issue. Journal of Cognition and Development, 24(2), 161-171.

Doborjeh, Z., Doborjeh, M., Taylor, T., Kasabov, N., Wang, G. Y., Siegert, R., & Sumich, A. (2019). Spiking neural network modelling approach reveals how mindfulness training rewires the brain. Scientific reports, 9(1), 6367.

Donaldson, S. I., Heshmati, S., Lee, J. Y., & Donaldson, S. I. (2021). Examining building blocks of well-being beyond PERMA and self-report bias. The Journal of Positive Psychology, 16(6), 811-818.

Duckworth, A. L., Milkman, K. L., & Laibson, D. (2018). Beyond willpower: Strategies for reducing failures of self-control. Psychological Science in the Public Interest, 19(3), 102-129.

Duncan, L. G., & Bardacke, N. (2010). Mindfulness-based childbirth and parenting education: Promoting family mindfulness during the perinatal period. Journal of Child and Family Studies, 19(2), 190-202.

Dunning, D. L., Griffiths, K., Kuyken, W., Crane, C., Foulkes, L., Parker, J., & Dalgleish, T. (2019). Research Review: The effects of mindfulness-based interventions on cognition and mental health in children

and adolescents–a meta-analysis of randomized controlled trials. Journal of Child Psychology and Psychiatry, 60(3), 244-258.

Dunning, D. L., et al. (2019). The relation of mindfulness and executive functioning in early childhood: A meta-analysis. Child Development, 90(6), 2143-2156.

Dunning, D., Ahmed, S., Foulkes, L., Griffin, C., Griffiths, K., Leung, J.T., Parker, J., Pi-Sunyer, B.P., Sakhardande, A., Bennett, M. and Haag, C., 2022. The impact of mindfulness training in early adolescence on affective executive control, and on later mental health during the COVID-19 pandemic: a randomised controlled trial. BMJ Ment Health, 25(3), pp.110-116.

Dunning, D., Tudor, K., Radley, L., Dalrymple, N., Funk, J., Vainre, M., ... & Dalgleish, T. (2022). Do mindfulness-based programmes improve the cognitive skills, behaviour and mental health of children and adolescents? An updated meta-analysis of randomised controlled trials. BMJ Ment Health, 25(3), 135-142.

Ecker, U. K., Lewandowsky, S., Oberauer, K., & Chee, A. E. (2010). The components of working memory updating: an experimental decomposition and individual differences. Journal of Experimental Psychology: Learning, Memory, and Cognition, 36(1), 170.

Eisenberg, J. M. (2002). Globalize the evidence, localize the decision: evidence-based medicine and international diversity. Health affairs, 21(3), 166-168.

Ellis, L. K., & Rothbart, M. K. (2001, April). Revision of the early adolescent temperament questionnaire. In Poster presented at the 2001 biennial meeting of the society for research in child development, Minneapolis, Minnesota.

EPPI-Centre. (2023). EPPI-Reviewer: Systematic review software (Version 6). UCL Social Research Institute. https://eppi.ioe.ac.uk/cms/er4/

Ergas, O., & Hadar, L. L. (2019). Mindfulness in and as education: A map of a developing academic discourse from 2002 to 2017. Review of Education, 7(3), 757-797.

Evans, S., Ling, M., Hill, B., Rinehart, N., Austin, D., & Sciberras, E. (2018). Systematic review of meditation-based interventions for children with ADHD. European Child & Adolescent Psychiatry, 27, 9-27.

Farley, J., Risko, E. F., & Kingstone, A. (2013). Everyday attention and lecture retention: the effects of time, fidgeting, and mind wandering. Frontiers in psychology, 4, 619.

Farrell, P., Woods, K., Lewis, S., Rooney, S., Squires, G., & O'Connor, M. (2006). A review of the functions and contribution of educational psychologists in England and Wales in light of "Every Child Matters: Change for Children". Nottingham: DfES Publications.

Felver, J. C., Celis-de Hoyos, C. E., Tezanos, K., & Singh, N. N. (2016). A systematic review of mindfulness-based interventions for youth in school settings. Mindfulness, 7, 34-45.

Felver, J. C., Doerner, E., Jones, J., Kaye, N. C., & Merrell, K. W. (2013). Mindfulness in school psychology: Applications for intervention and professional practice. Psychology in the Schools, 50(6), 531-547.

Ferguson, H. J., Brunsdon, V. E., & Bradford, E. E. (2021). The developmental trajectories of executive function from adolescence to old age. Scientific reports, 11(1), 1382.

Flook, L., Smalley, S. L., Kitil, M. J., Galla, B. M., Kaiser-Greenland, S., Locke, J., ... & Kasari, C. (2010). Effects of mindful awareness practices on executive functions in elementary school children. Journal of applied school psychology, 26(1), 70-95.

Folch, A., Gasol, L., Heredia, L., Vicens, P., & Torrente, M. (2021). Mindful schools: Neuropsychological performance after the implementation of a mindfulness-based structured program in the school setting. Current Psychology, 1-11.

Follmer, D. J. (2018). Executive function and reading comprehension: A meta-analytic review. Educational Psychologist, 53(1), 42-60.

Francis, G. (2012). Publication bias and the failure of replication in experimental psychology. Psychonomic Bulletin & Review, 19, 975-991.

Frank, J. L., Broderick, P. C., Oh, Y., Mitra, J., Kohler, K., Schussler, D. L., ... & Greenberg, M. T. (2021). The effectiveness of a teacher-delivered mindfulness-based curriculum on adolescent social-emotional and executive functioning. Mindfulness, 12, 1234-1251.

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. Cortex, 86, 186-204.

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.

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: a behavioral genetic analysis. Developmental psychology, 47(5), 1410.

Furlong, J., & Oancea, A. (2007). Assessing quality in applied and practice-based research in education: continuing the debate. Research Papers in Education, 22(2), 115-118.

Galante, J., Friedrich, C., Dawson, A. F., Modrego-Alarcón, M., Gebbing, P., Delgado-Suárez, I., ... & Jones, P. B. (2021). Mindfulness-based programmes for mental health promotion in adults in nonclinical settings: A systematic review and meta-analysis of randomised controlled trials. PLoS Medicine, 18(1), e1003481.

Gallant, S. N. (2016). Mindfulness meditation practice and executive functioning: Breaking down the benefit. Consciousness and cognition, 40, 116-130.

García-Rubio, C., & Jarillo, T. L. (2017). El programa" Crecer respirando". In Bienestar emocional y mindfulness en la educación (pp. 259-273). Alianza Editorial.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. Psychological bulletin, 134(1), 31.

Gemescui, M. (2019). Mindfulness-Based Couples and Family Therapy: a Literature Review. Journal of Experiential Psychotherapy/Revista de PSIHOterapie Experientiala, 22(3).

Genet, J. J., & Siemer, M. (2011). Flexible control in processing affective and non-affective material predicts individual differences in trait resilience. Cognition and emotion, 25(2), 380-388.

Geronimi, E. M., Arellano, B., & Woodruff-Borden, J. (2020). Relating mindfulness and executive function in children. Clinical child psychology and psychiatry, 25(2), 435-445.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. Exceptional children, 71(2), 149-164.

Gioia, G. A., Andrwes, K., & Isquith, P. K. (1996). Behavior rating inventory of executive function-preschool version (BRIEF-P). Odessa, FL: Psychological Assessment Resources.

Gioia, G. A., Collins, M., & Isquith, P. K. (2008). Improving identification and diagnosis of mild traumatic brain injury with evidence: psychometric support for the acute concussion evaluation. The Journal of head trauma rehabilitation, 23(4), 230-242.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2015). BRIEF: Behavior rating inventory of executive function.

Goldberg, S. B., Hanley, A. W., Baldwin, S. A., Bernstein, A., & Garland, E. L. (2020). Does mindfulness practice promote psychological functioning or is it the other way around? A daily diary study. Psychotherapy, 57(3), 310.

Goldin, P. R., & Gross, J. J. (2010). Effects of mindfulness-based stress reduction (MBSR) on emotion regulation in social anxiety disorder. Emotion, 10(1), 83.

Good, D. J., Lyddy, C. J., Glomb, T. M., Bono, J. E., Brown, K. W., Duffy, M. K., ... & Lazar, S. W. (2016). Contemplating mindfulness at work: An integrative review. Journal of management, 42(1), 114-142.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. Journal of child psychology and psychiatry, 38(5), 581-586.

Gonsalves, B. D., & Cohen, N. J. (2010). Brain imaging, cognitive processes, and brain networks. Perspectives on Psychological Science, 5(6), 744-752.

Gopnik, A., O'Grady, S., Lucas, C. G., Griffiths, T. L., Wente, A., Bridgers, S., ... & Dahl, R. E. (2017). Changes in cognitive flexibility and hypothesis search across human life history from childhood to adolescence to adulthood. Proceedings of the National Academy of Sciences, 114(30), 7892-7899.

Gough, D. (2007). Weight of evidence: A framework for the appraisal of the quality and relevance of evidence. Research Papers in Education, 22(2), 213-228.

Gough, D., Thomas, J., & Oliver, S. (2012). Clarifying differences between review designs and methods. Systematic reviews, 1(1), 1-9.

Gough, D. (2021). Appraising evidence claims. Review of Research in Education, 45(1), 1-26.

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. Journal of psychopathology and behavioral assessment, 26, 41-54.

Gu, J., Strauss, C., Bond, R., & Cavanagh, K. (2015). How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of mediation studies. Clinical psychology review, 37, 1-12.

Guise, J. M., Butler, M. E., Chang, C., Viswanathan, M., Pigott, T., Tugwell, P., & Workgroup, C. I. (2017). AHRQ series on complex intervention systematic reviews—paper 6: PRISMA-CI extension statement and checklist. Journal of clinical epidemiology, 90, 43-50.

Guise, J. M., Chang, C., Butler, M., Viswanathan, M., & Tugwell, P. (2017). AHRQ series on complex intervention systematic reviews—paper 1: an introduction to a series of articles that provide guidance and tools for reviews of complex interventions. Journal of clinical epidemiology, 90, 6-10.

Gupta, R., & Kar, B. R. (2009). Development of attentional processes in ADHD and normal children. Progress in brain research, 176, 259-276.

Haddaway, N. R., Page, M. J., Pritchard, C. C., & McGuinness, L. A. (2022). PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis Campbell Systematic Reviews, 18, e1230.

Harrer, M., Cuijpers, P., Furukawa, T. & Ebert, D. D. (2019). dmetar: Companion R Package For The Guide 'Doing Meta-Analysis in R'. R package version 0.1.0. URL http://dmetar.protectlab.org/.

Harrer, M., Cuijpers, P., Furukawa, T., & Ebert, D. (2021). Doing meta-analysis with R: A hands-on guide. Chapman and Hall/CRC.

Health and Care Professions Council. (2023). Standards of proficiency for practitioner psychologists. https://www.hcpc-uk.org/globalassets/resources/standards/standards-of-proficiency---practitioner-psychologists.pdf

Heyvaert, M., Maes, B., & Onghena, P. (2013). Mixed methods research synthesis: definition, framework, and potential. Quality & Quantity, 47, 659-676.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. Statistics in medicine, 21(11), 1539-1558.

Higgins, J. P., Savović, J., Page, M. J., Elbers, R. G., & Sterne, J. A. (2019). Assessing risk of bias in a randomized trial. Cochrane handbook for systematic reviews of interventions, 205-228.

Hilpert, J. C., & Marchand, G. C. (2018). Complex systems research in educational psychology: Aligning theory and method. Educational psychologist, 53(3), 185-202.

Himi, S.A., Bühner, M. and Hilbert, S., 2021. Advancing the understanding of the factor structure of executive functioning. Journal of Intelligence, 9(1), p.16.

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., ... & Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. Bmj, 348.

Hölzel, B. K., Lazar, S. W., Gard, T., Schuman-Olivier, Z., Vago, D. R., & Ott, U. (2011). How does mindfulness meditation work? Proposing mechanisms of action from a conceptual and neural perspective. Perspectives on psychological science, 6(6), 537-559.

Holas, P., & Jankowski, T. (2013). A cognitive perspective on mindfulness. International Journal of Psychology, 48(3), 232-243.

Hofmann, S. G., & Gómez, A. F. (2017). Mindfulness-based interventions for anxiety and depression. Psychiatric clinics, 40(4), 739-749.

Holas, P., & Jankowski, T. (2013). A cognitive perspective on mindfulness. International Journal of Psychology, 48(3), 232-243.

Holmboe, K., Bonneville-Roussy, A., Csibra, G., & Johnson, M. H. (2018). Longitudinal development of attention and inhibitory control during the first year of life. Developmental science, 21(6), e12690.

Hong, Q. N., Pluye, P., Bujold, M., & Wassef, M. (2017). Convergent and sequential synthesis designs: implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. Systematic reviews, 6(1), 1-14.

Hunnius, S., & Geuze, R. H. (2004). Developmental changes in visual scanning of dynamic faces and abstract stimuli in infants: A longitudinal study. Infancy, 6(2), 231-255.

Hutchinson, J. K., Huws, J. C., & Dorjee, D. (2018). Exploring experiences of children in applying a school-based mindfulness programme to their lives. Journal of Child and Family Studies, 27(12), 3935-3951.

Ibbotson, P. (2023). The development of executive function: Mechanisms of change and functional pressures. Journal of Cognition and Development, 24(2), 172-190.

Igelström, E., Campbell, M., Craig, P., & Katikireddi, S. V. (2021). Cochrane's risk of bias tool for non-randomized studies (ROBINS-I) is frequently misapplied: a methodological systematic review. Journal of Clinical Epidemiology, 140, 22-32.

Isbel, B., & Summers, M. J. (2017). Distinguishing the cognitive processes of mindfulness: Developing a standardised mindfulness technique for use in longitudinal randomised control trials. Consciousness and cognition, 52, 75-92.

Hwang, Y. S., Goldstein, H., Medvedev, O. N., Singh, N. N., Noh, J. E., & Hand, K. (2019). Mindfulness-based intervention for educators: Effects of a school-based cluster randomized controlled study. Mindfulness, 10, 1417-1436.

Jankowski, T., & Holas, P. (2014). Metacognitive model of mindfulness. Consciousness and cognition, 28, 64-80.

Janz P, Dawe S, & Wyllie M. (2019). Mindfulness-Based Program Embedded Within the Existing Curriculum Improves Executive Functioning and Behavior in Young Children: A Waitlist Controlled Trial. FRONTIERS IN PSYCHOLOGY, 10.

Jennings, P. A., Doyle, S., Oh, Y., Rasheed, D., Frank, J. L., & Brown, J. L. (2019). Long-term impacts of the CARE program on teachers' self-reported social and emotional competence and well-being. Journal of School Psychology, 76, 186-202.

John, Y. J., Caldwell, L., McCoy, D. E., & Braganza, O. (2023). Dead rats, dopamine, performance metrics, and peacock tails: proxy failure is an inherent risk in goal-oriented systems. Behavioral and Brain Sciences, 1-68.

Juliano, A. C., Alexander, A. O., DeLuca, J., & Genova, H. (2020). Feasibility of a school-based mindfulness program for improving inhibitory skills in children with autism spectrum disorder. Research in developmental disabilities, 101, 103641.

Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: past, present, and future.

Kabat-Zinn, J. (2009). Wherever you go, there you are: Mindfulness meditation in everyday life. Hachette UK.

Kabat-Zinn, J. (2013). Full catastrophe living, revised edition: how to cope with stress, pain and illness using mindfulness meditation. Hachette UK.

Kamm, K., Thelen, E., & Jensen, J. L. (1990). A dynamical systems approach to motor development. Physical therapy, 70(12), 763-775.

Karmiloff-Smith, A. (1992). Nature, nurture and PDP: Preposterous developmental postulates?. Connection Science, 4(3-4), 253-269.

Karmiloff-Smith, A. (2017). From constructivism to neuroconstructivism: The activity-dependent structuring of the human brain. In After Piaget (pp. 1-14). Routledge.

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. Psychological bulletin, 144(11), 1147.

Kang H, An SC, Kim NO, Sung M, Kang Y, Lee U, & Yang HJ. (2020). Meditative Movement Affects Working Memory Related to Neural Activity in Adolescents: A Randomized Controlled Trial. FRONTIERS IN PSYCHOLOGY, 11.

Kay, D. A., & Anglin, J. M. (1982). Overextension and underextension in the child's expressive and receptive speech. Journal of Child Language, 9(1), 83-98.

Kempster, S., & Parry, K. W. (2011). Grounded theory and leadership research: A critical realist perspective. The leadership quarterly, 22(1), 106-120.

Khoury, B., Sharma, M., Rush, S. E., & Fournier, C. (2015). Mindfulness-based stress reduction for healthy individuals: A meta-analysis. Journal of psychosomatic research, 78(6), 519-528.

Klatt, M., Norre, C., Reader, B., Yodice, L., & White, S. (2017). Mindfulness in motion: a mindfulness-based intervention to reduce stress and enhance quality of sleep in Scandinavian employees. Mindfulness, 8, 481-488.

Kloo, D., & Sodian, B. (2017). The developmental stability of inhibition from 2 to 5 years. British Journal of Developmental Psychology, 35(4), 582-595.

Koncz A, Koeteles F, Demetrovics Z, & Takacs ZK. (2021). Benefits of a Mindfulness-Based Intervention upon School Entry: A Pilot Study. INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH, 18(23).

Koncz, A., Kassai, R., Demetrovics, Z., & Takacs, Z. K. (2022). Short Mindfulness-Based Relaxation Training Has No Effects on Executive Functions but May Reduce Baseline Cortisol Levels of Boys in First Grade: A Pilot Study. Children, 9(2), 203.

Korecki, J. R., Schwebel, F. J., Votaw, V. R., & Witkiewitz, K. (2020). Mindfulness-based programs for substance use disorders: A systematic review of manualized treatments. Substance Abuse Treatment, Prevention, and Policy, 15, 1-37.

Kuyken, W., Weare, K., Ukoumunne, O. C., Vicary, R., Motton, N., Burnett, R., ... & Huppert, F. (2013). Effectiveness of the Mindfulness in Schools Programme: non-randomised controlled feasibility study. The British Journal of Psychiatry, 203(2), 126-131.

Kuyken, W., Nuthall, E., Byford, S., Crane, C., Dalgleish, T., Ford, T., ... & Williams, J. M. G. (2017). The effectiveness and cost-effectiveness of a mindfulness training programme in schools compared with normal school provision (MYRIAD): study protocol for a randomised controlled trial. Trials, 18(1), 1-17.

Kuyken, W., Ball, S., Crane, C., Ganguli, P., Jones, B., Montero-Marin, J., ... & MYRIAD Team. (2022). Effectiveness and cost-effectiveness of universal school-based mindfulness training compared with normal school provision in reducing risk of mental health problems and promoting well-being in adolescence: the MYRIAD cluster randomised controlled trial. BMJ Ment Health, 25(3), 99-109.

Klatt, M. D., Buckworth, J., & Malarkey, W. B. (2009). Effects of low-dose mindfulness-based stress reduction (MBSR-ld) on working adults. Health Education & Behavior, 36(3), 601-614.

Klingbeil, D. A., Renshaw, T. L., Willenbrink, J. B., Copek, R. A., Chan, K. T., Haddock, A., ... & Clifton, J. (2017). Mindfulness-based interventions with youth: A comprehensive meta-analysis of group-design studies. Journal of school psychology, 63, 77-103.

Kruger, H., & Kruger, J. L. (2017). Cognition and reception. The handbook of translation and cognition, 71-89.

Lam K & Seiden D. (2020). Effects of a Brief Mindfulness Curriculum on Self-reported Executive Functioning and Emotion Regulation in Hong Kong Adolescents. MINDFULNESS, 11(3), 627–642.

Lassander M, Hintsanen M, Suominen S, Mullola S, Fagerlund A, Vahlberg T, & Volanen SM. (2020). The Effects of School-based Mindfulness Intervention on Executive Functioning in a Cluster Randomized Controlled Trial. DEVELOPMENTAL NEUROPSYCHOLOGY, 45(7–8), 469–484.

Lawlor, M. S. (2014). Mindfulness in practice: Considerations for implementation of mindfulness-based programming for adolescents in school contexts. New directions for youth development, 2014(142), 83-95.

Lee, K., Bull, R., & Ho, R. M. (2013). Developmental changes in executive functioning. Child development, 84(6), 1933-1953.

Lee, J. W., Jones, P. S., Mineyama, Y., & Zhang, X. E. (2002). Cultural differences in responses to a Likert scale. Research in nursing & health, 25(4), 295-306.

