# DOES PRESCHOOL EXECUTIVE FUNCTION PREDICT SOCIAL, HEALTH AND BEHAVIORAL OUTCOMES? A META-ANALYSIS

by

Nicole J. Stucke
A Thesis
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Master of Arts
Psychology

Committee:

_________________________________________ Director

_________________________________________

_________________________________________

_________________________________________ Department Chairperson

_________________________________________ Dean, College of Humanities and Social Sciences

Date: ____________________________________ Spring Semester 2022
George Mason University
Fairfax, VA

# Does Preschool Executive Function Predict Social, Health and Behavioral Outcomes? A Meta-Analysis

A Thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts at George Mason University

by

Nicole J. Stucke
Bachelor of Science
University of Minnesota, 2015

Director: Sabine Doebel, Assistant Professor
Department of Psychology

Spring Semester 2022
George Mason University
Fairfax, VA

ACKNOWLEDGEMENTS

It takes a village, and this work would not have been possible without the support of my colleagues, family, and friends. First and foremost, I am forever grateful to my advisor, Sabine Doebel, for her thoughtful guidance and contagious enthusiasm and curiosity for child psychology and cognitive development. I am also deeply appreciative of my thesis committee members, David B. Wilson and Reeshad Dalal, for their helpful perspectives and questions. I want to acknowledge Sydney Foss, Songhan Pang, Isabela Alfaro, Mary Alphonse, Scarlett Bird-Guerra, and Jordan Hassani for their dedication and contentiousness in working on different aspects of this project with me. I also want to acknowledge the administrative staff in George Mason University's Psychology Department, specifically Tanie Boeddeker, who was immensely helpful in getting me to this point. Thank you all.

To my friends, to my family, to Ryan and my girls, for keeping me sane and for your unwavering support. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

DOES PRESCHOOL EXECUTIVE FUNCTION PREDICT SOCIAL, HEALTH AND BEHAVIORAL OUTCOMES? A META-ANALYSIS

Nicole J. Stucke, M.A.

George Mason University, 2022

Thesis Director: Dr. Sabine Doebel

Executive function is a widely studied psychological construct proposed to play a key role in healthy development and success in life. In children, executive function is often measured using particular behavioral laboratory tasks. Performance on these tasks robustly correlates with academic-related outcomes, yet they have also been claimed to predict a variety of outcomes outside the classroom, such as social skills, externalizing behaviors, and physical health. The evidence for these latter claims is less clear. Here, I report a meta-analysis testing the relation between executive function measured in preschool, and social, health, and behavioral outcomes measured concurrently and in later childhood and adolescence. Findings from 20 meta-analyses are reported. There were 853 usable effect sizes across the 115 included studies. A total of 104,827 children (*m* age = 55.62 months, *SD* = 6.07 months; 48.75% female) were included. For concurrent social outcomes, preschool executive function was positively related to social competence,

prosociality, peer acceptance, and emotion understanding and regulation. Effect magnitudes (expressed as *r*) ranged between 0.10 and 0.28, indicating small effects. For concurrent health outcomes, preschool executive function was negatively related to body mass indices (*r* = -0.15) but was not related to physical fitness. For concurrent behavioral outcomes, executive function was related to externalizing problems, lie understanding, adaptive classroom behaviors, and attention and hyperactivity symptoms, but not internalizing problems (*rs* between -0.04 to 0.25). Longitudinally, executive function was related to social competence, peer acceptance, adaptive classroom behaviors, externalizing problems, and attention and hyperactivity symptoms, but not prosociality, emotion understanding, or internalizing problems. Considering that few studies controlled for known covariates (e.g., verbal skills, age), we urge caution in interpreting these significant findings as support for the importance of executive function in social, health, and behavioral development over the lifespan. Future research can further explore these patterns to better understand the role of executive function in adaptive human functioning.

# INTRODUCTION

Executive function—the control and coordination of thought and action in the service of goals—is critical to human functioning and achievement. Early frameworks focused on top-down processes associated with complex problem solving (Zelazo et al., 1997; Zelazo & Cunningham, 2007). In recent decades, however, reductive models have been adopted, defining executive function in terms of two or more separable-but-related component processes (Miyake et al., 2000; Miyake & Friedman, 2012). Although theory and empirical findings have challenged these ideas (Doebel, 2020; Karr et al., 2018), a three-factor model including working memory/updating, inhibitory control, and cognitive flexibility/shifting, has emerged as a standard definition in developmental literature (Diamond, 2013). Early research also suggested that these components may support more complex executive functions like planning (Miyake et al., 2000) and it is widely claimed that these components are "building blocks" for self-control and regulation more broadly (e.g., Diamond, 2013; Miyake & Friedman, 2012).

Laboratory measures corresponding to these processes are now widely used to assess executive function in children and examine its relations with various skills and outcomes. Executive function was first studied in patients with injuries to the frontal lobe, and later several tasks were adapted for use with children (Diamond & Taylor, 1996). For example, in the peg tapping task, subjects are taught two rules. In the first, they are instructed first to copy the tapping actions of an experimenter — one tap from an experimenter is to be followed by one tap from the subject. In the second set of rules, subjects are instructed to tap twice when the experimenter taps once and tap once when the experimenter taps twice. Just as an adult with frontal lobe damage would perseverate (persist with an inappropriate response) by continuing to mimic the experimenter's tapping actions despite being told to follow the second set of rules, young children are said to perseverate by complying with only one of the two sets of rules or tapping many times regardless of the experimenter's actions (Diamond & Taylor, 1996). Numerous developmentally appropriate tasks have been developed and are now widely used with children, including the Dimensional Change Card Sort (DCCS; Zelazo, 2006), Day/Night Stroop (Gerstadt et al., 1994), Tower of London (Kochanska et al., 1996), Peg-Tapping (Diamond & Taylor, 1996), Go-NoGo task (Luria, 1959) and more (Carlson, 2005).

While the reliability of these tasks is not well studied[1], the tasks have nevertheless gained prominence in part because of their face validity as measures of the control needed to suppress or override a prepotent response. That is, failure to engage control is obvious when children respond impulsively to a stimulus despite continuous feedback that a different (nondominant) response is required. They have also been suggested to measure control more optimally than report or observational measures, which may be

[1] Many executive function measures for adult populations have demonstrated acceptable psychometric properties (e.g., Paap & Sawi, 2016); however, there are surprisingly few studies investigating the reliability of executive function measures for preschool-age children. Of the limited number of studies that did consider reliability of such measures in children, results have varied (e.g., Beck et al., 2011; Müller et al., 2012). Nevertheless, several measures have been shown to have good to excellent test-retest reliability in children, including Go/No-go, Head-Toes-Knees-Shoulders, and Peg Tapping (*rs* = .57 to .93; Karalunas et al., 2020), and the Corsi Block task (*ICC* = .90; Alloway & Passolunghi, 2011), as well as Self-Ordered Pointing, Day/Night Stroop, and Fruit Stroop (*rs* = .76 to .93; Archibald & Kerns, 1999).

more easily influenced by systematic biases (e.g., observer and parental biases; Friedman et al., 2020).

Performance on these tasks has been found to robustly predict education-related outcomes and academic achievement (e.g., Allan et al., 2014; Blair & Razza, 2007; Blankenship et al., 2019; Carlson & Moses, 2001; Deer et al., 2020). As a result, executive function is currently being prioritized in early childhood policy, educational research and practice, and prevention programs (e.g., Bierman & Torres, 2016; Zelazo et al., 2016). Moreover, a number of intervention and executive function training programs have been developed in an effort to promote adaptive academic functioning in children (e.g., Cartwright et al., 2020; Espinet et al., 2013).

Executive function is also frequently claimed to be crucial to a broader variety of social, behavioral, and health outcomes. For example:

> Inhibitory control early in life appears to be quite predictive of outcomes throughout life, including in adulthood. When 1,000 children born in the same city in the same year were followed for 32 years with a 96% retention rate, Moffitt et al. (2011) found that children who at ages 3 to 11 had better inhibitory control (e.g., were better at waiting their turn, less easily distracted, more persistent, and less impulsive) were more likely as teenagers to still be in school and were less likely to make risky choices or to be smoking or taking drugs. They grew up to have better physical and mental health (e.g., were less likely to be overweight or to have high blood pressure or substance abuse problems), earn more, and be more law-abiding as adults 30 years later than were those with

> worse inhibitory control as children, controlling for IQ, gender, social class, and their home lives and family circumstances growing up. They were also happier as adults" (Diamond, 2013, pp. 141-142).

However, claims such as these often involve research that did not measure executive function at all, but rather a related construct: self-control. A key study that is often cited to support the claim that executive function is important for many outcomes actually used a composite measure of self-control composed of observational ratings of lack of control, as well as parent-, teacher-, and self-reports of impulsive aggression, hyperactivity, lack of persistence, inattention, and impulsivity (Moffitt et al., 2011). Similarly, studies focusing on executive function frequently cite longitudinal research involving delay of gratification as evidence that executive function predicts a broader range of outcomes (e.g., Shoda et al., 1990). Constructs such as self-control and delay of gratification, although conceptually related to executive function, are operationalized and assessed rather differently and frequently do not correlate with it (e.g., Eisenberg et al., 2019; Saunders et al., 2018). Thus, these findings do not tell us anything about whether executive function predicts these kinds of outcomes.

Concurrent studies of executive function and social, behavioral, and health outcomes show inconsistent patterns. For example, in a study of 131 typically developing kindergarteners, researchers found that executive function predicted teacher-reported prosocial behaviors, even after controlling for sex, mother's education, and verbal and nonverbal IQ (Hubert et al., 2017). A more recent study of 171 typically developing preschoolers, however, found no such relation (Tan et al., 2020). Similarly, whereas some studies have found executive function predicts children's body mass indices and weight gain (e.g., Levitan et al., 2015), many studies report no such association (Beck et al., 2020; Blair et al., 2020; Hughes et al., 2015; Pieper & Laugero, 2013). In terms of behavioral outcomes, a similarly inconsistent pattern emerges. While some studies have found children with higher executive function had lower BMI percentiles and were more likely to be categorized as normal weight (e.g., Schmitt et al., 2019), others have failed to replicate such findings (Keye et al., 2021).

While there has been no comprehensive review of studies measuring executive function and social, health, and behavioral outcomes, there have been meta-analyses involving related constructs and involving at least some overlap in tasks. For example, a recent meta-analysis including 150 studies investigated self-regulation as it relates to 25 discrete outcomes, categorized as depicting achievement, interpersonal behaviors, mental health, and healthy living, both in childhood and in later life (Robson et al., 2020). Authors found that self-regulation measured in preschool was positively related to social competency, and negatively related to internalizing problems, peer victimization, and externalizing problems in the early school years; negatively associated with criminal behavior, internalizing problems, depressive symptoms, obesity, and cigarette smoking, and alcohol and illicit drug use in later school years; and negatively associated with unemployment, criminal behavior, depression and anxiety, obesity, cigarette smoking, alcohol and substance abuse, and symptoms of physical illness in adulthood. Another large systematic review and meta-analysis (Smithers et al., 2018) explored associations between attention, self-regulation, and perseverance and later psychosocial, cognitive, and health outcomes in children. Results indicated that self-regulation was associated with internalizing and externalizing problems, social skills, and intelligence, but was unrelated to body mass index. These meta-analyses suggest related constructs predict at least some non-academic outcomes concurrently and longitudinally; however, as discussed earlier, self-regulation is distinct from executive function. While there is debate in the field about the ways in which these constructs are related (e.g., overlapping, supporting), it is clear that executive function is measured differently. If executive function, as traditionally defined and measured, does predict important outcomes beyond academic skills, as is so often claimed in the literature — particularly longitudinal outcomes — this would bolster the idea that it is a building block for self-control broadly and a valuable target for interventions.

The current meta-analysis tests the relation between executive function as traditionally defined and measured in preschool and social, health, and behavioral outcomes assessed concurrently and longitudinally. We also explore whether associations are moderated by publication status (published vs. not), child age, and by the component of executive function that was tapped by the measure used (e.g., inhibitory control, working memory, cognitive flexibility/shifting). The latter, theoretically-based moderation analysis is motivated by a 'unity and diversity' view of executive function, which posits that executive function is a multifaceted construct with different components (e.g., inhibitory control) that — while correlated — are separable (Friedman et al., 2008; Miyake et al., 2000). Thus, we aim to better understand whether the relation between executive function and social, health, and behavioral outcomes varies based on whether the tasks used are classified as measures of inhibitory control, working memory, or cognitive flexibility/shifting. The results of this meta-analysis deepen our knowledge as to whether individual differences in executive function as measured by standard laboratory measures is important to healthy functioning and adaptation.

# METHOD

Recently, there have been calls for greater transparency and reproducibility in meta-analytic work, with more consistent reporting standards and practices (Polanin et al., 2020). All preregistered plans, data, and analyses for this meta-analysis will be made available on the Open Science Framework: https://osf.io/w942t/.

## Inclusion and Exclusion Criteria

Because we were interested in executive function as traditionally conceptualized and measured in developmental literature, only studies including behavioral laboratory tasks explicitly proposed and/or treated as a standard assessment of executive function were included. We did not include observational, self-report, parent- and teacher-report measures of executive function, measures typically characterized as assessing delayed gratification, such as the marshmallow test or gift delay, or measures of “hot” executive function (see Zelazo & Carlson, 2012 for review), such as measures involving treats or other rewards. Outcomes were restricted to non-laboratory social, health, and behavioral outcomes, such that measures of academic achievement (e.g., math, reading, vocabulary, language) and theory of mind or false beliefs were not included. The search was restricted to having a measure of executive function collected in the preschool years (between the ages of 3.00–5.99 years), determined by the mean age of the sample. Only studies including healthy and typically developing participants were included. To be included in this meta-analysis, a study was required to report at least one zero-order correlation between preschool executive function and a social, health, and behavioral outcome. The search terms were developed by NJS and SD and are included in Appendix A. There were no restrictions in the type of publication, but only papers written in English were included.

**Searching for Eligible Studies**

We comprehensively searched the literature for all eligible studies. The search followed a two-pronged approach. We first conducted a systematic search of three electronic databases, PsycINFO, Web of Science, and ERIC, to identify research published prior to October 2021. Then, in an effort to address potential publication selection bias toward inclusion of only positive results (Lipsey & Wilson, 2001), we employed several additional methods to locate unpublished literature. Figure 1 summarizes all search and screening procedures.

**Abstract and Full-text Screening**

The initial database and other methods search resulted in 8,824 records. After duplicates were removed, 5,101 records remained. Those 5,101 records were then uploaded to Abstrackr, a text-mining software designed specifically to streamline the abstract screening phase of systematic reviews (Wallace et al., 2012). To be used in conjunction with Abstrackr, we developed a title, keyword, and abstract screening tool (see Appendix B), that included seven single-barreled, clear and concise questions to which the answer can only be yes, no, or unsure[2]. Questions were organized

[2] This initial screening of title and abstracts was designed to determine *potential eligibility*, not *final* eligibility, meaning the goal was to simply identify whether the abstract indicated that a study was *clearly not eligible* or might be eligible upon closer inspection. Final eligibility was based on a final screening of the full text.

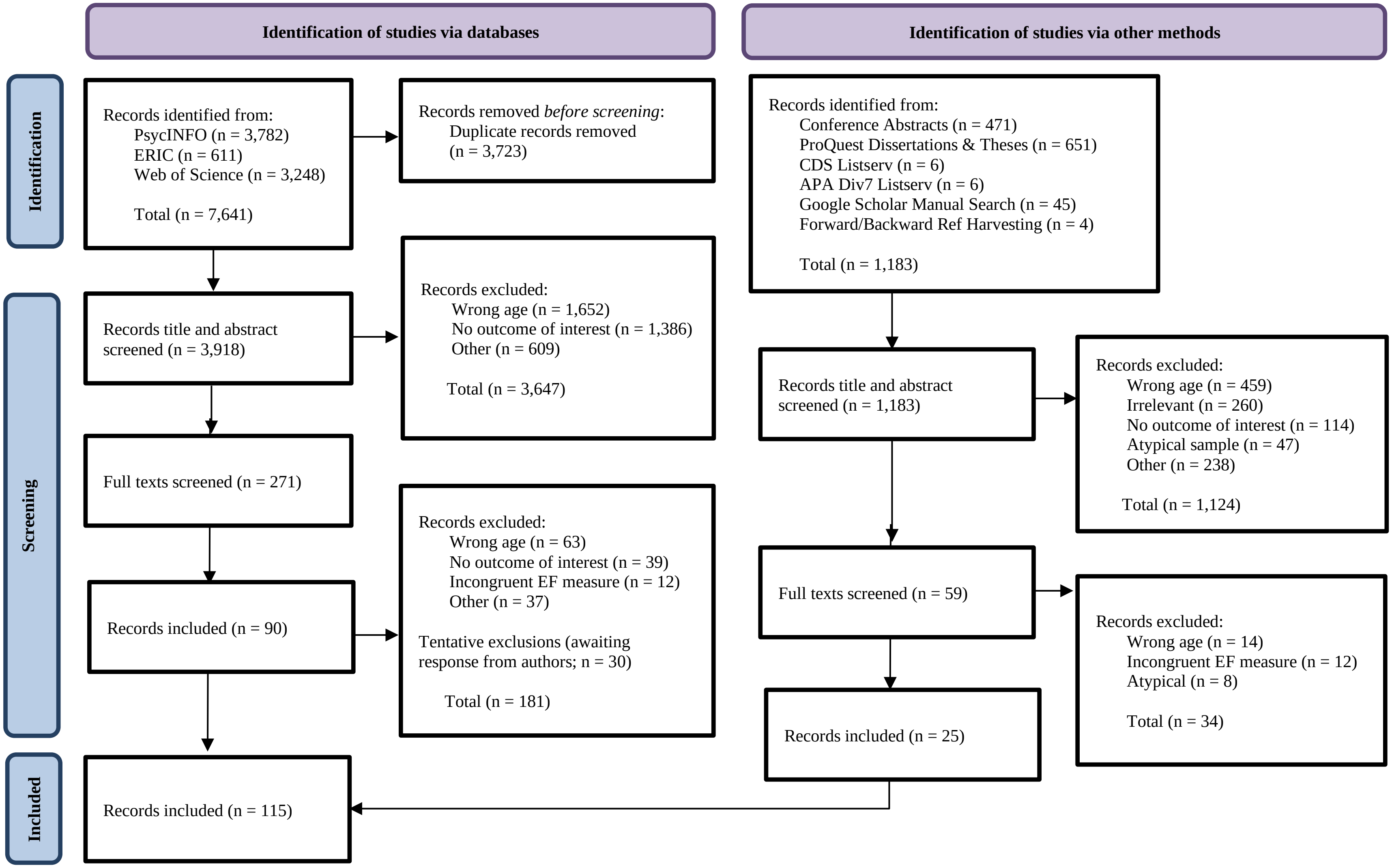



**Figure 1** PRISMA Flow Diagram of Search and Screening Procedures and Results

hierarchically, with the easiest questions at the beginning, allowing for a more efficient screening system[3]. Title, keyword, and abstract screening was conducted by a single researcher (NJS). If a study's eligibility could not be determined from the title and abstract, the full text was obtained and screened by a single researcher (NJS; $k$ = 330). Full texts for all studies were then screened by author NJS who determined final eligibility.

To ensure that no eligible studies were missed, we used several additional search strategies. We attempted to locate unpublished data by searching ProQuest Dissertations and Theses, Google Scholar, and the Open Science Framework Preprint repository, which includes various preprint databases such as PsyArXiv and EdArXiv. We also conducted a manual search of conference programs from the last 5 years of meetings for the American Psychological Association, Association for Psychological Science, Cognitive Development Society, and Society for Research in Child Development. Finally, researchers in the field were contacted via the Cognitive Development Society Listserv and the American Psychological Association Division 7 Listserv for any supplementary and/or unpublished data. A final total of $k$ = 115 studies were included.

**Data Extraction and Coding Procedures**

Once all eligible studies were identified, data were extracted and coded by two[4] independent researchers following guidelines for reporting outlined in the Preferred

---

[3] Once a study failed to meet the criteria described in the question, it was screened out. Once screened out, a study could not be screened back in. Addressing 'easy' (i.e., less subjective) questions early helped speed up the screening process (studies were eliminated right away and not later in the process).
[4] Data extraction and coding is complete by independent coder 1 (NJS) but is still in progress by the second independent coder.

Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher et al., 2009). Variables coded include information pertaining to study-level variables, such as authors, publication date, study design, methods, participant/sample demographic information (e.g., sample size, attrition rate, age, etc.), and sampling method (random vs. convenience); information pertaining to executive function task-level variables, such as the measures administered, the specific component of executive function authors were attempting to capture (e.g., inhibitory control, shifting, working memory), whether measures were aggregated into a composite or treated as a single measure, and whether there was risk of bias; information pertaining to outcome-level variables, such as the measures used, how the measures were administered, and whether there was evidence that the result was biased by missing outcome data; and information pertaining to effect sizes. All data coded and the codebook used in this meta-analysis will be made available on the Open Science Framework: https://osf.io/w942t/. Once independent double coding is complete, we will report the percentage of coder agreement.

### ***Extracting and Computing Effect Sizes***

It is important that the effect size of interest provides appropriate standardization of the studies that are being investigated (Lipsey & Wilson, 2001). Due to the nature of this literature, most studies reported a non-adjusted correlation between two continuous variables, thus a zero-order correlation coefficient (Pearson's *r*) was selected as the effect size statistic of interest. In cases in which a useable correlation coefficient was not reported or when partial correlations were reported, the study authors were contacted ($n$ = 29) to obtain the necessary statistics to calculate a zero-order correlation. In several instances ($n = 3$), a study reported a useable effect size, but the sample included sets of twins. Similar to when a useable correlation coefficient was not reported or when partial correlations were reported, the study author was contacted to obtain an effect size that reflects a sample of non-twin children only. If we received no response ($n = 18$), the study was not included in the main model (see Figure 1).

Many studies included multiple eligible executive function measures and outcomes. When we have multiple effect size estimates from a single study based on the same participants, we must take into account the fact that these correlations are not independent and share both variance and error (Borenstein et al., 2009). Therefore, to compute a precise estimate of the effect, we combined correlations into a summary or synthetic effect (one per study). Procedures followed are described in Chapter 24 of Borenstein et al., (2009). This same logic does not apply, however, when a study provides multiple correlations from separate samples of children. In this case, we calculated one effect size estimate for each group of children (i.e., data were treated as separate studies).

We judged effect sizes in context. That is, even if executive function is found to predict social, health, and behavioral outcomes, it is important to go beyond statistical significance and consider the magnitude of the effects, especially if executive function is being considered a target for interventions. Thus, when interpreting our effect sizes, we followed the recommendations outlined by Durlak (2009):

(a) consider whether the designs, types of outcomes, and methods of calculating effects are the same across studies.

(b) evaluate the magnitude of the effect based on the research context and its practical or clinical value.

(c) if effects from previous studies are not presented, strive to calculate some using the procedures described here and in the additional references.

(d) use Cohen's (1992) benchmarks ($r$ = .10 considered a small effect, $r$ = .30 considered a medium effect, and $r$ = .50 considered a large effect size) only if comparisons to other relevant research are impossible.

**Meta-Analytic Strategy**

Effect sizes were extracted, coded, prepared, pooled, and then analyzed using inverse-variance weighted random effects meta-analysis using the metafor package (Viechtbauer, 2010) in R (R Core Team, 2014). In psychology research, regression or multiple regression analyses are common practice, as they are designed to test the relationship between independent and dependent variable and one or more covariates or moderators (Borenstein et al., 2009). In meta-analysis, the same logic can be applied, except that covariates or moderators are at the study-level rather than the subject-level and the dependent variable represents the study effect size as opposed to subject scores (Borenstein et al., 2009). This method is advantageous because it assigns the effect sizes a weight that is equal to the inverse of its variance, meaning more precise studies will have greater overall weight (Borenstein et al., 2009; Hedges & Olkin, 1985). We analyzed several inverse-variance weighted random effects models of pooled mean effect sizes. Models included:

model 1 = all cross-sectional and concurrent data;

model 2 = outcomes in middle childhood, longitudinal data with a follow–up in middle childhood [ages 6.0 – 11.99 years]; and

model 3 = outcomes in late childhood, longitudinal data with a follow–up in late childhood [ages 12.0 – 17.99 years];

We required that each model include a minimum of $k = 2$ studies to be included.

### ***Missing Data***

Missing data were handled following the infer, initiate, and impute process described by Pigott and Polanin (2020). Following this process, we first inferred, meaning we made an educated inference based on what authors state in the article. For example, if a study did not report means, standard deviations, and ranges for child age, but instead reported the child's school grade, ages were approximated on the basis of knowledge of school attendance age in the country in which data were collected. If inferring was not possible, we initiated, meaning we contacted the primary author(s) to ask directly for the missing information. In cases in which we received no response from the primary authors (i.e., when infer and initiate fail), we imputed, or made a decision regarding the handling of the specific missing data (e.g., listwise deletion, maximum likelihood, multiple imputation, etc.) following recommendations by Pigott (in press). For example, in cases with missing moderator data, we conducted a multiple imputation, but in cases with other missing data, such as missing sex information, we conducted a mean-value imputation as recommended by Pigott (2019). We performed a sensitivity analysis to assess the extent to which results reflect or depend upon the way in which we handled missing data by including a "missingness" dummy variable to evaluate whether the studies for which that variable were missing produced systematically smaller or larger effect sizes (Lipsey & Wilson, 2001).

***Heterogeneity and Moderator Analyses***

To test for heterogeneity of effect size estimations to determine how effect sizes varied between studies, we report Cochran's $Q$ and the $I^2$ statistic. Cochran's $Q$ is estimated using the weighted sum of squared differences between the individual study effects and the pooled effect estimate across all studies (Cochran, 1954). The $I^2$ index, which is a more advisable procedure for quantifying the degree of heterogeneity in a meta-analysis (Huedo-Medina et al., 2006), was calculated by adding our $Q$ value to degrees of freedom, dividing by the $Q$ value itself, and then multiplying by 100 (Higgins & Thompson, 2002). Results are interpreted using percentages, meaning an $I^2$ index of 0 indicates that all variation in effect sizes is due to sampling error within studies (i.e., 0% of the variance among effect sizes is caused by true heterogeneity between studies), whereas an $I^2$ index of 50 indicates that 50% of the total variance among effect sizes is caused by true heterogeneity between studies (and not by sampling error or chance), and so on. When the $I^2$ index reaches or exceeds a moderate level (i.e., 50%) and the $Q$ statistic indicates a significant amount of heterogeneity, it is appropriate to conduct a moderator analysis to attempt to identify factors that may be contributing to the variance. In addition to examining the meta-analytic relations between preschool executive function and outcomes, we explored whether these relations varied as a function of child age (mean centered), publication status (published vs. unpublished dissertation or thesis), as well as by the executive function task category (inhibitory control, working memory, or cognitive flexibility/shifting). We required a minimum of $k = 5$ studies in each moderator category to preserve statistical power. In testing for differences in the strength of the association by publications status, we note that there were only two models in which there the minimum number of five studies in each of the two dichotomous categories (published vs. not published): social competence model 1 and externalizing behavior problems models 1. Results are reported below. For the moderator analyses looking at executive function task category, we coded each task as measuring working memory, inhibitory control, or cognitive flexibility using definitions and examples from Diamond (2013), Miyake et al. (2000), and Wiebe et al. (2008). Composite measures were not included in these analyses. Since we did not have any a priori hypotheses about the relative predictive strength of these task categories, we created a full set of contrast codes allowing us to conduct the equivalent of an omnibus ANOVA, testing whether there were any differences between the task categories in the strength of the relation between executive function and specific outcomes. Specifically, one variable coded the contrast between working memory (1) and inhibitory control (–1), and the second coded the contrast between the average of these two task categories (working memory = –1, inhibitory control = –1) and cognitive flexibility (–2). Significant omnibus tests were to be followed up with additional pairwise tests as needed. Notably, there was only one outcome model, social competence model 1, in which there were enough studies (again, minimum $k = 5$) with measures of working memory and inhibitory control to test for differences in the strength of the associations.

### ***Evaluating Potential Bias***

Bias is a well-documented issue in psychology research (e.g., Ferguson & Brannick, 2012) and is commonly thought to reflect issues related to missing studies (i.e., significant results having a higher probability of being published). A commonly used method of evaluating the possibility of publication bias is to test for differences in mean effects across published versus unpublished studies using a simple moderator analysis (Polanin et al., 2016). In addition to testing publication status as a moderator in our models, we addressed publication bias by visually examining of the symmetry of meta-analytic funnel plot to aid in the detection of bias or systematic heterogeneity. Funnel plots for each model and outcomes category are included in Appendix D. For several outcome models (namely for attention and hyperactivity symptoms model 1, externalizing behavior problems models 1, and emotion understanding model 1), studies were fairly symmetrical in terms of spacing around the summary effect size. Thus, it is less likely that publication bias was present. It is important to note, however, that for models where the total number of effects was much smaller ($k > 10$, which included several of our outcomes across social, health, and behavior domains), estimates of publication bias tend to be unreliable (Borenstein et al., 2009).

# RESULTS

Across social, health, and behavioral outcome domains, there were 12 outcomes: social competence, prosociality, peer acceptance, emotion understanding, and emotion regulation (social outcomes); body mass indices and physical fitness (health outcomes); and externalizing behavior problems, internalizing behavior problems, lie understanding, attention and hyperactivity symptoms, and adaptive classroom behaviors (behavioral outcomes).

Findings from inverse-variance weighted random effects meta-analyses are reported in Table 1. There were 853 usable effect sizes across the 115 included studies. Across all studies, a total of 104,827 children (*m* age = 55.62 months, *SD* = 6.07 months; range = 12.00 – 63.48 months; 48.75% female) were included. The majority of studies were conducted in the United States (*k* = 71), followed by Canada (*k* = 10), China (*k* = 7), Germany (*k* = 3), the United Kingdom (*k* = 3), Japan (*k* = 2), and Norway (*k* = 2). The remaining studies (*k* = 17) were conducted in various other countries. Overall, a number of different executive function measures were used (*k* = 112), with Day-Night Stroop being the most popular (*k* = 42), followed by the Dimensional Change Card Sort task (*k* = 40), Head-Toes-Knees-Shoulders (*k* = 26), Backwards Digit/Word Span (*k* = 16), the Continuous Performance Task (*k* = 13), and the Peg Tapping task (*k* = 11). Results of moderation analyses are included in tables 2 through 4. Forest and funnel plots for effect sizes included in each model are included in Appendices C and D.

Table 1 Results of Random-Effects Meta-analyses: Associations Between Preschool Executive Function and Social, Health, and Behavioral Outcomes

| | Model 1: Cross-sectional/concurrent | | | | | Model 2: Childhood outcomes | | | | | Model 3: Adolescent outcomes | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Outcome | *k* | *n* | *r* [95% CI] | *Q* | $I^2(\%)$ | *k* | *n* | *r* [95% CI] | *Q* | $I^2(\%)$ | *k* | *n* | *r* [95% CI] | $Q^2$ | $I^2(\%)$ |
| Social | | | | | | | | | | | | | | | |
| Social Competence | 31 | 10,887 | 0.18*** [0.12, 0.25] | 760.3 | 91.4 | 5 | 1,559 | 0.15* [0.02, 0.28] | 23.4 | 86.1 | -- | -- | -- | -- | -- |
| Prosociality | 13 | 2,239 | 0.21*** [0.12, 0.30] | 41.1 | 73.1 | 2 | 195 | 0.08 [-0.29, 0.45] | 6.5 | 84.7 | -- | -- | -- | -- | -- |
| Peer Acceptance | 4 | 1,406 | 0.10*** [0.05, 0.15] | 1.4 | 0.0 | 2 | 1,240 | 0.13*** [0.07, 0.18] | 0.3 | 0.0 | -- | -- | -- | -- | -- |
| Emotion Understanding | 10 | 1,881 | 0.28*** [0.22, 0.34] | 17.1 | 45.8 | 2 | 292 | 0.18 [-0.13, 0.49] | 2.7 | 62.5 | -- | -- | -- | -- | -- |
| Emotion Regulation | 3 | 464 | 0.10* [0.01, 0.19] | 0.6 | 0.0 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| Health | | | | | | | | | | | | | | | |
| Body Mass Indices | 3 | 224 | -0.15* [-0.28, -0.02] | 1.5 | 0.0 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| Physical Fitness | 4 | 291 | 0.18 [-0.01, 0.38] | 9.5 | 68.3 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| Behavioral | | | | | | | | | | | | | | | |
| Externalizing | 37 | 47,491 | -0.15*** [-0.18, -0.13] | 116.7 | 76.1 | 8 | 29,440 | -0.13*** [-0.19, -0.08] | 61.9 | 92.6 | -- | -- | -- | -- | -- |
| Internalizing | 10 | 13,584 | -0.04 [-0.12, 0.03] | 17.3 | 58.9 | -- | -- | -- | -- | -- | 2 | 1,309 | -0.03 [-0.08, 0.02] | 0.4 | 0.0 |
| Lie understanding | 6 | 653 | 0.21*** [0.14, 0.29] | 2.1 | 0.0 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| Attention & Hyperactivity | 18 | 4,872 | -0.19*** [-0.23, -0.14] | 36.4 | 54.5 | 6 | 5,945 | -0.22*** [-0.28, -0.16] | 36.6 | 81.5 | -- | -- | -- | -- | -- |
| Adaptive Classroom Behavior | 16 | 35,745 | 0.25*** [0.21, 0.29] | 105.8 | 88.6 | 6 | 27,016 | 0.23*** [0.19, 0.26] | 20.7 | 72.6 | -- | -- | -- | -- | -- |

*Note*. * $p$ <.05, ** $p$ <.01, *** $p$ <.001. $k$ = number of studies; $n$ = pooled sample size; $r$ = mean effect size; CI = confidence interval; $Q$ = Cochran's Q heterogeneity estimate using the weighted sum of squared differences between the individual study effects and the pooled effect estimate across all studies; $I^2$ = heterogeneity index expressed as a percentage.

Table 2 Results of Moderator Analysis Testing Child Age (mean centered)

| | | | | | | *95% CI* | |
|---|---|---|---|---|---|---|---|
| Model | $Q_M$ | *p* | *β* | *SE* | *z* | *LL* | *UL* |
| Social Outcomes | | | | | | | |
| Social Competence | | | | | | | |
| Model 1 | 0.80 | 0.37 | | | | | |
| Intercept | | | 0.19*** | 0.04 | 5.28 | 0.12 | 0.26 |
| Child Age | | | -0.00 | 0.00 | -0.90 | -0.01 | 0.00 |
| Prosociality | | | | | | | |
| Model 1 | 1.36 | 0.24 | | | | | |
| Intercept | | | 0.14* | 0.07 | 2.09 | 0.01 | 0.28 |
| Child Age | | | 0.01 | 0.01 | 1.16 | -0.01 | 0.03 |
| Behavior Outcomes | | | | | | | |
| Externalizing Behaviors | | | | | | | |
| Model 1 | 0.54 | 0.46 | | | | | |
| Intercept | | | -0.16*** | 0.01 | -11.58 | -0.19 | -0.13 |
| Child Age | | | -0.00 | 0.00 | -0.74 | -0.00 | 0.00 |
| Model 2 | 0.10 | 0.75 | | | | | |
| Intercept | | | -0.14** | 0.04 | -3.11 | -0.22 | -0.05 |
| Child Age | | | -0.00 | 0.00 | -0.31 | -0.01 | 0.01 |
| Internalizing Behaviors | | | | | | | |
| Model 1 | 0.00 | 0.98 | | | | | |
| Intercept | | | -0.05 | 0.04 | -1.19 | -0.13 | 0.03 |
| Child Age | | | -0.00 | 0.00 | -0.02 | -0.01 | 0.01 |
| Attention & Hyperactivity | | | | | | | |
| Model 1 | 0.23 | 0.63 | | | | | |
| Intercept | | | -0.18*** | 0.02 | -7.66 | -0.23 | -0.14 |
| Child Age | | | -0.00 | 0.00 | -0.48 | -0.01 | 0.00 |
| Model 2 | 0.02 | 0.89 | | | | | |
| Intercept | | | -0.22*** | 0.04 | -5.47 | -0.30 | -0.14 |
| Child Age | | | 0.00 | 0.01 | 0.14 | -0.02 | 0.02 |
| Classroom Behavior | | | | | | | |
| Model 1 | 0.25 | 0.62 | | | | | |
| Intercept | | | 0.25*** | 0.03 | 9.22 | 0.20 | 0.30 |
| Child Age | | | 0.00 | 0.00 | 0.50 | -0.00 | 0.01 |
| Model 2 | 3.04 | 0.08 | | | | | |
| Intercept | | | 0.19*** | 0.02 | 8.59 | 0.15 | 0.23 |
| Child Age | | | 0.06 | 0.04 | 1.74 | -0.01 | 0.13 |

*Note*. * $p$ <.05, ** $p$ <.01, *** $p$ <.001. $\beta$ = standardized coefficients. *SE* = standard error. *CI* = confidence interval. *LL* = lower limit. *UL* = upper limit. $Q_M$ = the omnibus test, which follows a chi-square distribution with *m* degrees of freedom (*m* denoting the number of coefficients tested) under the null hypothesis.

## Social Outcomes

For social outcomes there were 282 usable effect sizes across $k$ = 61 studies.

### ***Social Competence***

Executive function was positively and significantly associated with social competence, both concurrently, $k = 31$, $r = 0.18$ (95% CI [0.12, 0.25]), and longitudinally (model 2; ages 6 – 11.99 years), $k = 5$, $r = 0.15$ (95% CI [0.02, 0.28]). The heterogeneity output for models 1 and 2 showed $I^2 = 91.4\%$, $Q(30) = 760.3$ ($p < .001$) and $I^2 = 86.1\%$, $Q(4) = 23.4$ (p < .001), respectively, prompting a search for potential moderators; however, there was an insufficient number of effects to explore moderators in model 2, thus only moderators for model 1 were tested. There were no significant moderating effects for age ($p = 0.37$), publication status ($p = 0.30$), or executive function task category ($p = 0.12$).

Table 3 Results of Moderator Analysis Testing Publication Status

| Model | $Q_M$ | *p* | *β* | *SE* | *z* | *95% CI* *LL* | *95% CI* *UL* |
|---|---|---|---|---|---|---|---|
| Social Outcomes | | | | | | | |
| Social Competence | | | | | | | |
| Model 1 | 1.07 | 0.30 | | | | | |
| Intercept | | | 0.12* | 0.06 | 1.98 | 0.00 | 0.25 |
| Publication Status | | | 0.08 | 0.08 | 1.04 | -0.07 | 0.22 |
| Behavior Outcomes | | | | | | | |
| Externalizing Behaviors | | | | | | | |
| Model 1 | 0.20 | 0.66 | | | | | |
| Intercept | | | -0.17*** | 0.03 | -5.59 | -0.22 | -0.11 |
| Publication Status | | | 0.01 | 0.03 | 0.45 | -0.05 | 0.08 |

*Note*. * $p$ <.05, ** $p$ <.01, *** $p$ <.001. $\beta$ = standardized coefficients. *SE* = standard error. *CI* = confidence interval. *LL* = lower limit. *UL* = upper limit. $Q_M$ = the omnibus test, which follows a chi-square distribution with $m$ degrees of freedom ($m$ denoting the number of coefficients tested) under the null hypothesis. We required a minimum of $k = 5$ studies in each of the two moderator categories (published and unpublished) to preserve statistical power.

### ***Prosociality***

Similar to the findings for social competence, results from model 1 showed that children with better executive function tended to be more prosocial concurrently, $k = 13$,

$r = 0.21$ (95% CI [0.12, 0.30]), however this association was no longer significant when prosocial behavior was measured in later childhood, $k = 2$, $r = 0.08$ (95% CI [-0.29, 0.45]). There was considerable heterogeneity across both models 1 and 2, $I^2 = 73.1\%$, $Q(12) = 41.09$ ($p < .001$) and $I^2 = 84.7\%$, $Q(1) = 6.53$ ($p < .05$), respectively; however, there was an insufficient number of effects to test for moderators in model 2. The moderator analysis testing child age as moderating the strength of the relation between executive function and prosociality concurrently (model 1) was not significant ($p = 0.24$; see table 2), and there were not enough studies in each of the categories for publication status or executive function task category to test for moderation.

Table 4 Results of Moderator Analysis Testing Executive Function Task Category

| Model | $Q_M$ | *p* | *β* | *SE* | *z* | *95% CI* *LL* | *95% CI* *UL* |
|---|---|---|---|---|---|---|---|
| Social Outcomes | | | | | | | |
| Social Competence | | | | | | | |
| Model 1 | 3.68 | 0.06 | | | | | |
| Intercept | | | 0.23*** | 0.04 | 5.64 | 0.15 | 0.31 |
| WM vs. IC | | | 0.13 | 0.07 | 1.92 | -0.00 | 0.27 |

*Note*. * $p$ <.05, ** $p$ <.01, *** $p$ <.001. $\beta$ = standardized coefficients. IC = inhibitory control measures. WM = working memory measures. *SE* = standard error. *CI* = confidence interval. *LL* = lower limit. *UL* = upper limit. $Q_M$ = the omnibus test, which follows a chi-square distribution with *m* degrees of freedom (*m* denoting the number of coefficients tested) under the null hypothesis. We required a minimum of $k = 5$ studies in each of the moderator categories (inhibitory control, working memory, and/or cognitive flexibility) to preserve statistical power.