Lefebvre, C., Glanville, J., Wieland, L. S., Coles, B., & Weightman, A. L. (2013). Methodological developments in searching for studies for systematic reviews: past, present and future?. Systematic reviews, 2(1), 1-9.

Lerner, M. D., & Lonigan, C. J. (2014). Executive function among preschool children: Unitary versus distinct abilities. Journal of psychopathology and behavioral assessment, 36, 626-639.

Lertladaluck, K., Suppalarkbunlue, W., Moriguchi, Y., & Chutabhakdikul, N. (2021). School-based mindfulness intervention improves executive functions and self-regulation in preschoolers at risk. The Journal of Behavioral Science, 16(2), 58-72.

Levi, U., & Rosenstreich, E. (2019). Mindfulness and memory: A review of findings and a potential model. Journal of Cognitive Enhancement, 3(3), 302-314.

Li, W., Howard, M. O., Garland, E. L., McGovern, P., & Lazar, M. (2017). Mindfulness treatment for substance misuse: A systematic review and meta-analysis. Journal of substance abuse treatment, 75, 62-96.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., ... & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Annals of internal medicine, 151(4), W-65.

Lodha, S., & Gupta, R. (2022). Mindfulness, attentional networks, and executive functioning: A review of interventions and long-term meditation practice. Journal of Cognitive Enhancement, 6(4), 531-548.

Luiselli, J. K., Worthen, D., Carbonell, L., & Queen, A. H. (2017). Social validity assessment of mindfulness education and practices among high school students. Journal of applied school psychology, 33(2), 124-135.

Lutz, A., Jha, A. P., Dunne, J. D., & Saron, C. D. (2015). Investigating the phenomenological matrix of mindfulness-related practices from a neurocognitive perspective. American Psychologist, 70(7), 632.

MacKay, T., & Lindsay, G. (2015). Guest Editorial: Universal psychology and systemic approaches to practice. Educational & Child Psychology, 32(1), 7.

Magalhaes S, Nunes T, Soeiro I, Rodrigues R, Coelho A, Pinheiro M, Castro SL, Leal T, & Limpo T. (2022). A Pilot Study Testing the Effectiveness of a Mindfulness-Based Program for Portuguese School Children. MINDFULNESS, 13(11), 2751–2764.

Mak, C., Whittingham, K., Cunnington, R., & Boyd, R. N. (2018). Efficacy of mindfulness-based interventions for attention and executive function in children and adolescents—A systematic review. Mindfulness, 9, 59-78.

Makmee, P. (2022). Increasing Attention and Working Memory in Elementary Students Using Mindfulness Training Programs. FWU Journal of Social Sciences, 16(3), 107-119.

Manuel, J. A., Somohano, V. C., & Bowen, S. (2017). Mindfulness practice and its relationship to the Five-Facet Mindfulness Questionnaire. Mindfulness, 8, 361-367.

Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M., Thomas, M. S., & Westermann, G. (2007). Neuroconstructivism-I: How the brain constructs cognition. Oxford University Press.

Masten, A. S., Herbers, J. E., Desjardins, C. D., Cutuli, J. J., McCormick, C. M., Sapienza, J. K., ... & Zelazo, P. D. (2012). Executive function skills and school success in young children experiencing homelessness. Educational Researcher, 41(9), 375-384.

Maynard, B. R., Solis, M. R., Miller, V. L., & Brendel, K. E. (2017). Mindfulness-based interventions for improving cognition, academic achievement, behavior, and socioemotional functioning of primary and secondary school students. Campbell systematic reviews, 13(1), 1-144.

McAlister, A., & Peterson, C. (2007). A longitudinal study of child siblings and theory of mind development. Cognitive Development, 22(2), 258-270.

Meixner, T., Irwin, A., Wolfe Miscio, M., Cox, M., Woon, S., McKeough, T., & Milligan, K. (2019). Delivery of integra mindfulness martial arts in the secondary school setting: Factors that support successful implementation and strategies for navigating implementation challenges. School Mental Health, 11, 549-561.

Meltzer, L. (Ed.). (2018). Executive function in education: From theory to practice. Guilford Publications.

Messer, D., Kearvell-White, J., Danielsson, H., Faulkner, D., Henry, L., & Ibbotson, P. (2022). The structure of executive functioning in 11 to 14 year olds with and without special educational needs. British Journal of Developmental Psychology, 40(3), 453-470.

Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., & Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. BMC health services research, 14(1), 1-10.

Mettler, J., Khoury, B., Zito, S., Sadowski, I., & Heath, N. L. (2023). Mindfulness-based programs and school adjustment: A systematic review and meta-analysis. Journal of school psychology, 97, 43-62.

Milaré, C. A., Kozasa, E. H., Lacerda, S., Barrichello, C., Tobo, P. R., & Horta, A. L. D. (2021). Mindfulness-based versus story Reading intervention in public elementary schools: Effects on executive functions and emotional health. Frontiers in Psychology, 12, 576311.

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.

Moerman, D. E., & Jonas, W. B. (2002). Deconstructing the placebo effect and finding the meaning response. Annals of Internal medicine, 136(6), 471-476.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., ... & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. Proceedings of the national Academy of Sciences, 108(7), 2693-2698.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., ... & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic reviews, 4(1), 1-9.

Moon, K., & Blackman, D. (2014). A guide to understanding social science research for natural scientists. Conservation biology, 28(5), 1167-1177.

Moriguchi, Y., & Phillips, S. (2023). Evaluating the distinction between cool and hot executive function during childhood. Brain Sciences, 13(2), 313.

Morra, S., Panesi, S., Traverso, L., & Usai, M. C. (2018). Which tasks measure what? Reflections on executive function development and a commentary on Podjarny, Kamawar, and Andrews (2017). Journal of Experimental Child Psychology, 167, 246-258.

Muller C, Otto B, Sawitzki V, Kanagalingam P, Scherer JS, & Lindberg S. (2021). Short breaks at school: Effects of a physical activity and a mindfulness intervention on children's attention, reading comprehension, and self-esteem. TRENDS IN NEUROSCIENCE AND EDUCATION, 25.

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O'Reilly, R. C. (2011). A unified framework for inhibitory control. Trends in cognitive sciences, 15(10), 453-459.

Newman, M., & Gough, D. (2020). Systematic reviews in educational research: Methodology, perspectives and application. Systematic reviews in educational research: Methodology, perspectives and application, 3-22.

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.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. Psychological bulletin, 126(2), 220.

Nigg, J. T. (2017). Annual Research Review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. Journal of child psychology and psychiatry, 58(4), 361-383.

Oh-uchi, A., Kawahara, J. I., & Sugano, L. (2010). Attentional capture and metaattentional judgment: a study of young children, parents, and university students. Psychologia, 53(2), 114-124.

Panikratova, Y. R., Vlasova, R. M., Akhutina, T. V., Korneev, A. A., Sinitsyn, V. E., & Pechenkova, E. V. (2020). Functional connectivity of the dorsolateral prefrontal cortex contributes to different components of executive functions. International Journal of Psychophysiology, 151, 70-79.

Perone, S., Simmering, V. R., & Buss, A. T. (2021). A dynamical reconceptualization of executive-function development. Perspectives on Psychological Science, 16(6), 1198-1208.

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. Annual review of neuroscience, 35, 73-89.

Petticrew, M., & Roberts, H. (2008). Systematic reviews in the social sciences: A practical guide. John Wiley & Sons.

Petticrew, M. (2015). Time to rethink the systematic review catechism? Moving from 'what works' to 'what happens'. Systematic reviews, 4(1), 1-6.

Petticrew, M., Knai, C., Thomas, J., & Rehfuess, E. (2019). Implications of a complexity perspective for systematic reviews and guideline development in health decision making. BMJ Global Health, 4(Suppl 1), e000899.

Phan, M. L., Renshaw, T. L., Caramanico, J., Greeson, J. M., MacKenzie, E., Atkinson-Diaz, Z., ... & Nuske, H. J. (2022). Mindfulness-based school interventions: A systematic review of outcome evidence quality by study design. Mindfulness, 13(7), 1591-1613.

Pickerell, L. E., Pennington, K., Cartledge, C., Miller, K. A., & Curtis, F. (2023). The effectiveness of school-based mindfulness and cognitive behavioural programmes to improve emotional regulation in 7–12-year-olds: A systematic review and meta-analysis. Mindfulness, 14(5), 1068-1087.

Pierson, S., Goto, K., Giampaoli, J., Hart, S., & Wylie, A. (2019). Impacts of a mindful eating intervention on healthy food-related behaviors and mindful eating practices among elementary school children. Californian Journal of Health Promotion, 17(2), 41-50.

Polanin, J. R., Espelage, D. L., Grotpeter, J. K., Valido, A., Ingram, K. M., Torgal, C., ... & Robinson, L. E. (2020). Locating unregistered and unreported data for use in a social science systematic review and meta-analysis. Systematic Reviews, 9, 1-9.

Poldrack, R. A. (2011). Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. Neuron, 72(5), 692-697.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. Annual review of neuroscience, 13(1), 25-42.

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., ... & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. A product from the ESRC methods programme Version, 1(1), b92.

Posner, M. I., & Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. Annu. Rev. Psychol., 58, 1-23.

Price, Z. (2023). Exploring teachers', teaching assistants' and Educational Psychologists' understanding of Executive Functions and reasonable adjustments that work to support children, in the classroom, with executive function difficulties (Doctoral dissertation, UCL (University College London)).

Puljak, L., Ramic, I., Naharro, C. A., Brezova, J., Lin, Y. C., Surdila, A. A., ... & Salvado, M. S. (2020). Cochrane risk of bias tool was used inadequately in the majority of non-Cochrane systematic reviews. Journal of clinical epidemiology, 123, 114-119.

Purser, R. E., & Milillo, J. (2015). Mindfulness revisited: A Buddhist-based conceptualization. Journal of Management Inquiry, 24(1), 3-24.

Quach Dianna, Jastrowski Mano, Kristen E, & Alexander Kristi. (2016). A Randomized Controlled Trial Examining the Effect of Mindfulness Meditation on Working Memory Capacity in Adolescents. Journal of Adolescent Health, 58(5), 489–496.

Rea, S. E. (2021). The Effect of School-Based Mindfulness Intervention on Student Attention and Executive Functioning: A Meta-Analysis (Doctoral dissertation, Duquesne University).

Reangsing, C., Lauderman, C., & Schneider, J. K. (2022). Effects of mindfulness meditation intervention on depressive symptoms in emerging adults: a systematic review and meta-analysis. Journal of Integrative and Complementary Medicine, 28(1), 6-24.

Ricarte J J, Ros L, Latorre J M, & Beltrán M T. (2015). Mindfulness-Based Intervention in a Rural Primary School: Effects on Attention, Concentration and Mood. International Journal of Cognitive Therapy, 8(3), 258–270.

Rodgers, M., Sowden, A., Petticrew, M., Arai, L., Roberts, H., Britten, N., & Popay, J. (2009). Testing methodological guidance on the conduct of narrative synthesis in systematic reviews: effectiveness of interventions to promote smoke alarm ownership and function. Evaluation, 15(1), 49-73.

Roeser, R. W., Galla, B., & Baelen, R. N. (2022). Mindfulness in schools: Evidence on the impacts of school-based mindfulness programs on student outcomes in P–12 educational settings.

Roeser, R. W., Greenberg, M. T., Frazier, T., Galla, B. M., Semenov, A. D., & Warren, M. T. (2023). Beyond all splits: Envisioning the next generation of science on mindfulness and compassion in schools for students. Mindfulness, 14(2), 239-254.

Ropovik, I., Adamkovic, M., & Greger, D. (2021). Neglect of publication bias compromises meta-analyses of educational research. Plos one, 16(6), e0252415.

Rosen, M. L., Hagen, M. P., Lurie, L. A., Miles, Z. E., Sheridan, M. A., Meltzoff, A. N., & McLaughlin, K. A. (2020). Cognitive stimulation as a mechanism linking socioeconomic status with executive function: A longitudinal investigation. Child development, 91(4), e762-e779.

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. Child development, 72(5), 1394-1408.

Salmoirago-Blotcher Elena, Druker Susan, Meleo-Meyer Florence, Frisard Christine, Crawford Sybil, & Pbert Lori. (2019). Beneficial Effects of School-based Mindfulness Training on Impulsivity in Healthy Adolescents: Results from a Pilot Randomized Controlled Trial. Explore: The Journal of Science & Healing, 15(2), 160–164.

Scammacca, N., Roberts, G., & Stuebing, K. K. (2014). Meta-analysis with complex research designs: Dealing with dependence from multiple measures and multiple group comparisons. Review of educational research, 84(3), 328-364.

Schauer, J. M., & Hedges, L. V. (2021). Reconsidering statistical methods for assessing replication. Psychological Methods, 26(1), 127.

Schonert-Reichl Kimberly A, Oberie Eva, Stewart Lawlor, Molly, Abbott David, Thomson Kimberly, Oberlander Tim F, & Diamond Adele. (2015). Enhancing Cognitive and Social-Emotional Development Through a Simple-to-Administer Mindfulness-Based School Program for Elementary School Children: A Randomized Controlled Trial. Developmental Psychology, 51(1), 52–66.

Schweizer, S., Gotlib, I. H., & Blakemore, S. J. (2020). The role of affective control in emotion regulation during adolescence. Emotion, 20(1), 80.

Sedó, M. A. (2004). '5 digit test': a multilinguistic non-reading alternative to the Stroop test. Revista de neurologia, 38(9), 824-828.

Shapiro, S. L., Carlson, L. E., Astin, J. A., & Freedman, B. (2006). Mechanisms of mindfulness. Journal of clinical psychology, 62(3), 373-386.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. Annu. Rev. Clin. Psychol., 4, 1-32.

Shlomov, I., Levit-Binnun, N., & Horowitz-Kraus, T. (2023). Neurodevelopmental Effects of a Mindfulness and Kindness Curriculum on Executive Functions in Preschool Children—A Randomized, Active-Controlled Study. Mind, Brain, and Education, 17(2), 132-148.

Sims, S., Anders, J., Inglis, M., & Lortie-Forgues, H. (2022). Quantifying "promising trials bias" in randomized controlled trials in education. Journal of Research on Educational Effectiveness, 1-18.

Spiegel, J. A., Goodrich, J. M., Morris, B. M., Osborne, C. M., & Lonigan, C. J. (2021). Relations between executive functions and academic outcomes in elementary school children: A meta-analysis. Psychological bulletin, 147(4), 329.

Snel, E. (2015). Breathe through this: mindfulness for parents of teenagers. Shambhala Publications.

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.

Steinberg, L. (2005). Cognitive and affective development in adolescence. Trends in cognitive sciences, 9(2), 69-74.

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., ... & Higgins, J. P. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. bmj, 355.

Sterne, J. A., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., ... & Higgins, J. P. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. bmj, 366.

Sumantry, D., & Stewart, K. E. (2021). Meditation, mindfulness, and attention: A meta-analysis. Mindfulness, 12, 1332-1349.

Suarez-Garcia Z, Alvarez-Garcia D, Garcia-Redondo P, & Rodriguez C. (2020). The Effect of a Mindfulness-Based Intervention on Attention, Self-Control, and Aggressiveness in Primary School Pupils. International Journal of Environmental Research and Public Health, 17(7).

Takacs, Z. K., & Kassai, R. (2019). The efficacy of different interventions to foster children's executive function skills: A series of meta-analyses. Psychological bulletin, 145(7), 653.

Tang, Y. Y., Yang, L., Leve, L. D., & Harold, G. T. (2012). Improving executive function and its neurobiological mechanisms through a mindfulness-based intervention: Advances within the field of developmental neuroscience. Child development perspectives, 6(4), 361-366.

Tang, Y. Y., Hölzel, B. K., & Posner, M. I. (2015). The neuroscience of mindfulness meditation. Nature reviews neuroscience, 16(4), 213-225.

Taylor, S. J., Barker, L. A., Heavey, L., & McHale, S. (2015). The longitudinal development of social and executive functions in late adolescence and early adulthood. Frontiers in behavioral neuroscience, 9, 252.

Thierry, K. L., Bryant, H. L., Nobles, S. S., & Norris, K. S. (2016). Two-year impact of a mindfulness-based program on preschoolers' self-regulation and academic performance. Early Education and Development, 27(6), 805-821.

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. BMC medical research methodology, 8(1), 1-10.

Thomas, G., & Atkinson, C. (2016). Measuring the effectiveness of a mindfulness-based intervention for children's attentional functioning. Educational & Child Psychology, 33(1), 51-64.

Thomas, E. M., & Centeio, E. E. (2020). The benefits of yoga in the classroom: A mixed-methods approach to the effects of poses and breathing and relaxation techniques. International Journal of Yoga, 13(3), 250.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct?. Journal of child psychology and psychiatry, 54(2), 131-143.

Tooley, U. A., Park, A. T., Leonard, J. A., Boroshok, A. L., McDermott, C. L., Tisdall, M. D., ... & Mackey, A. P. (2022). The age of reason: Functional brain network development during childhood. Journal of Neuroscience, 42(44), 8237-8251.

Turoman, N., Tivadar, R. I., Retsa, C., Maillard, A. M., Scerif, G., & Matusz, P. J. (2021). The development of attentional control mechanisms in multisensory environments. Developmental cognitive neuroscience, 48, 100930.

Tudge, J. R., Mokrova, I., Hatfield, B. E., & Karnik, R. B. (2009). Uses and misuses of Bronfenbrenner's bioecological theory of human development. Journal of family theory & review, 1(4), 198-210.

Tudge, J. R., Payir, A., Merçon-Vargas, E., Cao, H., Liang, Y., Li, J., & O'Brien, L. (2016). Still misused after all these years? A reevaluation of the uses of Bronfenbrenner's bioecological theory of human development. Journal of Family Theory & Review, 8(4), 427-445.

Vago, D. R., & Silbersweig, D. A. (2012). Self-awareness, self-regulation, and self-transcendence (S-ART): a framework for understanding the neurobiological mechanisms of mindfulness. Frontiers in human neuroscience, 6, 296.

Van Dam, N. T., Van Vugt, M. K., Vago, D. R., Schmalzl, L., Saron, C. D., Olendzki, A., ... & Meyer, D. E. (2018). Mind the hype: A critical evaluation and prescriptive agenda for research on mindfulness and meditation. Perspectives on psychological science, 13(1), 36-61.

Veloso, A. S., Vicente, S. G., & Filipe, M. G. (2022). Assessment of 'cool'and 'hot'executive skills in children with ADHD: the role of performance measures and behavioral ratings. European Journal of Investigation in Health, Psychology and Education, 12(11), 1657-1672.

Verburgh, L., Königs, M., Scherder, E. J., & Oosterlaan, J. (2014). Physical exercise and executive functions in preadolescent children, adolescents and young adults: a meta-analysis. British journal of sports medicine, 48(12), 973-979.

Verhaeghen, P. (2021). Mindfulness as attention training: Meta-analyses on the links between attention performance and mindfulness interventions, long-term meditation practice, and trait mindfulness. Mindfulness, 12, 564-581.

Vickery CE & Dorjee D. (2016). Mindfulness Training in Primary Schools Decreases Negative Affect and Increases Meta-Cognition in Children. FRONTIERS IN PSYCHOLOGY, 6.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. Journal of statistical software, 36, 1-48.

Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. Research synthesis methods, 1(2), 112-125.

Völter, C. J., Reindl, E., Felsche, E., Civelek, Z., Whalen, A., Lugosi, Z., ... & Seed, A. M. (2022). The structure of executive functions in preschool children and chimpanzees. Scientific Reports, 12(1), 6456.

Vu, M. C., & Burton, N. (2023). Beyond the Inclusion–Exclusion Binary: Right Mindfulness and Its Implications for Perceived Inclusion and Exclusion in the Workplace. Journal of Business Ethics, 1-19.

Weare, K. (2012). Evidence for the impact of mindfulness on children and young people. The mindfulness in schools project in association with mood disorders centre.

Weare, K. (2023). Where have we been and where are we going with mindfulness in schools?. Mindfulness, 14(2), 293-299.

Wiebe, S. A., Sheffield, T. D., & Espy, K. A. (2012). Separating the fish from the sharks: A longitudinal study of preschool response inhibition. Child development, 83(4), 1245-1261.

Weisz, J. R., Kuppens, S., Eckshtain, D., Ugueto, A. M., Hawley, K. M., & Jensen-Doss, A. (2013). Performance of evidence-based youth psychotherapies compared with usual clinical care: A multilevel meta-analysis. JAMA psychiatry, 70(7), 750-761.