### *Peer Acceptance*

Children with higher executive function tended to be more accepted by their peers, both concurrently, $k = 4$, $r = 0.10$ (95% CI [0.05, 0.15]), and when children were followed up longitudinally in later childhood (age 6 – 11.99 years), $k = 2$, $r = 0.21$ (95% CI [0.07, 0.18]). The heterogeneity output showed that for both models, heterogeneity was in an acceptable range, $I^2 = 0.0\%$, $Q(3) = 1.44$ ($p = .70$) and $I^2 = 0.0\%$, $Q(1) = 0.34$ ($p = .56$), respectively.

### *Emotion Understanding*

For model 1, executive function was positively and significantly correlated with emotion understanding, $k = 10$, $r = 0.28$ (95% CI [0.22, 0.34]). This association was no longer significant, however, when emotion understanding was measured in later childhood, $k = 2$, $r = 0.18$ (95% CI [-0.13, 0.49]). There was an insufficient number of studies to test for moderators, and heterogeneity was in an acceptable range for both models 1 and 2, $I^2 = 45.8\%$, $Q(9) = 17.12$ ($p = .05$) and $I^2 = 62.5\%$, $Q(1) = 2.67$ ($p = .10$), respectively.

### *Emotion Regulation*

Children with higher executive function tended to have better emotion regulation concurrently, $k = 3$, $r = 0.10$ (95% CI [0.01, 0.19]). There was an insufficient number of effects to test for moderators, and heterogeneity was in an acceptable range, $I^2 = 0.0\%$, $Q(2) = 0.60$ ($p = .74$).

## Health Outcomes

Across all health outcomes there were a total of 56 usable effect sizes across $k = 8$ studies.

### *Body Mass Indices*

As shown in table 1, children with higher executive function in preschool tended to have lower body mass indices concurrently, $k = 3$, $r = –0.15$ (95% CI [–0.28, –0.02]).

Heterogeneity was acceptable, $I^2 = 0.0\%$, $Q(2) = 1.46$ ($p = .48$) and there was an insufficient number of effects to explore moderators.

### *Physical Fitness*

There was a positive association between preschool executive function and physical fitness, such that children with better executive function were more physically fit concurrently, however this association was not significant, $k = 4$, $r = 0.18$ (95% CI [–0.01, 0.38]). There was substantial heterogeneity, $I^2 = 68.3\%$, $Q(3) = 9.53$ ($p < .05$), however there was an insufficient number of effects to explore moderators.

## Behavior Outcomes

Across all behavioral outcomes we identified a total of 515 usable effect sizes from $k = 72$ studies.

### *Externalizing Behavior Problems*

Preschool executive function was negatively associated with externalizing behavior problems, both concurrently, $k = 37$, $r = –0.15$ (95% CI [–0.18, –0.13]) and when externalizing behavior problems were captured in later childhood (model 2), $k = 8$, $r = –0.13$ (95% CI [–0.19, –0.08]). The heterogeneity output for models 1 and 2 showed $I^2 = 76.1\%$, $Q(36) = 116.7$ ($p < .001$) and $I^2 = 92.6\%$, $Q(7) = 61.92$ ($p < .001$), respectively, prompting a search for potential moderators. In model 1, age and publication status were not significant predictors of the strength of the association, and there were not enough studies in each executive function task category to test for moderation. In model 2, child age was not a significant predictor of the strength of the negative correlation between executive function and externalizing behavior problems, and there were not enough studies in executive function task category or publication status to test for moderation.

### *Internalizing Behavior Problems*

Executive function was not associated with internalizing behavior problems concurrently (model 1), $k = 10$, $r = -0.04$ (95% CI [–0.12, 0.03]). When internalizing behavior problems were measured in adolescence (model 3), the negative association continued to be insignificant, $k = 2$, $r = -0.03$ (95% CI [–0.08, 0.02]). There was considerable heterogeneity in model 1, $I^2 = 58.9\%$, $Q(9) = 17.3$ ($p < .05$), prompting a search for potential moderators; however, there were no significant effects for age ($p = 0.98$) and an insufficient number of studies to test for moderation by publication status or executive function task category.

### *Lie Understanding*

Preschool executive function was positively related to children's lie understanding measured concurrently, $k = 6$, $r = 0.21$ (95% CI [0.14, 0.29]). Heterogeneity was acceptable, $I^2 = 0.0\%$, $Q(5) = 2.1$ ($p = .84$) and there was an insufficient number of effects to explore moderators.

### *Attention and Hyperactivity Symptoms*

Children with higher executive function tended to have fewer attention and hyperactivity symptoms, both when measured concurrently, $k = 18$, $r = -0.19$ (95% CI [–0.23, –0.14]), and when measured in later childhood, $k = 6$, $r = -0.22$ (95% CI [–0.28, –0.16]). There was substantial heterogeneity across both models 1 and 2, $I^2 = 54.5\%$, $Q(17) = 36.4$ ($p < .05$) and $I^2 = 81.5\%$, $Q(5) = 36.6$ ($p < .001$), respectively, prompting a search for potential moderators. Child age was not a significant predictor of the strength of the negative correlation between executive function and attention and hyperactivity symptoms problems concurrently ($p = 0.63$) or longitudinally ($p = 0.89$), and there were not enough studies in each executive function task category or publication status category to test for moderation of the association.

### ***Adaptive Classroom Behaviors***

Children with higher executive function tended to show more adaptive classroom behaviors, both when measured concurrently, $k = 16$, $r = 0.25$ (95% CI [0.21, 0.29]), and in later childhood (model 2), $k = 6$, $r = 0.23$ (95% CI [0.19, 0.26]). There was substantial heterogeneity across both models 1 and 2, $I^2 = 88.6\%$, $Q(15) = 105.8$ ($p < .001$) and $I^2 = 72.6\%$, $Q(5) = 20.7$ ($p < .001$), respectively, prompting a search for moderators. Child age was not a significant predictor of the strength of the positive correlation between executive function and adaptive classroom behaviors concurrently ($p = 0.62$) or longitudinally ($p = 0.08$), and there were not enough studies in each executive function task category or publication status category to test for moderation of the association.

## DISCUSSION

In this meta-analysis, we sought to shed light on whether executive function — as traditionally defined and measured in preschool — relates to important social, health, and behavioral outcomes assessed concurrently and longitudinally. Generally, the results are consistent with executive function being associated with several diverse, non-academic outcomes but limited longitudinal data leave open the question as to whether executive function is associated with outcomes in the long-term.

For social outcomes, children with higher executive function in preschool tended to have higher social competence, prosociality, peer acceptance, emotion understanding, and emotion regulation concurrently. This finding is consistent with previous meta-analyses reporting on similar constructs (i.e., attention, self-regulation, and perseverance; Robson et al., 2020; Smithers et al., 2018), and with theories positing that executive function skills support children's learning of and engagement in positive social interactions by enabling them to regulate their thoughts, actions, and emotions in social contexts (Obradović & Willoughby, 2019). Longitudinally, however, while higher executive function was associated with having higher social competence and peer acceptance in later childhood (ages 6 – 11.99 years), the associations between executive function and emotion understanding and prosociality were no longer statistically significant. Although effect sizes were small to medium in size, they were somewhat stronger for emotion understanding, prosociality, and social competence than for emotion regulation and peer acceptance. Notably, none of the moderators tested emerged as significant predictors of the strength of the association between executive function and social outcomes.

For health outcomes, children with better executive function in preschool tended to have lower body mass indices concurrently, providing support for transdisciplinary theories linking poor executive function with dysregulated eating and better executive function with an increased ability to recognize satiety cues and ignore distractions while eating (Harrist et al., 2012). This finding, however, is qualified by the fact that we found very few studies that explored this link, and within those studies, effects sizes were small. Executive function was not associated, on the other hand, with children's physical fitness, though the pattern of results did trend in the expected direction. While there is a growing body of literature demonstrating the benefits of physical activity for executive function and cognitive control in childhood (e.g., Keye et al., 2021), the majority of research to date has focused on older children and adolescents. In addition, many of the studies that we initially located involved interventions aimed at improving children's physical fitness, which we omitted from this meta-analysis to avoid biasing results.

For behavioral outcomes, children with higher executive function in preschool tended to have fewer externalizing behavior problems and attention and hyperactivity symptoms, and better lie understanding and adaptive classroom behaviors concurrently. Longitudinally, higher executive function was associated with more adaptive classroom behaviors, as well as lower instances of externalizing behavior problems and attention and hyperactivity symptoms in later childhood (ages 6 – 11.99 years). The significant findings for externalizing behaviors both concurrently and longitudinally are consistent with previous meta-analyses reporting on similar constructs (Robson et al., 2020; Smithers et al., 2018), but the magnitude of the effect sizes reported here are much smaller than those found in previous work. This was surprising given the finding reported by Robson et al. (2020) that the relation between self-regulation and externalizing behaviors was stronger in task-based assessments of self-regulation as compared to parent- or teacher-reports, observation measures, or self-reports. Because there was at least a small degree of overlap in the types of task-based measures used to capture self-regulation and those included as a measure of executive function in this meta-analysis (e.g., the Head-Toes-Knees-Shoulders task), we expected that effect sizes would be comparable to the findings reported previously, but this was not the case.

Even more surprisingly, executive function was unrelated to internalizing behavior problem both concurrently and longitudinally (measured in adolescence), in contrast with findings from previous research looking at conceptually related constructs (Robson et al., 2020; Smithers et al., 2018). These findings related to behavior outcomes concurrently and longitudinally lend support for the argument that while constructs like executive function and self-regulation may be conceptually related, they are not identical and may not support development in the same ways. Additional research is needed to further disentangle executive function from related constructs and determine the mechanisms by which each uniquely supports development.

We also explored behavioral outcomes related to children's lie understanding, attention and hyperactivity symptoms, and adaptive classroom behaviors. To our knowledge, no meta-analyses to date have explored executive function or related constructs as they correlate with these outcomes. While few studies explored the link between executive function and lie understanding, the results showed a positive, small to medium effect. This important new finding is consistent with recent theories suggesting that lie-telling and understanding requires that young children not only engage executive function to hold both the truth and lie in mind simultaneously, but inhibit any verbal or nonverbal information that would point toward the transgression (Evans et al., 2011). It is important to note, however, that because of the limited number of studies located, we were unable to test for any moderator effects. Another key new finding of this meta-analysis is the association between executive function and attention and hyperactivity symptoms. Children with higher executive function in early childhood tended to show lower instances of attention and hyperactivity symptoms, both concurrently and in later childhood (ages 6 – 11.99 years). Notably, effect sizes for these relations were small to medium in magnitude, which could suggest that the results are practically meaningful. Relatedly, we found that executive function was positively associated with adaptive classroom behaviors, both concurrently and prospectively. Effect sizes for these relations were, again, of small to medium magnitude, and the strength of the relation did not vary for child age. These results, alongside research showing that executive function plays a critical role in academic achievement (e.g., Allan et al., 2014), suggest that executive function may be especially important to education-related outcomes. This new finding also provides support for recent efforts to reprioritize executive function in early childhood education policy, practice, and prevention programs (e.g., Bierman & Torres, 2016; Zelazo et al., 2016).

## Implications

The results of this meta-analysis have important implications for future research, as well as for theory, measurement, and intervention. Future research is sorely needed to better understand the utility of lab-based executive function measures in supporting important non-academic outcomes, like health-related indices and emotion regulation, as well as outcomes that were not addressed in this meta-analysis, such as socioeconomic mobility, mental health disparities, and indices of quality of life. In particular, this meta-analysis was limited in determining whether executive function is associated with outcomes in the long-term and when known important covariates like verbal skills and age are accounted for. Future research exploring outcomes from preschool to adolescence, adulthood, and beyond would be of particularly important value.

Theoretically, these findings deepen our understanding as to whether individual differences in early executive function abilities — as measured by standard laboratory tasks — is broadly important to healthy functioning and adaptation. While effect sizes were small to medium in magnitude across outcomes, results provide a basis for which new and innovative theoretical frameworks and ideas about the development of executive function, how executive function is measured, and associations with important life outcomes can be developed.

The findings reported in this meta-analysis also have important implications for researchers, educators, caregivers, and clinicians who are all tasked with promoting adaptive functioning in young children. While more research is needed before we can make any confirmatory claims about the importance of executive function as a reliable predictor of a broad array of non-academic outcomes, these results do point toward the potential for executive function-based interventions in the promotion of several critical social, heath, and behavioral outcomes. However, it is imperative to note that a recent meta-analysis of lab-based executive function training had limited practical value, at least in obtaining far transfer effects (Kassai et al., 2019); thus, we recommend that interventions for improving social, heath, and behavioral outcomes focus on improving multiple developmental skill domains, rather than exclusively targeting executive function.

## Limitations

This meta-analysis had some limitations. First, in choosing zero-order correlation coefficients as our effect size of interest, we were unable to account for well-known covariates, such as verbal abilities and age, which has several key implications. For one, it is possible and likely that our effect sizes were inflated. In addition, while correlation and causation can exist simultaneously, it is important that our findings are not misconstrued as executive function causing outcomes. Our findings suggest that executive function may be related to a number of social, health, and behavioral outcomes, but the casual pathway is unknown. Related to this issue, several potentially important moderators could not be explored in this meta-analysis. For example, existing literature points toward to importance of several social and economic factors, such as income, education, and a child's home–environment quality, in the development of children's executive function skills (e.g., Dilworth-Bart et al., 2007; Hackman et al., 2014). It is possible that these uncontrolled, confounding variables could have accounted for substantial variance in the associations between executive function and social, health, and behavioral outcomes.

A third limitation is that many of the outcome models tested in this meta-analysis contained substantial heterogeneity, yet the moderator analyses conducted in an attempt to identify factors that may be contributing to the variance did not provide new information about the effect. It is possible that this could be a reflection of a limitation in the executive function construct itself, and not a limitation in this meta-analysis, however. That is, the heterogeneity could, in part, be due to the fact that behavioral lab tasks used to measure executive function vary widely in their task demands and reliability. While many executive function lab tasks require the inhibition of prepotent responses, others require recollection and manipulation of verbal or visuospatial information or attention to two simultaneous stimuli at the same time (Wiebe et al., 2008). It is also likely that the studies included in this meta-analysis differed from each other in terms of their designs and procedures, such that some studies were of lower quality than others. We did not employ any methods for characterizing study quality other than publication status, which may have impacted the interpretability of aggregated effects (such that studies of lower-quality may have influenced the summary effect).

An additional limitation related to study design is that when doing meta-analysis, it is often the case that multiple different measures of a construct are collected. When this happens, it is common practice to compute a combined, synthetic effect (Borenstein et al., 2009). Fundamentally, however, these effects are different from effects drawn from independent subgroups. For example, in their study of the cognitive correlates of children's early moral functioning, Tan et al. (2020) collected several measures of instrumental helping, comforting, and sharing behaviors, which they classified as measuring prosociality. In cases like this, where outcomes are conceptually related enough to be combined into a composite measure (e.g., prosocial behaviors), we computed a synthetic effect to preserve assumptions of independence; however, it is possible that despite these efforts, systematic differences between these effects and unadjusted effects derived from independent samples remained.

## Conclusion

To date, researchers have assumed that lab tasks measuring executive function component processes capture real-world functioning. There is good evidence that this is true in the domain of academic achievement, but evidence is less clear in other important developmental domains. The current meta-analysis clarifies that indeed, preschool behavioral measures of executive function do relate to social competence, peer acceptance, externalizing behavior problems, attention and hyperactivity symptoms, and adaptive classroom behaviors, both concurrently and longitudinally, but relations were generally small to moderate, ranging from 0.10 to 0.25. Evidence was more limited for longitudinal outcomes, leaving open the question as to whether executive function is associated with outcomes in the long-term. Considering that few studies controlled for potentially important confounds, we urge caution in interpreting these findings as conclusive support for the importance of executive function in social, health, and behavioral development over the lifespan; however, our results do provide new and exciting directions for future research to further explore these patterns and better understand the role of executive function in adaptive human functioning.

APPENDIX A

*Search Terms*

executive function* [*OR* executive control *OR* cognitive control *OR* cognitive flexibility *OR* impulse control *OR* inhibit* *OR* working memory]

*AND*

child* [*OR* early years *OR* k-12 *OR* prek *OR* preschool*]

*AND*

a-not-b [OR antisacc* OR “ambiguous figure” OR “ambiguous figures” OR “digit span” OR “word span” OR beads OR dragon OR big/little OR big-little OR stroop OR cat-mouse OR cat/mouse OR “design fluency” OR "semantic fluency" OR "category fluency" OR "verbal fluency" OR “paced auditory serial addition” OR ChiPASAT OR size-ordering OR “continuous performance” OR “corsi block*” OR “count and label” OR count-and-label* OR d-kefs OR dkefs OR “color-word interference” OR “color word interference” OR proverb* OR tower OR trail* OR “trail making” OR trail-making OR “twenty questions” OR “word context” OR day/night OR day-night OR “delayed alternation” OR “delayed response” OR “unusual use*” OR “differentials abilities” OR “card sort*” OR card-sort* OR dots OR “hearts and flowers” OR hearts-flowers OR hearts/flowers OR heart-flower OR heart/flower OR “executive function touch” OR “ef touch” OR go/no-go OR go-no-go OR no-go OR nogo OR farmer OR “pick the picture” OR pick-the-picture OR spatial-conflict OR "spatial conflict" OR flanker OR grass/snow OR grass-snow OR “grass snow” OR handgame OR “hand game” OR hand-game OR HTKS OR head-toes OR head/toes OR “head toes” OR KRISP OR “kansas reflection-impulsivity” OR “kansas reflection impulsivity” OR “kansas reflection/impulsivity” OR “less is more” OR less-is-more OR “location memory” OR “familiar figures” OR “minnesota executive function” OR “minnesota executive functioning” OR mefs OR “motor sequencing” OR “multilocation search” OR n-back OR “n back” OR barnyard OR NEPSY OR knock-tap OR “knock tap” OR knock/tap OR statue OR “visual attention” OR “boxes task” OR “boxes test” OR “noisy book” OR “peg tap*” OR “pencil tap*” OR pinball OR “preschool attentional switching task” OR “reverse cat*” OR “self-ordered point*” OR “shape school” OR “simon says” OR “simon task” OR “spatial working memory test” OR “spatial working memory task” OR “spin the pots” OR spin-the-pots OR “stop signal test” OR “stop signal task” OR “switching, inhibition, and flexibility task” OR swift OR “tower of hanoi” OR “tower of london” OR trucks OR “wide range assessment of memory and learning” OR “design memory” OR “picture memory” OR “wisconsin card” OR “windows task” OR “windows test” OR "counting span" OR “reading span" OR "sorting test" OR "Something’s the Same"]

Wherever possible, I will also employ the following search term restrictions:

*NOT*

Title: meta-analysis OR erratum OR correction OR systematic review OR meta-analytic OR autism OR disease* OR disorder* OR disabil* OR neurofibromatosis OR intervention OR training OR preterm OR pre-term OR trauma* OR impairment* OR institutionalized OR clinical* OR ADHD OR adopt* OR maltreat* OR "spina bifida" OR virus OR deaf OR cochlear OR schizophrenia OR patient* OR anorexia OR asperger OR syndrome OR epilep* OR phenylketonuria OR dyscalculia OR dyslexia OR dysgraphia OR inpatient OR diabetes OR HIV OR atrophy OR stutter* OR TBI OR cerebral palsy OR prematur* OR iron OR homeless* OR neglect* OR preeclampsia OR sclerosis OR tumor OR tumour OR leukemia OR AD/HD OR ASD OR injury OR stroke OR lesion OR abuse OR disability OR dystrophy OR PTSD OR anemia OR hyperplasia OR congenital OR “cocaine-expos*” OR “drug-expos*” OR “alcohol-expos*” OR “lead-expos*” OR “low birth weight” OR seizur* OR FASD OR dyslexic

*AND*

Population Group: Human

*AND*

Language: English

# APPENDIX B

*Meta-Analysis Screening Tool: citation, title, and abstract screening*

1. Is the **title or abstract** written in English?
   a. **Yes**, it indicates this: continue screening
   b. **Unsure**
   c. **No**, it's not written in English: stop screening
2. Does the **title or abstract** indicate that this is a primary research study (NOT a correction, erratum, systematic review, or meta-analysis)?
   a. **Yes**, it indicates this: continue screening
   b. **Unsure**
   c. **No**, it does not indicate this: stop screening
3. Does the **title or abstract** indicate that typically developing children are included in this study?
   a. **Yes**, it indicates this: continue screening
   b. **Unsure**
   c. **No**, it does not indicate this (i.e., study includes atypical children *only*): stop screening
4. Does the **abstract** indicate that participants were between 3.0 – 5.99 years at the time of participation?
   a. **Yes**, it indicates this: continue screening
   b. **Unsure**

c. **No**, it does not indicate this: stop screening

5. Does the **abstract** indicate that executive function was measured (specifically)?
   a. **Yes**, it indicates this: continue screening
   b. **Unsure**
   c. **No**, it does not indicate this: stop screening

6. Does the **abstract** indicate that executive function was measured via a lab-based behavioral assessment or task?
   a. **Yes**, it indicates this: continue screening
   b. **Unsure**
   c. **No**, it does not indicate this (e.g., says that executive function was measured using an observational or self-, parent- or teacher-report measure ONLY): stop screening

7. Does the **abstract** indicate that at the study includes at least one zero-order correlation between an assessment of executive function AND an outcome?
   a. **Yes**, it indicates this: continue screening
   b. **Unsure**
   c. **No**, it does not indicate this: stop screening

**Decision**: If, for ***any question***, 'No, stop screening' is selected, then we will NOT include the study in the meta-analysis. ‘Yes, continue screening' or ‘Unsure’ must be selected for ***all*** questions for the study to continue to a final full-text screening to determine definitive eligibility.

APPENDIX C

*Note*. For Figures 2 – 21, confidence intervals (95% CI; presented in square brackets) containing zero suggest a non-significant effect.

Figure 2 Forest Plot for Social Competence Model 1

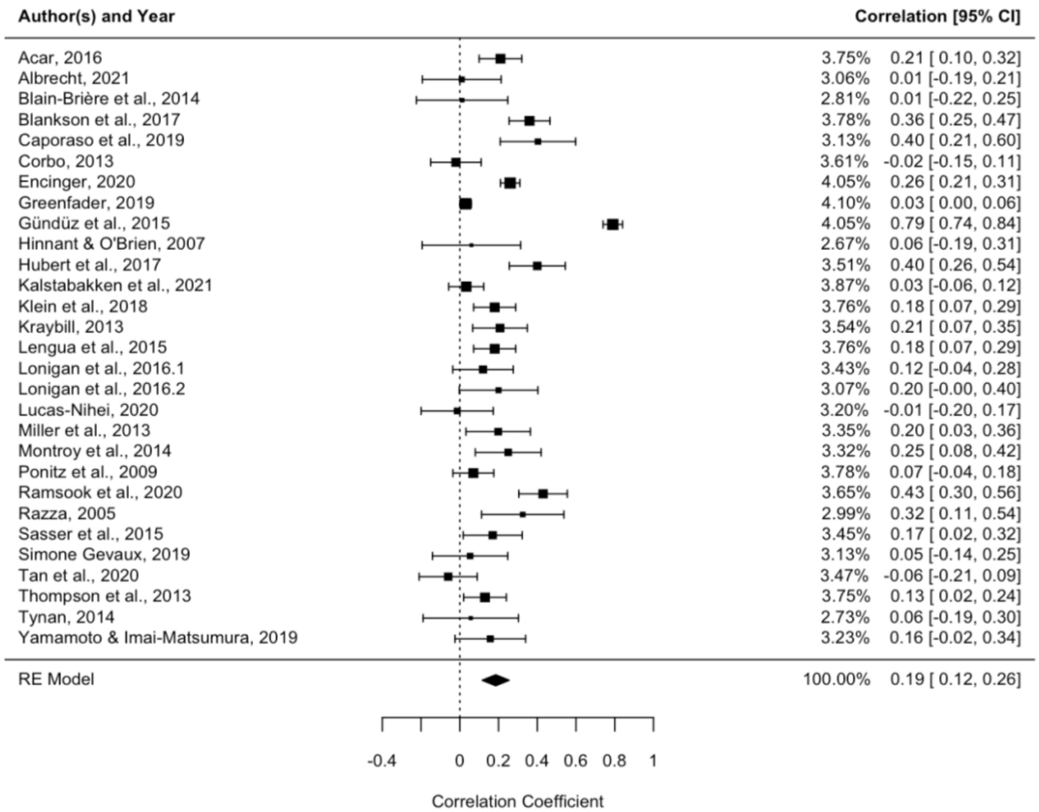



Figure 3 Forest Plot for Social Competence Model 2

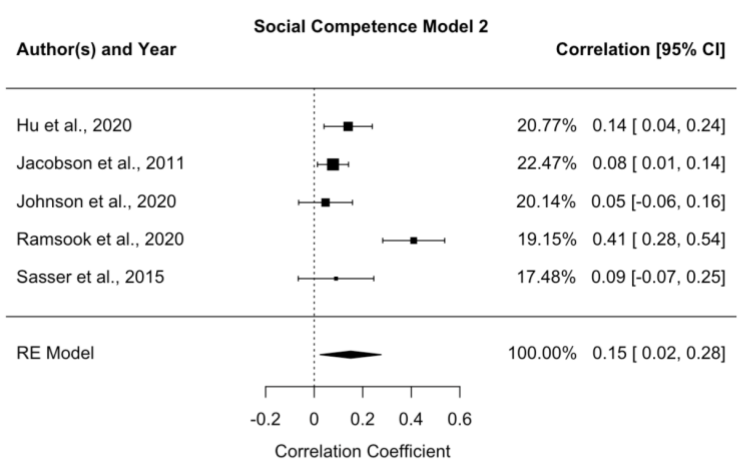


Figure 4 Forest Plot for Prosociality Model 1

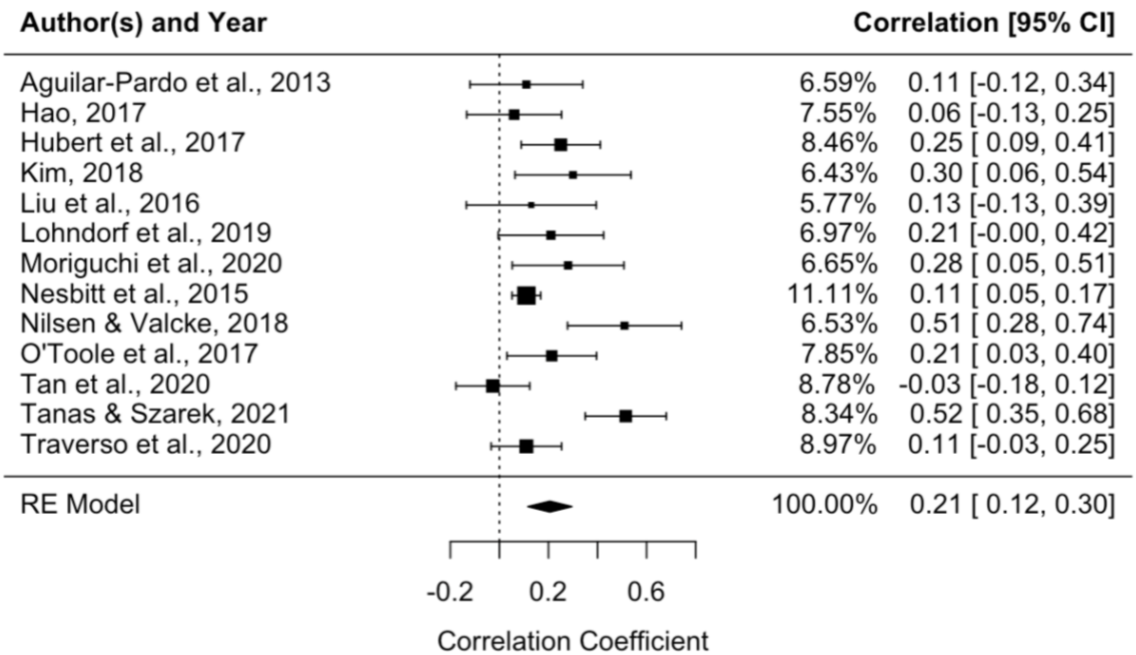



Figure 5 Forest Plot for Prosociality Model 2

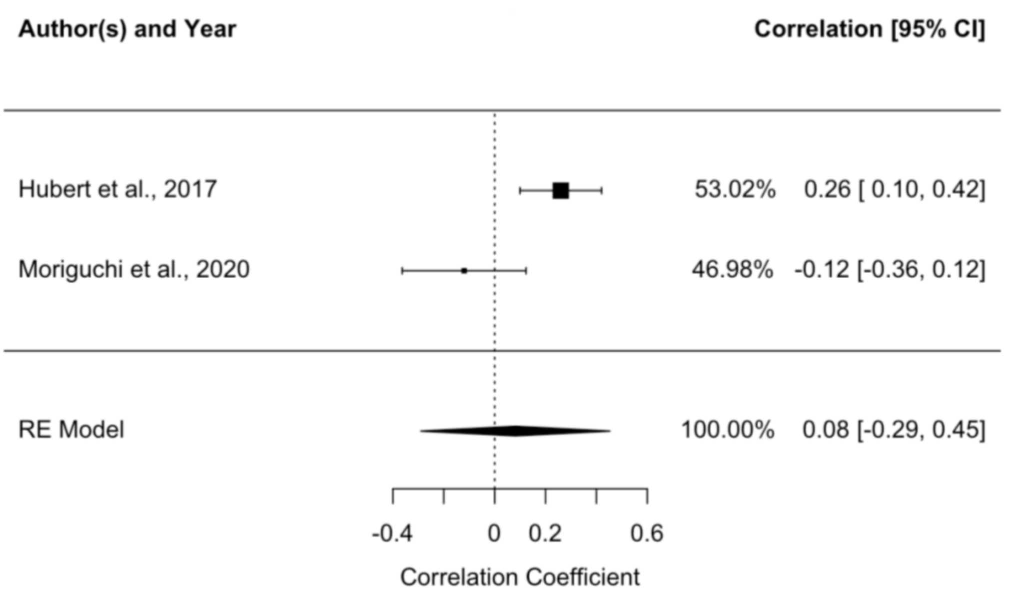


Figure 6 Forest Plot for Peer Acceptance Model 1

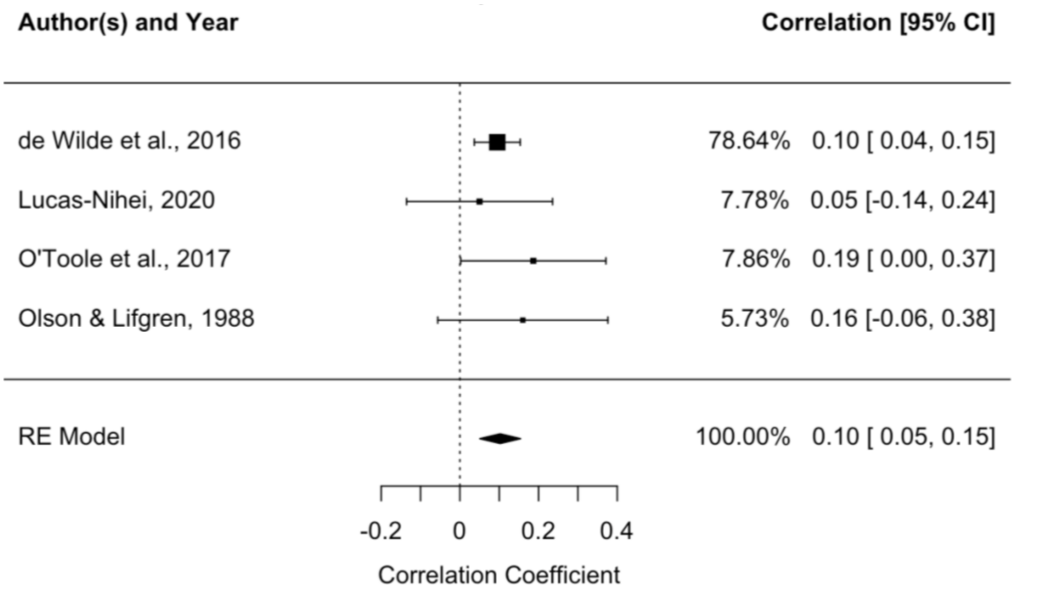



Figure 7 Forest Plot for Peer Acceptance Model 2

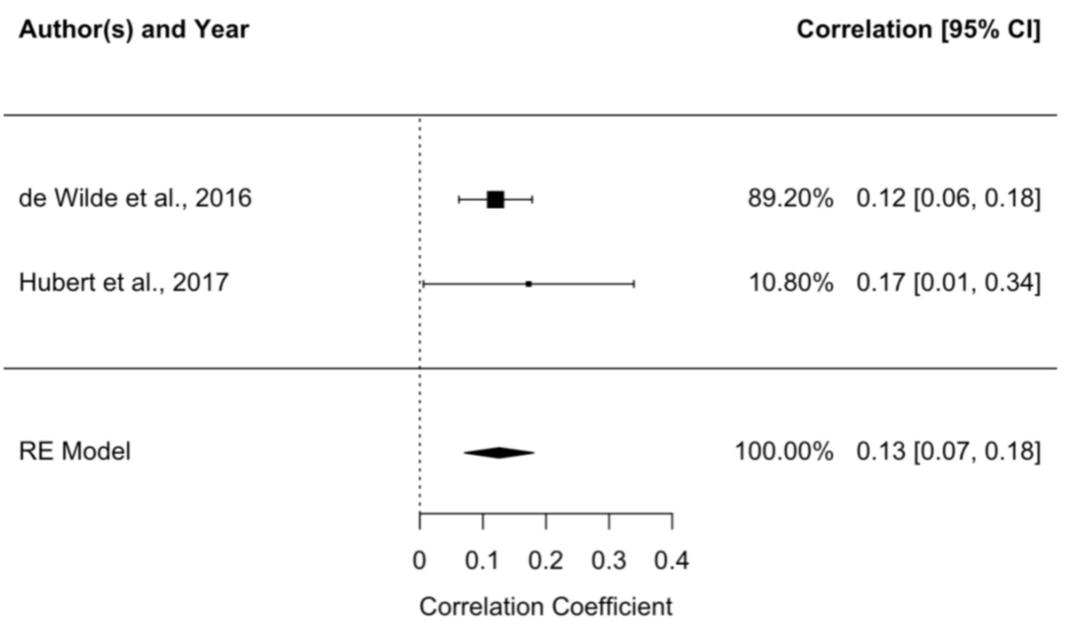


Figure 8 Forest Plot for Emotion Understanding Model 1

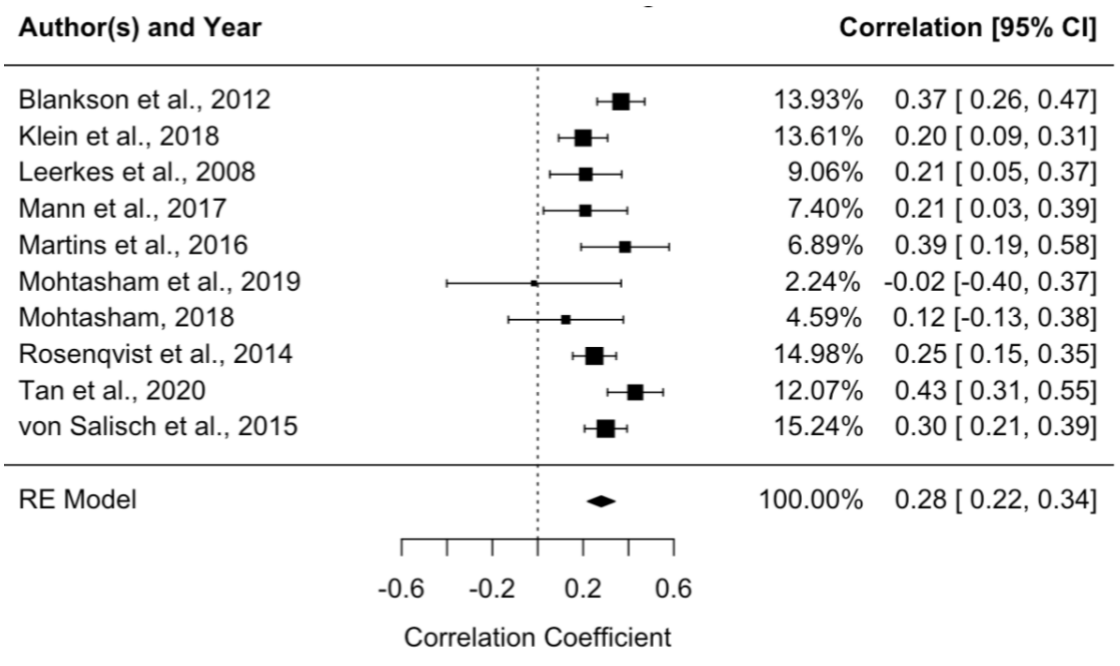



Figure 9 Forest Plot for Emotion Understanding Model 2

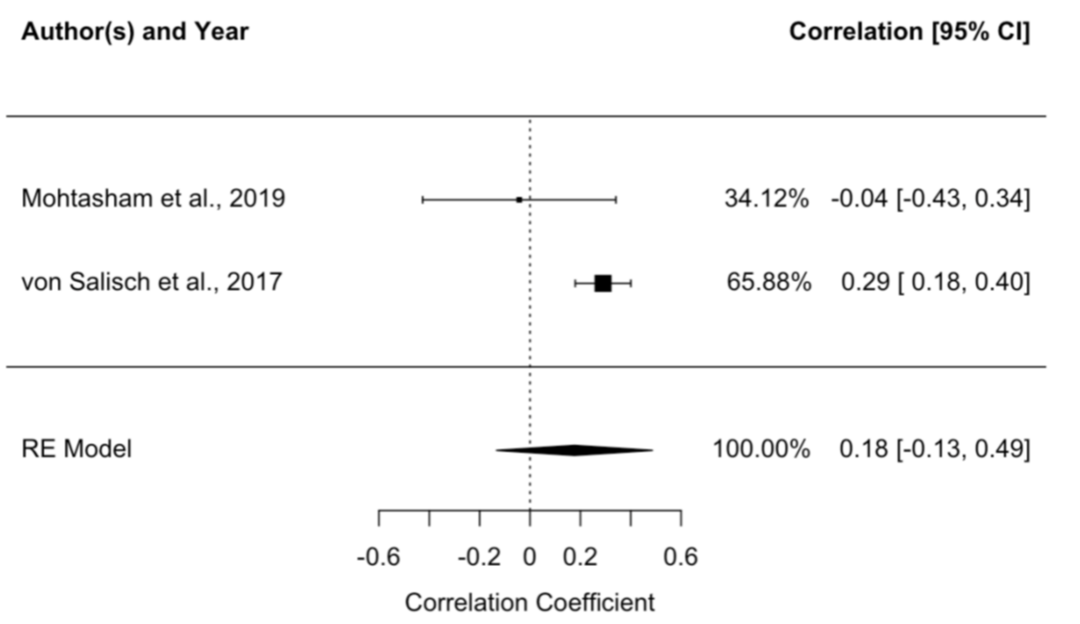


Figure 10 Forest Plot for Emotion Regulation Model 1

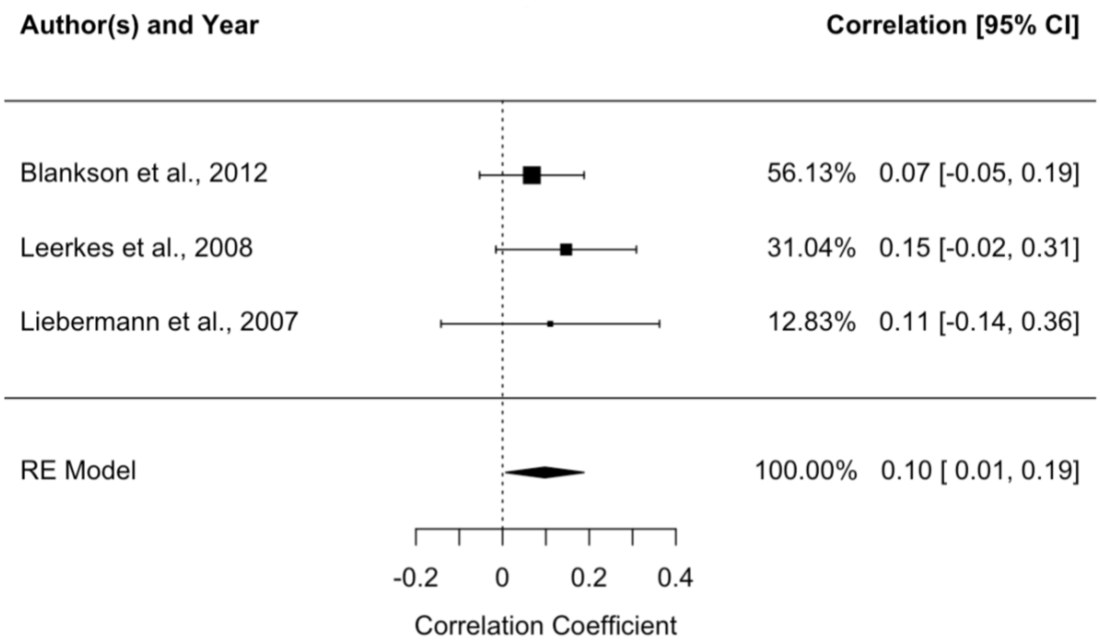



Figure 11 Forest Plot for Body Mass Indices Model 1

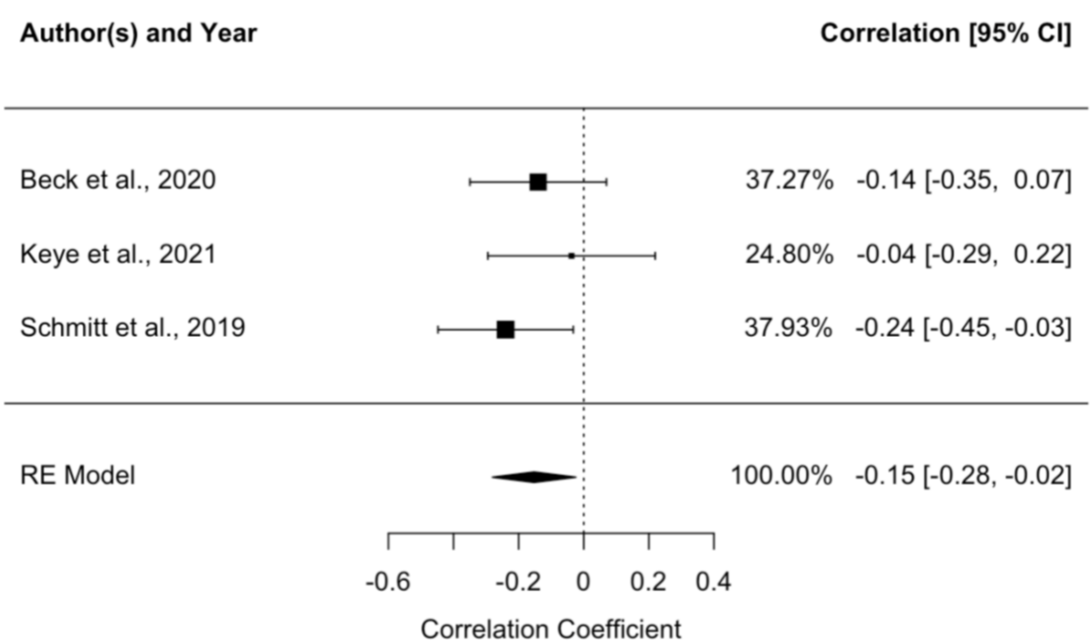


Figure 12 Forest Plot for Physical Fitness Model 1

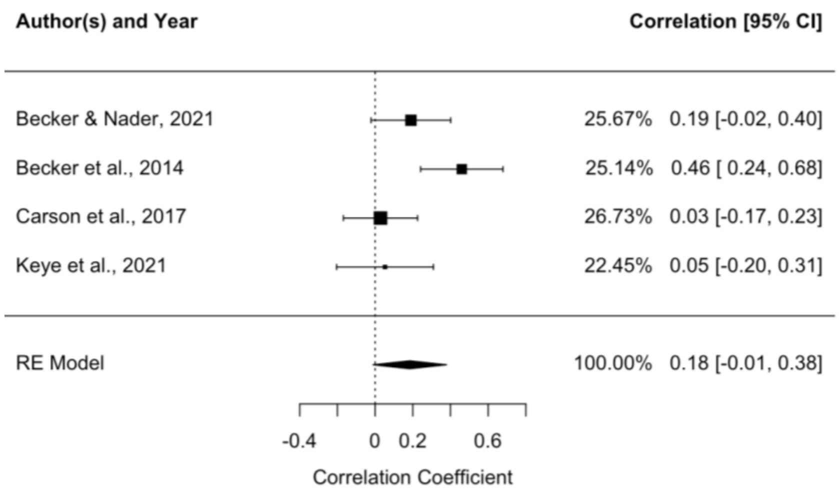