Welsh, M., & Peterson, E. (2014). Issues in the conceptualization and assessment of hot executive functions in childhood. Journal of the International Neuropsychological Society, 20(2), 152-156.

Wimmer L, Bellingrath S, & von Stockhausen L. (2016). Cognitive Effects of Mindfulness Training: Results of a Pilot Study Based on a Theory Driven Approach. FRONTIERS IN PSYCHOLOGY, 7.

Willoughby, M. T., & Blair, C. B. (2016). Measuring executive function in early childhood: A case for formative measurement. Psychological assessment, 28(3), 319.

Willoughby, M., Kupersmidt, J., Voegler-Lee, M., & Bryant, D. (2011). Contributions of hot and cool self-regulation to preschool disruptive behavior and academic achievement. Developmental neuropsychology, 36(2), 162-180.

Wood, L., Roach, A. T., Kearney, M. A., & Zabek, F. (2018). Enhancing executive function skills in preschoolers through a mindfulness-based intervention: A randomized, controlled pilot study. Psychology in the Schools, 55(6), 644-660.

Young, K. S., van der Velden, A. M., Craske, M. G., Pallesen, K. J., Fjorback, L., Roepstorff, A., & Parsons, C. E. (2018). The impact of mindfulness-based interventions on brain activity: A systematic review of functional magnetic resonance imaging studies. Neuroscience & Biobehavioral Reviews, 84, 424-433.

Zelazo, P. D. (2015). Executive function: Reflection, iterative reprocessing, complexity, and the developing brain. Developmental Review, 38, 55-68.

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. Child development perspectives, 6(4), 354-360.

Zelazo, P. D., & Carlson, S. M. (2020). The neurodevelopment of executive function skills: Implications for academic achievement gaps. Psychology & Neuroscience, 13(3), 273.

Zelazo, P. D., & Carlson, S. M. (2023). Reconciling the context-dependency and domain-generality of executive function skills from a developmental systems perspective. Journal of Cognition and Development, 24(2), 205-222.

Zelazo, P. D., Forston, J. L., Masten, A. S., & Carlson, S. M. (2018). Mindfulness plus reflection training: Effects on executive function in early childhood. Frontiers in psychology, 9, 208.

Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood: VI. The development of executive function: Cognitive complexity and control--revised. Monographs of the Society for Research in Child Development.

Zoogman, S., Goldberg, S. B., Hoyt, W. T., & Miller, L. (2015). Mindfulness interventions with youth: A meta-analysis. Mindfulness, 6, 290-302.

Zhang, D., Lee, E. K., Mak, E. C., Ho, C. Y., & Wong, S. Y. (2021). Mindfulness-based interventions: an overall review. British medical bulletin, 138(1), 41-57.

Zhang, X. F., Li, R. N., Deng, J. L., Chen, X. L., Zhou, Q. L., Qi, Y., ... & Fan, J. M. (2023). Effects of mindfulness-based interventions on cardiovascular risk factors: An umbrella review of systematic reviews and meta-analyses. Journal of Psychosomatic Research, 111586.

Zhou, X., Guo, J., Lu, G., Chen, C., Xie, Z., Liu, J., & Zhang, C. (2020). Effects of mindfulness-based stress reduction on anxiety symptoms in young people: A systematic review and meta-analysis. Psychiatry Research, 289, 113002.

Appendices

Appendix A

Articles excluded at full text screening

Excluded Articles	Exclusion Criteria and Rationale
Alampay, L. P., Galvez Tan, L. J. T., Tuliao, A. P., Baranek, P., Ofreneo, M. A., Lopez, G. D., ... & Guintu, V. (2020). A pilot randomized controlled trial of a mindfulness program for Filipino children. *Mindfulness*, *11*, 303-316.	8. Evaluation. The primary outcomes of this study were related to "depression, anxiety, and emotion dysregulation".
Baijal, S., Jha, A. P., Kiyonaga, A., Singh, R., & Srinivasan, N. (2011). The influence of concentrative meditation training on the development of attention networks during early adolescence. *Frontiers in Psychology*, *2*, 153.	6. Control. Participants attended schools where transcendental meditation had been part of the curriculum for years.
Bigelow, H., Gottlieb, M. D., Ogrodnik, M., Graham, J. D., & Fenesi, B. (2021). The differential impact of acute exercise and mindfulness meditation on executive functioning and psycho-emotional well-being in children and youth with ADHD. *Frontiers in Psychology*, *12*, 660845.	4. Setting. The intervention was administered in a non-educational setting.
Britton, W. B., Lepp, N. E., Niles, H. F., Rocha, T., Fisher, N. E., & Gold, J. S. (2014). A randomized controlled pilot trial of classroom-based mindfulness meditation compared to an active control condition in sixth-grade children. *Journal of School Psychology*, *52*(3), 263-278.	8. Evaluation. The primary outcomes of the study related to "clinical syndromes, affect, and mindfulness".
Brown, K. W., Goodman, R. J., Ryan, R. M., & Anālayo, B. (2016). Mindfulness enhances episodic memory performance: Evidence from a multimethod investigation. *PloS one*, *11*(4), e0153309.	5. Perspective. Participants were university Psychology students.

Cordeiro, C., Magalhães, S., Rocha, R., Mesquita, A., Olive, T., Castro, S. L., & Limpo, T. (2021). Promoting third graders' executive functions and literacy: A pilot study examining the benefits of mindfulness vs. relaxation training. *Frontiers in Psychology*, *12*, 643794.	4. Setting. The intervention was administered by psychologists in a non-educational setting.
Culang, H., Davis, E.R., Egan, L., Simmons, A., Vago, D.R., Finley, S. and Dennis-Tiwary, T., 2021. Changing Minds: A Pilot Feasibility Study of Mindfulness Training for At-Risk Adolescents. *National Youth Advocacy and Resilience Journal*, *5*(1), pp.3-26.	7. Control. Study did not include a control condition.
Di Girolamo, F., Fabris, V., Casadio, M., Capodieci, A., & Carretti, B. (2019). Mindfulness Based Attention Training for Children (MBATC). Un nuovo approccio per potenziare l'attenzione nei bambini di scuola primaria. *Psicologia clinica dello sviluppo*, *23*(3), 383-408.	3. Language. The body of the article is presented in Italian.
Dunning, D., Ahmed, S., Foulkes, L., Griffin, C., Griffiths, K., Leung, J. T., ... & MYRIAD Team. (2022). The impact of mindfulness training in early adolescence on affective executive control, and on later mental health during the COVID-19 pandemic: a randomised controlled trial. *BMJ Ment Health*, *25*(3), 110-116.	8. Evaluation. The primary outcomes of this study are related to affective executive function.
Flook, L., Goldberg, S. B., Pinger, L., & Davidson, R. J. (2015). Promoting prosocial behavior and self-regulatory skills in preschool children through a mindfulness-based Kindness Curriculum. *Developmental psychology*, *51*(1), 44.	6. Intervention. Kindness Curriculum included significant movement and gratitude elements.
Frenkel, M. O., Georg, A., & Plessner, H. (2020). Promoting Self-Regulation of Adolescents in School Through Mindfulness. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, *69*(4), 321-338.	8. Evaluation. Primary outcomes were related to mindfulness, chronic stress, and emotion regulation.
Fung, J., Kim, J. J., Jin, J., Chen, G., Bear, L., & Lau, A. S. (2019). A randomized trial evaluating school-based mindfulness intervention for ethnic minority youth: Exploring mediators and moderators of intervention effects. *Journal of abnormal child psychology*, *47*, 1-19.	8. Evaluation. Primary outcomes were related to internalising/externalising behaviour problems, stress, and emotional regulation.

García-Rubio, C., Herrero, M., Luna-Jarillo, T., Albert, J., & Rodríguez-Carvajal, R. (2023). Effectiveness and mechanisms of change of a mindfulness-based intervention on elementary school children: A cluster-randomized control trial. *Journal of School Psychology*, *99*, 101211.	8. Evaluation. Primary outcomes were related to mindfulness and wellbeing.
Greenberg, J., Romero, V. L., Elkin-Frankston, S., Bezdek, M. A., Schumacher, E. H., & Lazar, S. W. (2019). Reduced interference in working memory following mindfulness training is associated with increases in hippocampal volume. *Brain imaging and behavior*, *13*, 366-376.	4. Setting. Mindfulness intervention was administered in a clinical setting.
Jarraya, S., Jarraya, M., & Engel, F. A. (2022). Kindergarten-Based Progressive Muscle Relaxation Training Enhances Attention and Executive Functioning: A Randomized Controlled Trial. *Perceptual and Motor Skills*, *129*(3), 644-669.	6. Intervention. The progressive muscle relaxation does involve sustained attentional control but does not include an open monitoring style.
Juliano, A. C., Alexander, A. O., DeLuca, J., & Genova, H. (2020). Feasibility of a school-based mindfulness program for improving inhibitory skills in children with autism spectrum disorder. *Research in developmental disabilities*, *101*, 103641.	7. Control. Study did not include a control condition.
Kang H, An SC, Kim NO, Sung M, Kang Y, Lee U, & Yang HJ. (2020). Meditative Movement Affects Working Memory Related to Neural Activity in Adolescents: A Randomized Controlled Trial. *FRONTIERS IN PSYCHOLOGY*, *11*. https://doi.org/10.3389/fpsyg.2020.00931	6. Intervention. Based on DahnMuDo (meditative body exercises) and administered afterschool to volunteers.
Koncz Adam, Kassai Reka, Demetrovics Zsolt, & Takacs Zsofia K. (2022). Short Mindfulness-Based Relaxation Training Has No Effects on Executive Functions but May Reduce Baseline Cortisol Levels of Boys in First Grade: A Pilot Study. Children, 9(2), 203. https://doi.org/10.3390/children9020203	6. Intervention. Not universally targeted as small groups of boys only.
Lai, Y. J., & Chang, K. M. (2020). Improvement of attention in elementary school students through fixation focus training activity. *International Journal of Environmental Research and Public Health*, *17*(13), 4780.	6. Intervention. The Fixation Focus Training does not include an open monitoring style element.

Lawler, J. M., Esposito, E. A., Doyle, C. M., & Gunnar, M. R. (2019). A preliminary, randomized-controlled trial of mindfulness and game-based executive function trainings to promote self-regulation in internationally-adopted children. *Development and Psychopathology*, *31*(4), 1513-1525.	5. Perspective. Sample of internationally adopted children was deemed to introduce a potential confound after full-text reading.
Montero-Marin, J., Allwood, M., Ball, S., Crane, C., De Wilde, K., Hinze, V., ... & MYRIAD Team. (2022). School-based mindfulness training in early adolescence: what works, for whom and how in the MYRIAD trial?. *BMJ Ment Health*, *25*(3), 117-124.	8. Evaluation. This study investigated potential moderating effects on the efficacy of SBMP to improve mental health outcomes.
Moreno-Gómez, A. J., & Cejudo, J. (2019). Effectiveness of a mindfulness-based social–emotional learning program on psychosocial adjustment and neuropsychological maturity in kindergarten children. *Mindfulness*, *10*(1), 111-121.	8. Evaluation. The primary neuropsychological outcomes were not directly related to executive function (e.g. visual perception).
Oberle, E., Schonert-Reichl, K. A., Lawlor, M. S., & Thomson, K. C. (2012). Mindfulness and inhibitory control in early adolescence. *The Journal of Early Adolescence*, *32*(4), 565-588.	6. Intervention. This study explored dispositional mindfulness and did not involve a mindfulness-based intervention.
Osborne, E., Ainsworth, B., Hooper, N., & Atkinson, M. (2023). Experiences of using digital mindfulness-based interventions: Rapid scoping review and thematic synthesis. *Journal of Medical Internet Research*.	7. Control. This study used a longitudinal mediation analysis and therefore did not include a control condition.
Parker, A. E., Kupersmidt, J. B., Mathis, E. T., Scull, T. M., & Sims, C. (2014). The impact of mindfulness education on elementary school students: Evaluation of the Master Mind program. *Advances in School Mental Health Promotion*, *7*(3), 184-204.	8. Evaluation. Primary outcomes were related to self-regulation (i.e. emergent property of cognitive control and emotional arousal).
Sanger, K. L., & Dorjee, D. (2016). Mindfulness training with adolescents enhances metacognition and the inhibition of irrelevant stimuli: Evidence from event-related brain potentials. *Trends in Neuroscience and Education*, *5*(1), 1-11.	8. Evaluation. This study aims to measure the effects of an SBMP on higher order metacognitive processes, which is not readily comparable to EF.

Terjestam, Y., Bengtsson, H., & Jansson, A. (2016). Cultivating awareness at school. Effects on effortful control, peer relations and well-being at school in grades 5, 7, and 8. *School Psychology International*, *37*(5), 456-469.	8. Evaluation. The main outcome of this study was a measure of wellbeing.
Thierry, K. L., Vincent, R. L., & Norris, K. S. (2022). A Mindfulness-Based Curriculum Improves Young Children's Relationship Skills and Social Awareness. *Mindfulness*, *13*(3), 730-741.	8. Evaluation. The primary outcome of this study focused on "emotional awareness".
Vickery, C. E., & Dorjee, D. (2016). Mindfulness training in primary schools decreases negative affect and increases meta-cognition in children. *Frontiers in psychology*, *6*, 2025.	8. Evaluation. This study aims to measure the effects of an SBMP on higher order metacognitive processes, which is not readily comparable to EF.
Wimmer L, Isherwood KR, Parkinson J, & Dorjee D. (2023). Relating Dispositional Mindfulness and Long-Term Mindfulness Training with Executive Functioning, Emotion Regulation, and Well-Being in Pre-adolescents. *PSYCHOLOGICAL STUDIES*. https://doi.org/10.1007/s12646-023-00746-2	7. Control. The control group was exposed to a positive psychology intervention that included components that could be considered 'mindful'.

Appendix B

References of the 30 studies included in this review

1. Andreu, C. I., García-Rubio, C., Melcón, M., Schonert-Reichl, K. A., & Albert, J. (2023). The effectiveness of a school mindfulness-based intervention on the neural correlates of inhibitory control in children at risk: A randomized control trial. *Developmental Science*, *26*(6), e13403.
2. Baena-Extremera, A., del Mar Ortiz-Camacho, M., Sánchez, A. M. M., & Granero-Gallegos, A. (2021). Mejora de los niveles de atención y estrés en los estudiantes a través de un programa de intervención Mindfulness. *Revista de Psicodidáctica*, *26*(2), 132-142.
3. Berti, S., & Cigala, A. (2022). Mindfulness for preschoolers: Effects on prosocial behavior, self-regulation and perspective taking. *Early Education and Development*, *33*(1), 38-57.
4. Brann, L. S., Razza, R. A., & Smith, C. S. (2024). The Feasibility and Preliminary Effectiveness of a Mindfulness Intervention on Preschooler's Executive Function and Eating Behaviors. *Early Education and Development*, *35*(2), 395-412.
5. Crescentini, C., Capurso, V., Furlan, S., & Fabbro, F. (2016). Mindfulness-oriented meditation for primary school children: Effects on attention and psychological well-being. *Frontiers in psychology*, *7*, 805.
6. Crooks, C. V., Bax, K., Delaney, A., Kim, H., & Shokoohi, M. (2020). Impact of MindUP among young children: Improvements in behavioral problems, adaptive skills, and executive functioning. *Mindfulness*, *11*, 2433-2444.
7. Flook, L., Smalley, S. L., Kitil, M. J., Galla, B. M., Kaiser-Greenland, S., Locke, J., ... & Kasari, C. (2010). Effects of mindful awareness practices on executive functions in elementary school children. *Journal of applied school psychology*, *26*(1), 70-95.
8. Folch, A., Gasol, L., Heredia, L., Vicens, P., & Torrente, M. (2021). Mindful schools: Neuropsychological performance after the implementation of a mindfulness-based structured program in the school setting. *Current Psychology*, 1-11.
9. Frank, J. L., Broderick, P. C., Oh, Y., Mitra, J., Kohler, K., Schussler, D. L., ... & Greenberg, M. T. (2021). The effectiveness of a teacher-delivered mindfulness-based curriculum on adolescent social-emotional and executive functioning. *Mindfulness*, *12*(5), 1234-1251.
10. Janz, P., Dawe, S., & Wyllie, M. (2019). Mindfulness-based program embedded within the existing curriculum improves executive functioning and behavior in young children: A waitlist controlled trial. *Frontiers in psychology*, *10*, 2052.
11. Koncz, A., Köteles, F., Demetrovics, Z., & Takacs, Z. K. (2021). Benefits of a mindfulness-based intervention upon school entry: A pilot study. *International Journal of Environmental Research and Public Health*, *18*(23), 12630.

12. Lam, K., & Seiden, D. (2020). Effects of a brief mindfulness curriculum on self-reported executive functioning and emotion regulation in Hong Kong adolescents. *Mindfulness*, *11*(3), 627-642.
13. Lertladaluck, K., Suppalarkbunlue, W., Moriguchi, Y., & Chutabhakdikul, N. (2021). School-based mindfulness intervention improves executive functions and self-regulation in preschoolers at risk. *The Journal of Behavioral Science*, *16*(2), 58-72.
14. Lassander, M., Hintsanen, M., Suominen, S., Mullola, S., Fagerlund, Å., Vahlberg, T., & Volanen, S. M. (2020). The effects of school-based mindfulness intervention on executive functioning in a cluster randomized controlled trial. *Developmental neuropsychology*, *45*(7-8), 469-484.
15. Magalhães, S., Nunes, T., Soeiro, I., Rodrigues, R., Coelho, A., Pinheiro, M., ... & Limpo, T. (2022). A pilot study testing the effectiveness of a mindfulness-based program for portuguese school children. *Mindfulness*, *13*(11), 2751-2764.
16. Makmee, P. (2022). Increasing attention and working memory in elementary students using mindfulness training programs. *FWU Journal of Social Sciences*, *16*(3), 107-119.
17. Milaré, C. A., Kozasa, E. H., Lacerda, S., Barrichello, C., Tobo, P. R., & Horta, A. L. D. (2021). Mindfulness-based versus story Reading intervention in public elementary schools: Effects on executive functions and emotional health. *Frontiers in Psychology*, *12*, 576311.
18. Müller, C., Otto, B., Sawitzki, V., Kanagalingam, P., Scherer, J. S., & Lindberg, S. (2021). Short breaks at school: effects of a physical activity and a mindfulness intervention on children's attention, reading comprehension, and self-esteem. *Trends in Neuroscience and Education*, *25*, 100160.
19. Quach, D., Mano, K. E. J., & Alexander, K. (2016). A randomized controlled trial examining the effect of mindfulness meditation on working memory capacity in adolescents. *Journal of Adolescent Health*, *58*(5), 489-496.
20. Ricarte, J. J., Ros, L., Latorre, J. M., & Beltrán, M. T. (2015). Mindfulness-based intervention in a rural primary school: Effects on attention, concentration and mood. *International Journal of Cognitive Therapy*, *8*(3), 258-270.
21. Salmoirago-Blotcher, E., Druker, S., Meleo-Meyer, F., Frisard, C., Crawford, S., & Pbert, L. (2019). Beneficial effects of school-based mindfulness training on impulsivity in healthy adolescents: Results from a pilot randomized controlled trial. *Explore*, *15*(2), 160-164.
22. Schonert-Reichl, K. A., Oberle, E., Lawlor, M. S., Abbott, D., Thomson, K., Oberlander, T. F., & Diamond, A. (2015). Enhancing cognitive and social–emotional development through a simple-to-administer mindfulness-based school program for elementary school children: A randomized controlled trial. *Developmental psychology*, *51*(1), 52.
23. Shlomov, I., Levit-Binnun, N., & Horowitz-Kraus, T. (2023). Neurodevelopmental Effects of a Mindfulness and Kindness Curriculum on Executive Functions in Preschool Children—A Randomized, Active-Controlled Study. *Mind, Brain, and Education*, *17*(2), 132-148.

24. Suárez-García, Z., Álvarez-García, D., García-Redondo, P., & Rodríguez, C. (2020). The effect of a mindfulness-based intervention on attention, self-control, and aggressiveness in primary school pupils. *International journal of environmental research and public health*, *17*(7), 2447.
25. Thierry, K. L., Bryant, H. L., Nobles, S. S., & Norris, K. S. (2016). Two-year impact of a mindfulness-based program on preschoolers' self-regulation and academic performance. *Early Education and Development*, *27*(6), 805-821.
26. Thomas, G., & Atkinson, C. (2016). Measuring the effectiveness of a mindfulness-based intervention for children's attentional functioning. *Educational & Child Psychology*, *33*(1), 51-64.
27. Vickery, C. E., & Dorjee, D. (2016). Mindfulness training in primary schools decreases negative affect and increases meta-cognition in children. *Frontiers in psychology*, *6*, 2025.
28. Wimmer, L., Bellingrath, S., & von Stockhausen, L. (2016). Cognitive effects of mindfulness training: Results of a pilot study based on a theory driven approach. *Frontiers in psychology*, *7*, 1037.
29. Wood, L., Roach, A. T., Kearney, M. A., & Zabek, F. (2018). Enhancing executive function skills in preschoolers through a mindfulness-based intervention: A randomized, controlled pilot study. *Psychology in the Schools*, *55*(6), 644-660.
30. Zelazo, P. D., Forston, J. L., Masten, A. S., & Carlson, S. M. (2018). Mindfulness plus reflection training: Effects on executive function in early childhood. *Frontiers in psychology*, *9*, 208.

Appendix C

Summary of studies included in this review.

Setting	Research Design	Participants	Intervention	Implementation
Authors: Andreu et al. (2023). **Title:** The effectiveness of a school mindfulness-based intervention on the neural correlates of inhibitory control in children at risk: A randomized control trial **Country:** Chile (Santiago de Chile) **Setting:** Mainstream Primary School **Setting characteristics:** "school identified as having low socioeconomic status".	**Design:** Cluster Randomised controlled trial **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention (time not stated) **Follow Up:** n/a **Conditions:** 1. SMBP (n=24) 2. Control (n=22) completed a Skills-for-Life (SFL) programme (*Active-SEL*). **Allocation:** Cluster simple randomization based on coin flip.	**Sample Size:** 68 (46 analysed) **Age:** 9-11 years (M=9.80) **Gender:** SMBP = 62.5% female; Control = 63.6% female. **Ethnicity:** Not stated. **Participant characteristics:** Children "in a high-risk context" (Low SES).	**Mindfulness Programme:** GrowingUp Breathing (*Crecer* Respirando; García-Rubio & Luna-Jarillo, 2019). **Intervention Components:** 1. Attention Module: psychoeducation on mindfulness, attention, and attentional regulation. Core practices are "mindfulness of Breathing, mindfulness of sounds, mindful movement, body scan, and breathing space." 2. Self-Regulation Module: psychoeducation on self-regulation skills. Core practices are "mindful eating, mindfulness of thoughts and emotions, and breathing space." 3. Kindness Module: "children explore how to take care of oneself and others […] through loving-kindness and self-compassion meditation." **Theoretical Foundation:** Two-component model (Bishop et al., 2004). Mindfulness skills "increase attention and self-regulation processes" and "cultivate prosocial […] dispositions".	**Administrator**: "Experienced psychologist and mindfulness instructor" **Intervention Intensity:** 9 x 50-minute sessions (+ practice including twice-weekly mp3 audios and thrice-daily breathing space). **Intervention Frequency:** 9 weekly sessions with daily practice. **Replicability of Intervention:** Manualised programme promotes replicability; however, it has been adapted to "be implemented with children aged 9-11 years". These adaptations are not reported. **Intervention Incentives:** Headteacher received "$380 in compensation". **Adherence:** 50% of children attended all sessions, 33% missed one session, 12% missed two sessions, and 5% missed three sessions. Personal practice measured on 0-10 Likert scale (M = 6.69, SD = 3.95). **Acceptability:** Three-item 0-10 Likert scale (M = 8.45, SD = 2.78).
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Baena-Extremera et al. (2021).	**Design:** Quasi-experimental	**Sample Size:** 320	**Mindfulness Programme:** Novel programme based on "Breathe through this: Mindfulness for parents and teenagers" (Snel, 2015).	**Administrator**: Class teachers who had completed a 9-week mindfulness course.

Title: Improvement of attention and stress levels in students through a mindfulness intervention program. **Country:** Spain (Granada) **Settings:** Two state secondary schools; one state all-through school, one chartered all-through school. **Setting characteristics:** Mix of metropolitan and rural locations.	controlled pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention **Follow Up:** n/a **Conditions:** 1. SBMP (n=156) 2. Control (n=164) not exposed to intervention (*Passive-not specified).* **Allocation:** Cluster simple randomization based on coin flip.	**Age:** 10-16 years (SBMP: M=13.1; Control: M=13.0) **Gender:** Not stated. **Ethnicity:** Not stated. **Participant characteristics:** Heterogeneous SES and "family cultural level".	**Intervention Components:** Brief CD audio tracks presenting on perceiving thoughts and feelings; attentional breathing; emotional regulation; cognitive regulation; body scanning; self-acceptance; acceptance of others. Students encouraged to participate in daily homework practices of approx. 3 minutes involving breathing exercises; awareness of silence; mindful eating etc. **Theoretical Foundation:** More spiritual definition of mindfulness as being "aware of the present moment harmonizing mind and body", but drawing on Western research from MBSR, mindfulness-based CBT, and ACT. An exploration of the loss in translation from the Pali concept of *sati* into English *mindfulness,* or Spanish *atención/ conciencia* (attention/awareness).	**Intervention Intensity:** Each activity lasted between 4.45 and 16.26 minutes. Total duration of intervention is not reported. Homework tasks of approx. 3 minutes daily. **Intervention Frequency:** Daily over 6 weeks **Replicability of Intervention:** Full description following TIDIER (Hoffman et al., 2014) guidelines. Pre-recorded audio tracks facilitate replication. **Intervention Incentives:** None reported. **Adherence:** 23 participants did not complete the requisite 90% of the programme to be included in the final analysis (n=320). **Acceptability:** Not reported.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Berti & Cigala (2020) **Title:** Mindfulness for Preschoolers: Effects on Prosocial Behaviour, Self-Regulation and Perspective Taking **Country:** Italy (Northern)	**Design:** Quasi-experimental controlled pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention **Follow Up:** n/a **Conditions:** 1. SBMP (n=10)	**Sample Size:** 21 **Age:** 65.6 months **Gender:** 42.9% female **Ethnicity:** All with Italian citizenship, 90.5% with Italian parents.	**Mindfulness Programme:** Novel mindfulness-based programme **Intervention Components:** Sessions focused on exploring inner world through breathing and imagining "a peaceful garden"; emotions and body reaction; sensations including mindful eating; and connecting with others. Sessions took place in a wooden house in the preschool garden, decorated to promote comfort and intimacy. Visuals of the	**Administrator**: Instructor experienced with MBSR and working with pre-schoolers. **Intervention Intensity:** 15 sessions lasting 30 minutes each, plus nine brief (5-10 minute) meditation sessions. **Intervention Frequency:** Thrice weekly for six weeks **Replicability of Intervention:** Session contents are described in some detail

Setting: Kindergarten located in a comprehensive school complex **Setting characteristics:** Middle SES range.	2. Control (n=11) calming activities with their teachers (e.g. painting, storybooks) (*active-calming)* **Allocation:** Cluster simple randomization based on coin flip.	**Participant characteristics :** 52.4% of parents had a master's degree or higher.	mindfulness sessions were on the wall of the house. **Theoretical Foundation:** Mindfulness practices thought to activate brain areas involved in attention control, emotion regulation, self-awareness (Tang et al., 2015). MBIs share goals with SEL programs to enhance socio-emotional competencies but add focus on acceptance, resilience, compassion (Lawlor, 2016).	although no manual is provided. Fidelity not formally assessed. **Intervention Incentives:** None reported. **Adherence:** Instructor completed a grid to record CYP participation, compliance, and attention. Medium-high scores reported. **Acceptability:** Not reported.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Brann et al. (2022) **Title:** The feasibility and preliminary effectiveness of a mindfulness intervention on preschooler's executive function and eating behaviours. **Country:** United States (Northeastern) **Setting:** Two licensed childcare centres **Setting characteristics:** Similar size and demographics.	**Design:** Quasi-experimental controlled pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention (6 weeks) **Follow Up:** n/a **Conditions:** 1. SBMP (n=24) 2. Control (n=19) regular activities (*Passive-TAU).* **Allocation:** Centres were randomly assigned.	**Sample Size:** 43 **Age:** 4.4 (SD=0.6) **Gender:** 65% female **Ethnicity:** 95% White **Participant characteristics :** 94% of parents had college/grad school education and household income over $80000.	**Mindfulness Programme:** Novel mindful eating and yoga intervention. **Intervention Components:** Mindful eating focused on awareness of hunger/satiety, enjoyment of nourishing food, sensory exploration, and attunement. Yoga included age-appropriate meditations and movements for emotional literacy and mind-body connection. **Theoretical Foundation:** Draw on the work on Miyake et al. (2000) to define executive function as working memory, inhibitory control, and cognitive flexibility, as well as how these develop across the lifespan (Diamond, 2013). Mindfulness conceptualised through definition provided by Greenberg & Harris (2012) describing mindfulness as control over awareness.	**Administrator**: External experts including a dietitian, yoga instructor, and graduate student. **Intervention Intensity:** Two 30-minute sessions per week. **Intervention Frequency:** Weekly for 6 weeks (12 sessions). **Replicability of Intervention:** Activities described but no manual mentioned. Fidelity assessed via self-report logs. **Intervention Incentives:** Parents received $10. **Adherence:** All planned sessions were delivered. Average attendance of 76% for mindful eating, 64% for yoga. Some noted issues with engagement. **Acceptability:** Not reported.

Setting	Research Design	Participants	Intervention	Implementation
Authors: Crescentini et al. (2016) **Title:** Mindfulness plus reflection training: effects on executive function in early childhood **Country:** Italy **Setting:** Two Primary school classes **Setting characteristics:** Northeast Italy	**Design:** Quasi-experimental controlled pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention (8 days after) **Follow Up:** n/a **Conditions:** 1. SBMP (n=16) 2. Active (n=15) reading intervention (*Active-dissimilar).* **Allocation:** not stated	**Sample Size:** 31 **Age:** M=7.35 (SD=0.5) **Gender:** 50% female **Ethnicity:** Italian **Participant characteristics:** No children excluded for behavioural or learning difficulties.	**Mindfulness Programme:** Custom Mindfulness-oriented Meditation (MOM) training, adapted for children (shortened duration) **Intervention Components:** Sessions included three meditation exercises focused on breathing, mindfulness of body parts, and mindfulness of thoughts. Graduated increase in mediation duration over the 8-week period. Use of tools and mental images to support practice. **Theoretical Foundation:** Based on MOM interventions for adults, which is in turn adapted from Mindfulness-Based Stress Reduction (MBSR). Promotes awareness of breath, body, and thoughts with a non-judgemental orientation.	**Administrator**: Two experienced mindfulness instructors **Intervention Intensity:** Initially 30 minutes per week, building to 1.5 hours. **Intervention Frequency:** 3 sessions per week for 8 weeks **Replicability of Intervention:** Details of programme are provided but no manual linked. **Intervention Incentives:** None mentioned. **Adherence:** All children attended at least 6 of the 8 weeks. **Acceptability:** Not reported.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Crooks et al. (2020) **Title:** Impact of MIndUP among young children: Improvements in behavioural problems, adaptive skills, and executive functioning **Country:** Canada (Ontario)	**Design:** Quasi-experimental controlled pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention **Follow Up:** n/a **Conditions:** 1. SBMP (n=261)	**Sample Size:** 584 **Age:** M=4.37 years (SD=0.58) **Gender:** 51.54% female **Ethnicity:** 70.09% White **Participant characteristics:** Intervention	**Mindfulness Programme:** MindUP – A manualised mindfulness-based SEL programme. **Intervention Components:** Four units covering getting focussed (psychoeducation); sharpening your senses; it's all about attitude (positive psychology); and acting mindfully. **Theoretical Foundation:** "The MindUP curriculum is derived from psychological theory and informed by research in the fields of developmental neuroscience	**Administrator**: Classroom teachers trained for two days, one at start of programme and a midway extension day. **Intervention Intensity:** 15 lessons averaging 43 minutes each. **Intervention Frequency:** Weekly plus 3x daily core practice over the school year. **Replicability of Intervention:** Teachers followed a manualized curriculum and reported high adherence. **Intervention Incentives:** Teachers compensated for research activities.

Setting: Kindergarten classrooms **Setting characteristics:** Catholic school district.	2. Control (n=323) regular curriculum (*Passive-TAU).* **Allocation:** Non-random allocation by school district.	group had much higher social risk indicators.	(Diamond, 2009, 2012), contemplative science and mindfulness (Roeser & Zelazo, 2012), SEL (Greenberg et al., 2003), and positive psychology (Lyubomirsky, Sheldon, & Schkade, 2005)." Also covers elements of Positive Psychology through gratitude and performing acts of kindness.	**Adherence:** High self-reported adherence, but no formal measures of fidelity. **Acceptability:** Not reported.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Flook et al. (2010). **Title:** Effects of Mindful Awareness Practices on Executive Functions in Elementary School Children. **Country:** United States (Los Angeles) **Setting:** Single Elementary School **Setting characteristics:** On-campus setting.	**Design:** Quasi-experimental controlled pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention ("immediately after") **Follow Up:** n/a **Conditions:** 1. SBMP (n=32) 2. Control (n=32) Silent Reading (*Passive-Reading).* **Allocation:** Block randomization with stratification by classroom, gender, and age.	**Sample Size:** 64 **Age:** 7-9 years (M=8.23). **Gender:** 64.7% female **Ethnicity:** "45% Caucasian, 23% Latino, 14% Asian, 9% African American" **Participant characteristics :** n/a	**Mindfulness Programme:** Mindful Awareness Practices (MAPs). Custon programme developed by an author of the paper. **Intervention Components:** Programme uses games and exercises to promote awareness of sensory awareness, attentional regulation, and awareness of thoughts and feelings. Further components consider awareness of others, and awareness of environment. Most exercises involve interactions among students and with the instructor. Each session also includes meditation, lasting 3 minutes initially and increasing to 5 minutes. **Theoretical Foundation:** Based on "secular" mindfulness training for adults, adapted with age-appropriate exercises. Two-component model of mindfulness is outlined (Bishop et al., 2004) with a cognitive exploration of the links with EF.	**Administrator**: Not stated (most likely class teacher). **Intervention Intensity:** 16 sessions, 480 minutes total **Intervention Frequency:** Twice weekly, over eight weeks. **Replicability of Intervention:** Detailed appendix provided of intervention protocol, including exemplars of mindful activities. Highly replicable, although duration of sessions not explicitly stated. **Intervention Incentives:** None stated. **Adherence:** Not explicitly stated, although all participants included in data analyses. **Acceptability:** Not stated.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**

Authors: Folch et al. (2021). **Title:** Mindful schools: Neuropsychological performance after the implementation of a mindfulness-based structured program in the school setting. **Country:** Spain (Tarragona). **Setting:** Three 6th grade classes; and three 5th grade classes. **Setting characteristics:** Single primary school.	**Design:** Quasi-experimental control group pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention **Follow Up:** n/a **Conditions:** 1. SBMP with mindfulness trained teacher (n=36) 2. SBMP with non-trained teacher (n=33) 3. Control (n=31) not exposed to intervention (*Passive-None).* **Allocation:** Not stated.	**Sample Size:** 100 **Age:** 9-11 years (M=10.37) **Gender:** 64% female **Participant characteristics:** Exclusion criteria included poor school attendance.	**Mindfulness Programme:** Custom programme based on 'Inner Kids' that uses mindful games rooted in MBSR. **Intervention Components:** 1. Anchoring "in breathing, body, and movement or others. 2. "Open field/contemplative techniques" such as attending to the sounds of nature, or the use of bells as a guide. 3. Connection "with active interaction, related to positive affective states, such as loving-kindness and compassion meditations". **Theoretical Foundation:** Pragmatic justification of study based on previous SBMP research; little theoretical link made between mindfulness and executive function in this article.	**Administrator**: Teachers trained over "8 weekly 2-h sessions […] based on MBSR". **Intervention Intensity:** 65 x 5–10-minute sessions. **Intervention Frequency:** Daily over 13 weeks. **Replicability of Intervention:** Limited exploration of intervention procedure; no supplementary information relating to programme provided. "All instructions for the intervention exercises were previously audio recorded by the investigators to be listened in the classroom". **Intervention Incentives:** None stated. **Adherence:** Not stated. **Acceptability:** Not stated.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Frank et al. (2021). **Title:** The Effectiveness of a Teacher-Delivered Mindfulness-Based Curriculum on Adolescent Social-Emotional and	**Design:** Quasi-experimental control group pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention (one week after)	**Sample Size:** 255 (251 analysed) **Age:** 11th grade students (M=16 years) **Gender:** 43% female	**Mindfulness Programme:** Manualised Learning to BREATHE (L2B; Broderick & Metz, 2009) programme. **Intervention Components:** L2B covers the six themes of "Body; Reflections (thoughts); Emotions; Attention; Tenderness/take it like it is; and Habits for a healthy mind; and a final holistic 'E' representing overall goal of Empowerment.	**Administrator**: Class teacher who had received 14 hours of training in the L2B programme. During the intervention, five weekly coaching calls (each lasting 60 minutes) were provided. **Intervention Intensity:** 12 sessions of unspecified duration. **Intervention Frequency:** Twice weekly, for six weeks.

Executive Functioning. **Country:** United States (Northeast) **Setting:** Two suburban high schools **Setting characteristics:** Intervention administered during "required health education classes".	**Follow Up:** n/a **Conditions:** 1. SBMP (n=131) 2. Control (n=124) treatment-as usual SEL curriculum (*Passive-TAU).* **Allocation:** "Randomly assigned"	**Ethnicity:** Ethnically Diverse Cohort **Participant characteristics :** 23% "received free lunch" 55% "reside in two-biological-parent families".	Six sessions used PPT slides and multimedia clips to present mindfulness content. Opportunities to practice mindfulness activities were provided at the end of each session. These activities included: mindful listening; mindful eating; mindfulness of body, thoughts and emotions; drawing the effects of stress on the body etc. **Theoretical Foundation:** Integrates themes from MBSR (Kabat-Zinn, 1990) and is also informed by therapies targeting emotional regulation skills such as Acceptance and Commitment Therapy (ACT; Hays et al., 2016) and Mindfulness-Based Cognitive Therapy (Segal et al., 2013).	**Replicability of Intervention:** Homework practice procedures were digitalised (CD). Supplementary material is provided, outlining content and discussion topics, experiential activities and an overview of the curricula. Exact materials and processes used can be provide by L2B, presumably at cost. **Intervention Incentives:** Conflict of Interest with one co-author (P. Broderick) also the developer of L2B. No incentives given to participants. **Adherence:** Intervention fidelity was assessed by seven randomly allocated "independent coders". Average inter-rater coder agreement was 92%, reporting overall fidelity across sessions of 78.6%. Attrition reported at 1.6% (n=3). 66% of intervention students practiced mindfulness activities less than once a month. **Acceptability:** Not stated.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Janz et al. (2019). **Title:** Mindfulness-Based Program Embedded Within the Existing Curriculum Improves Executive Functioning and	**Design:** Quasi-experimental waitlist-controlled trial. **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention (last week of "term 4") T2= follow-up	**Sample Size:** 87 **Age:** 5-8 years (M=6.50) **Gender:** 49% female **Ethnicity:** 8% Indigenous Australian	**Mindfulness Programme:** Custom, manualised "CalmSpace" mindfulness programme, designed to be integrated into the Australian Curriculum. **Intervention Components:** Week 1: Construct of mindfulness is introduced through a children's book, "Mindful Monkey, Happy Panda".	**Administrator**: Class teachers with minimal prior mindfulness experience, all received a half-day training and CalmSpace programme manual. Weekly support also provided. **Intervention Intensity:** Three sessions of core practice daily, duration unspecified. Further optional activities, with teachers

Behavior in Young Children: A Waitlist Controlled Trial. **Country:** Australia **Setting:** A single school in Queensland **Setting characteristics:** In lowest quintile for SES nationally.	**Follow Up:** One term after intervention **Conditions:** 1. SBMP (n=51) 2. Waitlist (n=36) treatment-as usual SEL curriculum (*Passive-TAU).* **Allocation:** Cluster simple randomization based on hat draw.	**Participant characteristics :** 23% EAL speakers	A Core Practice "was undertaken three times a day in which children listened to a single resonant sound of a gong as they focused on the sound and slowed their breath". Further activities included body scans, munch and crunch time, watching clouds, rainbow walk, and glitter jar. **Theoretical Foundation:** Not explicitly stated, but mainly uses a cognitive perspective to justify benefits of intervention, following a two-component model (e.g. Bishop et al., 2004) focusing on self-regulation of attention and open orientation to experience.	reporting embedding an average of 2.5 a day. **Intervention Frequency:** Not stated. **Replicability of Intervention:** Mindfulness activities are described in supplementary material, alongside EF target of the activity. Full programme not outlined in article. Flexibility of approach makes it difficult to replicate. **Intervention Incentives:** None reported. **Adherence:** Fidelity monitored using a daily CalmSpace checklist given to teachers. Teachers reported implementing 100% of core content, as well as an average of two additional activities daily. **Acceptability:** Teacher report that children were "very attentive and engaged".