Figure 13 Forest Plot for Externalizing Behavior Problems Model 1

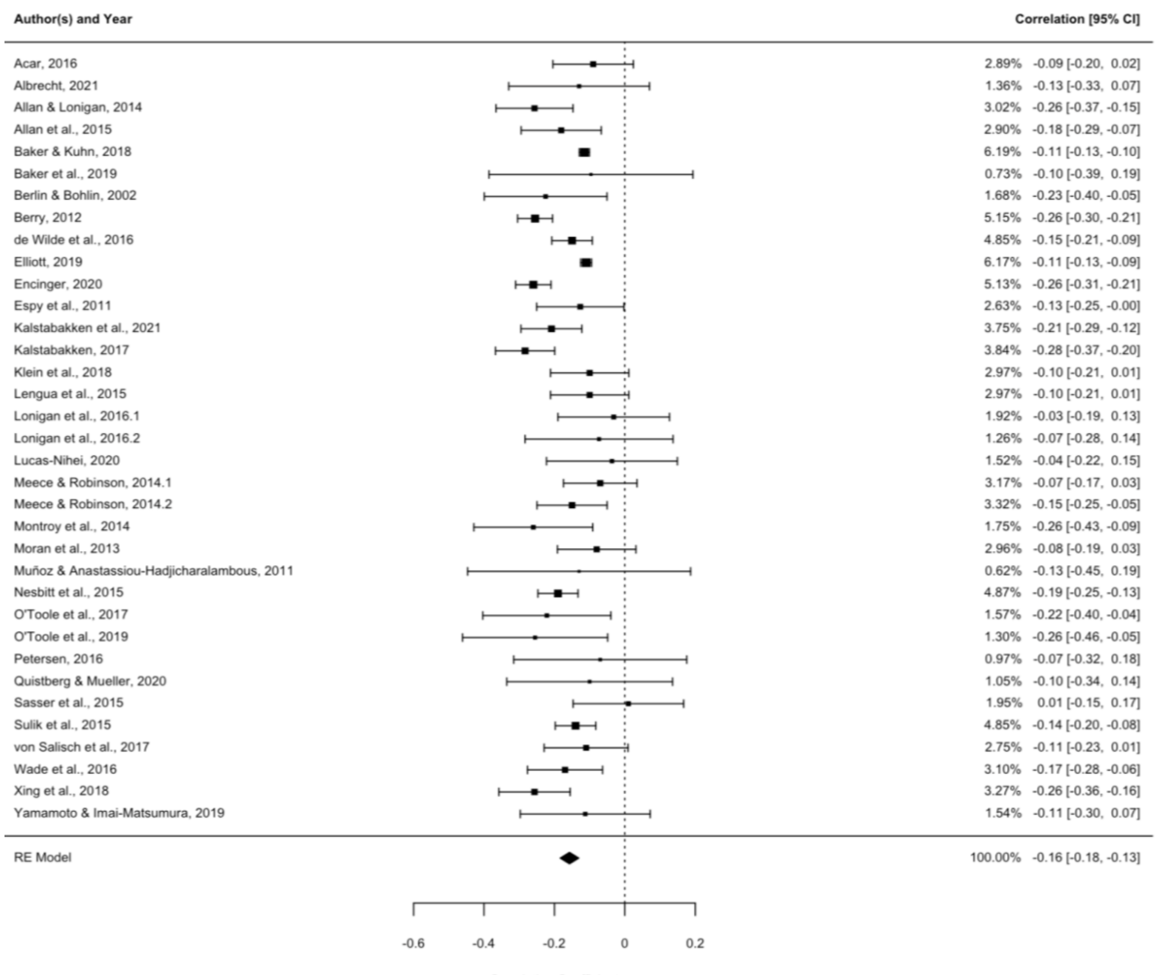


Figure 14 Forest Plot for Externalizing Behavior Problems Model 2

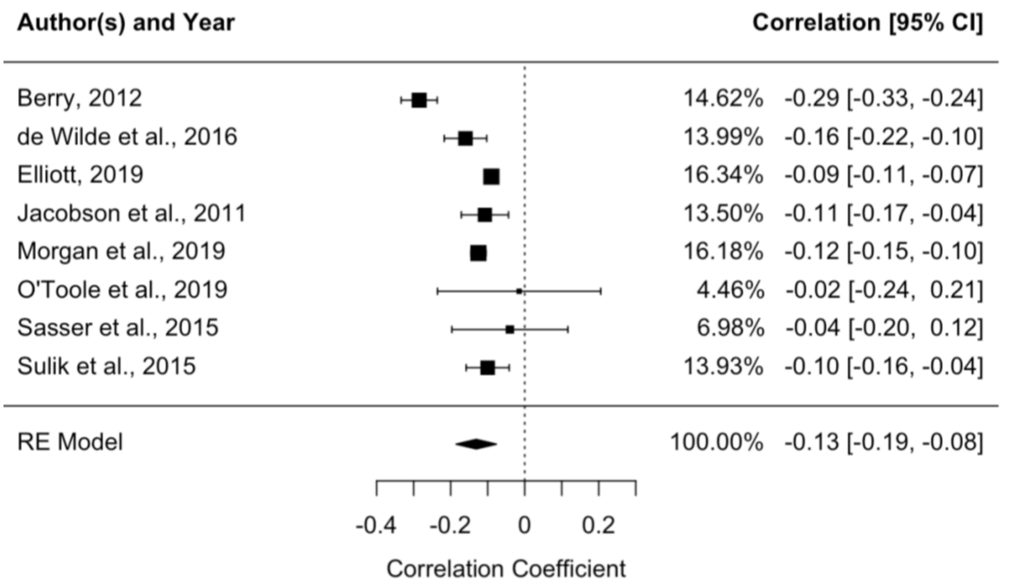



Figure 15 Forest Plot for Internalizing Behavior Problems Model 1

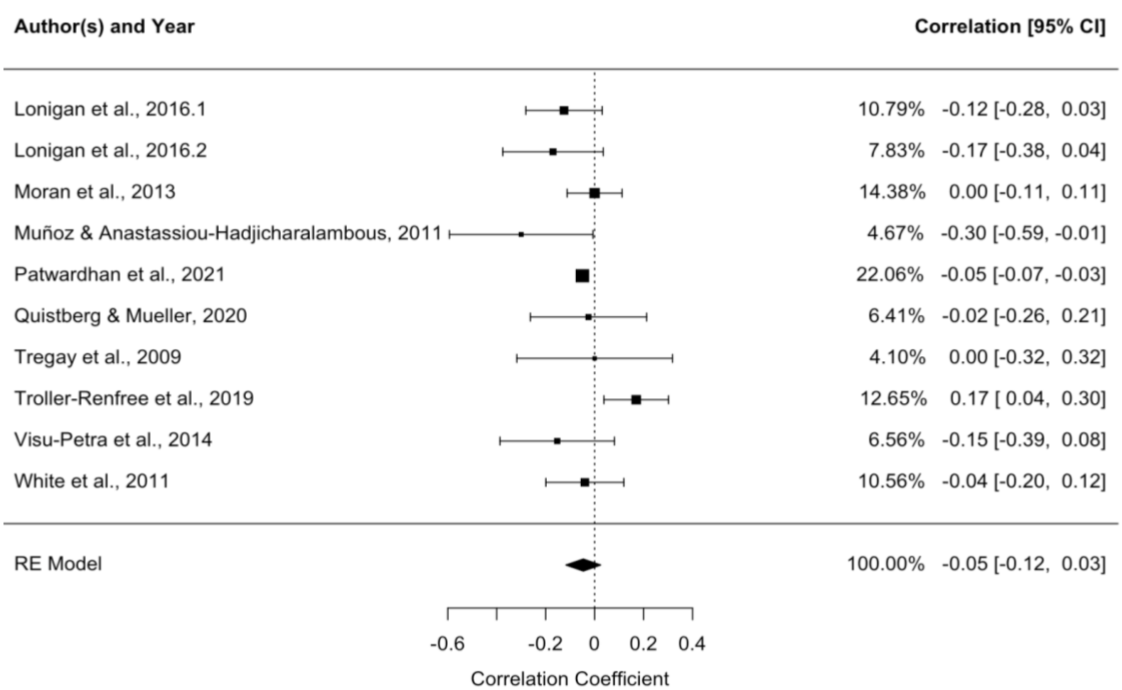


Figure 16 Forest Plot for Internalizing Behavior Problems Model 3

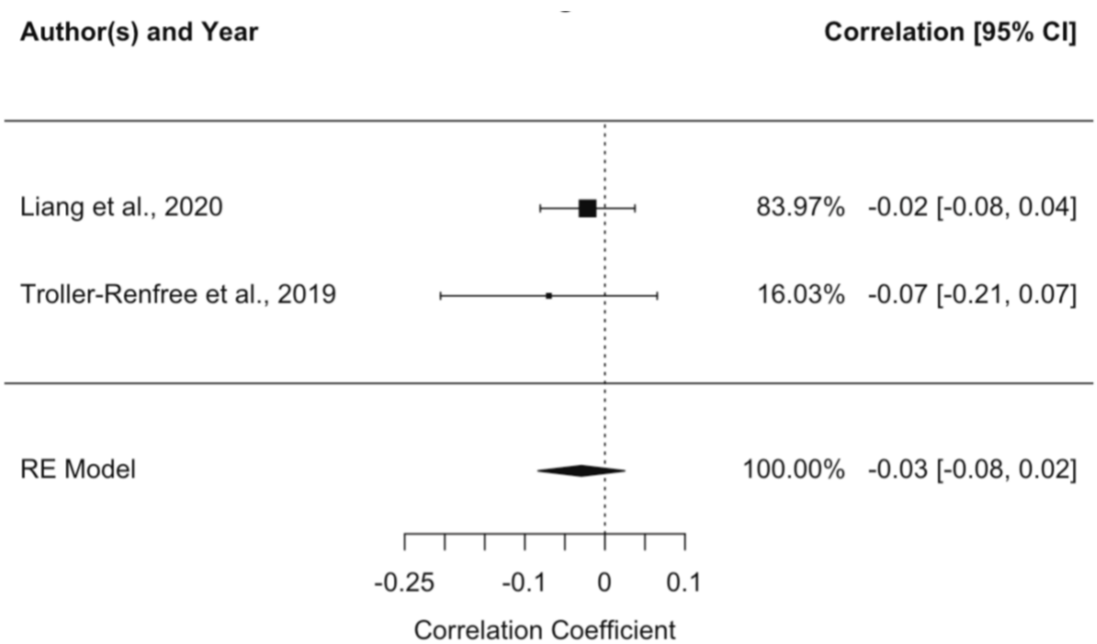



Figure 17 Forest Plot for Lie Understanding Model 1

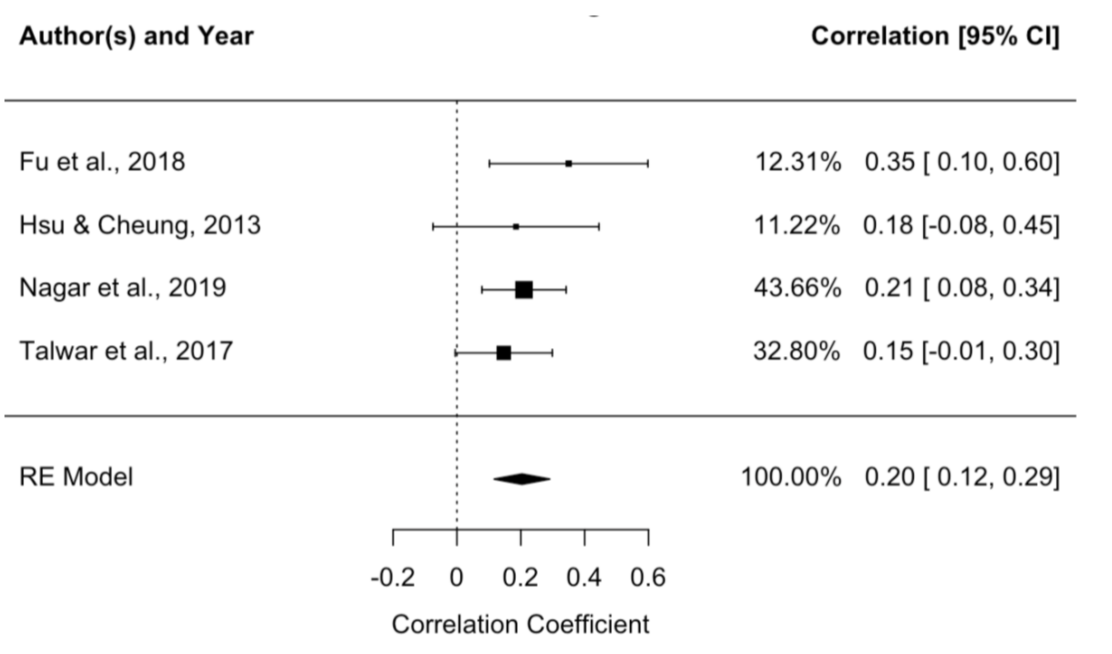


Figure 18 Forest Plot for Hyperactivity/Inattention Model 1

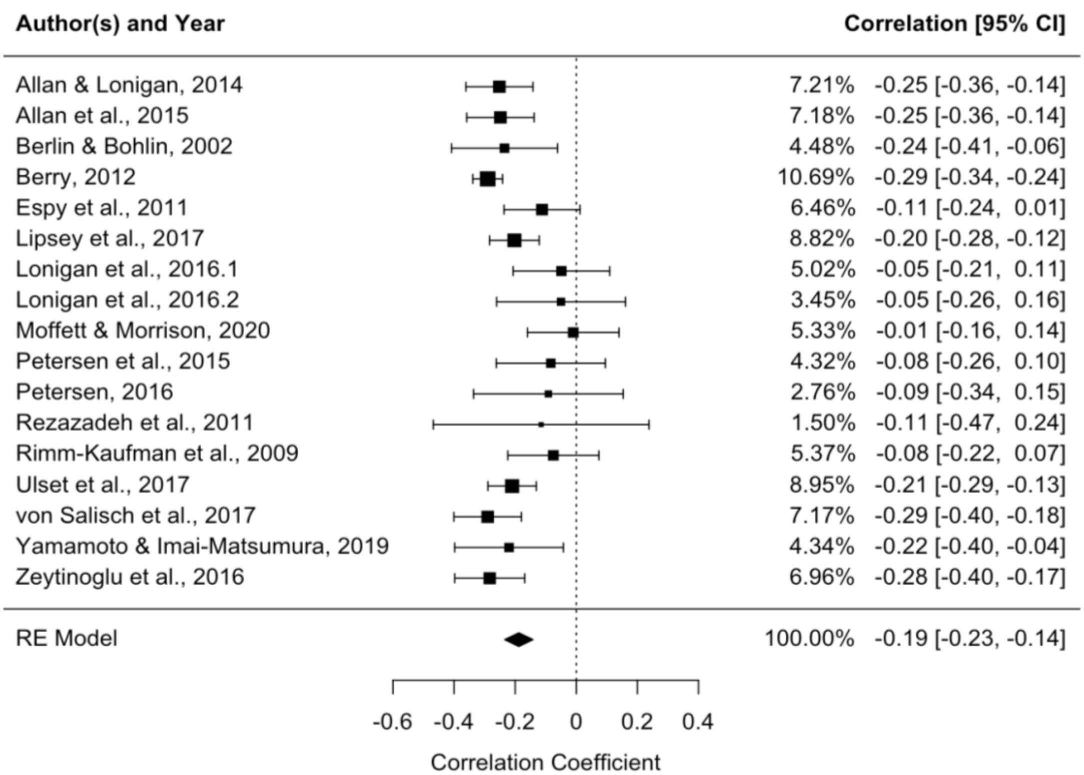



Figure 19 Forest Plot for Hyperactivity/Inattention Model 2

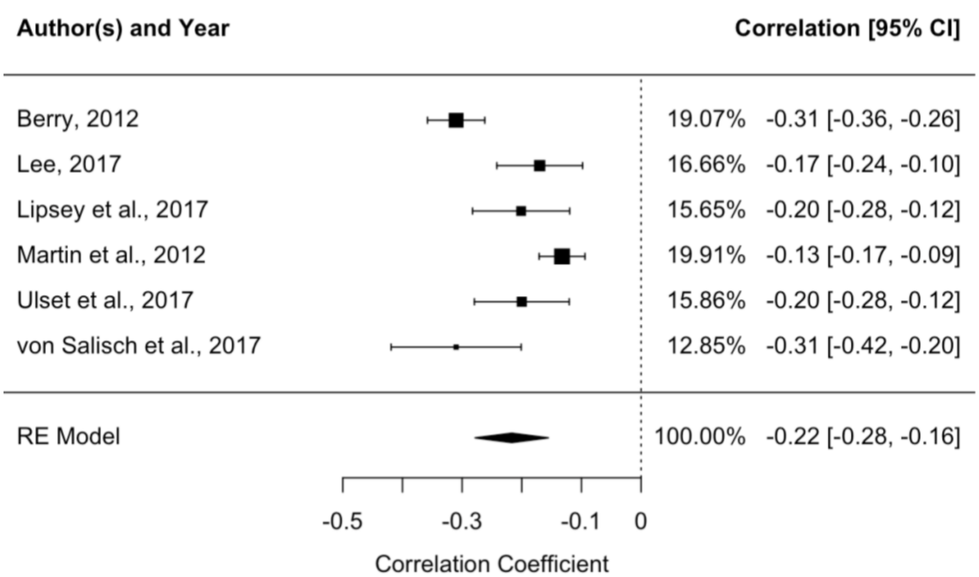


Figure 20 Forest Plot for Adaptive Classroom Behavior Model 1

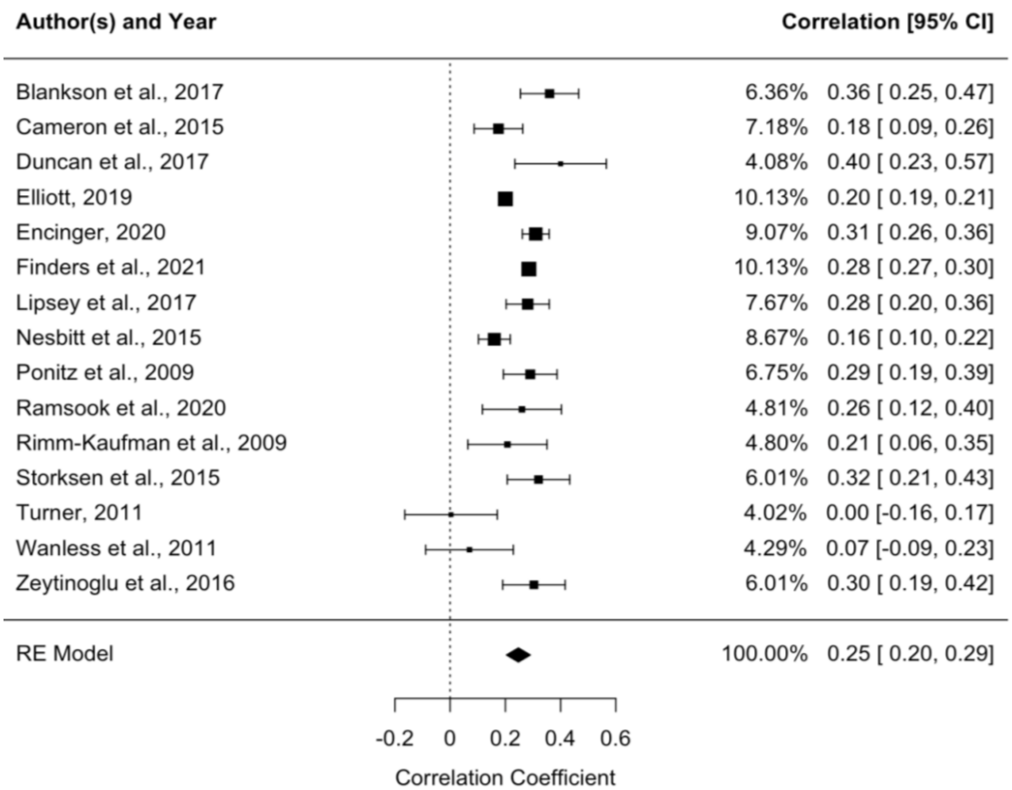
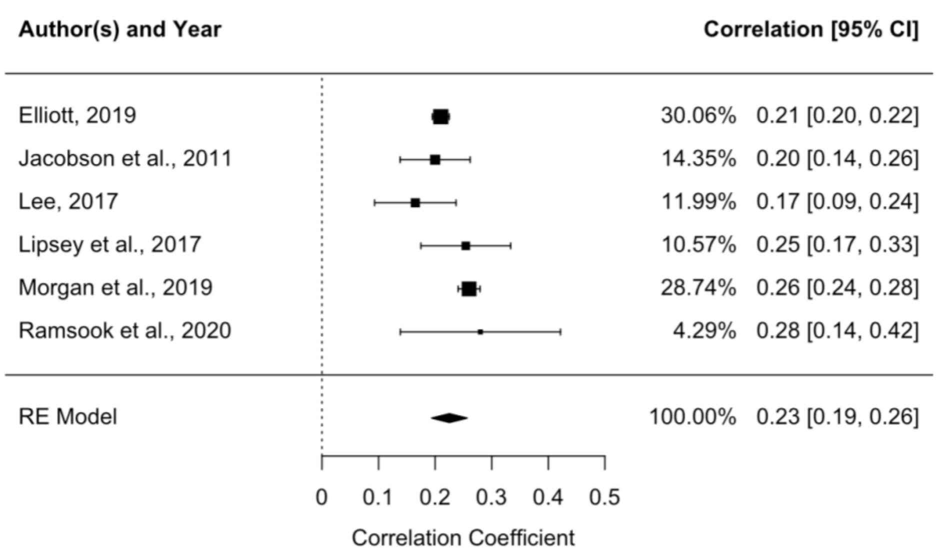



Figure 21 Forest Plot for Adaptive Classroom Behavior Model 2



APPENDIX D

*Note*. For Figures 22 – 41, funnel plot of correlations were used to evaluate the summary effect in the random effects models. The dotted, center vertical lines represent the overall mean effect for that model.