Setting	Research Design	Participants	Intervention	Implementation
Authors: Title: Koncz et al. (2021). **Title:** Benefits of a Mindfulness-Based Intervention upon School Entry: A Pilot Study. **Country:** Hungary (Budapest) **Setting:** A single primary school	**Design:** Randomised controlled pilot study **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention (one week after) **Follow Up:** n/a **Conditions:** 1. SBMP (n=31) 2. Control (n=30) (*Passive-None).*	**Sample Size:** 61 **Age:** SMBP: M=7.0, SD=.41; Control: M=7.1, SD=.46 **Gender:** 37.7% female **Ethnicity:** Not stated.	**Mindfulness Programme:** A story-based mindfulness programme, custom-developed by authors in a previous study. **Intervention Components:** This intervention uses commercially available books as stimuli to engage in mindfulness practices such as breathing meditation, muscle relaxation, sensory meditation and sitting meditation. **Theoretical Foundation:** Follows a cognitive model that somewhat rejects MBSR (e.g. mindfulness improves EF through stress reduction), and instead suggests a bottom-	**Administrator**: Trained research assistants under the supervision of a clinical child psychologist. **Intervention Intensity:** 6 x 45-minute sessions **Intervention Frequency:** Twice weekly, over three weeks **Replicability of Intervention:** Appendix provides overview of session, with links to content (mindful stories) used in each session. A link to a more detailed intervention protocol on which this

Setting characteristics: Four 1st grade classes	**Allocation:** "Full" (e.g. non-cluster) randomization via coin flip.	**Participant characteristics:** No diagnoses of "mental or somatic disorders" that could impact cortisol levels.	up processing model (Zelazo & Lyons, 2012), through the practice of monitoring attention. This study does not explicitly indicate that an 'open-monitoring style' is encouraged/developed.	study's intervention is based (same authors) is provided. **Intervention Incentives:** None stated. **Adherence:** Not stated. "No item had missing data greater than 3.5%". **Acceptability:** Not stated.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Lam & Seiden (2020). **Title:** Effects of a Brief Mindfulness Curriculum on Self-reported Executive Functioning and Emotion Regulation in Hong Kong Adolescents. **Country:** Hong Kong, PRC. **Setting:** Two state secondary schools **Setting characteristics:** Low-middle SES neighbourhood	**Design:** Quasi-experimental control group pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention (time not stated) **Follow Up:** n/a **Conditions:** 1. SBMP (n=45) 2. Control (n=51) treatment-as usual SEL curriculum (*Passive-TAU).* **Allocation:** Teacher asked to "randomly assign one academically stronger and one weaker class to each of the conditions".	**Sample Size:** 119 (initial); 96 (analytical). **Age:** 11-15 years (M=12.4) **Gender:** 34.8% female. **Ethnicity:** Chinese **Participant characteristics:** 90% aged 12-13.	**Mindfulness Programme:** Adapted, manualised Learning to BREATHE (L2B; Broderick & Metz, 2009) programme. **Intervention Components:** L2B covers the six themes of "Body; Reflections (thoughts); Emotions; Attention; Tenderness/take it like it is; and Habits for a healthy mind; and a final holistic 'E' representing overall goal of Empowerment. Each session involved psychoeducation through PPT slides and video/audio clips. This was followed by demonstration and practice of mindfulness activities. Activities included: awareness of emotions/stress; kindness to self; breathing exercises; gratitude; awareness of cognitions; stretching; attention on senses; observing emotions; and body scanning. **Theoretical Foundation:** L2B integrates themes from MBSR (Kabat-Zinn, 1990) and is also informed by therapies targeting emotional regulation skills such as Acceptance and Commitment Therapy (ACT;	**Administrator**: Clinical and school psychologist (first author) **Intervention Intensity:** 6 x 70-minute sessions; homework activity handouts after each intervention. **Intervention Frequency:** Monthly over five months. **Replicability of Intervention:** Outline of individual sessions provided, although adaptations to original L2B programme are not sufficiently described for replication. **Intervention Incentives:** None recorded. **Adherence:** 83.5% of participants had complete data. No more than two sessions were missed by any participant. 80% of students reported practicing homework at least once or twice during the programme. **Acceptability:** 52.8% of participants viewed the programme as useful. 66.0% of participants were satisfied with the programme. 41.5% of participants agreed

			Hays et al., 2016) and Mindfulness-Based Cognitive Therapy (Segal et al., 2013).	or strongly agreed that programme should be implemented in schools.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Lassander et al. (2020). **Title:** The Effects of School-based Mindfulness Intervention on Executive Functioning in a Cluster Randomized Controlled Trial. **Country:** Finland **Setting:** Four comprehensive schools **Setting characteristics:** Urban settings (Helsinki and Turku)	**Design:** Cluster randomised controlled trial **Data Collection Points:** T0= pre-intervention baseline T1= post-intervention (at 9 weeks) T2= follow-up **Follow Up:** 6 months after T0 **Conditions:** 1. SBMP (n=62) 2. Control (n=69) *Relax* programme providing "tools to relax and calm down" (*Active-SEL/relaxation).* **Allocation:** Random allocation sequence following grouping accounting for school location and SES.	**Sample Size:** 131 **Age:** Either "6th graders (median age 12)" or "8th graders (median age 15). M=13.53 **Gender:** SMBP = 56.5% female; Control = 44.9% female. **Ethnicity:** Not stated **Participant characteristics:** Universal student population (applicable to general school context).	**Mindfulness Programme:** Manualised .b (Stop & Breathe; Kuyken et al., 2013) programme. **Intervention Components:** "Nine group sessions and mindfulness practices designed to improve emotional awareness, sustained attention, and attentional and emotional regulation." "Sessions begin with an introduction to a themed lesson, including short formal or informal practices, group discussion, and ending with a longer practice." **Control Intervention Components:** 9x 45-minute sessions "aimed to produce relaxation skills and holistic wellbeing" including "relaxation exercises with practices available to listen at home." **Theoretical Foundation:** The .b programme integrates themes from MBSR and mindfulness-based cognitive therapy (MBCT).	**Administrator**: Trained and experienced facilitators who were not class teachers by profession. **Intervention Intensity:** 9 x 45 minutes sessions; and 1-15 minutes of daily practice. **Intervention Frequency:** Weekly over nine weeks. **Replicability of Intervention:** .b curriculum is a manualised with structure and materials available online. Some discrepancy between online programme and that presented here (e.g. 10 sessions (one is introductory), instead of 9 "Mindfulness interventions were available online to listen at home." **Intervention Incentives:** None stated **Adherence:** "90% of students took part in 7-9 sessions, 5% took part in 6 sessions, and 5% took part in 1-5 lessons". **Acceptability:** Not stated.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Lertladaluck et al. (2021)	**Design:** Randomised, controlled trial	**Sample Size:** 30	**Mindfulness Programme:** Custom-designed school-based mindfulness programme	**Administrator**: First author **Intervention Intensity:** 24 lessons of 40 minutes each

Title: School-based mindfulness intervention improves executive functions and self-regulation in preschoolers at risk **Country:** Thailand (Nonthaburi) **Setting:** Two child development centres **Setting characteristics:** Centres have a common curriculum.	**Data Collection Points:** T0= pre-intervention baseline T1= post-intervention (eight weeks after intervention) **Follow Up:** n/a **Conditions:** 1. SBMP (n=15) 2. Control (n=15) regular classroom activities (*passive-TAU*) **Allocation:** Not stated.	**Age:** M= 72.73 months (SD=2.57 months) **Gender:** 50% female **Ethnicity:** Not reported. **Participant characteristics:** Preschool children with lower total EF scores on a behavioural checklist.	**Intervention Components:** Daily activities (breathing, mindful eating) and 24 classroom lessons categorised into 4 units: mindful attention; mindful sensation; mindful movement; and mindful feelings. The emphasis is on attentional training and body/emotional awareness. **Theoretical Foundation:** Study draws on mindfulness from the Buddhist concept of sati. Mindfulness is maintenance of awareness without judgement (Narada, 2013). Reference to reperceiving/decentring (Fresco et al., 2007). Also reference Shapiro et al.'s (2006) conceptualisation of mindfulness as involving intention, attention, and attitude. Further reference to self-regulatory mechanisms (Holzel et a., 2011; Tang et al., 2015).	**Intervention Frequency:** Thrice weekly for 8 weeks **Replicability of Intervention:** Manualised curriculum with content validated by academics in the field. No fidelity/replicability information reported. **Intervention Incentives:** None reported. **Adherence:** Not reported. **Acceptability:** Not reported.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Magalhaes et al. (2022). **Title:** A Pilot Study Testing the Effectiveness of a Mindfulness-Based Program for Portuguese School Children. **Country:** Portugal **Setting:** Four 3rd grade classrooms	**Design:** Quasi-experimental control group pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1= post-intervention (week after intervention) **Follow Up:** n/a **Conditions:** 1. SBMP (n=28)	**Sample Size:** 57 **Age:** SBMP: M=7.8 (0.35) years; Control: M=7.78 (0.34) years. **Gender:** SMBP= 64% female; Control= 55% female)	**Mindfulness Programme:** Custom programme using audio-guided mindfulness training (authors' previous work; Bakosh et al., 2018). **Intervention Components:** Organised into 6 modules: Introduction, Senses, Body, Heart, Brain, and Consolidation. Each module includes psychoeducation and opportunities to learn and practice mindful activities. 5-minute meditation activities are facilitated by audio-guides. Control activity follows similar modular structure with practice activities, but	**Administrator**: 30-minute sessions delivered by an external psychologist, and 5-minute sessions delivered by class teacher. Intensive training (24h) to both teachers and psychologists. **Intervention Intensity:** 16 x 30-minute sessions; and 24 x 5-minute sessions. **Intervention Frequency:** Twice weekly for 30m sessions and thrice weekly for 5 min sessions, over eight weeks. **Replicability of Intervention:** Instruction kits with step-by-step implementation instructions, PPT files and multimedia files

from a single school	2. Control (n=29) Health-based programme (*active-SEL*) **Allocation:** Not stated.	**Participant characteristics :** Exclusion criteria included identified special educational needs	focusing on physical health (making healthy choices, stretching activities etc.). **Theoretical Foundation:** Not explicitly stated, but follows a two-component model (e.g. Bishop et al., 2004) focusing on self-regulation of attention and open orientation to experience. Also explores the link between mindfulness and emotional regulation.	provided to administrators. There are not provided as accessible supplementary material. **Intervention Incentives:** None stated **Adherence:** Impacted by COVID-19 pandemic restrictions, therefore some sessions were administered online. Session checklists indicated that independent observers "showed a complete agreement between psychologists/teachers". **Acceptability:** Not stated.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Makmee (2022). **Title:** Increasing Attention and Working Memory in Elementary Students Using Mindfulness Training Programs. **Country:** Thailand **Setting:** Single Primary School **Setting characteristics:** Small school of 112 students (not anonymised in study).	**Design:** Quasi-experimental controlled pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention **Follow Up:** n/a **Conditions:** 1. SBMP (n=30) 2. Control (n=30) treatment-as usual (*Passive-None).* **Allocation:** Not stated.	**Sample Size:** 60 (SMBP=30; Control=30). **Age:** Not stated other than "Grade one – six". **Gender:** Not stated **Ethnicity:** Thai **Participant characteristics :** Volunteer Sampling	**Mindfulness Programme:** Based on the Full Emotional Intelligence Programme (Ramos et al., 2012), as cited in Campillo et al. (2018). **Intervention Components:** "Employed audiovisual techniques for mindfulness training" where students were "instructed to focus their attention on specific images and sounds that corresponded to those images". Visual stimuli included colours, mandalas and landscapes. Audio stimuli included mantras. **Theoretical Foundation:** Based on Integrative Body-Mind Training (IBMT). Mindfulness can be supported through audiovisual stimuli (Campillo et al., 2018).	**Administrator**: Not stated **Intervention Intensity:** 12 x 30-minute sessions. **Intervention Frequency:** Thrice weekly, over four weeks. **Replicability of Intervention:** Intervention is inadequately presented in article, although is described in reasonable detail in Campillo et al. (2018). There is no information on how this has been adapted to this study's context. **Intervention Incentives:** None stated. **Adherence:** Not stated **Acceptability:** Not stated
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**

Authors: Milare et al. (2021). **Title:** Mindfulness-Based Versus Story Reading Intervention in Public Elementary Schools: Effects on Executive Functions and Emotional Health. **Country:** Brazil (Sao Paulo). **Setting:** Two state primary schools	**Design:** Non-randomised controlled pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention **Follow Up:** n/a **Conditions:** 1. SBMP (n=111) 2. Control (n=96) story reading intervention (*Active-SEL).* **Allocation:** Each school assigned a condition.	**Sample Size:** 207 **Age:** 8-9 years (M=8.95) **Gender:** SMBP = 46.8% female; Control = 55.2% female. **Participant characteristics :** Exclusion criteria included diagnosis of autism, Prader-Willi syndrome, and BPD.	**Mindfulness Programme:** Adapted programme "inspired by the K5 curriculum of the Mindful Schools (MS), with adaptations". Adaptations included use of physical objects such as hula-hoop, gratefulness ball, and generosity buttons. **Intervention Components:** Taught components are categorised into awareness sessions and heartfulness sessions. Awareness covers mindfulness of body, breathing, thoughts, emotions, and movement. Heartfulness sessions were organized around themes of generosity, gratitude, and kindness to others. **Theoretical Foundation:** Some grounding in MBSR, although theoretical positioning is not fully explored. Justification is often pragmatic, based on successes reported in previous studies.	**Administrator**: A certified external instructor **Intervention Intensity:** 16 x 30-minute sessions; plus asked to incorporate practice into daily routine. **Intervention Frequency:** Twice weekly for seven weeks, with a double session on the eighth week. **Replicability of Intervention:** Outline of the sessions is provided, but no supplementary material exploring the resources used to support. Inspired by a published programme. **Intervention Incentives:** None stated. **Adherence:** Not stated. **Acceptability:** Not stated.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Muller et al. (2021). **Title:** Short breaks at school: Effects of a physical activity and a mindfulness intervention on children's attention, reading comprehension, and self-esteem.	**Design:** Quasi-experimental controlled pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention **Conditions:** 1. SBMP (n=43) 2. Control (n=37) uninstructed	**Sample Size:** 79 **Age:** M=11.4 years (SD= 0.51) **Gender:** 48.1% female **Ethnicity:** Not stated. **Participant characteristics**	**Mindfulness Programme:** Custom, brief programme, predominately through short activities. **Intervention Components:** Following a two-minute psychoeducation brief, mindful exercises such as body scans; breathing exercises; and guided attention exercises were practiced for 10 minutes. **Theoretical Foundation:** Mindfulness is defined in MBSR terms (Kabat-Zinn), with some reference to executive function, although this is minimally discussed.	**Administrator**: External research students **Intervention Intensity:** 10 x 12-minute sessions **Intervention Frequency:** Daily, over two weeks. **Replicability of Intervention:** Materials and outline of sessions is not provided beyond cursory description of activities. Very difficult to replicate effectively. No supplementary materials to support. **Intervention Incentives:** No incentives provided.

Country: Germany (Frankfurt) **Setting:** One state elementary school **Setting characteristics:** n/a	classroom breaks *(Active-dissimilar activity)* **Allocation:** Cluster randomisation by class, method not stated.	**:** Voluntary sampling.	No explicit reference to open orientation to experience. Therefore, might be more 'meditative' than mindful, per se.	**Adherence:** Not stated. **Acceptability:** Not stated.

Setting	Research Design	Participants	Intervention	Implementation
Authors: Quach et al. (2016). **Title:** A Randomized Controlled Trial Examining the Effect of Mindfulness Meditation on Working Memory Capacity in Adolescents. **Country:** United States (Southwest) **Setting:** A "large public middle school" **Setting characteristics:** Predominately low-income households.	**Design:** Randomized Controlled Trial **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention **Follow Up:** n/a **Conditions:** 1. SBMP (n= 54) 2. Control: Hatha Yoga (n=65) (*active-yoga).* 3. Control: Waitlist (n=53), physical education lessons *(passive-waitlist)* **Allocation:** Randomly assigned, method not stated.	**Sample Size:** 198 **Age:** 12-17 years **Gender:** SBMP= 83.3% female; Control= 52.8% female. **Ethnicity:** 66.3% Latino **Participant characteristics :** English speaking and "without physical disabilities that would hinder full participation"	**Mindfulness Programme:** Manualised MBSR curriculum. **Intervention Components:** "Each mindfulness meditation session consisted of breathing techniques, formal meditation, and discussion. Participants learned new types of mindfulness meditation each week (Week 1, Breathing; Week 2, Being in the Body and Feelings; Week 3, Awareness and Leaves on a Stream; Week 4, Silent and Loving Kindness)." **Theoretical Foundation:** Based on MBSR, with developmentally appropriate modifications (e.g. minor word changes and shortening of session times).	**Administrator**: "Two female instructors with extensive training in mindfulness meditation". Both completed MBSR training. **Intervention Intensity:** 8 x 45-minute sessions. Participants encouraged to engaged in 20 x 15–30-minute home practice. **Intervention Frequency:** Twice weekly for school sessions, over four weeks. Daily for homework. **Replicability of Intervention:** Manualised approach is easy to replicate, as well as being ubiquitous. Adaptations not provided in supplementary material. Although some are minor, reduction in session length by 25% warrants explicit description. Home practice is guided through audio (CDs). **Intervention Incentives:** None stated. **Adherence:** Participant flowchart provided. Although participants were encouraged to monitor practice through a log, this data is not presented.

				Acceptability: Not stated.

Setting	Research Design	Participants	Intervention	Implementation
Authors: Ricarte et al. (2015). **Title:** Mindfulness-Based Intervention in a Rural Primary School: Effects on Attention, Concentration and Mood. **Country:** Spain **Setting:** One primary school **Setting characteristics:** Rural	**Design:** Quasi-experimental controlled pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention **Follow Up:** n/a **Conditions:** 1. SBMP (n=45) 2. Control (n=45) uninstructed classroom breaks *(Active-dissimilar activity)* **Allocation:** Randomly assigned, method not stated.	**Sample Size:** 90 **Age:** 6-13. SBMP: M=8.87, SD=2.04; Control: M=8.93, SD=1.92. **Gender:** SBMP: 35.6% female; Control: 55.6% female. **Ethnicity:** Spanish **Participant characteristics:** Low SES area: 98% of younger population only formally educated to primary level.	**Mindfulness Programme:** Adaptation of manualised Mindfulness Emotional Intelligence Training Program (Ramos et al., 2012). **Intervention Components:** Mindfulness training exercises covering: "breathing (Weeks 1 and 2): the senses (Weeks 3 and 4); and attention to the body (Weeks 5 and 6)." **Theoretical Foundation:** Mindfulness training improves capacity to sustain attention and ignore distractions through practice of mindfulness skills. Uses Bishop et al.'s (2004) conceptualization of mindfulness in relation to sustained attention, selective attention, and attention switching. Also references openness to experience through describing mindfulness as "a specific form of processing emotions and states of consciousness in human experience".	**Administrator**: Not explicitly stated, but presumably class teacher. **Intervention Intensity:** 30 x 15-minute sessions. **Intervention Frequency:** daily, over six weeks. **Replicability of Intervention:** Summary of weekly component instructions and outcomes provided within text. Based on a manualised approach outlined in cited literature (e.g. Ramos et al., 2012). No supplementary materials provided. **Intervention Incentives:** None reported. **Adherence:** Not stated. **Acceptability:** Informal interviews with a range of stakeholders (parents, students, teachers) suggested "positive feedback with a high level of acceptance of the mindfulness program".

Setting	Research Design	Participants	Intervention	Implementation

Authors: Salmoirago-Blotcher et al. (2019). **Title:** Beneficial Effects of School-based Mindfulness Training on Impulsivity in Healthy Adolescents: Results from a Pilot Randomized Controlled Trial. **Country:** United States (Mass.) **Setting:** Two state high schools. **Setting characteristics:** n/a	**Design:** Pilot Randomised Controlled Trial **Data Collection Points:** T0= pre-intervention baseline T1= post-intervention T2= follow-up **Follow Up:** 6 months after T1. **Conditions:** 1. SBMP (n=30) 2. Control (n=23) treatment-as usual SEL curriculum plus an attentional control intervention (*Active-Attentional Control).* **Allocation:** Cluster randomization, using a random numbers generator.	**Sample Size:** 53 **Age:** M=14.55 (SD=0.3). **Gender:** SBMP: 70% female; Control: 43.5% female. **Ethnicity:** SBMP: 30% Hispanic; 66.7% White. Control: 47.8% Hispanic; 52.2% White. **Participant characteristics:** English-speaking; no history of psychiatric condition or developmental delay.	**Mindfulness Programme:** Manualised MBSR curriculum. **Intervention Components:** "The curriculum was based on the standard Mindfulness-Based Stress Reduction (MBSR) program, adapted to the needs of adolescents. Students were trained to cultivate awareness of bodily sensations, sounds, visual objects, thoughts, and emotions and to practice mindful movement (mindful walking and standing yoga exercises)." **Theoretical Foundation:** Not explicitly stated but uses Kabat-Zinn definitions of mindfulness and describes mindfulness as being "the very opposite of impulsivity".	**Administrator**: A certified mindfulness instructor. **Intervention Intensity:** 8 x 45-minute sessions. Participants encouraged to engaged in 20 x 15-minute home practice. **Intervention Frequency:** Once weekly for school sessions, over eight weeks. Daily for homework. **Replicability of Intervention:** Manualised approach is easy to replicate, as well as being ubiquitous. Adaptations not provided in supplementary material. Although some are minor, reduction in session length by 25% warrants explicit description (same concern with Quach et al., 2016). Home practice is guided through audio (CDs). **Intervention Incentives:** None stated. **Adherence:** "Mindfulness sessions were digitally recorded and a random selection of 10% of all recordings was audited; health education teachers used a structured checklist to record whether or not planned topics were discussed during each session". **Acceptability:** Not stated.

Setting	Research Design	Participants	Intervention	Implementation
Authors: Schonert-Reichl et al. (2015). **Title:** Enhancing Cognitive and	**Design:** Cluster randomised controlled trial	**Sample Size:** 99 **Age:** 9-11.2 years	**Mindfulness Programme:** Manualised MindUP curriculum. **Intervention Components:** "The curriculum includes 12 lessons, and each component of	**Administrator**: Class Teacher **Intervention Intensity:** 12 x 45-minute sessions; alongside 180 x 3-minute daily practice.

Social-Emotional Development Through a Simple-to-Administer Mindfulness-Based School Program for Elementary School Children: A Randomized Controlled Trial. **Country:** Canada **Setting:** Four elementary schools **Setting characteristics:** Neighbourhoods from a "Large city" with average SES.	**Data Collection Points:** T0= pre-intervention baseline T1=post-intervention **Follow Up:** n/a **Conditions:** 1. SBMP (n=48) 2. Control (n=51) Active social responsibility program (*Active-SEL).* **Allocation:** Cluster randomisation by coin flip	(M=10.24, SD= 0.53) **Gender:** 44% female. **Ethnicity:** Not stated. **Participant characteristics :** 84% reported living in two-parent homes. 66% English as a native language; 25% East Asian language.	the program builds on previous skills learned, moving children from focusing on subjective sense-based experiences (e.g., mindful smelling, mindful tasting) to cognitive experiences (e.g., taking others' perspectives), to actions such as the practice of gratitude and the doing of kind things for others in the home, classroom, and community." This programme is more community focused than traditional MBSR and includes an eco-behavioural systems orientation. **Theoretical Foundation:** "The MindUP curriculum is derived from psychological theory and informed by research in the fields of developmental neuroscience (Diamond, 2009, 2012), contemplative science and mindfulness (Roeser & Zelazo, 2012), SEL (Greenberg et al., 2003), and positive psychology (Lyubomirsky, Sheldon, & Schkade, 2005)." Kabat-Zinn definition of mindfulness supporting two-component model (Bishop et al., 2004). This relationship is presented alongside self-regulation (e.g. Zelazo & Lyons, 2012). Developmental perspective is explored in detail. This is one of the most theory-driven studies included.	**Intervention Frequency:** Weekly sessions, for 12 weeks; thrice daily for practice. **Replicability of Intervention:** Highly replicable as uses manualised programme. **Intervention Incentives:** None reported. **Adherence:** To assess implementation dosage and quality, teachers implementing the MindUP program were asked to complete surveys. Dosage was assessed by teacher report and review of implementation diaries. **Acceptability:** Not stated.

Setting	Research Design	Participants	Intervention	Implementation
Authors: Shlomov et al. (2023)	**Design:** Cluster-randomised, controlled trial.	**Sample Size:** 51	**Mindfulness Programme:** Custom Mindfulness and Kindness Curriculum (MC) for preschoolers.	**Administrator**: Experienced mindfulness instructor.

Title: Neurodevelopmental effects of mindfulness and kindness curriculum on executive functions in preschool children- a randomised, active-controlled study. **Country:** Israel (Northern) **Setting:** Day-care facilities **Setting characteristics:** Middle-class families with parents having a bachelor's degree.	**Data Collection Points:** T0= pre-intervention baseline T1=post-intervention **Follow Up:** n/a **Conditions:** 1. SBMP (n=22) 2. Control (n=29) dialogic reading programme (*Active-dissimilar activity).* **Allocation:** Random assignment	**Age:** M= 4.9 (SD=0.69) **Gender:** 37.3% female **Ethnicity:** Not stated **Participant characteristics:** Hebrew as home language	**Intervention Components:** Focusing on cultivating mindful attention, self-regulation and pro-social skills through three main modules: 1. Directing attention to present moment (body, breath, senses, movement) 2. Directing attention inwards (feelings, thoughts, sensations, imagination) 3. Direction attention towards others (developing kind and caring attitude). Each module includes experiential activities and story-based discussions. **Theoretical Foundation:** Specific reference to the two-factor model for mindfulness (Bishop et al., 2004) and suggest mindfulness acts to support for bottom-up processes that can interfere with attention (Zelazo & Lyons, 2012).	**Intervention Intensity:** 24 sessions of 30 minutes each. **Intervention Frequency:** Thrice weekly for eight weeks. **Replicability of Intervention:** Curriculum described, but fidelity measures not reported. **Intervention Incentives:** Children received a small gift for participation. **Adherence:** Not reported. **Acceptability:** Not reported.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Suarez-Garcia et al. (2020). **Title:** The Effect of a Mindfulness-Based Intervention on Attention, Self-Control, and Aggressiveness in Primary School Pupils.	**Design:** Quasi-experimental "switching replications" design. **Data Collection Points:** T0= pre-intervention baseline T1= post-intervention T2= follow-up	**Sample Size:** 73 **Age:** 7-10 years old (M=8.1 years, SD=0.49) **Gender:** Not stated **Ethnicity:** Not stated	**Mindfulness Programme:** Manualised programme, *Mindkeys Training*, designed for primary school students. **Intervention Components:** "The 8 sessions had the following generic structure: (1) Sounding a singing bowl (to start the session) and silence; (2) The explanation of a breathing technique; (3) An attention to sound exercise; (4) Reading a story and a debate; (5) A specific activity and an	**Administrator**: Mindfulness Expert administered sessions. Class teachers, trained by a mindfulness expert over two 1-hour sessions. **Intervention Intensity:** 8 x 60-minute sessions; 32 x 10-minute teacher-directed practice. **Intervention Frequency:** 1 session and 4 practice activities weekly, over eight weeks.

Country: Spain (Asturias) **Setting:** Three state schools **Setting characteristics:** Urban, middle SES status schools.	**Follow Up:** 8 weeks after T1 **Conditions:** 1. EG1: Intervention (n=40) 2. EG2: Waitlist controlled (n=33) (*passive-waitlist*) **Allocation:** Clustered, based on class.	**Participant characteristics :** n/a	explanation of the week's challenge; and (6) Sounding the singing bowl (to end the session). On the other days of each week, the teachers directed reinforcement sessions for the pupils, each lasting 10 min" **Theoretical Foundation:** No exploration of mindfulness theory; justification for using SMBP is pragmatic and based on positive results reported in prior studies. They suggest the evidence base indicates that SMBPs have 'clearer' effects on cognitive dimensions. Some reference to developmental psychology, although not linked to mindfulness explicitly.	**Replicability of Intervention:** Supplementary material providing details on programme outline, main objectives, and activities. Manualised programme easier to replicate. **Intervention Incentives:** None reported. **Adherence:** Not stated. **Acceptability:** Not stated.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Thierry et al. (2016) **Title:** Two-year impact of a mindfulness-based programme on preschoolers' self-regulation and academic performance. **Country:** United States (Southwestern) **Setting:** An urban elementary school.	**Design:** Quasi-experimental pre-test/post-test design **Data Collection Points:** T0= pre-intervention T1= post-intervention **Follow Up:** n/a **Conditions:** 1. SBMP (n=23) 2. Business as usual (n=24) (*passive-TAU*) **Allocation:** Non-random; cohort design.	**Sample Size:** 47 **Age:** 4.55 years **Gender:** 49% female **Ethnicity:** 85% Hispanic **Participant characteristics :** 72% qualify for free school meals.	**Mindfulness Programme:** Based on MindUP programme. **Intervention Components:** Lessons on mindful awareness, focused breathing, mindful sensing, perspective-taking, optimism, gratitude. 3x daily core breathing exercises. **Theoretical Foundation:** Cognitive perspective arguing that mindfulness training engages top-down (prefrontal) self-regulatory processes and calms bottom-up (subcortical) stress-arousal. Background knowledge on the brain is included.	**Administrator**: Class teachers receiving a full day training. **Intervention Intensity:** 15 lessons of 20-30 minutes plus core practices. **Intervention Frequency:** Not explicitly stated, but possibly around 37.5 weeks. **Replicability of Intervention:** Manualised curriculum. **Intervention Incentives:** None mentioned. **Adherence:** Teacher surveys indicated all 15 lessons were delivered and breathing practiced consistently.

Setting characteristics: School serves predominately low-income Hispanic families.				**Acceptability:** Teachers rated high levels of student engagement during the lessons.

Setting	Research Design	Participants	Intervention	Implementation
Authors: Thomas & Atkinson (2016). **Title:** Measuring the effectiveness of a mindfulness-based intervention for children's attentional functioning. **Country:** United Kingdom (NW England) **Setting:** Single comprehensive primary school **Setting characteristics:** "Mixed and ethnically diverse"	**Design:** Randomised Controlled Trial with quasi-experimental intervention cross-lag. **Data Collection Points:** T0= pre-intervention baseline T1= post-intervention T2= 1st follow-up T3= 2nd follow-up **Follow Up:** T2= 6 weeks after T1 T3= 12 weeks after T2 **Conditions:** 1. Group 1(n= 16) 2. Group 2 (n=14) Waitlist Control (*Passive-TAU*). **Allocation:** Cluster randomisation via coin-toss.	**Sample Size:** 30 **Age:** 8-9 years. SMBP: M= 8 years, 10 months; Control: M=8 years, 9 months. **Gender:** SMBP: 50% female; Control: 57% female. **Ethnicity:** "ethnically diverse" **Participant characteristics:** All EAL pupils, predominance of Urdu and	**Mindfulness Programme:** Established, manualised "Paws .b" programme (from MiSP) **Intervention Components:** The six sessions explore psychoeducational elements such as introduction to the brain; 'searchlight' of attention; philosophy of mindfulness etc. Activities included breathing exercises; discussion of wobbly feelings; discussion on how to avoid reacting badly to situations. **Theoretical Foundation:** Strong grounding in cognitive science, exploring the work of Petersen & Posner (2012) on attentional networks. Neuropsychological exploration with references to Luria etc.	**Administrator**: Not explicitly stated in article, but MiSP materials indicate that it is teacher-led. **Intervention Intensity:** 6 x 60-minute sessions **Intervention Frequency:** Once weekly, over six weeks. **Replicability of Intervention:** An overview provided in supplementary material. Manualised programme has resources to support replicability (e.g. PPT presentation) **Intervention Incentives:** None reported. **Adherence:** "Implementation checks were carried out by the first researcher throughout the two intervention periods". The process/results are not outlined. Change of class teacher in waitlist group created a confounding variable. **Acceptability:** Not stated.

		Bengali as 1st language.		
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Vickery & Dorjee (2016). **Title:** Mindfulness Training in Primary Schools Decreases Negative Affect and Increases Meta-Cognition in Children. **Country:** United Kingdom (North Wales) **Setting:** Three primary schools **Setting characteristics:** Volunteer schools	**Design:** Quasi-experimental controlled pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1= post-intervention T2= follow-up **Follow Up:** 3 months after T1. **Conditions:** 1. SMBP (n=33) 2. Control (n=38) (*passive-TAU*). **Allocation:** Non-random, based on "volunteer interest".	**Sample Size:** 71 **Age:** 7-9 years (M=7.9; SD=0.64) **Gender:** 49.3% female **Ethnicity:** Not stated. **Participant characteristics :** n/a	**Mindfulness Programme:** Established, manualised "Paws .b" programme (from MiSP) **Intervention Components:** "Teachers delivered the program as part of PSE lessons in the classroom setting to approximately 30 pupils. The Paws b program aimed to support children to develop more mindful and less automatic relating to their present moment experiences in the classroom. The six themes covered in the Paws b program (i.e., 'Our Amazing Brain,' 'Puppy Training,' 'Finding a Steady Place,' 'Dealing with Difficulty,' 'The Story Telling Mind,' and 'Growing Happiness') can be flexibly delivered to suit school demands using 1 h or ½ h lessons (12 half hour lessons overall)." **Theoretical Foundation:** Suggests that links between mindfulness and cognitive control could extend to emotion regulation. They reference Zelazo & Lyons (2012), with a focus on bottom-up processing, as well as top-down processing (e.g. metacognition). Overall, it appears to advocate for a self-regulation framework. Concepts of EF and metacognition are used interchangeably (e.g. Authors use full BRIEF as a measure of meta-cognition).	**Administrator**: Class teachers, trained through the ".b Foundations" course, and assessed 6 months later by an expert mindfulness trainer. **Intervention Intensity:** 4 x 60-minute sessions and 4 x 30-minute sessions; teachers led 8 x 5–10-minute informal practices. **Intervention Duration:** 60-minute and 30-minute sessions delivered weekly, over eight weeks. Informal practices delivered weekly. **Replicability of Intervention:** Brief (single paragraph) description of programme. Manualised programme has resources to support replicability (e.g. PPT presentation). Audio guides help support fidelity of independent practice sessions. **Intervention Incentives:** None reported. **Adherence:** "participants were asked how frequently they practiced mindfulness outside of school for which 21.2% responded 'never,' 39.4% 'rarely,' 30.3% 'often,' and 9.1% 'everyday.'" **Acceptability:** An acceptability questionnaire found that 76% of young people reported 'liking', 'liking a lot', or 'extremely liking' the mindfulness

				programme. 18.1% reported 'disliking' or 'extremely disliking' programme.
Setting	**Research Design**	**Participants**	**Intervention**	**Implementation**
Authors: Wimmer et al. (2016). **Title:** Cognitive Effects of Mindfulness Training: Results of a Pilot Study Based on a Theory Driven Approach. **Country:** Germany (Essen) **Setting:** One state school **Setting characteristics:** Gymnasium (selective)	**Design:** Quasi-experimental controlled pre-test/post-test design **Data Collection Points:** T0= pre-intervention baseline T1= post-intervention **Follow Up:** n/a **Conditions:** 1. SBMP (n=16) 2. Active Control (n= 8) Concentration training (*active-attention)* 3. Passive Control (n=10) (*passive-none*). **Allocation:** Randomly assigned to either SBMP or Active Control via lot-draw. No randomisation of passive control (parallel class in same school).	**Sample Size:** 34 **Age:** M= 10.8 years (SD=0.53). **Gender:** 52.9% female **Ethnicity:** Not stated. **Participant characteristics :** Students at an academically selective school.	**Mindfulness Programme:** The mindfulness training was based on the well-established MBSR method (Kabat-Zinn, 2005). We also drew on an adapted version for children by Greenland (2010). **Intervention Components:** "The intervention comprised two essential exercises, sitting meditation, and the body scan. In sitting meditation, the aim is to constantly focus on one's own breath while letting go of arising thoughts or emotions. The training started with 3 min practicing times. Later, it was extended to 10 min. During the body scan, learners slowly guided their attention through the whole body, from the toes to the top of the skull. As it became apparent that the children were overtaxed with a complete scan, the instruction was split into an upper and a lower body part and these two were practiced in turns. Duration of these partial body scans varied between 5 and 15 min." Active Control group were administered a manualized programme popular in Germany, the German Marburg Concentration Training. **Theoretical Foundation:** Detailed theoretical basis drawing on the two-component model	**Administrator**: Five tutors and one study author, trained by the authors of study. Each session was led by at least two instructors. **Intervention Intensity:** Difficult to ascertain, but "children received approximately 150 minutes of treatment each week". **Intervention Frequency:** One 60-minute session and one 90-minute session a week, over 16 weeks. **Replicability of Intervention:** Although based on manualised programme, no supplementary material provides details on how these have been adapted. **Intervention Incentives:** "In exchange for their participation these children received a book voucher worth €25 after finishing the second series of cognitive tests." **Adherence:** Fidelity was supported through the regular presence of one of the study authors and by weekly supervision. **Acceptability:** Not stated.

			of mindfulness and attentional networks (Petersen & Posner, 2012).	

Setting	Research Design	Participants	Intervention	Implementation
Authors: Wood et al. (2017) **Title:** Enhancing executive function skills in preschoolers through a mindfulness-based intervention: A randomized, controlled pilot study. **Country:** United States (Southwestern) **Setting:** Two child development centres **Setting characteristics:** On-campus settings linked with university	**Design:** Randomised, waitlist-controlled design. **Data Collection Points:** T0= pre-intervention baseline T1=post-intervention **Follow Up:** n/a **Conditions:** 1. SBMP (n=12) 2. Waitlist Control group (n=15) (passive-waitlist) **Allocation:** Randomly assigned with teachers blinded.	**Sample Size:** 27 **Age:** M=3.75 years (SD=0.6) **Gender:** 59.3% female **Ethnicity:** 84.2% white **Participant characteristics :** Majority of parents holders of a postgraduate degree.	**Mindfulness Programme:** Custom mini-mind, manualised programme designed for preschool-aged children. **Intervention Components:** Each session included components such as: brief yoga; mindful breathing; compassion jar; interactive, concrete activities for developing awareness (e.g. taste, sight, sound etc.). The programme also included suggested home practices for parents and children based on the week's sessions. **Theoretical Foundation:** Uses definition of mindfulness provided by Kabat-Zinn, suggesting links with MBSR. Authors explicitly mention both regulation of attention and openness towards experience, aligning with the two-factor model proposed by Bishop et al. (2024).	**Administrator**: Three graduate students **Intervention Intensity:** twelve 25-minute sessions **Intervention Frequency:** Twice-weekly sessions over 6 weeks. **Replicability of Intervention:** Manualised programme. **Intervention Incentives:** None stated. **Adherence:** Treatment integrity data shows 98% of critical intervention components were implemented. **Acceptability:** Rated as highly acceptable by all stakeholders. Children scored intervention 4.04/5 on average; parents scored 4.41/5.

Setting	Research Design	Participants	Intervention	Implementation
Authors: Zelazo et al. (2018)	**Design:** Cluster-randomised controlled trial.	**Sample Size:** 218	**Mindfulness Programme:** Custom-designed curriculum combining mindfulness and reflection training.	**Administrator**: Class teachers trained in the curriculum.

Title: Mindfulness plus reflection training: Effects on executive function in early childhood. **Country:** United States (Texas & Washington D.C.) **Setting:** Two Preschools **Setting characteristics:** Serving predominately low-income families	**Data Collection Points:** T0= pre-intervention baseline T1= post-intervention T2= follow-up **Follow Up:** 4-6 weeks after T1 **Conditions:** 1. SBMP (n=72) 2. Literacy training (n=76) 3. Business as usual (n=68) (passive-TAU) **Allocation:** Randomly allocated.	**Age:** M=4.77 years **Gender:** 53.7% female **Ethnicity:** Mixed (55% White, 32% more than one race in Houston; 100% African American in D.C.) **Participant characteristics:** Lower SES as measured by median reported family income.	**Intervention Components:** 30 daily small-group sessions over 6 weeks. Each session included brief mindfulness and relaxation practices, along with 3 EF-challenging games. **Theoretical Foundation:** Draws on MBSR (Kabat-Zinn, 2003) to suggest that mindfulness can reduce stress. Mindfulness practices are designed to help children calm down, regulate stress, and become aware of the present moment. Reflection exercises are adjunct to the mindfulness, and draw on the Iterative Reprocessing model (Zelazo, 2015).	**Intervention Intensity:** 30 sessions of 24 minutes each. **Intervention Frequency:** Daily sessions over 6 weeks. **Replicability of Intervention:** Teachers had some training in delivering manualised curriculum. **Intervention Incentives:** None mentioned. **Adherence:** Not assessed. **Acceptability:** Not reported.