Figure 22 Funnel Plot for Social Competence Model 1

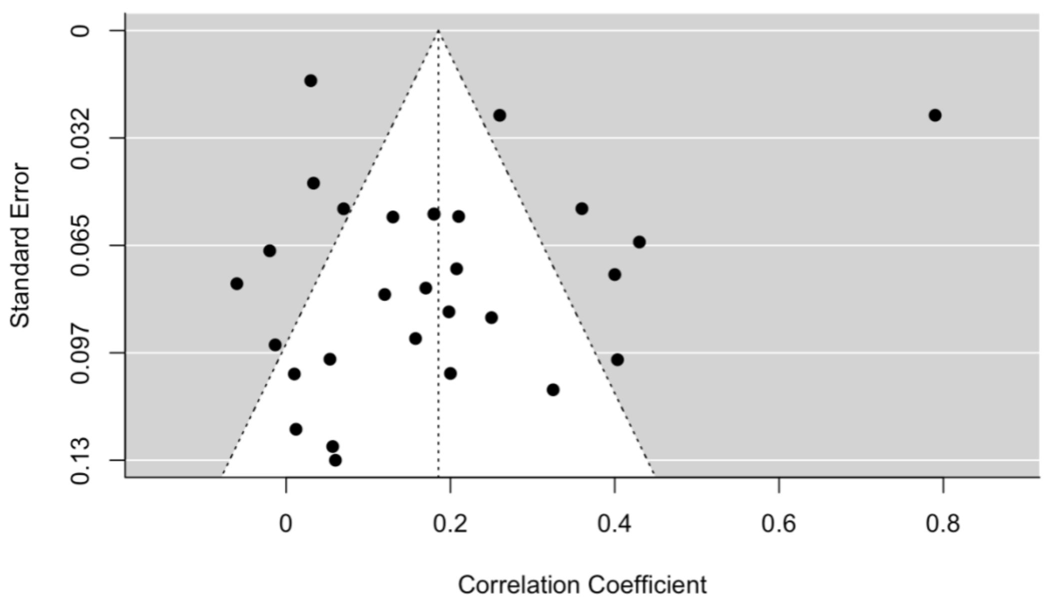



Figure 23 Funnel Plot for Social Competence Model 2

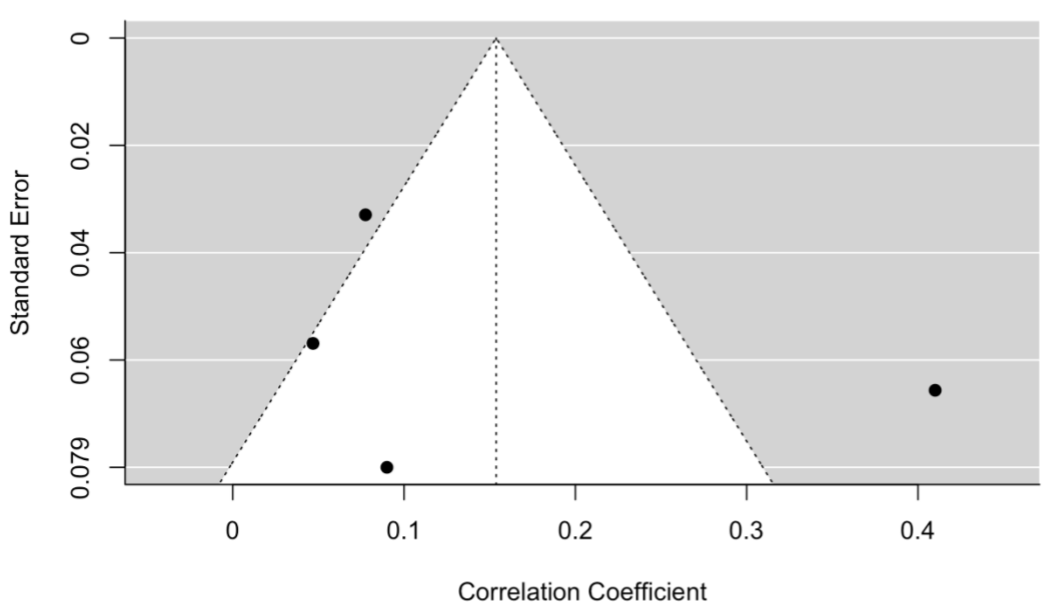


Figure 24 Funnel Plot for Prosociality Model 1

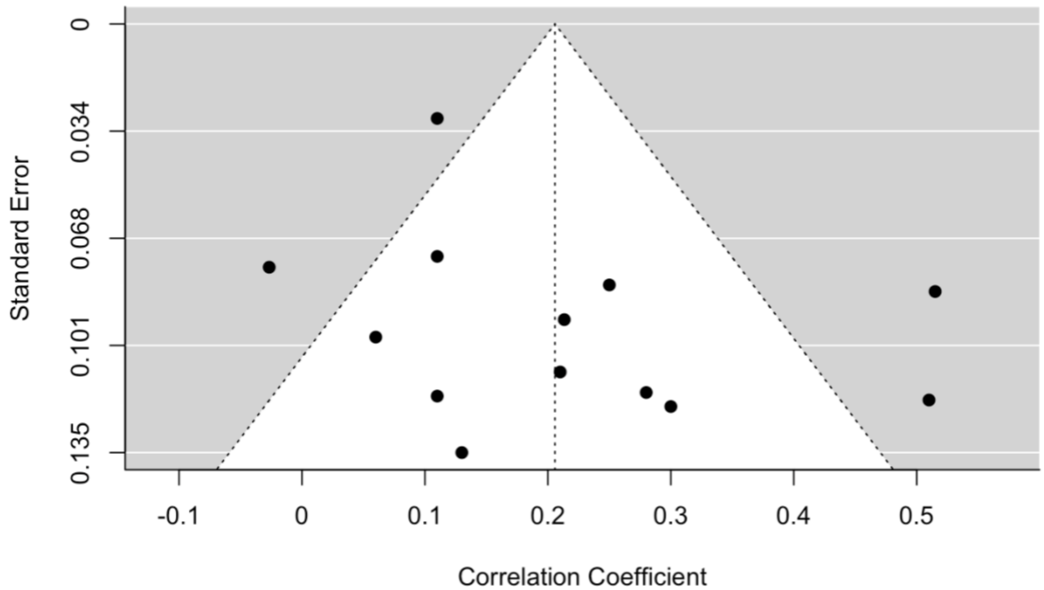



Figure 25 Funnel Plot for Prosociality Model 2

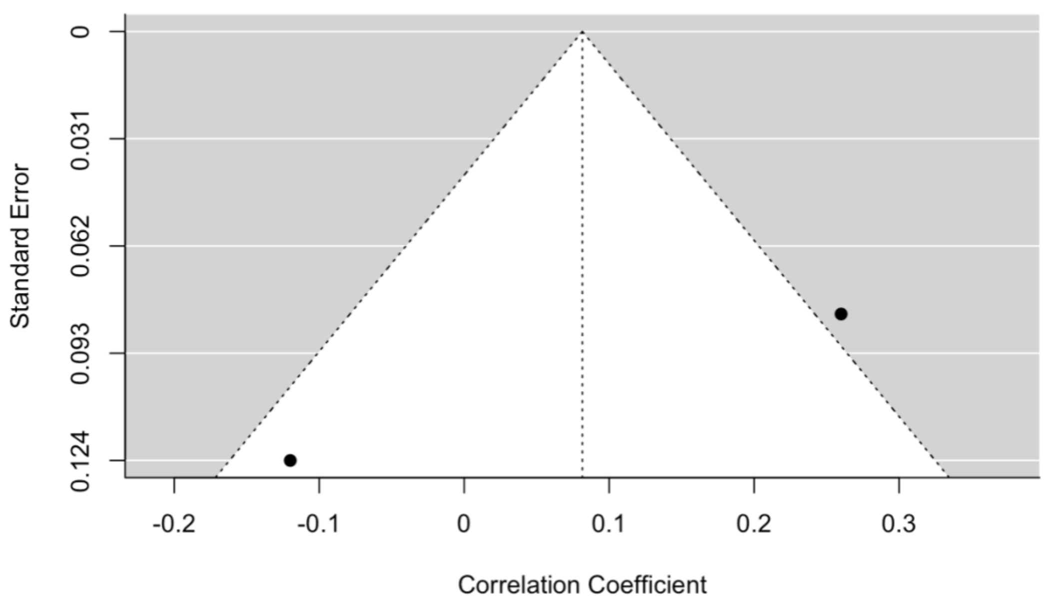


Figure 26 Funnel Plot for Peer Acceptance Model 1

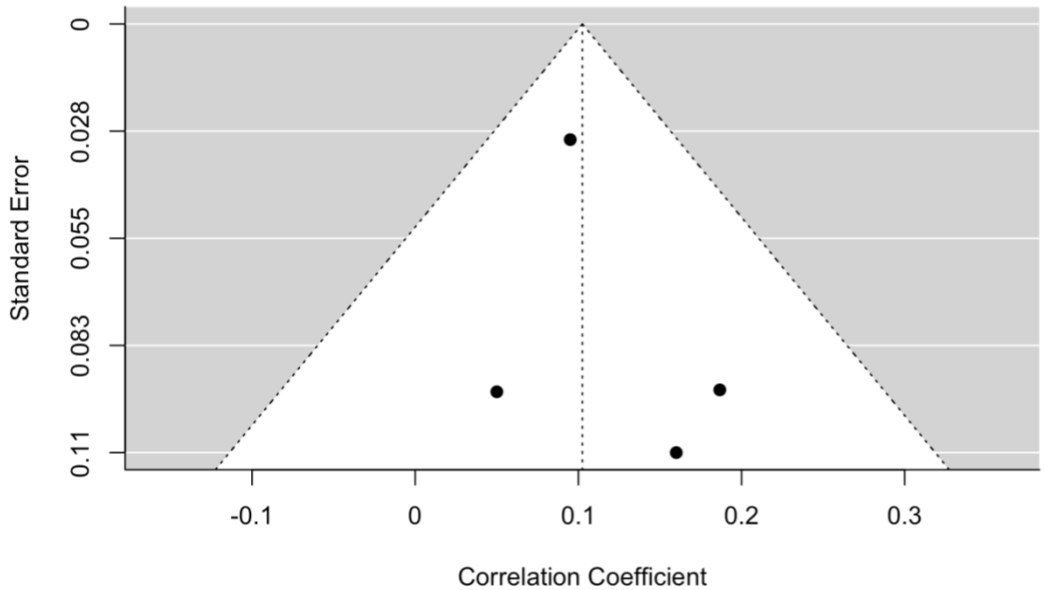



Figure 27 Funnel Plot for Peer Acceptance Model 2

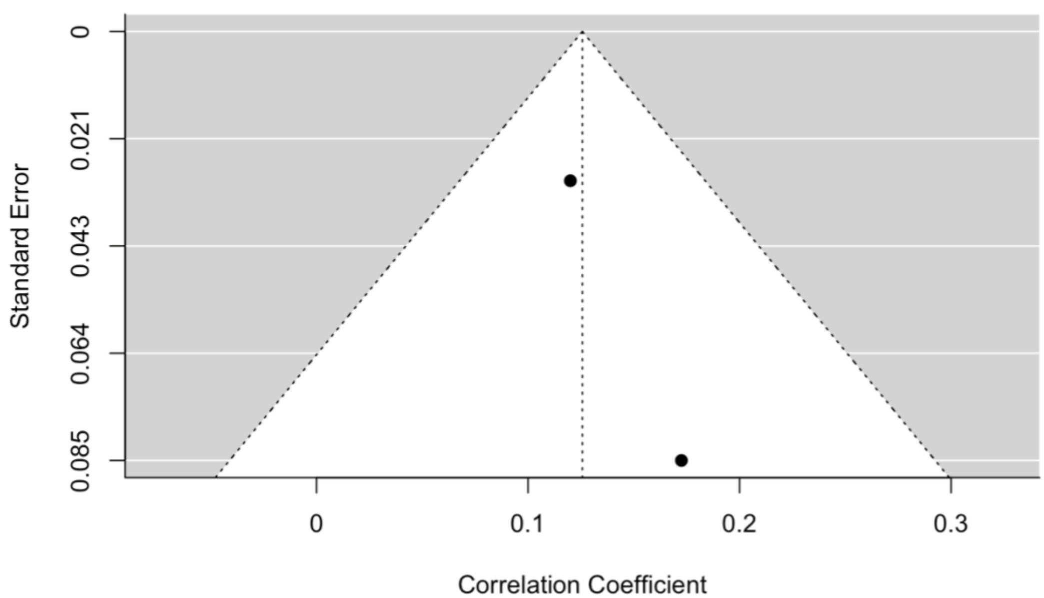


Figure 28 Funnel Plot for Emotion Understanding Model 1

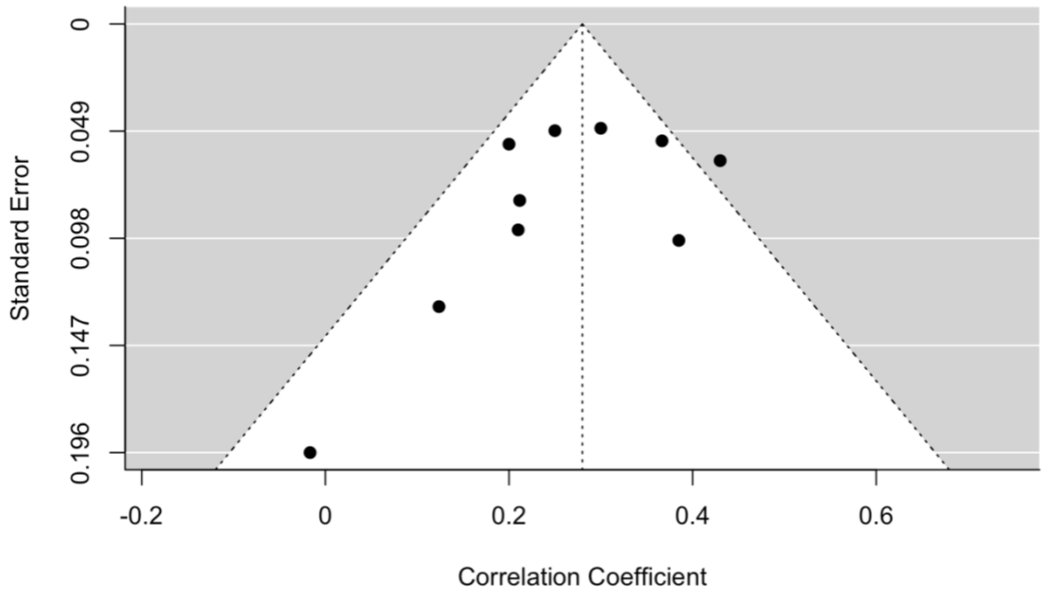



Figure 29 Funnel Plot for Emotion Understanding Model 2

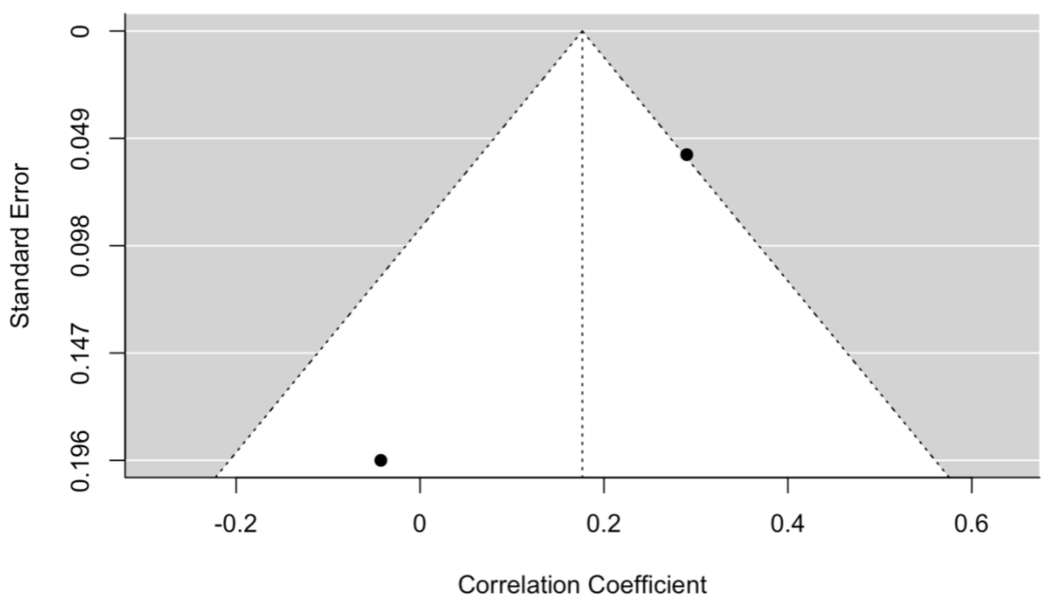


Figure 30 Funnel Plot for Emotion Regulation Model 1

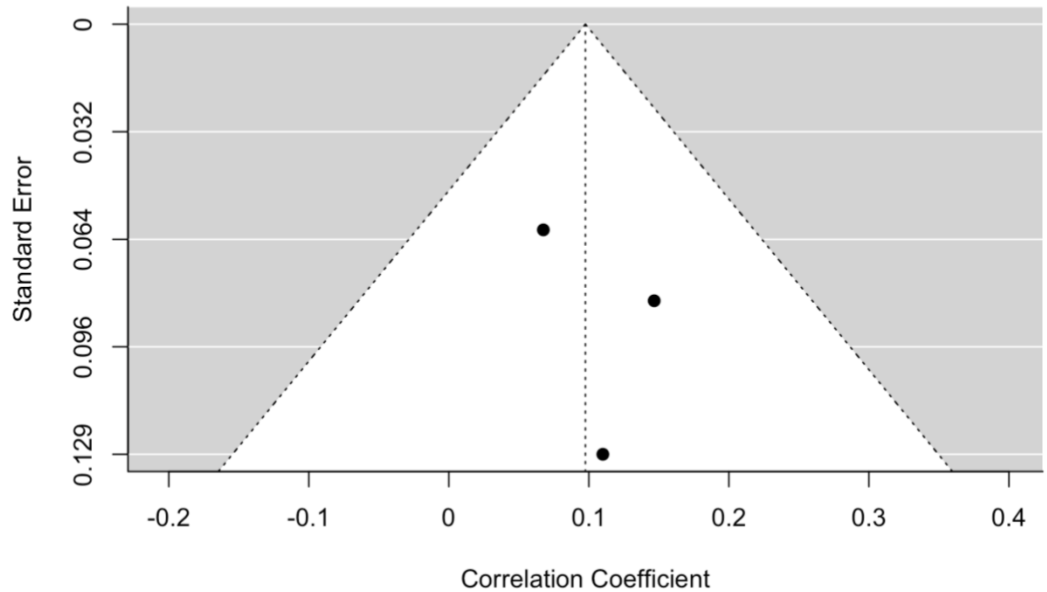



Figure 31 Funnel Plot for Body Mass Indices Model 1

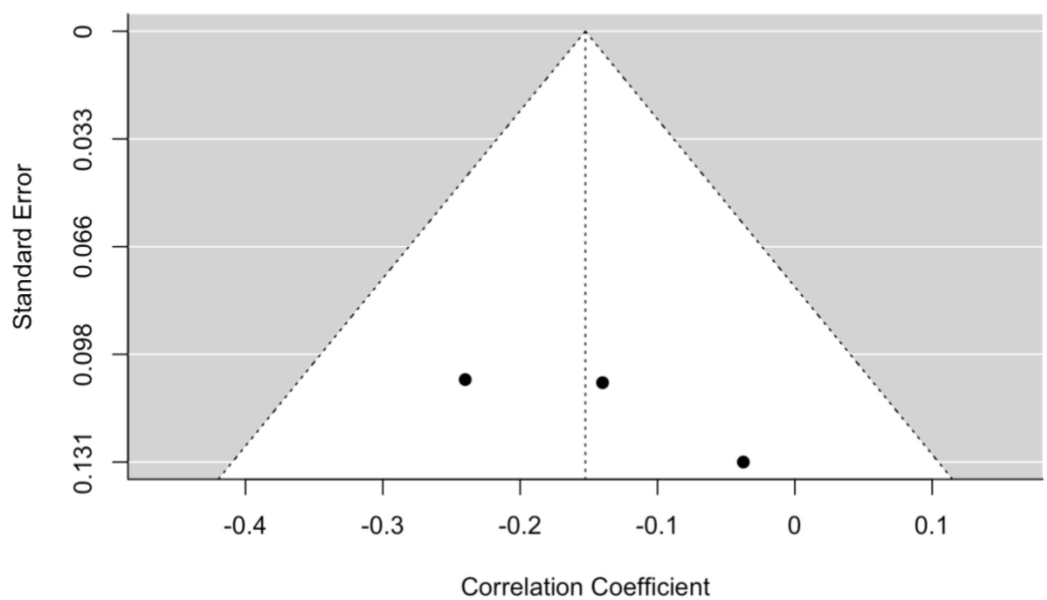


Figure 32 Funnel Plot for Physical Fitness Model 1

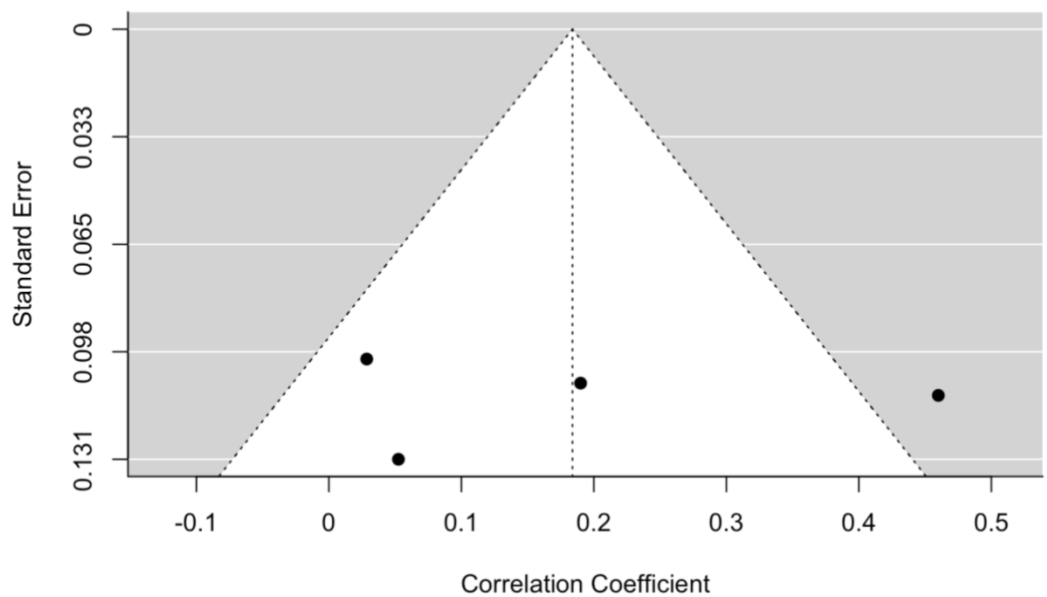



Figure 33 Funnel Plot for Externalizing Behavior Problems Model 1

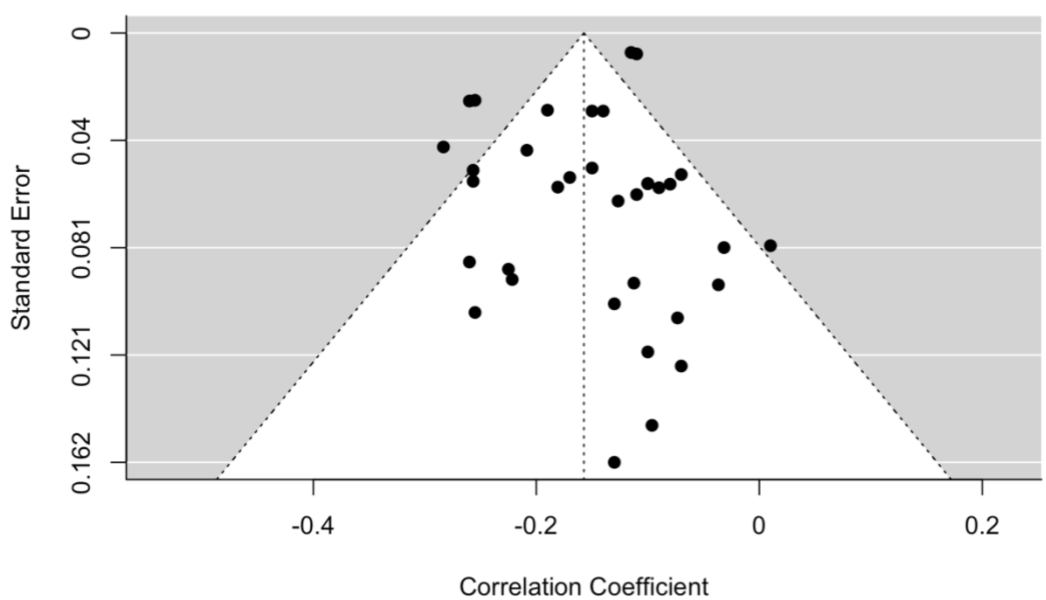


Figure 34 Funnel Plot for Externalizing Behavior Problems Model 2

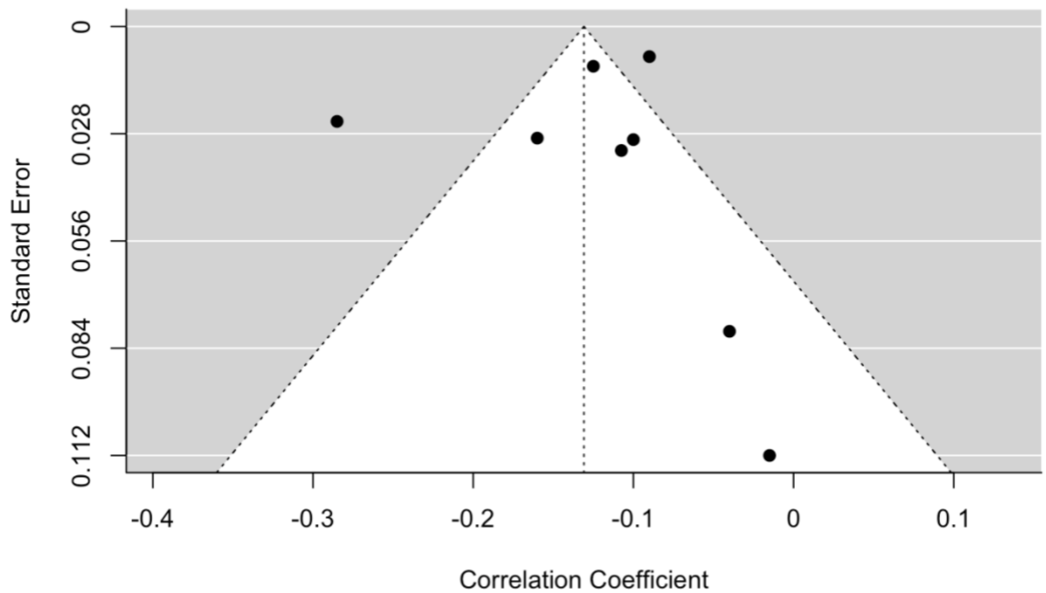



Figure 35 Funnel Plot for Internalizing Behavior Problems Model 1

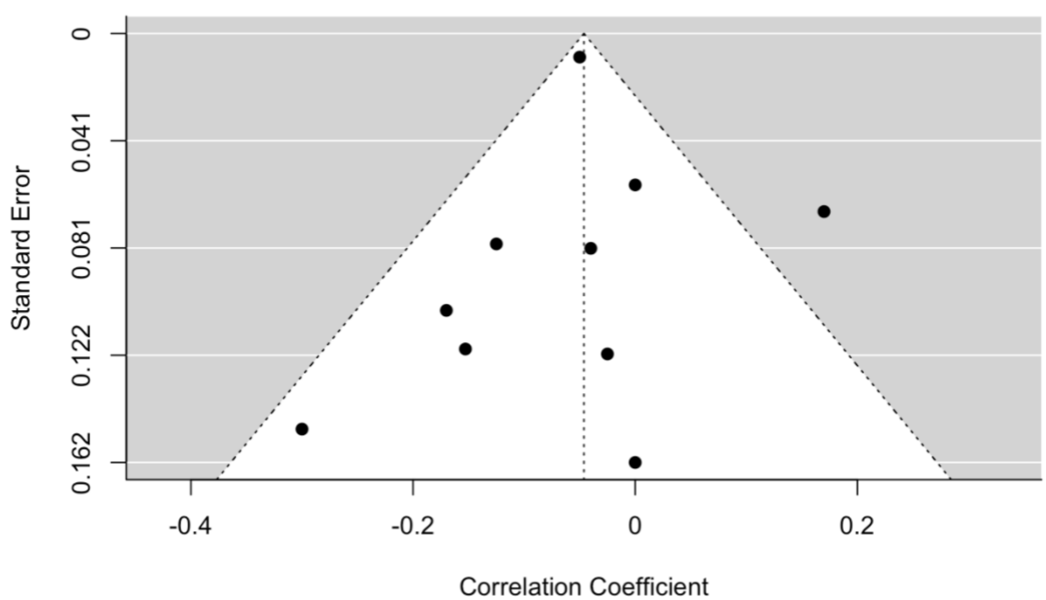


Figure 36 Funnel Plot for Internalizing Behavior Problems Model 3

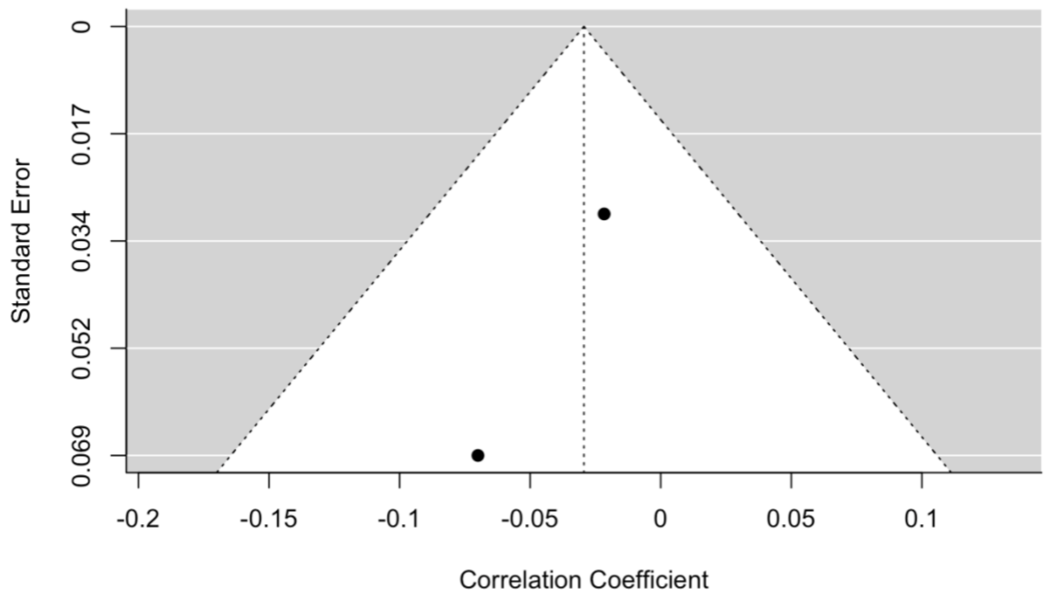



Figure 37 Funnel Plot for Lie Understanding Model 1

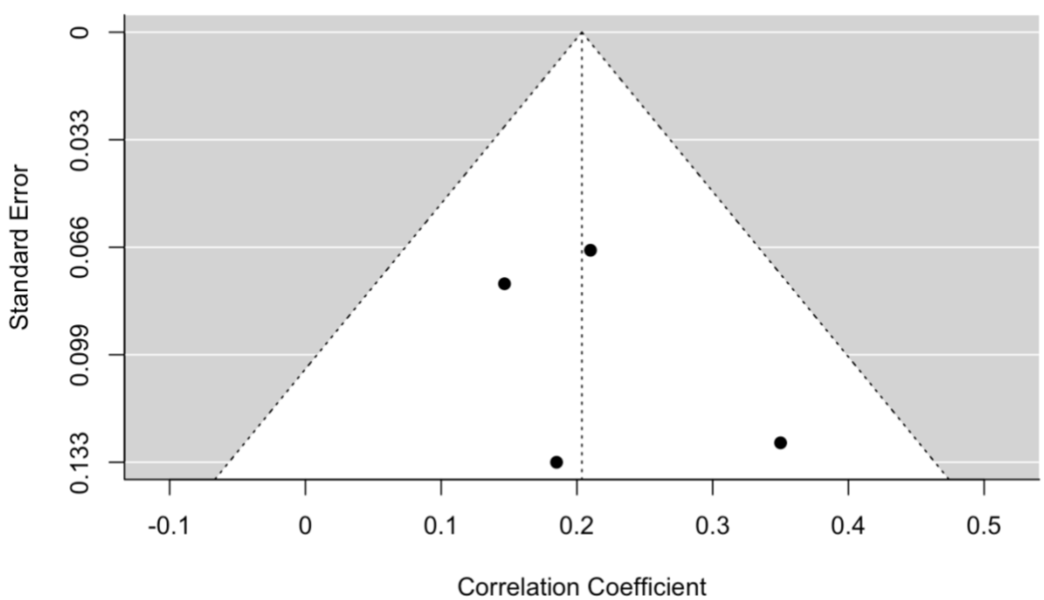


Figure 38 Funnel Plot for Attention & Hyperactivity Symptoms Model 1

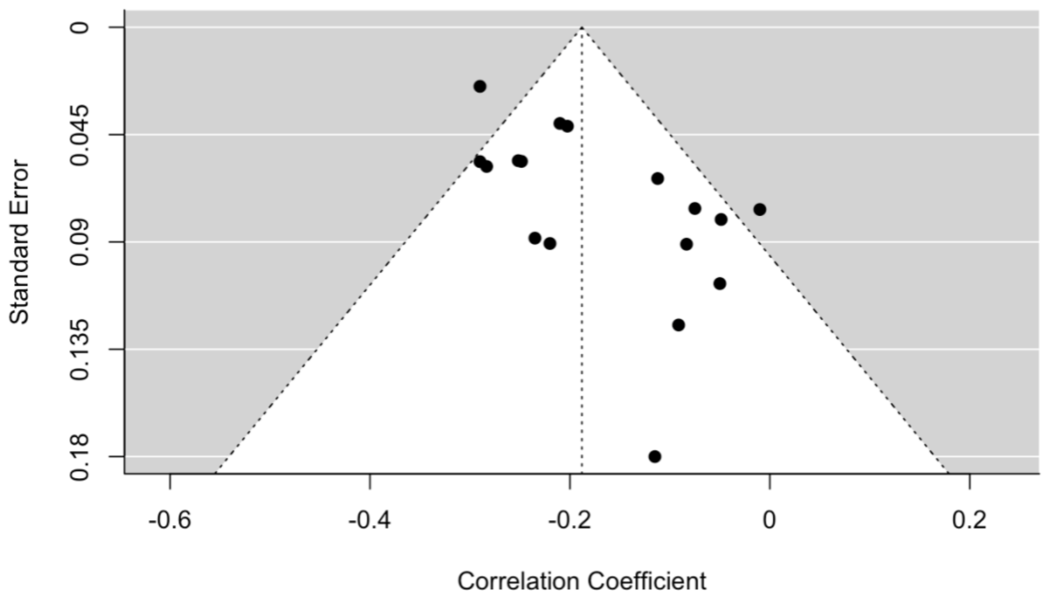



Figure 39 Funnel Plot for Attention & Hyperactivity Symptoms Model 2

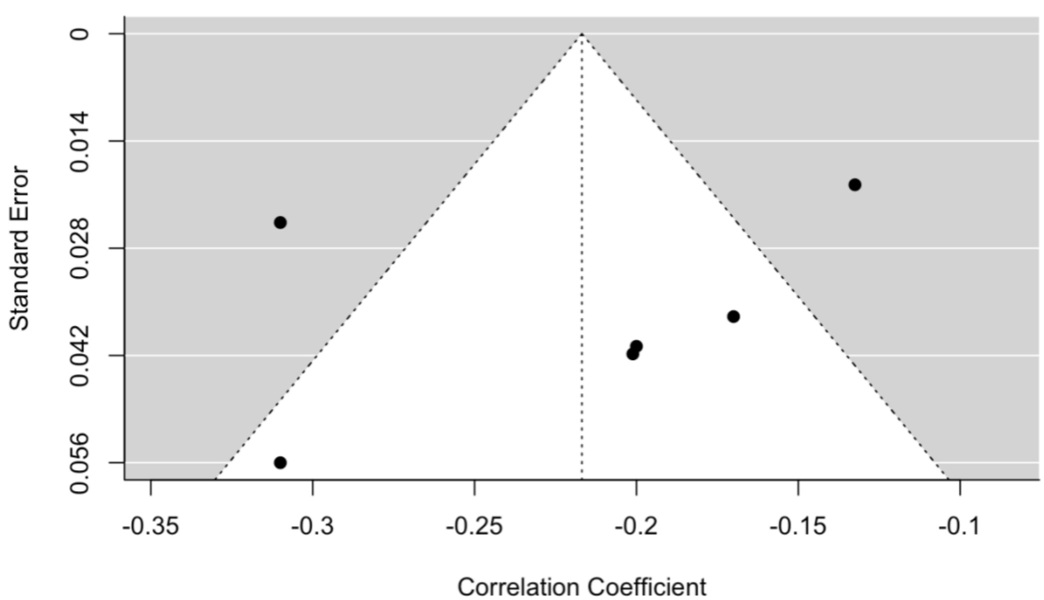


Figure 40 Funnel Plot for Adaptive Classroom Behavior Model 1

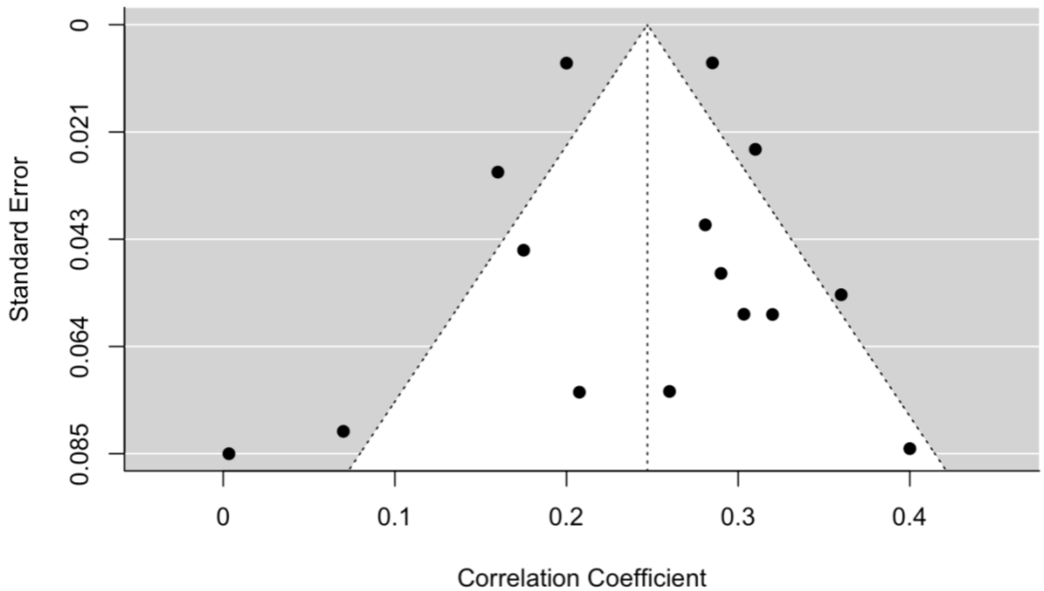
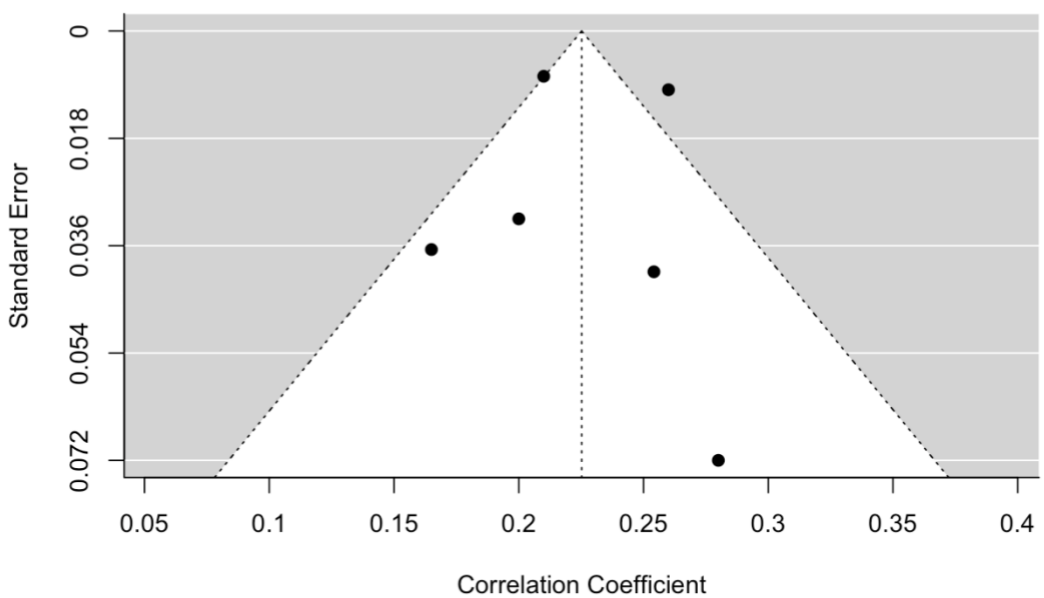



Figure 41 Funnel Plot for Adaptive Classroom Behavior Model 2



# REFERENCES

Allan, N. P., Hume, L. E., Allan, D. M., Farrington, A. L., & Lonigan, C. J. (2014). Relations between inhibitory control and the development of academic skills in preschool and kindergarten: A meta-analysis. *Developmental Psychology*, *50*(10), 2368–2379. https://doi.org/10.1037/a0037493

Alloway, T. P., & Passolunghi, M. C. (2011). The relationship between working memory, IQ, and mathematical skills in children. *Learning and Individual Differences*, *21*(1), 133–137. https://doi.org/10.1016/j.lindif.2010.09.013

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

Beck, D. M., Eales, L., & Carlson, S. M. (2020). Hot and cool executive function and body mass index in young children. *Cognitive Development*, *54*, 100883. https://doi.org/10.1016/j.cogdev.2020.100883

Bierman, K. L., & Torres, M. (2016). Promoting the development of executive functions through early education and prevention programs. In J. A. Griffin, P. McCardle, & L. S. Freund (Eds.), *Executive function in preschool-age children: Integrating measurement, neurodevelopment, and translational research.* (pp. 299–326). American Psychological Association. https://doi.org/10.1037/14797-014

Blair, C., Kuzawa, C. W., & Willoughby, M. T. (2020). The development of executive function in early childhood is inversely related to change in body mass index:

Evidence for an energetic tradeoff? *Developmental Science*, *23*(1). https://doi.org/10.1111/desc.12860

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*(2), 647–663. https://doi.org/10.1111/j.1467-8624.2007.01019.x

Blankenship, T. L., Slough, M. A., Calkins, S. D., Deater-Deckard, K., Kim-Spoon, J., & Bell, M. A. (2019). Attention and executive functioning in infancy: Links to childhood executive function and reading achievement. *Developmental Science*, *22*(6). https://doi.org/10.1111/desc.12824

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons.

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

Cartwright, K. B., Bock, A. M., Clause, J. H., Coppage August, E. A., Saunders, H. G., & Schmidt, K. J. (2020). Near- and far-transfer effects of an executive function intervention for 2nd to 5th-grade struggling readers. *Cognitive Development*, *56*, 100932. https://doi.org/10.1016/j.cogdev.2020.100932

Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics*, *10*(1), 101–129.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155.

Deer, L. K., Hastings, P. D., & Hostinar, C. E. (2020). The role of childhood executive function in explaining income disparities in long-term academic achievement. *Child Development*, *91*(5), 1046–1063.

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

Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do." *Developmental Psychobiology*, *29*(4), 315–334.

Dilworth-Bart, J., Khurshid, A., & Vandell, D. (2007). Do maternal stress and home environment mediate the relation between early income-to-need and 54-months attentional abilities? *INFANT AND CHILD DEVELOPMENT*, *16*(5), 525–552. https://doi.org/10.1002/icd.528

Doebel, S. (2020). Rethinking executive function and its development. *Perspectives on Psychological Science, 15*(4), 942–956. https://doi.org/10.1177/1745691620904771