Appendix D

Summary of Weight of Evidence A (WoE A) Ratings

Study	Essential Quality Indicators				Total Essential Indicators (max. 10)	Desirable Quality Indicators (max. 8)	WoE A
	Participant (max. 3)	Intervention (max. 3)	Outcome Measures (max. 2)	Data Analysis (max. 2)			
Andreu et al. (2023)	2	2	2	2	8	3	1 (low)
Baena-Extremera et al. (2021)	3	2	2	2	9	4	3 (high)
Berti & Cigala (2023)	3	3	2	1	9	4	3 (high)
Brann et al. (2023)	3	3	2	2	10	3	3 (high)
Crescentini et al. (2016)	2	2	1	2	7	2	1 (low)
Crooks et al. (2020)	3	3	2	1	9	2	2 (medium)
Flook et al. (2010)	2	2	2	2	8	3	1 (low)
Folch et al. (2021)	2	1	1	1	5	1	0 (very low)
Frank et al. (2021)	3	3	2	2	10	6	3 (high)
Janz et al. (2019)	2	2	2	2	8	4	2 (medium)
Koncz et al. (2021)	3	2	2	1	8	2	1 (low)
Lam & Seiden (2020)	3	2	2	2	9	4	3 (high)
Lassander et al. (2020)	3	3	2	2	10	4	3 (high)

Lertladaduck et al. (2021)	2	1	2	2	7	1	1 (low)
Magalhaes et al. (2022)	3	3	1	2	9	4	3 (high)
Makmee et al. (2022)	0	0	1	2	3	0	0 (very low)
Milare et al. (2021)	1	2	1	2	6	2	0 (very low)
Muller et al. (2021)	3	2	1	2	8	2	1 (low)
Quach et al. (2016)	3	3	1	2	9	4	3 (high)
Ricarte et al. (2015)	2	2	3	2	9	1	2 (medium)
Salmoirago et al. (2019)	3	3	1	2	9	4	3 (high)
Schonert-Reichl et al. (2015	3	3	2	2	10	3	2 (medium)
Shlomov et al. (2023)	3	2	2	2	9	1	2 (medium)
Suarez-Garcia et al. (2020).	2	1	2	2	7	3	1 (low)
Thierry et al. (2016)	3	3	2	2	10	4	3 (high)
Thomas & Atkinson (2016)	3	3	2	2	10	2	2 (medium)
Vickery & Dorjee (2016)	2	2	1	2	7	3	1 (low)
Wimmer et al. (2016)	3	3	2	2	10	1	2 (medium)
Wood et al. (2017)	3	2	0	2	7	1	1 (low)
Zelazo et al. (2018)	3	3	2	2	10	6	3 (high)

Appendix E

Amendments to WoE A Quality Indicator Coding Protocol (Gersten et al., 2005).

Essential Quality Indicators
Describing Participants
1. Was sufficient information provided to {describe the demographics of the study sample?} ~~determine/confirm whether the participants demonstrated the disability(ies), or difficulties presented?~~
[Rationale: The original Gersten et al. (2005) protocol was designed for group design studies exploring the impact of interventions that target specific SEN populations, whereas SBMPs are typically designed as universal interventions that target whole-school populations. As certain population characteristics are hypothesized to moderate the effect of SBMPs (e.g. age, gender), this criterion has been adapted to reflect the importance of including demographic information that may be relevant to the generalizability of the results.]
2. Were appropriate procedures used to increase the likelihood that relevant characteristics of participants in the sample were comparable across conditions?
3. Was sufficient information given characterizing the interventionists or teachers provided? Did it indicate whether they were comparable across conditions?
Implementation of the Intervention and Description of Comparison Conditions
1. Was the intervention clearly described and specified?
2. Was the fidelity of implementation described and assessed?
3. Was the nature of services provided in comparison conditions {clearly} described?
[Rationale: A clear description of comparison conditions helps to identify potential confounding variables that may contribute to Type I error risk.]
Outcome Measures

1. Were multiple measures used to provide an appropriate balance between measures closely aligned with the intervention ~~and measures of generalized performance?~~

[Rationale: Measures of generalized performance are not within the scope of this review.]

2. Were outcomes for capturing the intervention's effect measured at the appropriate times?

Data Analysis

1. Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the unit of analysis in the study?

2. Did the research report include not only inferential statistics but also effect size calculations? {In multiple hypothesis studies, has the familywise error rate been controlled?}

[Rationale: Corrections to account for Type I errors deemed essential in studies with multiple hypotheses].

Desirable Quality Indicators

---

1. Was data available on attrition rates among intervention samples? Was severe overall attrition documented? If so, is attrition comparable across samples? Is overall attrition less than 30%?

2. Did the study provide ~~not only internal consistency reliability but also~~ test–retest reliability and interrater reliability (when appropriate) for outcome measures? Were data collectors and/or scorers blind to study conditions and equally (un)familiar to examinees across study conditions?

[Rationale: Save for "Ambiguous Figures Task" outcome measure employed by Wimmer et al. (2016), all studies used standardized outcome measures which have been independently tested for internal consistency reliability].

3. Were outcomes for capturing the intervention's effect measured beyond an immediate post-test?

4. Was evidence of the criterion-related validity and construct validity of the measures provided?

5. Did the research team assess not only surface features of fidelity implementation (e.g., number of minutes allocated to the intervention or teacher/interventionist following procedures specified), but also examine the quality of implementation?

6. Was any documentation of the nature of instruction or series provided in comparison conditions?

7. Did the research report include actual audio or videotape excerpts that capture the nature of the intervention?

8. Were results presented in a clear, coherent fashion?

{10. Is the conceptualisation of the outcome measure based on well-designed studies and does it reflect the scope of extant knowledge?}

[Rationale: The heterogeneity of how EF is conceptualised and operationalised in the wider literature necessitates clear working definitions in included studies].

Appendix F

Example of WoE A appraisal using Gersten et al. (2005) coding protocol.

Study: Koncz A, Koeteles F, Demetrovics Z, & Takacs ZK. (2021). Benefits of a Mindfulness-Based Intervention upon School Entry: A Pilot Study. *INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH*, *18*(23). https://doi.org/10.3390/ijerph182312630		
Essential Quality Indicators		**Notes**
Describing Participants		
1. Was sufficient information provided to describe the demographics of the study sample?	☒ Yes ☐ No	*Sufficient demographic data provided, including age and gender.*
2. Were appropriate procedures used to increase the likelihood that relevant characteristics of participants in the sample were comparable across conditions?	☒ Yes ☐ No	*Experimental and control conditions were matched (based on age, gender, and pre-test scores).*
3. Was sufficient information given characterising the interventionists or teachers provided? Did it indicate whether they were comparable across conditions?	☒ Yes ☐ No	*Administrators are described as trained research assistants with undergraduate degrees in Psychology, supervised by a clinical child psychologist.*
Implementation of the Intervention and Description of Comparison Conditions		
1. Was the intervention clearly described and specified?	☒ Yes ☐ No	*The intervention is clearly described, outlining the SBMP, individual session content, and modifications made to the programme by the researchers.*
2. Was the fidelity of implementation described and assessed?	☐ Yes ☒ No	*No information is provided relating to how fidelity of implementation has been assessed.*
3. Was the nature of services provided in comparison conditions clearly described?	☒ Yes ☐ No	*The nature of services in the control condition is clearly described as "free play".*

Outcome Measures		
1. Were multiple measures used to provide an appropriate balance between measures closely aligned with the intervention?	☒ Yes ☐ No	*Multiple performance-based tasks (Corsi tests, Go/No-Go, Heart & Flowers Task) have been used to measure executive function.*
2. Were outcomes for capturing the intervention's effect measured at the appropriate times?	☒ Yes ☐ No	*Outcomes were measured at pre-test, post-test, and follow-up. Post-test data collection point one week after intervention is appropriate.*
Data Analysis		
1. Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the unit of analysis in the study?	☒ Yes ☐ No	*The statistical methods (ANOVA, Mann-Whitney U-tests) are appropriate. However, the clustering of participants into classroom groups is not addressed in the analysis.*
2. Did the research report include not only inferential statistics but also effect size calculations? In multiple hypothesis studies, has the familywise error rate been controlled?	☐ Yes ☒ No	*Inferential statistics are provided, alongside partial eta squared statistics as a measure of effect size. No correction for familywise error rates despite multiple hypotheses.*
Desirable Quality Indicators		
1. Was data available on attrition rates among intervention samples? Was severe overall attrition documented? If so, is attrition comparable across samples? Is overall attrition less than 30%?	☐ Yes ☒ No	*No explicit report on attrition rates is provided, although flow diagrams indicate that attrition was well below 30%.*
2. Did the study provide test–retest reliability and interrater reliability (when appropriate) for outcome measures? Were data collectors and/or scorers blind to study conditions and equally (un)familiar to examinees across study conditions?	☐ Yes ☒ No	*No explicit mention of reliability metrics (e.g. internal consistency/test-retest etc.) for outcome measures, although these are available through provided references.*
3. Were outcomes for capturing the intervention's effect measured beyond an immediate post-test?	☒ Yes ☐ No	*A follow-up data collection point (for cortisol level outcome measure only) was included a month after post-test.*

4. Was evidence of the criterion-related validity and construct validity of the measures provided?	☐ Yes ☒ No	*Study does not provide explicit information relating to the construct validity of the outcome measures.*
5. Did the research team assess not only surface features of fidelity implementation (e.g., number of minutes allocated to the intervention or teacher/interventionist following procedures specified), but also examine the quality of implementation?	☐ Yes ☒ No	*No detailed assessment of the quality of implementation is provided.*
6. Was any documentation of the nature of instruction or series provided in comparison conditions?	☐ Yes ☒ No	*No information on the nature of control condition activities is provided.*
7. Did the research report include actual audio or videotape excerpts that capture the nature of the intervention?	☐ Yes ☒ No	*None are provided. Although links to the materials used (mindfulness stories in the original Hungarian) are provided in an appendix.*
8. Were results presented in a clear, coherent fashion?	☒ Yes ☐ No	*Descriptive statistics are presented in a table that facilitates secondary data analysis.*

Appendix G

Rationale for WoE B criteria selection

Appraisal category and rationale	Score Criteria		
Research Design	3	2	1
Although Risk of Bias tools were considered to overly penalize research conducted on SBMPs as typical randomization processes are complicated by the classroom-based nature of the intervention, rigorous attempts to control confounding variables are necessary to isolate the intervention effects specific to SBMPs.	A cluster-randomized controlled trial that uses an equivalence active control group. Rigorous control of confounding variables.	A quasi-experimental design that uses an active control group OR a cluster-randomized controlled trial with a passive control group. Some attempt to control confounding variables.	A quasi-experimental design that uses a passive control group. Minimal attempt to control confounding variables.
Outcome measures	3	2	1
Measures of cognitive constructs such as EF and related processes present multiple challenges. Given the heterogeneity, standardised approaches that have good construct validity for the target age group is preferred, as are assessment batteries that can assess multiple EF components (working memory, cognitive flexibility, inhibition) in a	Comprehensive, well-validated, and sensitive psychological assessments specifically designed to measure EF, using both PBTs and BRSs. Examples include the BRIEF for BRS measures of EF in or PBT batteries like the CANTAB.	Specific cognitive tests that assess certain EF components may not cover the full spectrum of EF or lack sensitivity to the changes expected from mindfulness training. They may be designed to measure EF or similar construct, like digit span	The study employs a narrow set of outcome measures that are unstandardised or have not been used extensively with CYP. Intended use of outcome measure may not align with

detailed and sensitive manner, as well as a more general construct.		backwards, but not capture other aspects (e.g. visual WM).	operationalisation of said measure in the study.
Fidelity of Implementation This review conceptualizes SBMPs as complex interventions, and the implementation of the intervention is hypothesized to be a significant moderating variable. Studies that fail to present information on implementation are less well-placed to answer questions related to process.	3 A clear and detailed description of how the SBMP was implemented is provided. This includes protocols, delivery, adherence measures, and fidelity assessment.	2 Adequate description of the intervention with some details on implementation and adherence, though lacking a comprehensive quality appraisal.	1 Minimal exploration of fidelity of implementation. Unclear protocols outlined.
Moderator Variables As several moderator variables will be explored in the meta-analysis through meta-regression, appropriate reporting of these variables is necessary. Sufficient reporting on intervention dosage (including intensity and session duration); role of administrator (including any training school staff receives); and the number and type of intervention components.	3 Detailed information provided on participant age. The study records thorough information on the administrator's training in mindfulness. Clear description of dosage with a rationale based on theory/ empirical evidence.	2 Sufficient information is provided on participant age. Administrator role described; some training or qualifications mentioned without detail. Dosage described to the extent that total duration in minutes can be calculated.	1 Insufficient information is provided relating to participant age. Administrator role is not clearly defined, no information on training or qualifications. Dosage of intervention not specified or vaguely described.
Statistical Analysis	3	2	1

Methodological concerns specific to SBMPs suggest the need for appropriate statistical power. SBMP conditions are nested within classrooms and therefore assumption of independence is violated. Many studies report multiple outcomes relating to EF, and the FWER should be controlled to minimise the risk of Type I or M error.	Rigorous statistical analysis that is sufficiently powered to detect changes, and controls for potential confounders (e.g. clustering of groups; family-wise error rate).	Appropriate statistical analysis that is linked to the research questions but does not account for potential clustering confounders.	Statistical methods are inadequately described, inappropriate, or insufficiently powered to detect meaningful changes.

Appendix H

Summary of Weight of Evidence B (WoE B; methodological appropriateness) ratings

Study	*Criterion Category*					*WoE B*
	Research Design	*Outcome measures*	*FIdelity of Implementation*	*Moderator Variables*	*Statistical Analysis*	
Andreu et al. (2023)	2	3	1	2	2	2 (medium)
Baena-Extremera et al. (2021)	1	2	2	1	3	1.8 (medium)
Berti & Cigala (2023)	1	2	2	2	1	1.6 (low)
Brann et al. (2023)	2	2	2	2	2	2 (medium)
Crescentini et al. (2016)	3	1	1	2	3	2 (medium)
Crooks et al. (2020)	2	2	2	2	1	1.8 (medium)
Flook et al. (2010)	3	3	1	1	3	2.2 (medium)
Folch et al. (2021)	1	2	1	2	2	1.6 (low)
Frank et al. (2021)	2	2	2	1	3	2 (medium)
Janz et al. (2019)	2	2	2	1	2	1.8 (medium)
Koncz et al. (2021)	1	2	2	2	2	1.8 (medium)

Lam & Seiden (2020)	2	3	2	3	2	2.4 (high)
Lassander et al. (2020)	3	2	2	2	3	2.4 (high)
Lertladaduck et al. (2021)	1	2	1	2	1	1.4 (low)
Magalhaes et al. (2022)	2	2	2	2	2	2 (medium)
Makmee et al. (2022)	1	2	1	1	2	1.4 (low)
Milare et al. (2021)	2	1	2	2	2	1.8 (medium)
Muller et al. (2021)	2	1	1	2	2	1.6 (low)
Quach et al. (2016)	3	1	3	2	3	2.4 (high)
Ricarte et al. (2015)	1	2	1	1	2	1.4 (low)
Salmoirago et al. (2019)	3	1	2	2	3	2.2 (medium)
Schonert-Reichl et al. (2015	3	2	2	2	2	2.2 (medium)
Shlomov et al. (2023)	2	2	1	2	3	2 (medium)
Suarez-Garcia et al. (2020).	2	2	1	2	2	1.8 (medium)
Thierry et al. (2016)	2	3	1	2	2	2 (medium)
Thomas & Atkinson (2016)	2	2	2	2	1	1.8 (medium)

Vickery & Dorjee (2016)	1	3	2	2	2	2 (medium)
Wimmer et al. (2016)	2	2	2	2	3	2.2 (medium)
Wood et al. (2017)	2	1	2	1	2	1.6 (low)
Zelazo et al. (2018)	3	3	2	2	2	2.4 (high)

Appendix I

Rationale for WoE C criteria selection

Appraisal category and rationale	Score Criteria		
Theoretical Framework Underpinning Intervention SBMPs should align with the core principles and mechanisms of mindfulness. These studies are more relevant for determining whether SBMPs have a positive intervention effect due to mindfulness components.	3 Full exploration of how the SBMP aligns with theoretical literature on mindfulness and EF. The intervention has been specifically designed/adapted to focus on EF.	2 Some exploration of how mindfulness intervention includes elements of paying attention to the present moment and cultivating a nonjudgemental perspective. May include some adaptation of protocol to reflect EF focus.	1 Little exploration of how the SBMP aligns with core mindfulness principles (e.g. only focusing on awareness of the present moment without also focusing on emotional acceptance).
Analysis of cognitive effects This research has a specific focus on the potential differential effect of SBMPs on ECPs, whilst also holding that EF is a higher-order, emergent construct. As such, studies that analyse both separable cognitive processes associated with EF, as well as evaluating EF as a holistic	3 The study provides a comprehensive analysis of the impact of SBMPs on a range of specific cognitive processes associated with EF (e.g., response inhibition, working memory, cognitive flexibility) and overall EF.	2 The study includes an analysis of the impact of SBMPs on overall EF or examines several but not all cognitive processes associated with EF. There might be some depth in the exploration of these processes, but the analysis does	1 The study focuses only on one ECP (e.g., working memory) overall EF. The analysis is narrow, potentially overlooking the multifaceted nature of EF improvements.

construct, have more relevance to this review's aims.		not cover the full spectrum of cognitive components of EF.	
Relevance to UK Context	3	2	1
This research is submitted in partial completion of a UK Doctorate in Professional Educational Adolescent and Child Psychology. Studies that are conducted in similar contexts to the UK are considered more relevant, especially with regard to how results can be applied by EPs.	The interventions take place in a mainstream school setting that is broadly generalizable to a UK context (e.g. English-speaking; broadly similar SES/EAL rates; Western culture).	The setting is broadly congruent with a UK mainstream school context but may have specific factors that limit generalizability (e.g. rural setting only; areas of somewhat low SES; western cultures that are non-English speaking).	Population or setting variables make it difficult to generalize to mainstream school contexts (e.g. special school setting; non-Western culture more familiar with mindfulness practices).
Replicability of SBMP in schools	3	2	1
A key aim of this research is to support the implementation of mindfulness programmes in schools. As such, research that facilitates the transfer of findings to real-world contexts is deemed more relevant to this review. Studies should report interventions in detail as SBMPs,	The study provides comprehensive details on how the SBMP aligns with and complements existing school curricula, including evidence of successful past integrations. There is clear including detailed guidance for schools	The study outlines some alignment between the SBMP and school curricula, with examples or suggestions for integration. There is an acknowledgment of the need for alignment with educational practices supporting	There is minimal discussion of how the SBMP could align with school curricula or educational practices. Lack of detail on programme components, training requirements for facilitators etc. Schools would struggle to replicate the SBMP without

ideally according to TIDIER guidelines.	on implementing the program with minimal external support.	EF, but the guidance for schools on replication is somewhat vague.	significant adaptation or external support.

Appendix J

Summary of Weight of Evidence C (WoE C; evidence relevance) ratings

Study	*Criterion Category*				*WoE C*
	Theoretical Framework	*Cognitive Functions*	*Contextual Relevance*	*Replicability in schools*	
Andreu et al. (2023)	3	2	2	2	2.25 (medium)
Baena-Extremera et al. (2021)	2	1	2	3	2 (medium)
Berti & Cigala (2023)	2	2	2	1	1.75 (medium)
Brann et al. (2023)	2	2	3	1	2 (medium)
Crescentini et al. (2016)	2	1	2	2	1.75 (medium)
Crooks et al. (2020)	2	2	3	2	2.25 (medium)
Flook et al. (2010)	3	2	3	3	2.75 (high)
Folch et al. (2021)	2	2	2	2	2 (medium)
Frank et al. (2021)	2	3	3	2	2.5 (high)
Janz et al. (2019)	2	3	2	3	2.5 (high)
Koncz et al. (2021)	2	3	2	2	2.25 (medium)
Lam & Seiden (2020)	2	2	1	2	1.75

					(medium)
Lassander et al. (2020)	2	3	2	2	2.25 (medium)
Lertladaduck et al. (2021)	2	2	1	2	1.75 (medium)
Magalhaes et al. (2022)	2	1	2	2	1.75 (medium)
Makmee et al. (2022)	2	1	1	1	1.25 (low)
Milare et al. (2021)	2	2	2	2	2 (medium)
Muller et al. (2021)	2	2	2	1	1.75 (medium)
Quach et al. (2016)	3	1	2	2	2 (medium)
Ricarte et al. (2015)	2	2	1	2	1.75 (medium)
Salmoirago et al. (2019)	2	1	3	2	2 (medium)
Schonert-Reichl et al. (2015	2	2	2	3	2.25 (medium)
Shlomov et al. (2023)	2	2	2	2	2 (medium)
Suarez-Garcia et al. (2020).	2	2	2	2	2 (medium)
Thierry et al. (2016)	2	2	3	3	2.5 (high)
Thomas & Atkinson (2016)	2	2	3	2	2.25 (medium)

Vickery & Dorjee (2016)	2	2	3	3	2.5 (high)
Wimmer et al. (2016)	2	2	2	2	2 (medium)
Wood et al. (2017)	2	1	3	2	2 (medium)
Zelazo et al. (2018)	2	3	3	1	2.25 (medium)

Appendix K

Table summarising all EF-related outcome measures included in the meta-analysis

Study	N	EF-related outcome measure (mode)	function	effect size	descriptor	WoE D
Andreu et al. (2023)	46	Go/NoGo Task: NoGo Infrequent (**PBT**)	Inhibition	0.25	Small	1.75
		Early Adolescent Temperament Questionnaire-Revised (EATQ-R) (**BRS-SR**)	Global EF	0.83	Large	
		Behaviour Rating Inventory of Executive Function, 2nd Ed. (**BRS-TR**)	Global EF	0.68	Moderate	
Baena-Extremera et al. (2021)	320	d2 Test of Attention (**PBT**): Total Effectiveness	Attention	0.68	Moderate	2.18
		d2 Test of Attention (**PBT**): Commissions	Inhibition	0.07	Negligible	
Berti & Cigala (2020)	21	Go/NoGo Task (**PBT**)	Inhibition	1.64	Large	2.12
Brann et al. (2022)	33	Head-Toes-Knees-Shoulders Task (**PBT**)	Global EF	0.21	Small	2.33
		BRIEF-P (**BRS-PR**): Inhibit Scale	Inhibition	0.33	Small	
		BRIEF-P (**BRS-PR**): Shift Scale	Flexibility	0.40	Small	
		BRIEF-P (**BRS-PR**): Working Memory Scale	WM	0.06	Negligible	
Crescentini et al. (2016)	31	Revised Conners Teaching Rating Scale (CTRS-R, **BRS-TR**): Inattention	Attention	0.34	Small	1.58
Crooks et al. (2020)	323	BRIEF, Preschool Version (BRIEF-P, **BRS-TR**): Global Executive Composite	Global EF	0.34	Small	2.02
Flook et al. (2010).	64	BRIEF: Teacher Global Executive Composite (**BRS-PR**)	Global EF	0.12	Negligible	1.98
		BRIEF: Parent Global Executive Composite (**BRS-TR**)	Global EF	0.07	Negligible	
Folch et al. (2021)	100	Rey Auditory Verbal Learning Test (**PBT**)	WM	0.29	Small	1.2
Frank et al. (2021)	230	Stroop Task (**PBT**): Reaction Time	Inhibition	0.27	Small	2.5

		Emotional Faces N-back Task (**PBT**): nback hits	Inhibition	-0.1*	Negligible	
		Difficulties in Emotion Regulation Scale (DERS; **BRS-SR**): Impulse Control	Inhibition	0.19	Negligible	
		Difficulties in Emotion Regulation Scale (DERS; **BRS-SR**): Goal	Global EF	0.23	Small	
Janz et al. (2019).	87	Flanker Task (**PBT**)	Inhibition	0.36	Small	2.1
		Dimensional Change Card Sort (DCCS) Test (**PBT**)	Flexibility	0.55	Moderate	
		Strengths and Difficulties Questionnaire (SDQ; **BRS-TR**): Hype. /Att.	Global EF	0.96	Large	
Koncz et al. (2021)	56	Corsi Block Task (**PBT**): Backwards	WM	0.39	Small	2.02
		Go/NoGo Task (**PBT**): Commission Error	Inhibition	0.18	Small	
		Heart and Flowers Task (**PBT**): Error	Flexibility	0.29	Small	
Lam & Seiden (2020)	96	BRIEF-2 (**BRS-SR**): Working Memory Scale	WM	0.68	Moderate	2.38
		Difficulties in Emotion Regulation Scale (DERS; **BRS-SR**): Goal	Global EF	0.3	Small	
		Youth Self Report (YSR) (**BRS-SR**): Attention Syndrome Scale	Attention	0.19	Small	
		BRIEF-2 (**BRS-S**): Working Memory Scale	WM	0.68	Moderate	
Lassander et al. (2020)	117	WISC-IV Backward digit span subtest (**PBT**)	WM	0.27	Small	2.55
		NEPSY-II Inhibition A (shapes) subtest (**PBT**)	Inhibition	-0.27*	Small	
		D-KEFS Trail Making Test (**PBT**)	Flexibility	0.01	Negligible	
Lertladaluck et al. (2021)	30	Dimensional Change Card Sort (DCCS) (**PBT**)	Flexibility	0.80	Large	1.38
		Bear and Lion Task (**PBT**)	Inhibition	0.70	Large	
		Peg Tapping Task (**PBT**)	Inhibition	0.62	Moderate	
		Missing Scan Task (**PBT**)	WM	0.38	Small	
Magalhaes et al. (2022)	57	Vanderbilt ADHD Diagnostic Teacher Rating Scale (**BRS-TR**): Attention	Attention	0.37	Small	2.25
		Attentional Network Task (**PBT**): Orienting	Flexibility	0.17	Small	

Milare et al. (2022)	207	Five Digit Test (**PBT**): Inhibition	Inhibition	0.62	Moderate	1.33
		Five Digit Test (**PBT**): Flexibility Score	Flexibility	0.62	Moderate	
Muller et al. (2021)	79	d2 Test of Attention, Revised (d2-R, **PBT**)	Attention	-0.09*	Negligible	1.45
Quach et al. (2016)	103	Automated Operation Span Task (AOSPAN; **PBT**)	WM	0.65	Moderate	2.47
Ricarte et al. (2015)	90	WISC Digit Span (**PBT**): Backwards	WM	0.38	Small	1.72
		Trail Making Task (**PBT**): Part A	Attention	0.23	Small	
		Trail Making Task (**PBT**): Part B	Flexibility	0.09	Negligible	
Schonert-Reichl et al. (2015)	99	Flanker Task (**PBT**): Switch	Inhibition	-0.08*	Negligible	2.15
		Heart and Flowers Task (**PBT**)	Flexibility	-0.21	Small	
Shlomov et al. (2023)	28	Bear & Lion Task (**PBT**)	Inhibition	1.55	Large	2
		Dimension Change Card Sort (DCCS, **PBT**)	Flexibility	0.27	Small	
		Missing Scan Task (**PBT**)	WM	1.05	Large	
Suarez-Garcia et al. (2020)	73	Evaluation System for Children and Adolescents (ESCA; **BRS-TR**): Attention	Attention	0.54	Moderate	1.6
		ESCA (**BRS-TR**): Hyperactivity-Impulsivity Subscale	Inhibition	0.92	Large	
Thierry et al. (2016)	47	BRIEF-P (**BRS-TR**): Inhibit	Inhibition	0.35	Small	2.5
		BRIEF-P (**BRS-TR**): Shift	Flexibility	0.25	Small	
		BRIEF-P (**BRS-TR**): Working Memory	WM	1.01	Large	
Vickery & Dorjee (2016)	36	Behaviour Rating Inventory of Executive Function (BRIEF) (**BRS-TR**)	Global EF	0.43	Moderate	1.83

Wimmer et al. (2016)	26	German Test of Attentional Performance (GTAP; **PBT**): Moving Bar Task	Attention	0.32	Small	2.07
		Reversible Figures Task (**PBT**)	Flexibility	0.02	Negligible	
		Wisconsin Card Sorting Test (64; **PBT**)	Flexibility	0.26	Small	
		Stroop Colour-Word Inference Test (**PBT**)	Inhibition	1.43	Large	
Wood et al. (2017)	22	Researcher-designed Likert scale (**BRS-TR**): Total	Global EF	0.33	Small	1.53
		Likert Scale (**BRS-TR**): Attention	Attention	0.25	Small	
		Likert Scale (**BRS-TR**): Working Memory	WM	0.41	Moderate	
		Likert Scale (**BRS-TR**): Inhibition	Inhibition	0.40	Small	
		Likert Scale (**BRS-TR**): Shifting	Flexibility	0.34	Small	
Zelazo et al. (2018)	137	McClosey Executive Functions Scale (MEFS; **PBT**)	Global EF	-0.02	Negligible	2.55
		Heads-Toes-Knees-Shouldters (HTKS; **PBT**) Task	Global EF	0.16	Small	
		Children's Behaviour Questionnaire (CBQ; **BRS-PR**): Effortful Control	Inhibition	0.16	Small	
		Peg Tapping Task (**PBT**)	Inhibition	0.17	Small	

Note. Effect sizes have been calculated using hedge's *g* from descriptive statistics.

Appendix L

Overview of the frequency and duration of sessions in the planned intervention

Authors	Number of Sessions	Session Duration (mins)	Total Dosage (mins)	Duration of Trial Phase (weeks)	Intervention Minutes per week
Andreu et al. (2023)	9	50	450	9	50
Baena-Extremera et al. (2021)	30	4.45-16.26	±310	6	± 51.7
Berti & Cigala (2020)	15	24	360	6	60
Brann et al. (2023)	6	60	360	6	60
Crescentini et al. (2016)	24	±19.5	468	8	58.5
Crooks et al. (2020)	15	12.5	187.5	39*	variable
Flook et al. (2010)	16	30	480	8	60
Folch et al. (2021)	65	5-10	487.5	13	37.5
Frank et al. (2021)	12	60	720	6	120
Janz et al. (2019) **	-	-	-	-	-
Koncz et al. (2021)	6	45	270	3	90
Lam & Seiden (2020)	6	70	420	22	19.1

Lassander et al. (2020)	9	45	405	9	45
Lertladaluck et al. (2021)	24	40	960	8	120
Magalhaes et al. (2022)	16/24	30/5	600	8	75
Makmee (2022)	12	30	360	4	90
Milare et al. (2021)	16	30	480	8	60
Muller et al. (2021)	10	12	120	2	60
Quach et al. (2016	8	45	360	4	90
Ricarte et al. (2015)	30	15	450	6	75
Salmoirago et al. (2019)	8	45	360	8	45
Schonert-Reichl et al. (2015)	12	45	540	12	45
Shlomov et al., (2019)	24	30	720	8	90
Suarez-Garcia et al. (2020)	8/32	60/10	800	8	100
Thierry et al. (2016)	15	25	375	39*	variable
Thomas & Atkinson (2016)	6	60	360	6	60
Vickery & Dorjee (2016)	8	30/60	360	8	45
Wimmer et al. (2016)	25	60/90	1200	16	75

Wood et al. (2017)	12	25	300	6	50
Zelazo et al. (2018)	30	24	720	6	120

Note. *The intervention was delivered over the course of a school year, estimated to be 39 weeks based on a typical Canadian/US school calendar.

**Janz et al. (2019) did not report on intervention dosage.

Appendix M

An overview of behaviour rating scales included in the meta-analysis.

Measure	Description	Studies
	Attentional Control (n=5)	
YSR	A self-report questionnaire measuring emotional and behavioural problems in CYP.	Lam & Seiden (2020)
CTRS-R	A teacher-report rating scale measuring attention problems in CYP.	Crescentini et al. (2016)
SENA (Attention)	Comprehensive assessment tool that includes a measure of attention problems.	Suarez-Garcia et al. (2020)
Vanderbilt TSR	A teacher-report rating scale measuring behaviours associated with ADHD.	Magalhaes et al. (2022)
Likert (Attention)	Custom scale designed to measure attention through parent-report.	Wood et al. (2017)
	Cognitive Flexibility (n=4)	
BRIEF* (Shift)	Rating scale designed to measure flexible adaptation to changing situations and demands.	Brann et al. (2023); Lam & Seiden (2020); Thierry et al. (2016);
Likert (Shifting)	Custom scale designed to measure cognitive flexibility through parent-report.	Wood et al. (2017)
	Global EF (n=10)	
BRIEF*	An overall composite score of global executive functioning that includes questionnaire data targeting multiple EF components.	Andreu et al. (2023); Crooks et al. (2020); Flook et al. (2021); Vickery (2016)
DERS	Assesses various components associated with emotional regulation, taken as a proxy for EF.	Frank et al. (2021); Lam & Seiden (2020)
EATQ-R	Measures temperament dimensions relating to self-regulation and EF.	Andreu et al. (2023)

SDQ	Behavioural screening questionnaire that includes a subscale to assess attention and hyperactivity.	Janz et al. (2019)
Likert (total)	Composite score derived from three custom Likert scales.	Wood et al. (2017)
	Inhibition (n=7)	
BRIEF* (Inhibit)	Rating scale designed to measure impulsivity and ability to stop behaviour when needed.	Brann et al. (2023); Thierry et al. (2016)
DERS (Impulse)	Rating scale designed to assess problems with impulse control.	Frank et al. (2021)
CBQ (Emotional Control)	This rating scale is designed to measure CYP's self-control over emotions and behaviour.	Zelazo et al. (2018)
SENA (Hyperactivity)	Comprehensive assessment tool that includes a measure of hyperactivity problems.	Suarez-Garcia et al. (2020)
Likert (Inhibit)	Custom scale designed to measure inhibition through parent-report.	Wood et al. (2017)
	Working Memory (n=4)	
BRIEF*	Rating subscale designed to measure working memory. Based on either parent, teacher, or self-report.	Brann et al., 2023; Lam & Seiden 2020; Thierry et al., 2016
Likert (WM)	Custom scale designed to measure working memory through parent-report.	Wood et al. (2017)

Note. n indicates the number of BRS effect sizes included in the meta-analysis for each cognitive process. *BRIEF category includes BRIEF, BRIEF-2, and BRIEF-P outcome measures. **Flook et al. (2021) provided two outcome measures from the global executive composite of the BRIEF-2 (teacher and parent report).

Appendix N

An overview of performance-based tasks included in the meta-analysis.

Measure	Description	Studies
	Attentional Control (n=4)	
D2 (total)	A measure of sustained attention. Participants cross out target stimuli.	Baena-Extremera et al., 2021; Muller et al., 2021.
TMT A	Trail Making Test Part A. Participants connect numbered circles to form a trail.	Ricarte et al., 2015
GTAP moving bar (RA hits)	German Test of Attentional Performance. Participants respond to a moving stimulus.	Wimmer et al., 2016
	Cognitive Flexibility (n=11)	
DCCS	Dimensional Change Card Sort. Participants sort cards according to changing rules.	Janz et al., 2019; Lertladaluck et al., 2021; Shlomov et al., 2023
Heart & Flowers	Stroop-based task where CYP are shown hearts, flowers, or a combination of both.	Koncz et al., 2021; Schonert-Reichl et al., 2015;
D-KEFS Trial	Delis-Kaplan Executive Function System Trial Making Test. Participants switch between number and letter sequences.	Lassander et al., 2020
ANT Orienting	Attentional Network Task. Participants switch attention between changing spatial locations.	Magalhaes et al., 2022
5D Flexibility	Stroop-based task. Participants switch between automatic and controlled processing.	Milare et al., 2021
TMT B	Trial Making Test B. Participants switch between connecting numbers and letters to form trails.	Ricarte et al., 2015
Reversible Figures	Participants switch between perceptions of ambiguous figure images.	Wimmer et al., 2016

WCST	Wisconsin Card Sorting Test. Participants sort cards according to changing rules.	Wimmer et al., 2016
	Global EF (n=3)	
HTKS	Head-Toes-Knees-Shoulders (HTKS) Task. CYP perform opposites action to administrator.	Brann et al., 2023; Zelazo et al., 2018
MEFS	Minnesota Executive Function Scale. Battery of PBTs designed to measure global EF.	Zelazo et al., 2018
	Inhibition (n=13)	
Go/NoGo	Participants respond to "go" stimuli while inhibiting responses to "nogo" stimuli.	Andreu et al., 2023; Berti & Cigala, 2023; Koncz et al., 2021
Bear and Lion	Participants perform actions that are either congruent or opposite to storyteller's commands.	Lertladaluck et al., 2021; Shlomov et al., 2023
Peg Tapping	Participants tap twice when administrator taps once, and vice-versa.	Lertladaluck et al., 2021; Zelazo et al., 2018
D2 (error)	Test of Attention error score records number of commission errors made.	Baena-Extremera et al., 2021
Stroop	Participants name the colour of a coloured word whilst inhibiting reading the printed word.	Frank et al., 2021; Wimmer et al, 2016
Flanker	Participants focus on a stimulus while inhibiting the influence of nearby (flanking) stimuli.	Janz et al., 2019;
NEPSY-II Inhibition	Participants name shapes or arrows while inhibiting	Lassander et al., 2020
5D Inhibition	Stroop-based task. Participants count number of digits and inhibit reading the full number.	Milare et al., 2021
	Working Memory (n=8)	
Digit span backwards	Participants listen to a string of numbers and repeat them backwards.	Lassander et al., 2020; Ricarte et al., 2015
Missing scan	Participants identify a missing segment briefly presented in a grid.	Lertladaluck et al., 2021; Shlomov et al., 2023

RAVLT	Rey Auditory Verbal Learning Test. Participants recall a list of words after varying delays.	Folch et al., 2021
EFN-back	Emotional Faces N-back task. Participants identify emotional faces matching those presented n-trials before.	Frank et al., 2021
Corsi	Block-tapping task where participants reproduce a sequence of block taps in same order.	Koncz et al., 2021
AOSPAN	Automated Operation Span Task. Participants solve maths problems while remembering letters.	Quach et al., 2016

Note. n indicates the number of BRS effect sizes included in the meta-analysis for each ECP.

Appendix O

Summary of subgroup analyses performed on all 69 outcome measures aggregated within 27 studies

Covariate	k	*g*	95%CI	*p*	I^2	$p_{subgroup}$
WoE D rating[1]						0.01
low	8	0.42	0.18, 0.65	0.004	30.2%	
medium	14	0.36	0.21, 0.51	0.0002	34.6%	
high	5	0.15	0.02, 0.29	0.04	0.0%	
Key Stage						0.47
EYFS	7	0.34	0.22, 0.46		0.0%	
KS1	3	0.73	-0.87, 2.34		65.8%	
KS2	11	0.34	0.16, 0.52		30.6%	
KS3	5	0.25			55.6%	
KS4	1	0.15			n/a	
Administrator						0.59
External	14	0.34	0.13, 0.55		55.2%	
Teacher	11	0.30	0.20, 0.39		0.0%	
Mixed	2	0.52	-2.33, 3.38		37.0%	

SES[2]						0.41
Low	8	0.33	0.14, 0.51		16.2%	
Middle	9	0.33	0.07, 0.59		59.3%	
High	4	0.21	0.02, 0.40		0.0%	
Culture Group						<0.0001
Anglophone	11	0.28	0.13, 0.43	0.002	23.6%	
Central & Northern Europe	4	0.07	-0.20, 0.63	0.47	0.0%	
East Asia	2	0.28	-2.17, 2.74	0.38	10.0%	
Latin America	2	0.61	0.32, 0.90	0.02	0.0%	
Mediterranean	8	0.42	0.20, 0.35	0.003	32.5%	
Outcome Measure						0.76
BRS	10	0.35	0.24, 0.45	0	0.0%	
PBT	17	0.32	0.16, 0.49	0.04	50.7%	

[1]As the size of all subgroups in the WoE D Rating subgroup analysis is greater than 5 (k_g>5), a common estimate of between-study heterogeneity (τ^2) was included (Borenstein et al., 2011). For all other subgroup analyses, some groups have less than 5 studies, indicating that a common estimate of between-study heterogeneity (τ^2) would be imprecise (Borenstein et al., 2011). Instead, groups use a pooled τ^2.

[2]to explore the potential moderator of socio-economic status, a subgroup analysis on aggregate data was performed in which studies lacking SES data (n=6) were removed.