Durlak, J. A. (2009). How to Select, Calculate, and Interpret Effect Sizes. *Journal of Pediatric Psychology*, *34*(9), 917–928. https://doi.org/10.1093/jpepsy/jsp004

Eisenberg, I. W., Bissett, P. G., Zeynep Enkavi, A., Li, J., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Uncovering the structure of self-regulation through data-driven ontology discovery. *Nature Communications*, *10*(1), 2319. https://doi.org/10.1038/s41467-019-10301-1

Espinet, S. D., Anderson, J. E., & Zelazo, P. D. (2013). Reflection training improves executive function in preschool-age children: Behavioral and neural effects. *Developmental Cognitive Neuroscience*, *4*, 3–15. https://doi.org/10.1016/j.dcn.2012.11.009

Evans, A. D., Xu, F., & Lee, K. (2011). When all signs point to you: Lies told in the face of evidence. *Developmental Psychology*, *47*(1), 39–49. https://doi.org/10.1037/a0020787

Ferguson, C. J., & Brannick, M. T. (2012). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods*, *17*(1), 120–128. https://doi.org/10.1037/a0024445

Friedman, N. P., Hatoum, A. S., Gustavson, D. E., Corley, R. P., Hewitt, J. K., & Young, S. E. (2020). Executive functions and impulsivity are genetically distinct and independently predict psychopathology: Results from two adult twin studies. *Clinical Psychological Science*, *8*(3), 519–538. https://doi.org/10.1177/2167702619898814

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*(2), 201–225. https://doi.org/10.1037/0096-3445.137.2.201

Gardner, F. (2000). Methodological issues in the direct observation of parent–child interaction: Do observational findings reflect the natural behavior of participants? *Clinical Child and Family Psychology Review*, *3*(3), 185–198.

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 312–7 years old on a stroop-like day-night test. *Cognition*, *53*(2), 129–153. https://doi.org/10.1016/0010-0277(94)90068-X

Hackman, D. A., Betancourt, L. M., Gallop, R., Romer, D., Brodsky, N. L., Hurt, H., & Farah, M. J. (2014). Mapping the trajectory of socioeconomic disparity in working memory: Parental and neighborhood factors. *Child Development*, *85*(4), 1433–1445. https://doi.org/10.1111/cdev.12242

Harrist, A. W., Topham, G. L., Hubbs-Tait, L., Page, M. C., Kennedy, T. S., & Shriver, L. H. (2012). What Developmental Science Can Contribute to a Transdisciplinary Understanding of Childhood Obesity: An Interpersonal and Intrapersonal Risk Model. *Child Development Perspectives*, *6*(4), 445–455. https://doi.org/10.1111/cdep.12004

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*, 1539–1558. https://doi.org/10.1002/sim.1186

Hubert, B., Guimard, P., & Florin, A. (2017). Cognitive self-regulation and social functioning among French children: A longitudinal study from kindergarten to first grade: Self-regulation and social functioning. *PsyCh Journal*, *6*(1), 57–75. https://doi.org/10.1002/pchj.160

Huedo-Medina, T. B., Sanchez-Meca, J., Marın-Martınez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychological Methods*, *11*(2), 193–206. https://doi.org/10.1037/1082-989X.11.2.193

Hughes, S. O., Power, T. G., O'Connor, T. M., & Fisher, J. O. (2015). Executive functioning, emotion regulation, eating self-regulation, and weight status in low-income preschool children: How do they relate? *Appetite*, *89*, 1–9. https://doi.org/10.1016/j.appet.2015.01.009

Karalunas, S. L., Bierman, K. L., & Huang-Pollock, C. L. (2020). Test–Retest Reliability and Measurement Invariance of Executive Function Tasks in Young Children With and Without ADHD. *Journal of Attention Disorders*, *24*(13), 1891–1904. https://doi.org/10.1177/1087054715627488

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

Kassai, R., Futo, J., Demetrovics, Z., & Takacs, Z. K. (2019). A meta-analysis of the experimental evidence on the near- and far-transfer effects among children's executive function skills. *Psychological Bulletin*, *145*(2), 165–188. https://doi.org/10.1037/bul0000180

Keye, S., Walk, A., Cannavale, C., Iwinski, S., McLoughlin, G., Steinberg, L., & Khan, N. (2021). Six-Minute Walking Test Performance Relates to Neurocognitive

Abilities in Preschoolers. *Journal of Clinical Medicine*, *10*(4). https://doi.org/10.3390/jcm10040584

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development*, *67*, 490–507.

Levitan, R. D., Rivera, J., Silveira, P. P., Steiner, M., Gaudreau, H., Hamilton, J., Kennedy, J. L., Davis, C., Dube, L., Fellows, L., Wazana, A., Matthews, S., & Meaney, M. J. (2015). Gender differences in the association between stop-signal reaction times, body mass indices and/or spontaneous food intake in pre-school children: An early model of compromised inhibitory control and obesity. *International Journal of Obesity*, *39*(4), 614–619. https://doi.org/10.1038/ijo.2014.207

Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-analysis*. SAGE publications, Inc.

Luria, A. R. (1959). The directive function of speech in development and dissolution. *Word*, *15*(2), 453–464. https://doi.org/10.1080/00437956.1959.11659709

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

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., Sears, M. R., Thomson, W. M., & 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. https://doi.org/10.1073/pnas.1010076108

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, *151*, 264–269. https://doi.org/10.1136/bmj.b2535

Najman, J. M., Williams, G. M., Nikles, J., Spence, S., O’Callaghan, M., Brocque, R. L., Andersen, M. J., & Shuttlewood, J. (2001). Bias influencing maternal reports of child behaviour and emotional state. *Social Psychiatry and Psychiatric Epidemiology*, 12.

Obradović, J., & Willoughby, M. T. (2019). Studying Executive Function Skills in Young Children in Low- and Middle-Income Countries: Progress and Directions. *Child Development Perspectives*, *13*(4), 227–234. https://doi.org/10.1111/cdep.12349

Paap, K. R., & Sawi, O. (2016). The role of test-retest reliability in measuring individual and group differences in executive functioning. *Journal of Neuroscience Methods*, *274*, 81–93. https://doi.org/10.1016/j.jneumeth.2016.10.002

Pieper, J. R., & Laugero, K. D. (2013). Preschool children with lower executive function may be more vulnerable to emotional-based eating in the absence of hunger. *Appetite*, *62*, 103–109. https://doi.org/10.1016/j.appet.2012.11.020

Pigott, T. D. (2019). *Missing data in meta-analysis* (3rd ed.). Russell Sage Foundation.

Polanin, J. R., Hennessy, E. A., & Tsuji, S. (2020). Transparency and reproducibility of meta-analyses in psychology: A meta-review. *Perspectives on Psychological Science, 15*(4), 1026–1041. https://doi.org/10.1177/1745691620906416

Polanin, J. R., Tanner-Smith, E. E., & Hennessy, E. A. (2016). Estimating the Difference Between Published and Unpublished Effect Sizes: A Meta-Review. *Review of Educational Research*, *86*(1), 207–236. https://doi.org/10.3102/0034654315582067

R Core Team. (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Robson, D. A., Allen, M. S., & Howard, S. J. (2020). Self-regulation in childhood as a predictor of future outcomes: A meta-analytic review. *Psychological Bulletin*, *146*(4), 324–354. https://doi.org/10.1037/bul0000227

Rosenthal, R. (1979). The “File Drawer Problem” and Tolerance for Null Results. *Psychological Bulletin*, *86*(3), 638–641.

Saunders, B., Milyavskaya, M., Etz, A., Randles, D., & Inzlicht, M. (2018). Reported Self-control is not Meaningfully Associated with Inhibition-related Executive Function: A Bayesian Analysis. *Collabra: Psychology*, *4*(1), 39. https://doi.org/10.1525/collabra.134

Schmitt, S. A., Korucu, I., Jones, B. L., Snyder, F. J., Evich, C. D., & Purpura, D. J. (2019). Self-regulation as a correlate of weight status in preschool children. *Early Child Development and Care*, *189*(1), 68–78. https://doi.org/10.1080/03004430.2017.1299715

Shoda, Y., Mischel, W., & Peake, P. K. (1990). Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: Identifying diagnostic conditions. *Developmental Psychology*, *26*(6), 978–986.

Smithers, L. G., Sawyer, A. C. P., Chittleborough, C. R., Davies, N. M., Davey Smith, G., & Lynch, J. W. (2018). A systematic review and meta-analysis of effects of early life non-cognitive skills on academic, psychosocial, cognitive and health outcomes. *Nature Human Behaviour*, *2*(11), 867–880. https://doi.org/10.1038/s41562-018-0461-x

Tan, E., Mikami, A. Y., Luzhanska, A., & Hamlin, J. K. (2020). The Homogeneity and Heterogeneity of Moral Functioning in Preschool. *Child Development*, *92*(3), 959–975. https://doi.org/10.1111/cdev.13458

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.

Wallace, B. C., Small, K., Brodley, C. E., Lau, J., & Trikalinos, T. A. (2012). Deploying an interactive machine learning system in an evidence-based practice center: Abstrackr. *Proceedings of the ACM International Health Informatics Symposium*, 819–824.

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

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

Zelazo, P. D., Blair, C. B., & Willoughby, M. T. (2016). *Executive function: Implications for education* (NCER 2017–2000; p. 148). U.S. Department of Education, Institute of Education Sciences, National Center for Education Research.

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, *6*(4), 354–360. https://doi.org/10.1111/j.1750-8606.2012.00246.x

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology*, *1*(2), 198–226. https://doi.org/10.1037/1089-2680.1.2.198

Zelazo, P. D., & Cunningham, W. A. (2007). Executive Function: Mechanisms Underlying Emotion Regulation. In *In J. J. Gross (Ed.), Handbook of emotion regulation* (pp. 135–158). The Guilford Press.

# BIOGRAPHY

Nicole J. Stucke grew up in Cedarburg, Wisconsin. As the oldest of four children, she was often the “helper” to her mother, Sheila, who ran an in-home daycare center. Nicole attributes much of her curiosity about cognitive developmental science to her upbringing, as she was almost always surrounded by children of all ages. She often wondered why some children struggle in areas others flourish (e.g., academics, social functioning, empathy, self-control), and her research is centered on further understanding what can be done to promote adaptive functioning in these areas. Nicole graduated from Cedarburg High School in 2012. She received her Bachelor of Science in Child Psychology and Neuroscience from the University of Minnesota in 2015. Nicole received her Master of Arts in Applied Developmental Psychology in 2022, and currently, she is pursuing a Ph.D. in Applied Developmental Psychology at George Mason University.